

BREAST CANCER PREDICTION

IMAGE SEGMENTATION AND CLASSIFICATION WITH DEEP LEARNING

By

Huiqiao Feng

Nadimul Hasan

Solmaz Heidar Nassab

School of Software Design and Data Science

SENECA COLLEGE

2025

© 2025 Huiqiao Feng, Nadimul Hasan, Solmaz Heidar Nassab

Table of Contents

Abstract.....	2
Introduction.....	3
Problem Statement.....	3
Dataset.....	.4
Objectives.....	4
Challenges.....	5
Methodology.....	6
Data Preparation.....	6
Model Implementation.....	6
Evaluations.....	8
Comparative Analysis.....	.8
Implementation.....	10
Conclusion.....	11
Reference.....	11

Abstract

Breast cancer diagnosis through ultrasound imaging is a crucial yet challenging task, it is critical for improving survival rates and treatment outcomes. This study explores a series of deep learning-based approaches to improve classification accuracy by integrating image segmentation. Using the Breast Ultrasound Images Dataset from Kaggle, we trained the EfficientNetB3 classifier with UNet segmentation implemented. This segmentation-based neuron network classifier provided a promising output, yielding an accuracy of 86.39% in tumour classification. Our results highlight the importance of image preprocessing and region localization in improving medical image classification outcomes.

Introduction

Breast cancer remains a leading cause of mortality among women worldwide. Early detection significantly improves treatment outcomes, making accurate diagnostic tools essential. Ultrasound imaging is a widely used, non-invasive method for breast cancer screening. However, interpreting ultrasound images is challenging due to the presence of speckle noise. Traditional diagnostic methods rely heavily on the expertise of radiologists, which can lead to variability in interpretation, with time and efficiency concerns.

Recent advancements in deep learning have shown promise in automating and enhancing medical image analysis. Convolutional Neural Networks (CNNs), in particular, have demonstrated high accuracy in image classification tasks. In the context of breast ultrasound imaging, integrating segmentation and classification models can potentially improve diagnostic accuracy by focusing on regions of interest.

In this study, we propose a deep learning pipeline that combines UNet-based segmentation with EfficientNetB3-based classification to improve breast cancer detection from ultrasound images. We utilize the Breast Ultrasound Images Dataset from Kaggle, which comprises 780 images categorized into benign, malignant, and normal classes, along with corresponding tumor segmentation masks. The dataset presents challenges such as class imbalance and limited size, which we address through data augmentation techniques.

Our methodology involves training an EfficientNetB3 model on raw ultrasound images as a baseline, followed by enhancing the images using UNet-generated segmentation masks to highlight tumor regions. The refined images are then used to retrain the EfficientNetB3 classifier. Evaluation results indicate a significant improvement in classification accuracy from 42.01% with the baseline model to 86.39% with the segmentation-enhanced model. These findings underscore the effectiveness of combining segmentation and classification models in improving breast cancer diagnostics from ultrasound images.

Problem Statement

Dataset

This study utilizes the Breast Ultrasound Images Dataset introduced by Al-Dhabyani et al. (2019), which is publicly available on Kaggle. The dataset consists of 780 grayscale ultrasound images categorized into three classes: benign (437 images), malignant (210 images), and normal (133 images). Each image is accompanied by a ground truth segmentation mask that outlines the tumor region, making the dataset suitable for both classification and segmentation tasks.

