A Dual-Approach to Benign and Malignant Tumor Detection: Classification and Segmentation Using Advanced Deep Learning Models

Caitlin Dosch
Santa Clara High School
Santa Clara, United States
email: caitlindosch@gmail.com

Shilpi Shaw
AIClub Research Institute
Mountain View, United States
email: shilpi.shaw@pyxeda.ai

***Abstract*—Breast cancer remains a global health concern, with a 13.1% lifetime diagnosis rate among women. Early and accurate diagnosis plays a critical role in improving patient outcomes. Traditional diagnostic methods, such as MRIs, ultrasounds, CT scans, and mammograms, are widely used for detecting and characterizing breast lesions. In recent years, Artificial Intelligence has shown great promise in enhancing diagnostic accuracy, with models such as K-Nearest Neighbors (KNN), Random Forest Classifier (RFC), and Convolutional Neural Networks (CNN) being applied to breast cancer diagnosis. In this study, we explore the application of deep learning models, specifically MobileNetV2 and ResNet50, for breast cancer detection using ultrasound images from The Cancer Image Archive. A dataset comprising 522 breast lesion images was used, split into training, validation, and test sets. We implemented both image classification and segmentation tasks, optimizing hyperparameters such as learning rate and number of epochs. Our comparative analysis aims to evaluate the efficiency and diagnostic performance of the two models. We highlight key insights into their effectiveness in breast cancer detection and provide recommendations based on their application to ultrasound imaging. The findings of this study contribute to the ongoing efforts to improve AI-based diagnostic tools for breast cancer.**

Keywords-breast lesions; deep learning; image classification

I. INTRODUCTION

Breast cancer remains one of the most prevalent and deadly cancers among women worldwide, accounting for approximately 13.1% of women during their lifetime. Breast cancer is the most frequently diagnosed cancer, constituting 30% of all new cancer diagnoses in women and posing a significant threat to women's health.

It remains a significant health concern in the United States, with an estimated 310,720 new cases of invasive and 56,500 new cases of non-invasive breast cancer anticipated in 2024. Despite over 4 million breast cancer survivors, the disease is expected to cause 42,250 deaths this year. About 1 in 8 women will develop breast cancer in their lifetime, making it the most common cancer among American women, accounting for 30% of all new female cancer diagnoses. The risk factors include but are not limited to family history and younger age at diagnosis, with variations in incidence and outcomes across different racial and ethnic groups.

The segmentation of breast ultrasound images into various tissue types is valuable for tumor localization, measuring breast density, and evaluating treatment responses, which are critical for the clinical diagnosis of breast cancer. Manual segmentation is labor-intensive and relies heavily on the skill and experience of radiologists, making it prone to subjective interpretations and time-consuming due to the need to review numerous images [1].

Outwardly, the presence of breast lesions or lumps, discoloration, and irregularities in breast shape often characterize breast cancer cases. Common clinical signs include irritation, flaking, dimpling, discharge, and swelling of the breast. Early detection and treatment are crucial to minimizing potential complications and improving patient outcomes. Various diagnostic methods are employed to detect and assess breast cancer, including physical examinations, Magnetic Resonance Imaging (MRI), ultrasounds, CT scans, lab tests, and mammograms. Physical examinations aim to determine the location and severity of tumors.

MRI is commonly used to diagnose or measure the size of breast cancer tumors [2]. Ultrasounds can confirm a breast cancer diagnosis, while mammograms are essential for detecting cancers not visible through physical examination. The field of medical diagnostics has increasingly adopted Artificial Intelligence to enhance accuracy and efficiency. AI techniques, such as machine learning and deep learning, have been applied to analyze different types of diagnostic data. Various machine learning algorithms, including KNN, RFC, YOLO, CNN, Support Vector Machines (SVM), and Decision Trees, have been studied, each yielding diverse results.

