Spectrum of Engineering Sciences
ISSN (e) 3007-3138 (p) 3007-312X
Volume 3, Issue 12, 2025

U-NET++ BASED BREAST CANCER DETECTION

Hamza Shahid Gill[*1], Sohail Masood[2], Fawad Nasim[3], Muhammad Usman[4]

[*1,2,3,4]*Faculty of Computer Science and Information Technology, The Superior University, Lahore*

[*1]hamzashahidgill@gmail.com, [2]sohailmasood@superior.edu.pk, [3]fawad.nasim@superior.edu.pk, [4]mu1783495@gmail.com

DOI: https://doi.org/10.5281/zenodo.18051608

Keywords
Breast tumor, Ultrasound Images, Deep Learning, U-Net++

Article History
Received: 11 October 2025
Accepted: 21 November 2025
Published: 25 December 2025

Corresponding Author: *
Hamza Shahid Gill

Abstract
This research investigates the application of the U-Net++ deep learning architecture for the automated segmentation of breast tumors in ultrasound images. Accurate tumor segmentation is crucial for the effective diagnosis and treatment planning of breast cancer. The methodology encompasses a detailed preprocessing pipeline, U-Net++ model implementation, and rigorous evaluation. Initially, ultrasound images and corresponding masks were resized to 196x196 pixels to standardize input dimensions. Class distribution analysis revealed an imbalance. The U-Net++ architecture, known for its nested skip connections, was employed for semantic segmentation. The encoder extracts features through convolutional blocks with ELU activation, SpatialDropout2D, and batch normalization. The decoder reconstructs the segmentation using transposed convolutions. Model training utilized the AdamW optimizer, Dice loss function, and evaluation metrics, including IoU and accuracy. Callbacks, including ModelCheckpoint and EarlyStopping, were implemented to optimize training and prevent overfitting. The model achieved the following results: Training Dataset (IoU: 0.5383, Accuracy: 0.8880, Loss: 0.5749) and Test Dataset (IoU: 0.5667, Accuracy: 0.9124, Loss: 0.5963). While the model demonstrates promise, the moderate IoU scores suggest that further refinement is necessary for enhanced segmentation precision and clinical applicability.

INTRODUCTION

A tumor happens when a breast contains too many out-of-control cells is known as breast cancer. If not treated, it spreads all over the body, thereby threatening to take life. The cancer often starts in the milk ducts or lobules, and in its most early stage, wherein it is considered in situ, the disease is not fatal and can be cured right away. However, once the cancerous cells invade the surrounding breast tissue, tumors may develop that could result in lumps or thickening in the breast. The invasive kind of breast cancer can further metastasize to lymph nodes or other organs, which might be fatal. In 2022, according to a WHO report, an estimated 2.3 million worldwide cases of breast cancer resulted in 670,000 deaths[1]. No country is spared; it may develop at any age after puberty, with higher incidence rates in older groups.

Treatments differ with the individual, the type of cancer, and how far the cancer has spread. Usually, a patient will be treated with surgery, radiation therapy, and medications. The early detection of tumors with accuracy is the key to better outcomes, yet current screening methods continue to face challenging situations. Mammography stands out as the standard; however, some limitations exist, including inter-reader variability among radiologists and possible suboptimal detection of subtle lesions. Deep learning can hence offer automatic image analysis that might enhance outcomes in breast cancer detection by improving the segmentation of mammograms. This research focuses on deep learning potential, especially on the U-Net++ architecture, when analyzing mammographic images. Several deep learning models have been proposed so far for performing image segmentation in medical diagnosis: FCNs, U-Net, and U-Net++. This research describes the so-called application of breast cancer detection. Of them, the most interesting one is called U-Net++, because some features like inception modules and deep supervision will most likely yield an improvement in the segmentation results[2].

Diagnosis in breast cancer typically relies largely on conventional imaging methods, including mammography, ultrasound techniques, and MRI, along with confirmation by biopsy. Nevertheless, these have been crucial for early diagnosis and treatment and thus improved prognosis. However, they are rather time-consuming, and expensive, and may sometimes lead to false-positive or false-negative results, which raises the demand for novel and more precise diagnostic tools.