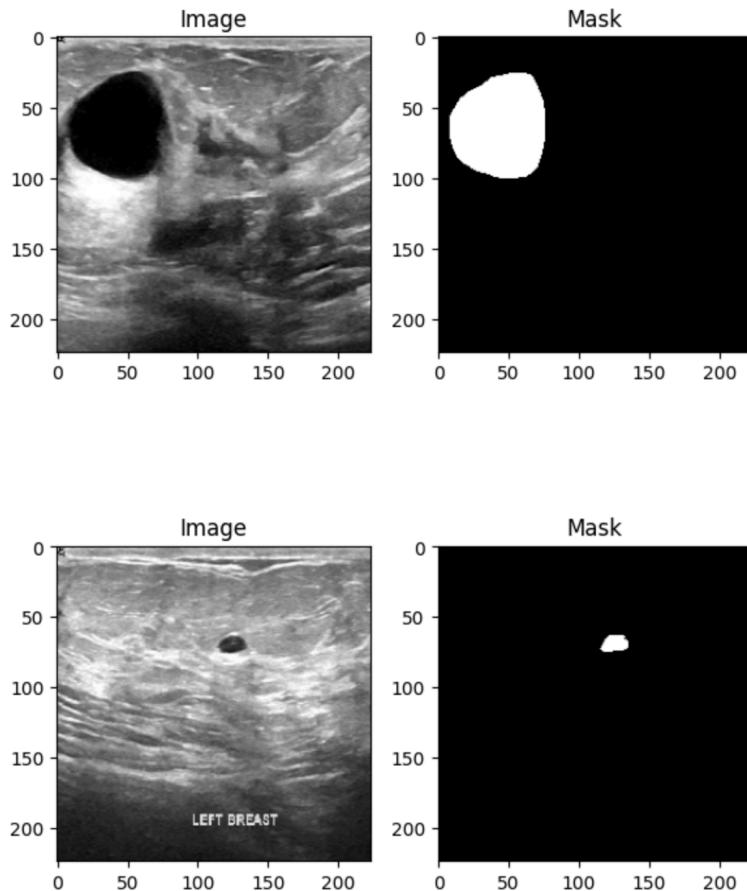


Figure 1: Sample Data with Raw Image and Provided Mask

Due to the limited size and class imbalance, we implemented various data augmentation techniques—including rotation, shear, and horizontal flipping—to enrich training data and address underrepresentation. Despite its limitations, this dataset provides a valuable foundation for exploring deep-learning approaches in medical imaging.

Objectives

Breast cancer remains one of the leading causes of cancer-related deaths among women worldwide. Early and accurate diagnosis through medical imaging is critical to improving patient

outcomes. However, interpreting breast ultrasound images poses significant challenges due to noise, low contrast, and the subtle appearance of tumors. Manual interpretation is time-consuming and prone to variability among radiologists. To address this, there is a growing need for automated, accurate, and scalable diagnostic tools.

This project investigates the potential of deep learning techniques to enhance the classification and segmentation of breast ultrasound images. Specifically, we aim to:

- **Develop and implement segmentation model (UNet)** to accurately localize tumor regions in breast ultrasound scans.
- **Train classification models (EfficientNetB3)** to categorize images into **benign**, **malignant**, or **normal** classes.
- **Evaluate performance** by comparing classification results on raw versus segmented images to determine if tumor localization improves diagnostic accuracy.

Our central research questions include:

- How accurately can UNet segment potential tumor regions in breast ultrasound images?
- How effectively can EfficientNetB3 classify raw ultrasound images compared to those enhanced by segmentation?
- Does the inclusion of segmentation improve classification metrics such as accuracy, precision, and recall?

Challenges

Throughout the project, we encountered various challenges—some of which we were able to address successfully, while others required alternative approaches or remain unresolved due to technical limitations or the need for domain expertise.

- **Limited Dataset Size:** With only 780 images available, the dataset posed a risk of overfitting and restricted the model's ability to generalize. We partially mitigated this through targeted data augmentation, especially for underrepresented classes, but acquiring more data remains a critical need for robust performance.
- **Severe Class Imbalance:** The dataset contains a disproportionate number of benign (437), malignant (210), and normal (133) cases. This imbalance negatively impacted classification performance, especially for the minority class. We addressed this by augmenting samples from the underrepresented categories, yet perfect balance is difficult to achieve without additional data or synthetic generation.
- **Compute Power Constraints:** Training deep learning models like UNet and EfficientNetB3 is computationally intensive. With our available resources, training a single model took several hours, limiting our ability to conduct comprehensive experiments, such as full-scale hyperparameter tuning or model ensembling.

- **Hyperparameter Tuning Limitations:** Due to the long training times and resource constraints, we could not perform exhaustive hyperparameter tuning. We used reasonable defaults and minor adjustments but recognize that further optimization could improve model performance.
- **Segmentation Mask Validation:** Although the dataset provides ground truth masks, potential inconsistencies or inaccuracies in expert annotations could affect model training and evaluation. Validating these masks would require access to medical professionals, which was outside the scope of our current capabilities.

Methodology

Data Preparation

The dataset used in this project is the publicly available Breast Ultrasound Images Dataset (Al-Dhabyani et al., 2019), which contains 780 grayscale ultrasound images categorized into three diagnostic classes: benign (437), malignant (210), and normal (133). On top of the raw medical images, it also provides expert-annotated ground truth segmentation masks for tumor regions.