Despite these advancements, challenges remain in breast lesion classification and segmentation. High variability in lesion appearance, dense breast tissues, and the need for large annotated datasets present significant barriers. To address these challenges, we utilized a new dataset from The Cancer Image Archive in this study. This dataset consists of 522 images from 256 subjects, featuring 266 segmented benign and malignant lesions [3]. We have worked to address these challenges by utilizing a dataset of breast lesion ultrasounds to perform both classification and segmentation tasks. Accurate classification allows clinicians to distinguish between benign and malignant lesions, while precise segmentation aids in the localization and quantification of tumor regions, which are essential for treatment planning and monitoring. For the classification task, we employed MobileNetV2 and ResNet50 models, chosen for their efficiency and accuracy in image analysis. For segmentation, we used the EfficientNetB2 model due to its ability to capture intricate details in medical imaging [3]. Our dual approach for the identification of breast lesions

is optimal, helping to improve accuracy and save time.

By comparing the performance of these models, we aim to identify the most applicable and accurate approaches for each task, thereby contributing to improved diagnostic methods for breast cancer.

Section II discusses related work and methods. Section III lays out the dataset selection. Section IV provides methodology. Section V discusses experiment results. Section VI covers discussion and evaluation. In Section VII, conclusions and future work are drawn.

II. Related work

Current research has focused significantly on improving breast cancer diagnosis using mammograms. The process usually involves three steps: initial screening, segmenting the images, and diagnosing the case.

One study used the Breast Cancer Dataset from the University of California, Irvine (UCI), which included 669 clinical cases. This dataset had 11 attributes, but they used nine key features including clump thickness, cell size, and shape to determine whether a tumor was benign or malignant. They tested two machine learning algorithms, Naïve Bayesian Classifier (NBC) and KNN, using K-fold cross-validation to check their results [4]. The KNN algorithm performed the best, achieving an accuracy of 97.51%. This process typically occurs through screening, segmentation, and diagnosis of a case.

Another study focused on detecting breast cancer using mammogram images. They used segmentation techniques alongside Max-Mean and Least-Variance methods to improve the models' performance. This shows that using advanced image processing techniques can help achieve more accurate results, although specific accuracy numbers were not provided [5].

A separate study used a database of CT Scan images from 2 hospitals in Norway, with each containing 100 patients. Images were of left sided breast cancer patients. The study used scores for clinical usability and dosage levels used for treatment for some of their data. For model scoring, they used Dice similarity coefficient and Hausdorff difference [6].

Researchers also focused on using CNN to automatically segment breast ultrasound images into four main tissue types: skin, glandular tissue, tumors, and fatty tissue. They worked with three-dimensional ultrasound images to accomplish this. The performance of their segmentation method was evaluated using various quantitative metrics, such as Accuracy, Precision, and Recall, all of which exceeded 80%. Additionally, they used the Jaccard Similarity Index (JSI) to measure the overlap between the predicted and actual segments, achieving an 85.1% score. This represented an improvement over their previous method, which employed the watershed algorithm and resulted in a JSI score of 74.54%. The findings suggest that their CNN-based approach could effectively support clinical breast cancer diagnosis by providing reliable tissue segmentations from ultrasound images [7].

Recent advancements in breast ultrasound image segmentation have focused on improving region of interest (ROI) extraction to differentiate between malignant and benign abnormalities effectively. One notable approach involves a model built on local pixel information combined with a neural network, comprising two stages: training and testing. During the training stage, the model is trained with batches from both ROI and background regions. In the testing stage, a fixed-size window scans the image to detect the ROI, followed by a distance transform to refine the ROI by eliminating non-ROI areas. This method was tested on a dataset of 250 ultrasound images, achieving a high success rate of 95.4% for breast contour extraction. Such innovations help reduce false positives and enhance the accuracy of breast ultrasound diagnostics, demonstrating a significant improvement over traditional segmentation techniques [1].

In another study, authors propose a Dual CNN for mammogram image processing. Two paths were utilized, with a Locality Preserving Learning (LPL) and a Conditional Graph Learner (CGL). The model (DualCoreNet) achieved a 92.27 DI coefficient [8].

A different technique, Saliency-Guided Morphology-Aware U-Net (SMU-Net) was used for breast cancer detection in ultrasound images. It contains a main network, auxiliary network, and a middle stream [9].

Separately, a study developed a Deep-Learning based method for diagnosis of breast cancer using ultrasound imaging. The automation of image segmentation is important for breast ultrasound images. A database of 221 images was used. This model achieved a dice coefficient of 0.825 [10].

