Deep Learning Techniques for Breast Ultrasound Image Classification and Segmentation

Ryan Jimenez, Meng Xu (Advisor), Ph.D.
Department of Computer Science and Technology, Kean University, Union, NJ 07083
{jimenrya, mexu}@kean.edu

Abstract

Checking for breast cancer is one of the most time-consuming and expensive processes, despite it being one of the leading causes of cancer-related deaths globally. This study explores the application of **deep learning techniques** within the frameworks of **classification** and **segmentation** to improve the efficiency and accuracy of breast cancer detection. We compared multiple approaches for **single-task** learning models and ultimately ended with the approach of a **multitasking learning model** that performed both tasks simultaneously. The results shown are from the models we experimented with; although the results aren't pixel-perfect, they still highlight the potential of deep learning and its connection to the breast cancer diagnostic process. If successfully implemented, it has the potential to decrease the time and cost it takes to perform these tests, prompting individuals to get checked more often.

Methods

We utilized deep learning techniques for both image classification and segmentation of breast ultrasound images. Models tested include a custom CNN, ResNet-50, VGG16, and transformer-based architectures. The multitask learning (MTL) model combined ResNet-50 for classification (left side) and U-Net for segmentation (right side).

Dataset and Preprocessing

The dataset comprised 647 breast ultrasound images (210 malignant, 437 benign) with paired ground truth masks for segmentation. Preprocessing included resizing, normalization, and custom weights to address class imbalance. Data augmentation techniques such as flips, rotations, and zooms were applied.

Training

Binary cross-entropy and BCEWithLogits losses were used for classification and segmentation. SGD and Adam optimizers were experimented with, alongside varying learning rates. The number of epochs were adjusted depending on performance per model to prevent overfitting or plateauing losses.

