# Breast Cancer Ultrasound Image Classification and Segmentation

## Introduction to the modeling problem

Breast cancer is a leading cause of mortality among women globally, with early detection significantly improving survival rates. While ultrasound imaging is a safe and widely accessible diagnostic tool, its interpretation is challenging due to the inherent complexity and variability of breast lesions. This project proposes a deep learning-based system for breast mass classification and segmentation, providing radiologists with a reliable second opinion to enhance early diagnosis and accurate cancer detection.

## Related Work

Breast cancer classification using convolutional neural networks (CNNs) and image fusion with ultrasound images demonstrates that VGG-19 performs well for classification tasks, highlighting its potential in accurately identifying breast masses. Similarly, ResNet-50, applied to brain tumor classification using transfer learning, proves effective due to its residual connections. These connections enable deeper network training, mitigate vanishing gradients, and facilitate the extraction of both fine details and broader contextual features, making it suitable for complex classification tasks.

In segmentation, VGG-19 shows limited effectiveness, even with preprocessing enhancements, indicating the need for more specialized architectures. U-Net, designed specifically for biomedical image segmentation, has been shown to achieve high accuracy with relatively limited training data, making it a strong candidate for breast mass segmentation. Additionally, the use of atrous convolution in semantic segmentation offers an alternative to U-Net by enabling multi-scale feature extraction and improving segmentation accuracy, addressing some of the inherent challenges in segmenting complex medical images.

## Overview of Breast Cancer Classification and Segmentation

Ultrasound imaging offers a cost-effective and safe method for breast cancer screening, enabling efficient classification and segmentation with minimal human intervention. This study employs model adjustments, ensemble strategies, and rigorous validation to enhance the accuracy of breast cancer detection from ultrasound images.

## Data Deep Dive

### Dataset Overview

This project utilizes the Breast Ultrasound Dataset, comprising 780 500x500-pixel images categorized as benign, malignant, or normal. Collected from 600 patients aged 25–75 at Baheya Hospital, Cairo, Egypt, each image includes a corresponding lesion boundary mask for segmentation training.

## Dataset Preparation and Exploration

The dataset was divided into training, validation, and testing subsets, with augmentations applied only to the training set to enhance generalization and prevent bias during testing. Key preprocessing steps included loading image-mask pairs for alignment, grayscale conversion to simplify analysis, resizing to 256x256 pixels for consistency, and normalization. Training data was augmented with rotations, scaling, and flips. Finally, the data was converted into PyTorch tensors, batched, and formatted for deep learning model requirements, ensuring uniformity and reliability across all stages of training, validation, and testing.

## Data Insights

The dataset contains an imbalanced distribution across normal, benign, and malignant categories, visualizations created to analyze sample diversity and distribution.
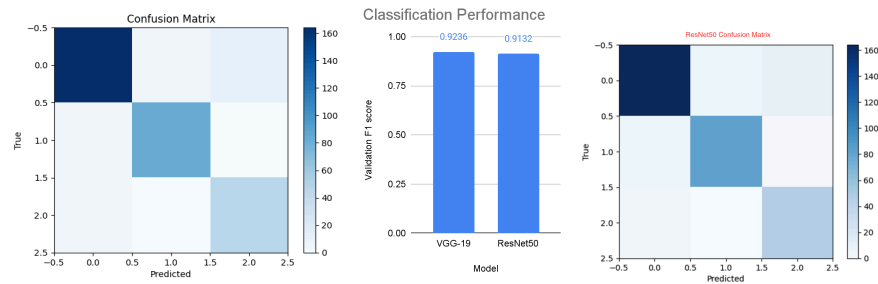
# Classification Task

## Model Choice and Architecture

For classification, VGG-19 was selected as the primary model, achieving a weighted F1 score of 0.9236. Evaluation metrics included the F1 score and confusion matrix, providing a comprehensive assessment of model performance. VGG-19 demonstrated superior performance compared to ResNet50 and EfficientNet due to its capacity to capture fine-grained texture details, which are critical for distinguishing between normal, benign, and malignant breast lesions. VGG-19's architecture incorporates fully convolutional layers and pooling layers approximately every 2-4 layers, effectively balancing feature extraction and computational efficiency. ResNet50 was also tested, leveraging skip connections to mitigate vanishing gradients and employing a fully connected layer as the final layer. Despite its robustness, ResNet50 did not outperform VGG-19 on this dataset.

## Training and Fine-tuning

The models were fine-tuned on the Breast Ultrasound Dataset, with specific hyperparameter tuning to optimize accuracy. Key adjustments included optimizing the learning rate, batch size, and regularization parameters to improve generalization and prevent overfitting. Training involved careful preprocessing and augmentation strategies to enhance the model's ability to generalize across diverse data, ensuring robust performance on unseen test samples.