Another study used volumetric heart segmentation for detecting breast cancer from CT scans. It was trained on manual heart segmentations, from a dataset of 5677 breast cancer patients who had undergone radiation therapy at the Dana-Farber/Brigham and Women's Cancer Center from 2008 - 2018 [11].

III. Methods

This study aims to enhance breast cancer diagnosis by leveraging advanced deep-learning methodologies.

We conducted dual experiments on this dataset: one using image classification with MobileNetV2 and ResNet50, and another using image segmentation with U-Net (EfficientNetB2). The object detection models are image-based, designed both to classify and segment images. MobileNetV2 was chosen for its efficiency and lightweight architecture, suitable for deployment on devices with limited computational power. In contrast, ResNet50 was selected for its depth and ability to capture intricate features through residual learning, making it apt for complex classification tasks. For segmentation, U-Net with EfficientNetB2 was employed due to its superior performance in achieving high accuracy in medical image segmentation by effectively capturing both spatial and contextual information. The second experiment was evaluated using Intersection over Union (IoU) scores, which are based upon the overlap of ground truth and predicted values [12].

While the use of a Vision Transformer (ViT) model was considered, it was ultimately decided against. ViTs are compu-

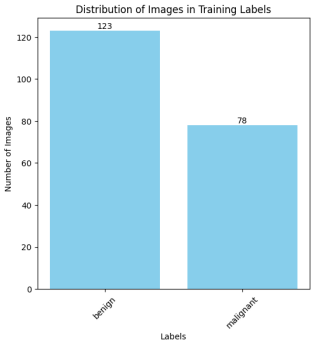

Figure 1. Dataset Distribution

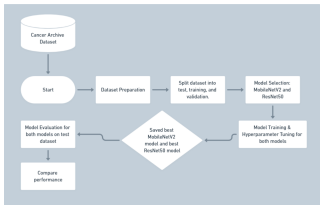

Figure 2. Classification Model Flowchart

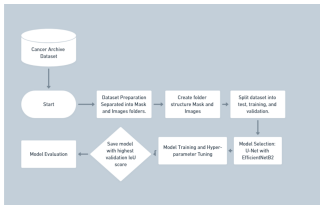

Figure 3. Segmentation Model Flowchart

tationally expensive and require powerful GPUs to train. They process images differently than CNNs, using self-attention, which makes them slower and more resource-intensive. This may not be practical for many real-world medical applications, and thus the decision was made to utilize MobileNetV2 and ResNet50. In comparison to previous studies, our study made use of segmentation techniques on ultrasound imaging, with the application of segmentation in tandem with experimentation on ResNet50 and MobileNetV2.

IV. Dataset

The dataset utilized in this study was sourced from The Cancer Image Archive [3]. It consisted of a total of 522 ultrasound images, with 256 total subjects, and 266 benign and malignant segmented lesions. Accuracy of the labels in the study was verified through follow up care. The entirety of the dataset was anonymized to protect patients identities.

Features included Image_filename, Mask_tumor_filename, Mask_other_filename, Pixel_size, Age, Tissue_composition, Signs, Symptoms, Shape, Margin, Echogenicity, Posterior_features, Halo, Calcifications, Skin_thickening, Interpretation, BIRADS, Verification, Diagnosis, and Classification. Tumors were labeled by freehand annotation with the associated BIRADS features. The distribution of benign to malignant data is displayed in Figure 1.

The dataset for this study was chosen carefully to ensure it is suitable for both classification and segmentation of breast tumors. We selected a dataset that includes a good mix of benign and malignant cases, making the model more reliable for real-world use. High-quality labels were an important factor, as they helped train the model accurately. Since deep learning models work best with clear and detailed images, we made sure the dataset had high resolution and the right type of medical images. Additionally, we considered the balance between benign and malignant cases to avoid bias and ensure fair and accurate results. The dataset includes a diverse set of benign and malignant cases, ensuring variability in tumor characteristics. However, future studies may incorporate additional datasets.

A. Data Processing

Data processing for image classification began by reading the clinical data excel file, and removal of null data in the Mask_tumor_filename feature. Within these, data was put into subfolders for malignant and benign cases. Once complete, data was split using the Split-folders library, into a training set (80%), validation set (10%), and test set (10%).