U-Net, U-Net++, and U-Net3++ are considered some of the most popular deep-learning models used in segmenting breast images. These models utilized a convolutional neural network architecture to learn representations from both the spatial and contextual information of the input images. The overall architecture consists of an encoder network that extracts features from the input image and a decoder network that reconstructs a segmented image from these features. It was originally proposed by Ronneberger et al. in the year 2015. Since then, U-Net models have been one of the broadly used models in segmentation applications for most medical images, including breast imaging. It uses skip connections between the encoder and decoder to maintain high-resolution information during up-sampling[3].

U-Net++ and U-Net3++ extend the original U-Net to enhance segmentation accuracy and efficiency. U-Net++ introduces nested skip connections, capturing multi-scale information. Attention gates are added in U-Net3++, which emphasizes the relevant features in the input image. Although quite effective, several challenges remain concerning the segmentation of breast images, including noise and artifact handling. The following article documents an in-depth analysis of the segmentation study of the images of the breast using U-Net, U-Net++, and U-Net3++. It compares the performance based on the dataset of histopathological images of breast cancer. It deals with the capability of such models in terms of accuracy and efficiency of deep models, interpretability, and robustness against noise. The impact of various post-processing methods is studied on segmentations. These results provide new insights into the development of more accurate and more efficient segmentation methods for automatic breast image diagnostics and treatment[4].

There are several challenges to be addressed with these techniques before they can realize their full potential. The generalization power, majorly of deep learning models to varied diversities of mammograms in routine clinical practices, is one major limitation. Other significant factors, which are of essential concern, involve computational efficiency and handling imbalanced datasets. Problems of this sort do call for future studies to focus on generalizability studies, augmentation data methods, and the integration of AI techniques. If these challenges can be overcome, deep learning can eventually be done to revolutionize breast tumor detection and offer earlier diagnosis with implications for better patient outcomes.

The rest of the research is presented as follows: The literature review has been presented in the section "literature review", the methodology of the whole paper will be in the section "methodology", the proposed attention-driven u-net++ model in the section "proposed U-Net++ model", findings and discussion in section "results and discussion", and the conclusion in the section "conclusion and future work".

LITERATURE REVIEW:

Artificial intelligence, expert systems, and convolutional neural networks have been employed in the diagnosis of breast cancer to refine medical image segmentation with enhanced accuracy[5]. Various models and techniques are engineered to help improve the effectiveness of segmentation in cases of breast cancer: U-Net, SegNet, PSPNet, and Attention U-Net, among others. While deep learning is developing continuously, different models have been constructed for automatic breast cancer segmentation[6]. Several segmentation networks were designed by researchers to deal with various related challenges in ultrasound imaging of the breasts. For the precise segmentation of small tumors, Pramanik et al. proposed the small tumor perception network, which embedded multi-scale convolutional blocks for incorporating context information regarding breast tumors with high-resolution feature information for improving segmentation accuracy[7]. They also suggested a boundary-regularized deep convolutional encoder-decoder network for complete breast ultrasound image segmentation. They talked about a deep CNN including a global guide block and a boundary detection module, enhancing breast lesion segmentation. And presented the boundary rendering network, which leverages a differentiable boundary selection module coupled with a GCN-based rendering; accurate detection of the boundary is still an open challenge when it is located in areas characterized by complex scenes/objects or low light conditions.

Drioua et al., refined the segmentation further by replacing the convolutional blocks with residual blocks in AttentionU-Net. They further extended this to add more layers for feature extraction over varying receptive fields and proposed a multistage segmentation approach that combined the segmentation with the classification of ultrasound images, using RFS-UNet for images classified as abnormal[6]. In this work, they summarize the common limitations of existing U-shaped networks for a new breast tumor segmentation model, proposing the reconstruction of a convolution block and enhancing the skip connection for better accuracy[8].