During initial inspection, we identified two major issues: a relatively small dataset size and a severe class imbalance. To address this, we applied targeted data augmentation to the underrepresented classes (normal and malignant) using transformations such as horizontal flipping, shearing, rotation, and zooming. This effectively increased the diversity and quantity of samples while maintaining clinical relevance.

To meet the architectural input requirements of models like EfficientNetB3 and UNet, all images and corresponding masks were resized to 255×255 pixels. We then split the data into training, validation, and testing sets using a stratified sampling strategy to preserve class distribution across all subsets. Data augmentation was applied only to the training set to avoid leakage and maintain fair evaluation conditions. The entire preprocessing pipeline was implemented using TensorFlow and Keras.

Model Implementation

The methodology followed a three-phase deep learning pipeline involving baseline classification, segmentation, and refined classification.

In the first phase, we trained a baseline image classification model using the EfficientNetB3 architecture. Leveraging transfer learning, we initialized the model with ImageNet weights and replaced the top layer with a three-class softmax classifier. Dropout layers and L2 regularization were added to prevent overfitting, and categorical cross-entropy was used as the loss function.

Despite these efforts, the classification accuracy on raw images remained low, with a test set accuracy of 42.01%. This was largely attributed to noisy background data and limited attention to tumor regions.

In the second phase, we aimed to enhance classification by introducing tumor localization. First, we used the provided ground truth segmentation masks to overlay and mask the raw images, effectively highlighting tumor regions and reducing irrelevant background information. These masked images were then used to train a UNet model for further refinement and generalization of tumor region prediction. UNet, known for its encoder-decoder architecture and skip connections, is particularly well-suited for biomedical image segmentation. It was trained to generate segmentation masks from raw input images, learning to identify tumor regions without relying on the original ground truth masks during inference. The UNet model achieved a segmentation accuracy of 95.81% on the test set, successfully generalizing tumor detection.