In turn, data processing for image segmentation began similarly with the reading of the clinical data excel file, and removal of null data in the Mask_tumor_filename feature. Images were then copied into their respective mask and images folders. Then, data was split using the Split-folders library, into a training set (80%), validation set (10%), and test set (10%).

B. Classification Task

Two different deep learning models, MobileNetV2 and ResNet50 were tested. MobileNetV2 is a lightweight CNN model with prioritized speed and balanced accuracy. Meanwhile, ResNet50 is a 50-deep-layer residual neural network, with slower but higher performance.

For the first experiment, with binary image classification, validation accuracy was utilized to rate model performance. In the second experiment, with image segmentation, IoU value was used to rate model performance. Figure 2 is the flowchart for the binary image classification experiment.

In the first experiment, image classification was utilized. Models were trained on the training image set. Various hyperparameters were inputted, including: learning rate (lr), epochs, and optimizer (Adam). Hyperparameter tuning was performed, comparing learning rate and epoch values alongside their impacts on accuracy. Learning rates tested included 0.000001, 0.00001, 0.0001, 0.001, 0.005, 0.01, and 0.05. Epochs tested were 10, 20, 30, 40, and 50. These experiments consisted of 35 total tests, comparing the effects of these hyperparameters on accuracy. Initial tests showed similar scoring between the two deep learning models as seen in Table I and Table IV, with MobileNetV2 selected for continuation.

C. Segmentation Task

In the second experiment, we performed image segmentation to compare accuracy levels. We utilized U-Net Architecture with EfficientNetV2 as a backbone combining U-Net's strong spatial localization ability with EfficientNetB2's advanced feature extraction. For this model, hyperparameter tuning was also performed. The results from these tests suggested that the optimal hyperparameters were 20 epochs, and a 0.005 learning rate. Figure 3 displays the segmentation process flowchart.

Our study integrates both classification and segmentation, where classification serves as an initial diagnostic step, and segmentation further refines tumor localization. This dual-stage approach strengthens interpretability, assisting clinical decision-making.

V. Results

The results of this study demonstrate the effectiveness of using deep learning models for both classification and segmentation tasks in the context of breast cancer diagnosis.

In this study, two experiments were conducted using the MobileNetV2 and ResNet50 models with 35 hyperparameter tuning tests performed obtaining varying results. The performance of each model was evaluated based on accuracy, precision, and F1-score across two classes (label 0 and label 1). Table I displays the results of our study.

TABLE I
Summary of Deep CNN Model Evaluation on test data

Algorithm	Best Accuracy	F1 Score	Precision
MobileNetV2	63%	60%	0.61
ResNet50	66%	63%	0.65

A. Classification task results - MobileNetV2

Hyperparameter tuning was conducted to identify the optimal combination of epochs and learning rate for the MobileNetV2 model. The tuning grid included various learning rates (0.000001, 0.00001, 0.0001, 0.001, 0.005, 0.01, and 0.05) and epochs (10, 20, 30, 40, and 50). The objective was to maximize the validation accuracy. The optimal combination identified was 50 epochs with a learning rate of 0.0001, yielding the highest validation accuracy of 0.8333. This model configuration was subsequently saved for further evaluation. The performance of the MobileNetV2 model, configured with the optimal hyperparameters (50 epochs and learning rate of 0.0001), was evaluated on the test dataset. The classification metrics, including precision, recall, F1-score, and support, are presented in Table II.

TABLE II
Classification Report for the MobileNetV2 Model

Precision	Recall	F1-score	Support
61%	60%	60%	27

The multiline plot for MobileNetV2 can be seen in Figure 4. The plot shows that accuracy improves with lower learning rates

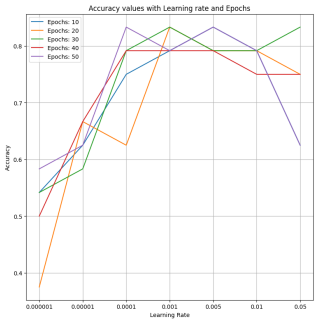

Figure 4. Multiline Plot for MobileNetV2 Model.