The most frequent type of cancer in females is breast cancer, and early diagnosis increases the possibility of a cure. Xia et al. have explained that early-stage breast tumors are small with indistinct edges; this usually creates false or missed detections. They present a Shape Enhanced U-Net with a Transformer encoder layer for automatic breast tumor segmentation. The Transformer encoder layer introduces a global self-attention mechanism that enhances the accuracy of tumor segmentation. This model also incorporates a new shape-enhanced branch, which learns from the boundary of the tumor with additional supervision to help the algorithm focus on relevant edges and raise the bar for small tumor segmentations[9].

Michael et al. has given the only way to improve survival rates of patients with breast cancer worldwide is through early detection, it has an added advantage of reducing treatment costs when one is diagnosed at an early stage. According to them, mammograms are one of the most common detection tools used for early detection, and the effectiveness of mammograms in detecting tumors is highly dependent on the results of segmentation techniques[10]. Some of the important aspects of image analysis include detection, feature extraction, classification, and treatment planning, with segmentation playing a key role. It further aids a physician in quantifying the volume of tissue in the breast for treatment planning[11]. They have classified the segmentation methodologies into three categories: classical segmentation, which also emphasizes region-, threshold-, and edge-based techniques; machine learning segmentation; and deep learning segmentation, both supervised and unsupervised. Their results showed that region-based segmentation, especially region growing, is commonly used in the class of classical methods, and that the MIAS database is a frequently used resource as well. Median filtering has also proved to be useful in noise removal[12].

Chen et al. have highlighted that segmentation in the domain of machine learning is more into unsupervised methods. Based on this context, for mammogram image segmentation, U-Net has become well known as it requires less number of annotated images in comparison with other deep learning models. Their study further suggests that deep learning models, such as U-Net, can be effectively trained without extensive preprocessing or postprocessing through ongoing advances in high-performance GPU computing by enabling the training of deeper networks[13]. These multi-scale adversarial networks and enhanced U-Nets demonstrated excellent performance in capturing multi-scale contextual information and, thus, enabled much improved segmentation of breast masses. Variants of U-Net++ further improve the quality of breast cancer segmentation with its refined feature extraction techniques, multiscale approach, and attention mechanisms. These advances will, therefore, remain very useful in medical imaging for the diagnosis and treatment planning of breast cancer, translating into higher accuracy and robustness across diverse datasets[14].

Samudrala and Mohan in their work represents a semantic segmentation methodology for early detection of cancer tumors, which overcomes the drawbacks related to the classification accuracy and decision-making capability of previous methods[15]. This approach introduces hybrid semantic segmentation networks, initiated with pre-processing the images through Adaptive Local Gamma Correction to enhance the contrast. Segmentation was drawn by a hybrid network that combined DenseNet-121 with an Attention-based pyramid scene parsing network, Att-PSPnet, for feature extraction and scene parsing. It also involves the Attention Gate mechanism, which amplifies high-dimensional features by highlighting key information and suppressing useless information with noise[16]. A pyramid dilated convolution module-PDM-widens the reception field for a better capture of global features, with a global average pooling layer-GAP-refining the output. The proposed method was tested on the histologically confirmed dataset in the Google Colab environment and outperformed existing approaches, FCN, U-Net, and PSPNet, thereby obtaining a prediction accuracy of 94.68%[17].

Methodology

This research employs a U-Net++ architecture for the semantic segmentation of breast tumors in ultrasound images. The methodology encompasses data preprocessing, model architecture design, and implementation details, which are delineated in the subsequent subsections.

1. Data Preprocessing

The dataset, sourced from [*You should cite your dataset here*], comprises ultrasound images and corresponding ground truth masks annotated for breast tumor regions. The dataset exhibits variability in image dimensions, necessitating a uniform input size for the neural network.

1.1 Image Resizing

All images and their corresponding masks were resized to 196x196 pixels to ensure homogeneity and computational efficiency. This resizing operation, while essential for batch processing and model compatibility, was carefully considered to mitigate potential distortions. The rationale for this specific resolution (196x196) was predicated on a balance between maintaining sufficient anatomical detail for accurate segmentation and constraining computational demands.