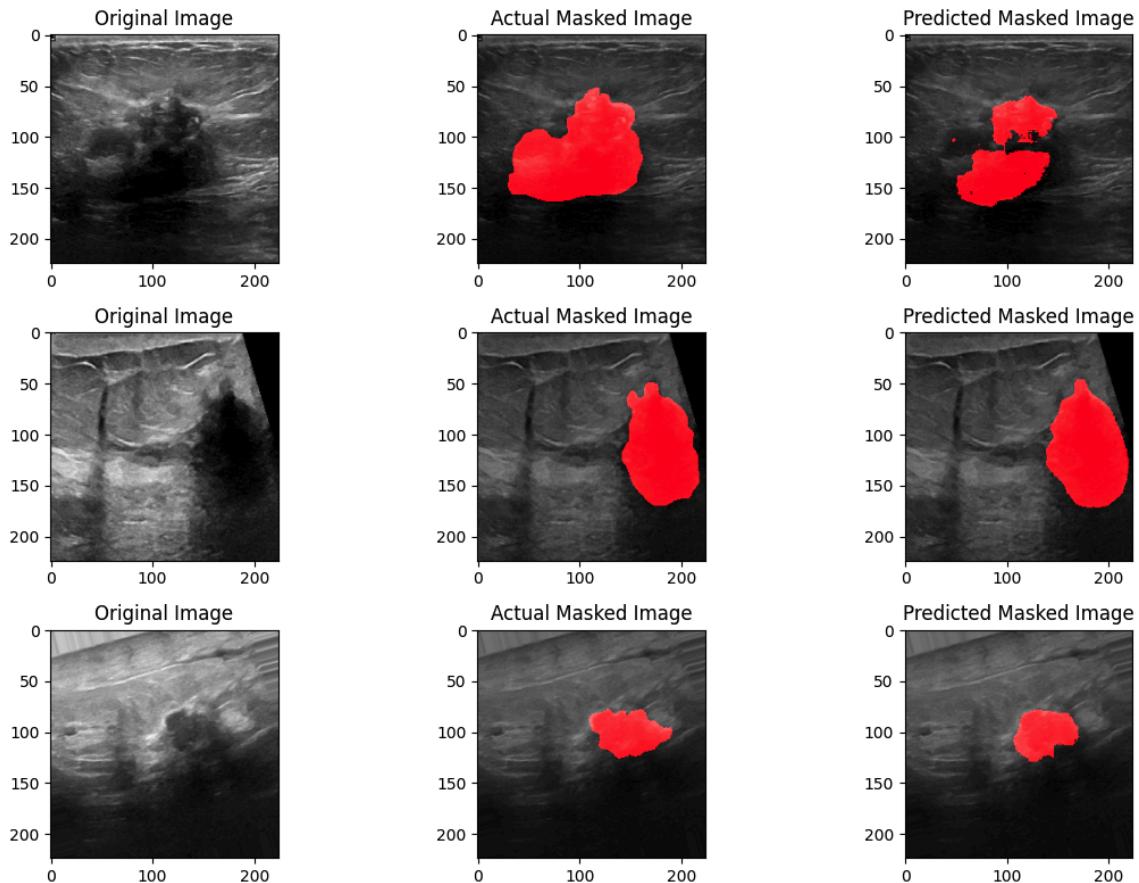


Figure 2: Example of Original Image Overlaid with Actual Mask and UNet Predicted Mask

In the final phase, we leveraged the segmentation output to improve classification performance. The segmented masks were overlaid on the original images to suppress background noise and emphasize tumor boundaries. These enhanced images were then used to retrain the EfficientNetB3 classifier under the same configuration as the baseline. This approach

significantly improved the classifier's performance, increasing the test accuracy to 86.39%. The improvement was observed across all evaluation metrics, particularly for the underrepresented classes. This demonstrates that incorporating segmentation as a preprocessing step can substantially enhance diagnostic accuracy in medical imaging tasks.

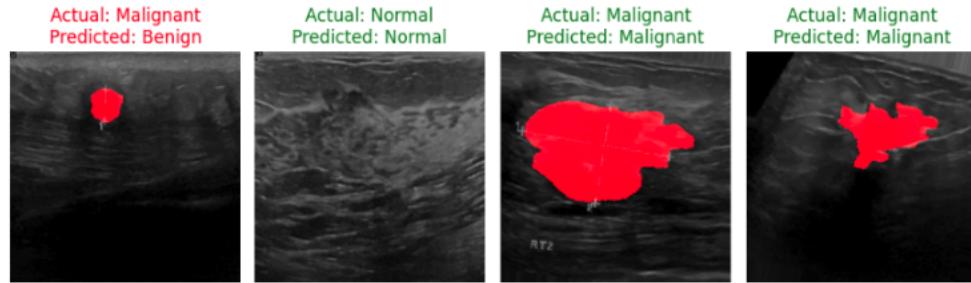


Figure 3: Example of Final Model Prediction

Evaluations

EfficientNetB3 and U-Net were chosen for their complementary strengths in handling medical image analysis, particularly for classification and segmentation tasks. EfficientNetB3 was selected due to its computational efficiency, which allows it to achieve high accuracy while maintaining a manageable computational cost. This makes it well-suited for medical imaging applications, especially when resources are limited. Additionally, EfficientNet models have demonstrated strong performance on multiple image classification benchmarks, including those involving medical image datasets. Another key advantage is the availability of pre-trained weights from ImageNet, which facilitates faster convergence and improves performance, particularly when working with a limited amount of labelled training data.

U-Net, on the other hand, is a well-established architecture specifically designed for biomedical image segmentation. Its encoder-decoder structure with skip connections enables it to effectively capture both global context and fine-grained local details, which are essential for accurate segmentation. U-Net has been widely applied to medical imaging tasks, including tumor detection and segmentation in ultrasound images, demonstrating its effectiveness in delineating structures within complex medical scans. By combining EfficientNetB3 for feature extraction and classification with U-Net for segmentation, the model leverages the strengths of both architectures to enhance medical image analysis, achieving robust and efficient performance in tasks such as tumor detection and classification.

Comparative Analysis

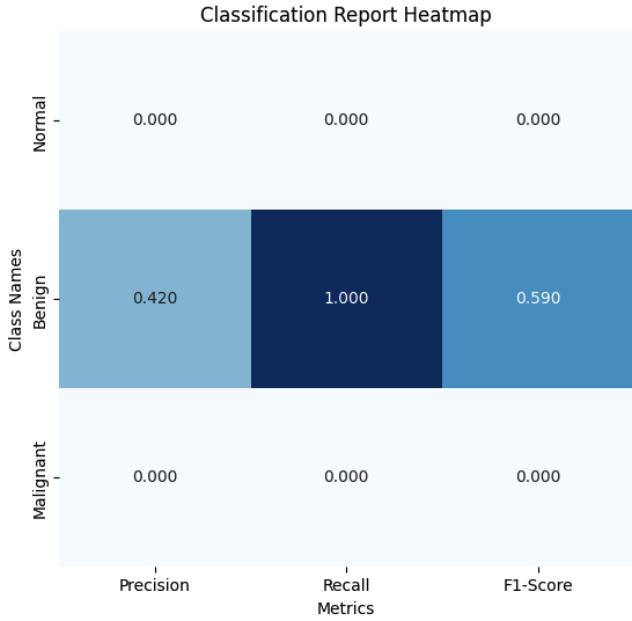


Figure 4: Classification Report on the baseline model (EfficientNetB3 on Raw Images)