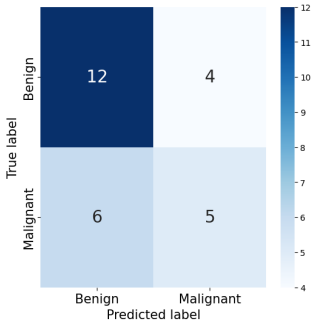

Figure 5. Test Confusion Matrix for the MobileNetV2 Model.

(<0.001) and more epochs, peaking at a learning rate of 0.001 with 30 epochs. Beyond this rate, accuracy declines, especially for larger learning rates (>0.005), indicating training instability. This highlights the need to fine-tune learning rates and epochs, with 0.001 and 30 epochs providing the best balance.

The overall accuracy achieved was 62.96%. For benign cases (Label 0), the model achieved a precision of 66.67%, recall of 75%, and an F1-score of 70.59%. For malignant cases (Label 1), the precision was 55.56%, recall was 45.45%, and F1-score was 50%. The macro average of these metrics indicates balanced performance across classes, while the weighted average reflects performance adjusted by the number of samples in each class. The confusion matrix for MobileNetV2 is shown in Figure 5.

B. Classification task results - ResNet50

Hyperparameter tuning was conducted to identify the optimal combination of epochs and learning rate for the ResNet50 model. The tuning grid included various learning rates (0.000001, 0.00001, 0.0001, 0.001, 0.005, 0.01, and 0.05) and epochs (10, 20, 30, 40, and 50). The objective was to maximize the validation accuracy.

The optimal combination identified was 30 epochs with a learning rate of 0.001, yielding the highest validation accuracy of 0.8333. This model configuration was subsequently saved for further evaluation.

The performance of the ResNet50 model, configured with the optimal hyperparameters (30 epochs and learning rate of

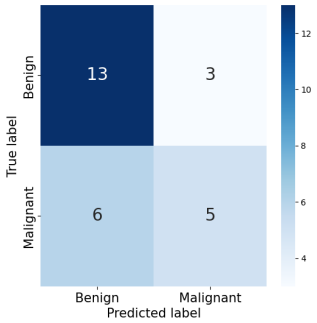

Figure 6. Test Confusion Matrix for the ResNet50 Model.

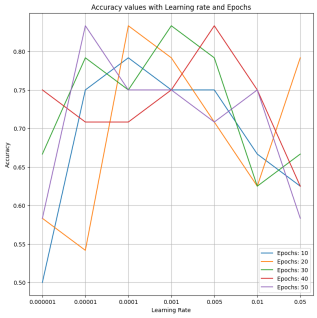

Figure 7. Multiline Plot for the ResNet50 Model.

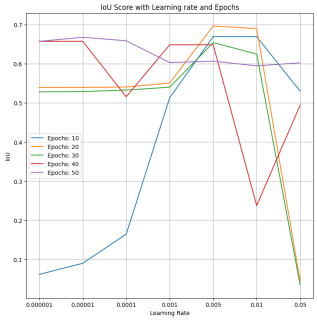

Figure 8. Multiline Plot for the U-Net Model (segmentation).

0.001), was evaluated on the test dataset. The classification metrics included precision, recall, F1-score, and support.

The overall accuracy achieved was 66.67%. For benign cases (Label 0), the model achieved a precision of 68.42%, recall of 81.25%, and an F1-score of 74.29%. For malignant cases (Label 1), the precision was 62.50%, recall was 45.45%, and F1-score was 52.63%. The macro average of these metrics indicates balanced performance across classes, while the weighted average reflects performance adjusted by the number of samples in each class. The confusion matrix for ResNet50 is shown in the Figure 6. As seen, the highest accuracy scores were 0.83, achieved repeatedly throughout testing.

The ResNet50 model outperformed the MobileNetV2 model in overall accuracy, achieving 66.67% compared to MobileNetV2's 62.96%. ResNet50 also demonstrated higher precision and F1-score for both labels, indicating its superior performance in this experiment.

For the ResNet50 model, the classification metrics, including precision, recall, F1-score, and support, are presented in Table III.