1.2 Class Distribution Analysis

Before model training, an analysis of the class distribution within the dataset was conducted. The number of images and masks for each category is as follows:

- 'Normal': 133 images, 133 masks
- 'Benign': 437 images, 454 masks
- 'Malignant': 210 images, 211 masks

This class distribution highlights a degree of imbalance, particularly with the 'benign' category having a higher representation. As we see benign and malignant have more masks as compared to images. Because some breasts have one or more tumor in a breast. Therefore, we combined the tumor into a single image. When two tumors are added, the range can increase by 0-2. So we will reduce it again to the range 0-1, and data is converted into dtype='float32'.

1.3 Data Partitioning

The dataset was partitioned into training and testing sets to rigorously evaluate the model's performance and generalization capability. The partition ratio is 85:15. For this study, we exploited the bigger size of data for the training purposes, and a small portion of data was used for the test set with a 15% ratio.

2. Model Architecture

The core of this methodology is the U-Net++ architecture, a deep convolutional neural network renowned for its efficacy in medical image segmentation. U-Net++ builds upon the original U-Net by introducing nested and dense skip pathways. These architectural modifications facilitate feature aggregation at multiple semantic scales, enhancing the model's ability to capture both low-level spatial details and high-level contextual information, crucial for precise tumor boundary delineation.

2.1 Encoder Pathway

The encoder pathway, responsible for feature extraction, comprises five encoding blocks. Each block consists of a series of convolutional layers, Exponential Linear Unit (ELU) activation functions, SpatialDropout2D layers, and batch normalization layers. Specifically, each encoder block is structured as follows:

Two consecutive 2D convolutional layers with a kernel size of 2 and 'same' padding. The number of filters in the convolutional layers increases progressively through the encoder pathway.

ELU activation function applied after each convolutional layer.

A SpatialDropout2D layer with a dropout rate of 0.06 is incorporated to regularize the model and mitigate overfitting. The rationale for employing SpatialDropout2D, as opposed to conventional Dropout, lies in its capacity to promote feature map independence by dropping entire feature maps rather than individual elements. This is particularly advantageous in image segmentation, where spatial relationships between pixels are paramount.

Batch normalization is applied to stabilize training and accelerate convergence.

An additional 2D convolutional layer with a kernel size of 2, 'same' padding, and a filter count increased by a factor of 1.1 (approximately) is included.

ELU activation function follows.

Max pooling layers, typically used for downsampling in U-Net, are notably absent in this implementation. The downsampling is implicitly achieved through the strided convolutions within the encoder blocks. The filter sizes in the encoder blocks are [64, 128, 256, 336, 512]. In the bottleneck, we used only one CNN block with 786 filters, as mentioned above, without a pooling layer

2.2 Decoder Pathway

The decoder pathway reconstructs the segmented image from the encoded feature maps. It mirrors the encoder in structure but employs transposed convolution layers for upsampling. The decoder consists of five decoding blocks. Each decoding block receives input from the preceding decoder block and the corresponding encoder block via skip connections. The structure of each decoder block is as follows:

- Two consecutive 2D convolutional layers with a kernel size of 2 and 'same' padding.
- ELU activation function applied after each convolutional layer.
- Batch Normalization.

2.3 Skip Connections and Transposed Convolutions

U-Net++ leverages a series of nested, dense skip connections to bridge the encoder and decoder pathways. These connections facilitate the flow of information across different network depths, mitigating the vanishing gradient problem and enabling the decoder to recover fine-grained spatial information.

Transposed convolution layers, rather than conventional upsampling layers, are utilized for increasing the spatial dimensions of the feature maps in the decoder. Transposed convolutions learn the 2D upsampling operation, allowing the network to optimize the upsampling process for the specific task of tumor segmentation. This learned upsampling contributes to more precise localization of tumor boundaries compared to fixed interpolation methods.