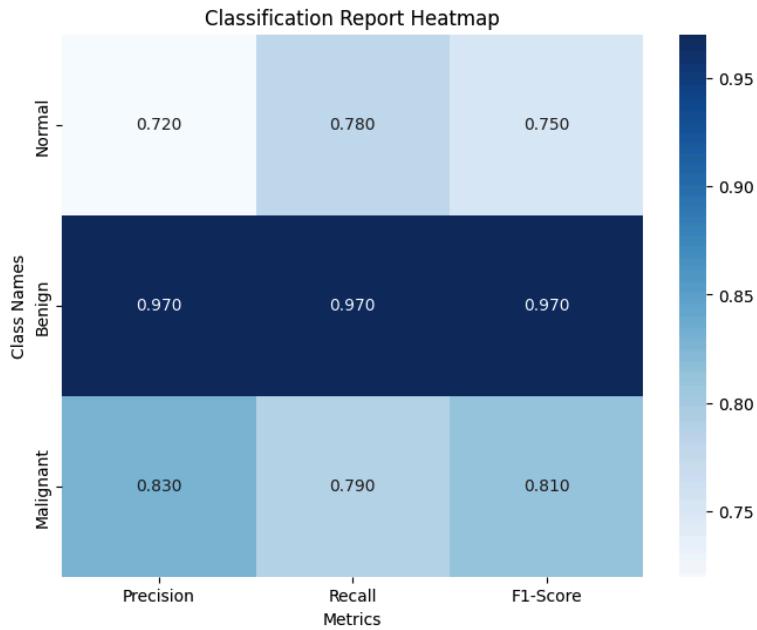


Figure 5: Classification Report on the refined model (EfficientNetB3 on UNet segmented Images)

Above are the classification reports for the baseline model (EfficientNetB3 trained on raw images), and the refined model (EfficientNetB3 on UNet segmented images). The same parameters were used on both models. By comparing the classification reports, only 36.69% accuracy with 0 precision and recall for normal and malignant classes were found on the baseline model. On the improved model, the accuracy was significantly improved, to 86.39%, with balanced performance across all classes. The UNet segmentation model achieved 95.81%

accuracy in detecting and segmenting tumor regions, enhancing the input data quality, and leading to improvements in the refined model performance.

When comparing the precision, recall and f1-score, the refined model achieved much higher precision across all classes, especially for benign and normal classes, which were completely missed in the baseline model. For Normal Class, precision increased to 0.72, recall to 0.78, and F1-score to 0.75, showing substantial recovery in detecting normal tissue. For the Benign Class, The baseline model achieved a perfect recall of 1.00, but only 0.42 precision, leading to a modest F1-score of 0.59. This implies that while it correctly identified all benign cases, it misclassified many non-benign cases as benign. In the refined model, all precision, recall and F1-score achieved to 0.97, showing reliable and confident prediction of benign cases. Malignant Class, similar to Normal, completely failed to detect in the baseline model, with all metrics showing 0.00. In the final model, precision increased to 0.83, 0.79 for recall, and an F1-score of 0.81, showing significant success in recognizing malignant tumors, which is critical for clinical relevance.

The baseline model struggled with severe misclassification and failed to detect two of the three classes. After incorporating UNet segmentation, the refined model demonstrated strong and balanced performance across all categories. This comparison validates the effectiveness of combining segmentation with classification for medical image analysis, particularly in challenging datasets with high-class imbalance and subtle visual cues.

Implementation

Based on the model we designed, a flask-based web application was developed that allows users to upload medical images. The application classifies each image into three categories: benign, malignant, or normal. It also recommends the nearest clinic or hospital for medical consultation based on the user's location, which can be detected automatically or entered manually if preferred. This will be helpful for patients to get an initial assessment of their scans since the wait time for doctors is a few weeks in Canada. Based on the severity of their results they can request a quicker meeting with a specialist.