Additionally, Figure 7 represents multiline plot for ResNet50. The plot shows the relationship between learning rate, training epochs, and model accuracy. Accuracy generally improved at lower learning rates (<0.001) as epochs increase, peaking near a learning rate of 0.001 for 30 epochs. However, accuracy declined sharply for higher learning rates (>0.005) across all epoch values, indicating instability during training. We can optimize both learning rate and epoch count, with 0.001 and 30 epochs offering a balance between performance and stability.

TABLE III
CLASSIFICATION REPORT FOR THE RESNET50 MODEL

Precision	Recall	F1-score	Support
65%	63%	63%	27

C. Segmentation task results

In the second experiment, the segmentation model was tested with 20 epochs and a learning rate of 0.005, resulting in a validation IoU score of 0.697 and test IoU score of 0.629. Table IV shows the summary of the segmentation model's results. Figure 8 shows IoU scores for different learning rates and epochs. This plot demonstrates the IoU score's variation with learning rate and the number of epochs. The IoU score generally increases with learning rates up to 0.001, particularly for 20 and 30 epochs, where the scores peak around 0.7. For higher learning rates (>0.005), the IoU score drops significantly across all epoch values, indicating unstable segmentation performance. The results emphasize that a learning rate of 0.001 and 30 epochs provide the most consistent and optimal segmentation accuracy.

Results from segmentation experiments are shown in Figure 9. The optimal IoU score was achieved at a learning rate of 0.005.

TABLE IV
SUMMARY OF THE SEGMENTATION MODEL RESULTS

Model	Validation IoU	Test IoU
U-Net	70%	63%

VI. DISCUSSION | EVALUATION

The hyperparameter tuning for MobileNetV2 and ResNet50 revealed that the optimal settings for both models resulted in a validation accuracy of 0.8333, but with different configurations (50 epochs and a learning rate of 0.0001 for MobileNetV2; 30 epochs and a learning rate of 0.001 for ResNet50). These settings were selected based on their performance metrics.

In the classification task, the ResNet50 model outperformed the MobileNetV2 model. ResNet50 achieved an overall accuracy of 66.67%, compared to MobileNetV2's 62.96%.

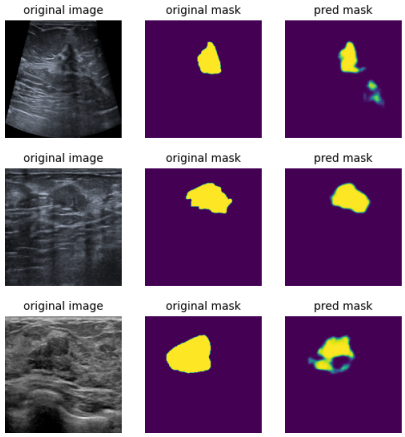

Figure 9. Segmented Mask Images

ResNet50 also showed better precision, recall, and F1-scores for benign cases, reflecting a more balanced and effective classification capability. However, both models exhibited lower recall for malignant cases, which indicates a need for further improvement in detecting malignant lesions.

The results indicate that the MobileNetV2 model is more effective at correctly identifying benign images compared to malignant ones. However, there is a notable trade-off between the precision and recall, especially for malignant cases, which suggests areas for future improvement.

Our results also indicate that the ResNet50 model demonstrates improved overall accuracy compared to the MobileNetV2 model, particularly in identifying benign images. However, the recall for malignant cases still shows room for improvement, indicating that some malignant images are not being correctly identified by the model.

In the application of these models, the differences in false predictions cannot be overlooked. While a prediction of malignant in a case of benign cancer is certainly undesirable, it is especially concerning if a prediction of benign is made in the case of a malignant tumor. Such errors could lead to a lack of testing and treatment for a patient.

For the segmentation task, the U-Net model with EfficientNetB2 as the backbone achieved a test IoU score of 0.629. This result indicates the model's strong ability to accurately segment breast lesions, providing valuable information for tumor localization and quantification. The segmentation results are promising, given the complexity of the task and the variability in lesion appearance.