2.4 Output Layers

The decoder pathway culminates in three output layers, a design choice stemming from the implementation of deep supervision. Each output layer is a 2D convolutional layer with a 1x1 kernel and a sigmoid activation function to produce a probability map representing the likelihood of each pixel belonging to the tumor class and deep_supervision = False.

3. Model Training

The U-Net++ model was trained to segment breast tumors using the following protocol:

Optimizer: The model's weights were optimized using the AdamW optimizer, a variant of Adam that incorporates weight decay for improved regularization. The learning rate for the AdamW optimizer was set to 0.00006.

Loss Function: The loss function employed was the Dice loss. The Dice loss is a region-based loss function that directly optimizes the Dice Similarity Coefficient

ISSN (e) 3007-3138 (p) 3007-312X Volume 3, Issue 12, 2025

(DSC), a metric commonly used for evaluating segmentation performance. The Dice loss was calculated as:

$$dice\ loss = 1 - \frac{2 * intersection + smooth}{sum(y\_true) + sum(y\_pred) + smooth}$$

where *y_true* represents the ground truth, *y_pred* the model's prediction, and 'smooth' is a small constant (1e-6) added to the numerator and denominator to prevent division by zero.

Metrics: In addition to the loss function, the model's performance was monitored using accuracy and the Mean Intersection over Union (IoU) metric. The IoU was calculated for 2 classes.

Training Procedure: The model was trained using a batch size of 2. The reason to choose the batch size is 2 owing to GPU memory constraints, regularization effect and training stability. The training and testing datasets were prepared using the tensorflow.data.Dataset.from_tensor_slices and batched and prefetched for optimized data loading. The model was trained for 60 epochs.

Callbacks: Several callbacks were employed during training to enhance the training process and prevent overfitting:

Model Checkpoint: This callback saved the model's weights at the end of each epoch if the validation IoU improved, ensuring that the best-performing model (based on validation IoU) was preserved.

Early Stopping: This callback monitored the validation IoU and stopped the training process if the validation IoU did not improve for 10 consecutive epochs, preventing unnecessary training and mitigating overfitting. The restore_best_weights argument ensured that the model's weights were restored to those of the best epoch.