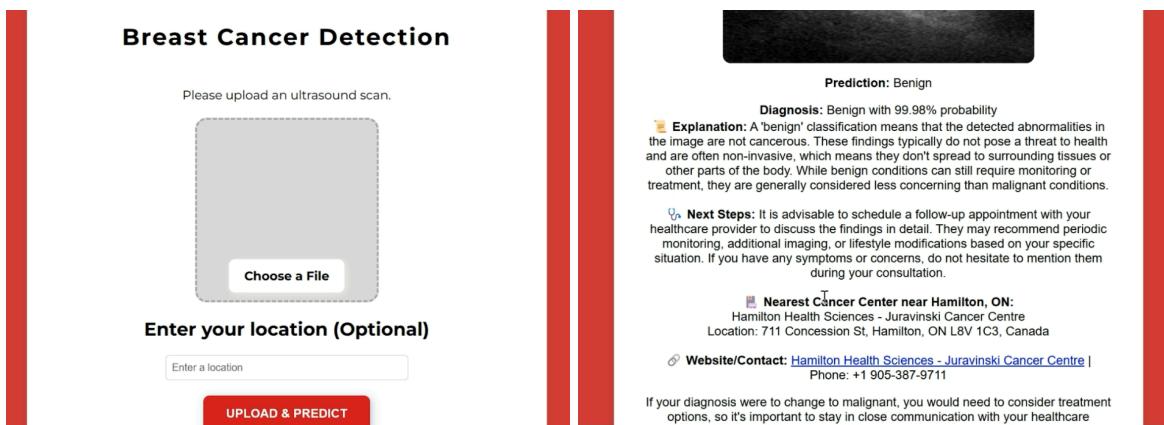


Figure 6: Web Interface Implementation

Conclusion

This project demonstrates the potential of deep learning techniques in enhancing breast cancer diagnosis from ultrasound imagery. Initially, our baseline classification model, EfficientNetB3 trained on raw ultrasound images, suffered from poor performance due to challenges such as small dataset size, class imbalance, and lack of localized tumor information.

To overcome these limitations, we integrated a UNet-based segmentation model to extract tumor regions and used the resulting segmented images for a second round of classification. This refined approach significantly improved the model's diagnostic accuracy—from 42.01% in the baseline model to 86.39% in the refined model. All major classification metrics (precision, recall, F1-score) saw substantial improvements across all three classes (normal, benign, malignant), especially in recognizing malignant tumors, which is critical for clinical applications.

These results support our hypothesis that segmenting the regions of interest before classification enhances model performance. While computing limitations and dataset size posed constraints on model tuning and experimentation, the improvements achieved through segmentation demonstrate a promising direction for medical image analysis. Future work should focus on expanding the dataset, further validating segmentation accuracy, and optimizing model architectures and hyperparameters to build a more robust, generalizable diagnostic tool.

Reference

- Abdelrahman, A., & Viriri, S. (2023). EfficientNet family U-Net models for deep learning semantic segmentation of kidney tumors on CT images. *Frontiers in Computer Science*, 5. <https://doi.org/10.3389/fcomp.2023.1235622>
- Al-Dhabyani, W., Gomaa, M., Khaled, H., & Fahmy, A. (2019). Dataset of breast ultrasound images. *Data in Brief*, 28, 104863. <https://doi.org/10.1016/j.dib.2019.104863>
- Goodfellow, I., Bengio, Y., & Courville, A. (2016). *Deep learning*. MIT Press.
- Litjens, G., Kooi, T., Bejnordi, B. E., Setio, A. A., Ciompi, F., Ghafoorian, M., ... & van der Laak, J. A. (2017). A survey on deep learning in medical image analysis. *Medical Image Analysis*, 42, 60–88. <https://doi.org/10.1016/j.media.2017.07.005>
- Pérez, L., & Wang, J. (2017). The effectiveness of data augmentation in image classification using deep learning. *Convolutional Neural Networks for Visual Recognition*, 11(1), 1–8.

- Ronneberger, O., Fischer, P., & Brox, T. (2015). U-Net: Convolutional networks for biomedical image segmentation. In *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015* (pp. 234–241). Springer, Cham.
https://doi.org/10.1007/978-3-319-24574-4_28
- Tan, M., & Le, Q. V. (2019). EfficientNet: Rethinking model scaling for convolutional neural networks. In *Proceedings of the 36th International Conference on Machine Learning* (pp. 6105–6114). PMLR.
- World Health Organization (WHO). (2022). *Breast cancer: Early detection*.
<https://www.who.int/news-room/fact-sheets/detail/breast-cancer>