VII. Conclusion and Future Work

The findings from this study underscore the potential of deep learning models in enhancing the accuracy and efficiency of breast cancer diagnosis. The ResNet50 model's superior performance in classification tasks suggests its suitability for diagnostic applications where accurate classification of lesions is critical. On the other hand, the U-Net model with EfficientNetB2 demonstrated robust segmentation capabilities, which are essential for precise tumor localization and treatment planning.

These results align with existing literature that highlights the efficacy of deep learning models in medical imaging tasks. The use of EfficientNetB2 as a backbone for the U-Net model has shown to be particularly effective in capturing intricate details in ultrasound images, which is crucial for accurate segmentation. The precise segmentation provided by the models means that treatment can be better tailored to each patient's specific condition, enhancing the effectiveness of treatment plans and potentially improving survival rates.

The deployment of these deep learning models in clinical settings can automate and enhance the breast cancer screening process, enabling early detection of cancerous growths with higher accuracy. This early detection is key to improving patient outcomes. By accurately classifying and segmenting breast cancer images, these models can significantly reduce the diagnostic workload of radiologists and pathologists.

Our future research could focus on further refining and optimizing the U-Net deep learning models to enhance their accuracy and efficiency for segmentation tasks, possibly through the integration of more advanced architectures or ensemble techniques. Improvement of IoU score would also be central to development as it provides detailed and precise insights into medical imaging data. Additionally, conducting studies with larger and more diverse datasets would help validate the general applicability of the models, ensuring their applicability across different populations and imaging conditions.

References

[1] Y. Xu *et al.*, "Medical breast ultrasound image segmentation by machine learning", *Ultrasonics*, vol. 91, pp. 1–9, 2019.

[2] S. Radhakrishna *et al.*, "Role of magnetic resonance imaging in breast cancer management", *South Asian journal of cancer*, vol. 7, no. 02, pp. 069–071, 2018.

[3] A. Pawłowska *et al.*, "Curated benchmark dataset for ultrasound based breast lesion analysis", *Scientific Data*, vol. 11, no. 1, p. 148, 2024.

[4] M. Amrane, S. Oukid, I. Gagaoua, and T. Ensari, "Breast cancer classification using machine learning", in *2018 electric electronics, computer science, biomedical engineerings' meeting (EBBT)*, IEEE, 2018, pp. 1–4.

[5] A. K. Singh and B. Gupta, "A novel approach for breast cancer detection and segmentation in a mammogram", *Procedia Computer Science*, vol. 54, pp. 676–682, 2015.

[6] S. S. Almberg *et al.*, "Training, validation, and clinical implementation of a deep-learning segmentation model for radiotherapy of loco-regional breast cancer", *Radiotherapy and Oncology*, vol. 173, pp. 62–68, 2022.

[7] A. Kaur, M. Rashid, A. K. Bashir, and S. A. Parah, "Detection of breast cancer masses in mammogram images with watershed segmentation and machine learning approach", in *Artificial intelligence for innovative healthcare informatics*, Springer, 2022, pp. 35–60.

[8] R. Khaled, J. Vidal, J. C. Vilanova, and R. Martí, "A u-net ensemble for breast lesion segmentation in dce mri", *Computers in Biology and Medicine*, vol. 140, p. 105 093, 2022.

[9] G. Piantadosi, S. Marrone, A. Galli, M. Sansone, and C. Sansone, "Dce-mri breast lesions segmentation with a 3tp u-net deep convolutional neural network", in *2019 IEEE 32nd International Symposium on Computer-Based Medical Systems (CBMS)*, IEEE, 2019, pp. 628–633.

[10] H. Li, D. Chen, W. H. Nailon, M. E. Davies, and D. I. Laurenson, "Dual convolutional neural networks for breast mass segmentation and diagnosis in mammography", *IEEE Transactions on Medical Imaging*, vol. 41, no. 1, pp. 3–13, 2021.

[11] R. Zeleznik *et al.*, "Deep-learning system to improve the quality and efficiency of volumetric heart segmentation for breast cancer", *NPJ digital medicine*, vol. 4, no. 1, p. 43, 2021.

[12] Y. Zhang *et al.*, "Automatic detection and segmentation of breast cancer on mri using mask r-cnn trained on non–fat-sat images and tested on fat-sat images", *Academic Radiology*, vol. 29, S135–S144, 2022.