4. Results and discussion

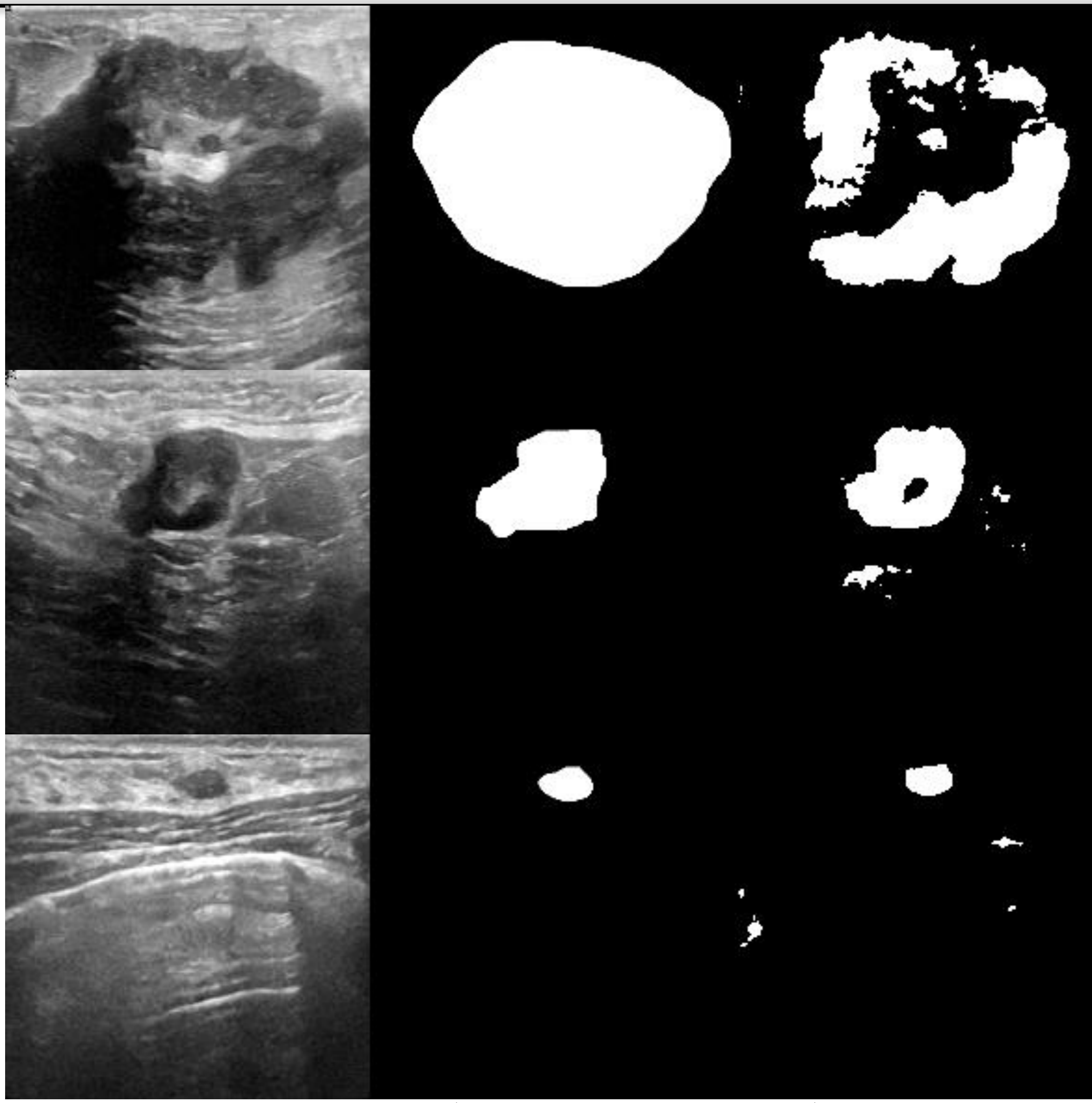
The U-Net++ model demonstrated a reasonable capacity for breast tumor segmentation, with the evaluation scores summarized in Table 1.

Table 1: Evaluation Scores for Training and Test Datasets

Dataset	IoU	Accuracy	Loss
Training	0.5383	0.8880	0.5749
Test	0.5667	0.9124	0.5963

Export to Sheets

The model exhibited slightly improved performance on the test dataset compared to the training dataset, suggesting effective generalization and controlled overfitting. However, the moderate IoU scores indicate that while the model shows promise, further refinement is needed to achieve more precise and clinically reliable segmentation results.



The first image column is images breast, 2nd column is true mask and the 3rd column is predicted tumor.

Conclusion

In conclusion, this research demonstrates the potential of the U-Net++ architecture for the automated segmentation of breast tumors within ultrasound imagery. The model achieved reasonable segmentation performance, indicating its capacity to learn and identify tumor regions. However, the results also highlight the necessity for continued research and development to improve segmentation accuracy and robustness. Future work should focus on addressing class imbalance, exploring advanced data augmentation techniques, and potentially refining the model architecture to achieve clinically viable segmentation precision.

REFERENCE:

[1] "Breast cancer." Accessed: Oct. 22, 2024. [Online]. Available: https://www.who.int/news-room/fact-sheets/detail/breast-cancer

[2] A. Sulaiman *et al.*, "Attention based UNet model for breast cancer segmentation using BUSI dataset," *Sci. Rep.*, vol. 14, no. 1, p. 22422, Sep. 2024, doi: 10.1038/s41598-024-72712-5.

[3] Z. Zhu, Y. Sun, and B. Honarvar Shakibaei Asli, "Early Breast Cancer Detection Using Artificial Intelligence Techniques Based on Advanced Image Processing Tools," *Electronics*, vol. 13, no. 17, Art. no. 17, Jan. 2024, doi: 10.3390/electronics13173575.