## Classification Results



VGG-19 demonstrated a higher validation F1 score compared to ResNet50, indicating slightly better performance. The marginal difference in F1 scores (0.0104) suggests that both models perform similarly, with VGG-19 showing a slight advantage. Consequently, VGG-19 was selected as the preferred model due to its superior validation performance.

# Segmentation Task

## U-Net Model: Design and Advantages

The U-Net model was selected and further optimized to enhance segmentation accuracy, leveraging its unique skip connection design. U-Net is particularly suited for medical imaging due to its ability to preserve both low-level features, such as edges and textures, and high-level semantic information, enabling precise boundary detection. Its fast convergence rate and high accuracy in identifying small structures and lesions make it ideal for breast cancer segmentation. The skip connections in U-Net effectively combine low- and high-level features, resulting in superior segmentation performance, especially when training data is limited.

## Modified U-Net Architecture

To enhance U-Net's depth and boundary sensitivity, the architecture was modified to include:

- **Deeper Encoder:** Based on VGG-19 for improved feature extraction.
- **Enhanced Decoder:** Designed to mirror the encoder with transposed convolutions and a final 1x1 convolution layer for precise segmentation outputs.
- **Leaky ReLU Activation:** Replacing ReLU with Leaky ReLU allowed finer gradient flow, emphasizing subtle detail capture crucial for segmenting small lesions.

These modifications improved the U-Net model's ability to detect small and intricate structures in medical imaging applications, such as breast cancer detection.
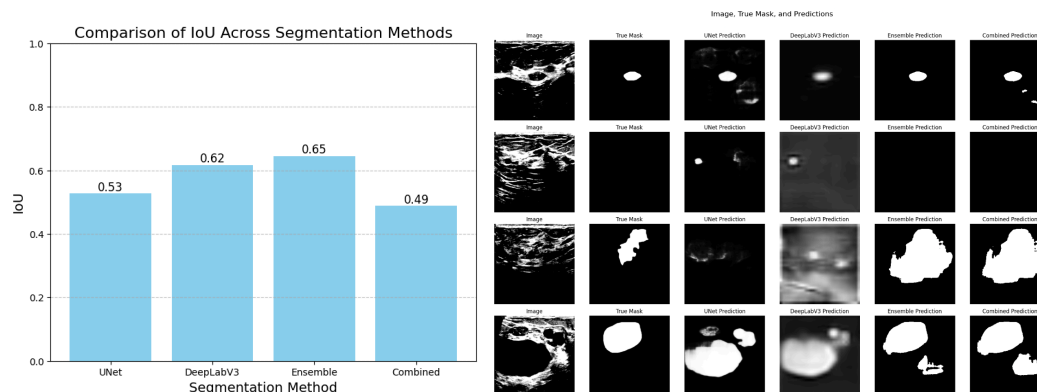
## Additional Segmentation Methods

Complementary approaches included thresholding, morphological operations, DeepLabV3 with ResNet101, and ensemble techniques to enhance accuracy and scalability.

## Data Preprocessing and Training

Preprocessing involved loading image-mask pairs, grayscale conversion, resizing to 256x256 pixels, and batching. Training utilized IoU and Dice coefficients to evaluate performance, with improvements tracked via validation metrics. This streamlined approach improved segmentation accuracy and robustness, making the modified U-Net model well-suited for breast cancer detection.

## Segmentation Results and Implications



The Ensemble method achieved the highest IoU of 0.645, highlighting the benefits of model combination. DeepLabv3, with an IoU of 0.617, emerged as the best-performing single model, albeit with higher computational demands. The standard U-Net provided a stable baseline with an IoU of 0.528, while the Combined approach underperformed with an IoU of 0.488, indicating the need for refinement in its weighting strategy. These findings emphasize the effectiveness of ensemble learning and the need for careful evaluation of combination techniques, guiding future optimization efforts.

# Future Work

To enhance segmentation, integrating DeepLabv3+ and exploring advanced methods will refine model accuracy. Intelligent weighted combination techniques and experiments with diverse architectures, including transformer-based models such as SegFormer and Segmenter, will optimize ensemble performance. Future research will focus on extensive testing of new models, comparing results across segmentation and classification to validate improvements and assess performance gains.

# Conclusion

This project developed a deep learning-based system for breast cancer classification and segmentation using ultrasound images. VGG-19 achieved an F1 score of 0.9236, demonstrating strong classification performance across benign, malignant, and normal cases. For segmentation, a modified U-Net and DeepLabv3 provided high boundary precision and accuracy, with DeepLabv3 showing superior IoU but requiring higher computational resources. Ensemble methods further enhanced stability and generalization. These advancements in classifying and segmenting complex tissue structures support early and accurate breast cancer detection and highlight pathways for further improvement through model refinement.