[4] M. A. Ortega-Ruíz, C. Karabağ, E. Roman-Rangel, and C. C. Reyes-Aldasoro, "DRD-UNet, a UNet-Like Architecture for Multi-Class Breast Cancer Semantic Segmentation," *IEEE Access*, vol. 12, pp. 40412–40424, 2024, doi: 10.1109/ACCESS.2024.3377428.

[5] A. Lagree *et al.*, "A review and comparison of breast tumor cell nuclei segmentation performances using deep convolutional neural networks," *Sci. Rep.*, vol. 11, no. 1, p. 8025, Apr. 2021, doi: 10.1038/s41598-021-87496-1.

[6] W. R. Drioua, N. Benamrane, and L. Sais, "Breast Cancer Histopathological Images Segmentation Using Deep Learning," *Sensors*, vol. 23, no. 17, p. 7318, Aug. 2023, doi: 10.3390/s23177318.

[7] P. Pramanik, R. Pramanik, F. Schwenker, and R. Sarkar, "DBU-Net: Dual branch U-Net for tumor segmentation in breast ultrasound images," *PLOS ONE*, vol. 18, no. 11, p. e0293615, Nov. 2023, doi: 10.1371/journal.pone.0293615.

[8] C. Qin *et al.*, "Joint Dense Residual and Recurrent Attention Network for DCE-MRI Breast Tumor Segmentation," *Comput. Intell. Neurosci.*, vol. 2022, p. 3470764, 2022, doi: 10.1155/2022/3470764.

[9] Y. Xia, W. Liu, D. Yang, H. Wang, H. Wang, and M. Jiang, "Early stage tumor segmentation in breast MRI using shape enhanced U-Net," *Biomed. Signal Process. Control*, vol. 93, p. 106198, Jul. 2024, doi: 10.1016/j.bspc.2024.106198.

[10] "Segmentation of Breast Cancer on Ultrasound Images using Attention U-Net Model - ProQuest." Accessed: Oct. 22, 2024. [Online]. Available: https://www.proquest.com/openview/e4b97146954cd492cade336033e0b725/1?pq-origsite=gscholar&cbl=5444811

[11] M. R. Islam *et al.*, "Enhancing breast cancer segmentation and classification: An Ensemble Deep Convolutional Neural Network and U-net approach on ultrasound images," *Mach. Learn. Appl.*, vol. 16, p. 100555, Jun. 2024, doi: 10.1016/j.mlwa.2024.100555.

[12] "Breast Cancer Segmentation Methods: Current Status and Future Potentials - Michael - 2021 - BioMed Research International - Wiley Online Library." Accessed: Oct. 22, 2024. [Online]. Available: https://onlinelibrary.wiley.com/doi/10.1155/2021/9962109

[13] J. Chen, L. Chen, S. Wang, and P. Chen, "A Novel Multi-Scale Adversarial Networks for Precise Segmentation of X-Ray Breast Mass," *IEEE Access*, vol. 8, pp. 103772–103781, 2020, doi: 10.1109/ACCESS.2020.2999198.

[14] X.-X. Yin, L. Sun, Y. Fu, R. Lu, and Y. Zhang, "U-Net-Based Medical Image Segmentation," *J. Healthc. Eng.*, vol. 2022, p. 4189781, Apr. 2022, doi: 10.1155/2022/4189781.

[15] S. Samudrala and C. K. Mohan, "Semantic segmentation of breast cancer images using DenseNet with proposed PSPNet," *Multimed. Tools Appl.*, vol. 83, no. 15, pp. 46037–46063, May 2024, doi: 10.1007/s11042-023-17411-5.

[16] E. Ramalakshmi, L. Gunisetti, and L. Sumalatha, "Different U-Net Variants for Segmentation of Histological Breast Images: An Analytical Comparison," Jul. 19, 2024. doi: 10.21203/rs.3.rs-4554519/v1.

[17] Y. Guo, X. Duan, C. Wang, and H. Guo, "Segmentation and recognition of breast ultrasound images based on an expanded U-Net," *PloS One*, vol. 16, no. 6, p. e0253202, 2021, doi: 10.1371/journal.pone.0253202.