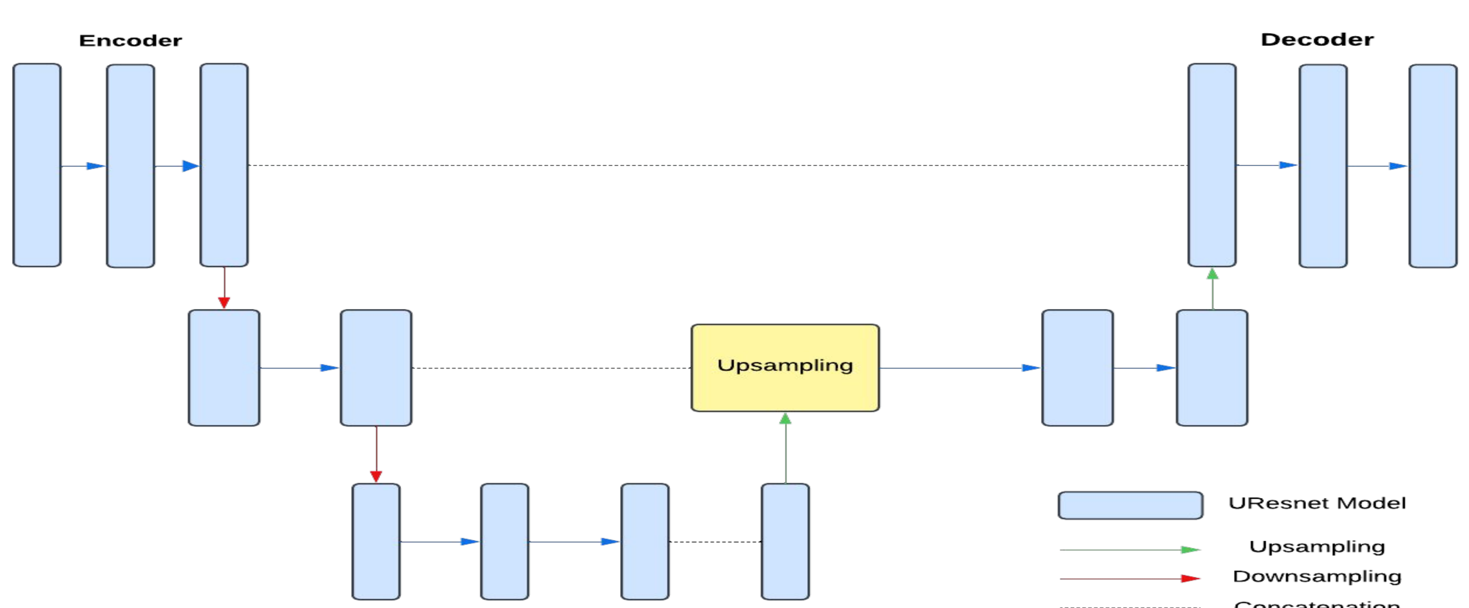

Figure 1: Flow Diagram of UResNet MTL Model: Single Layer Processing

Results

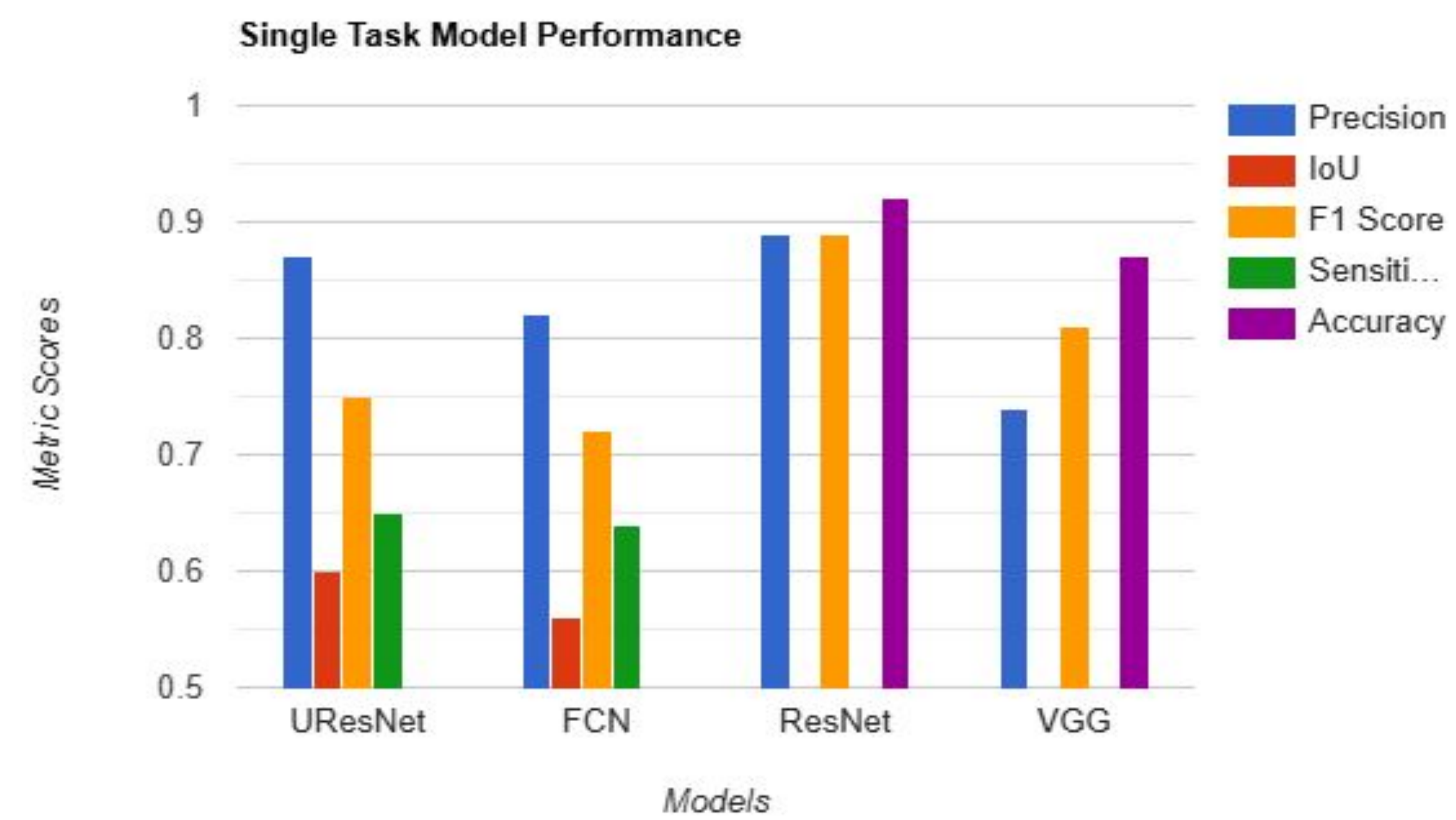

Figure 2: Bar Graph Showing Single-Task Model Performance

Metric	Classification(Score)	Segmentation(Score)
Accuracy	0.908	N/A
Precision	0.886	0.842
Recall(Sensitivity)	0.821	0.821
F1-Score	0.838	0.831
Intersection over Union(IoU)	N/A	0.711
Dice Coefficient	N/A	0.831

Table 1: Metric Table For MTL Model

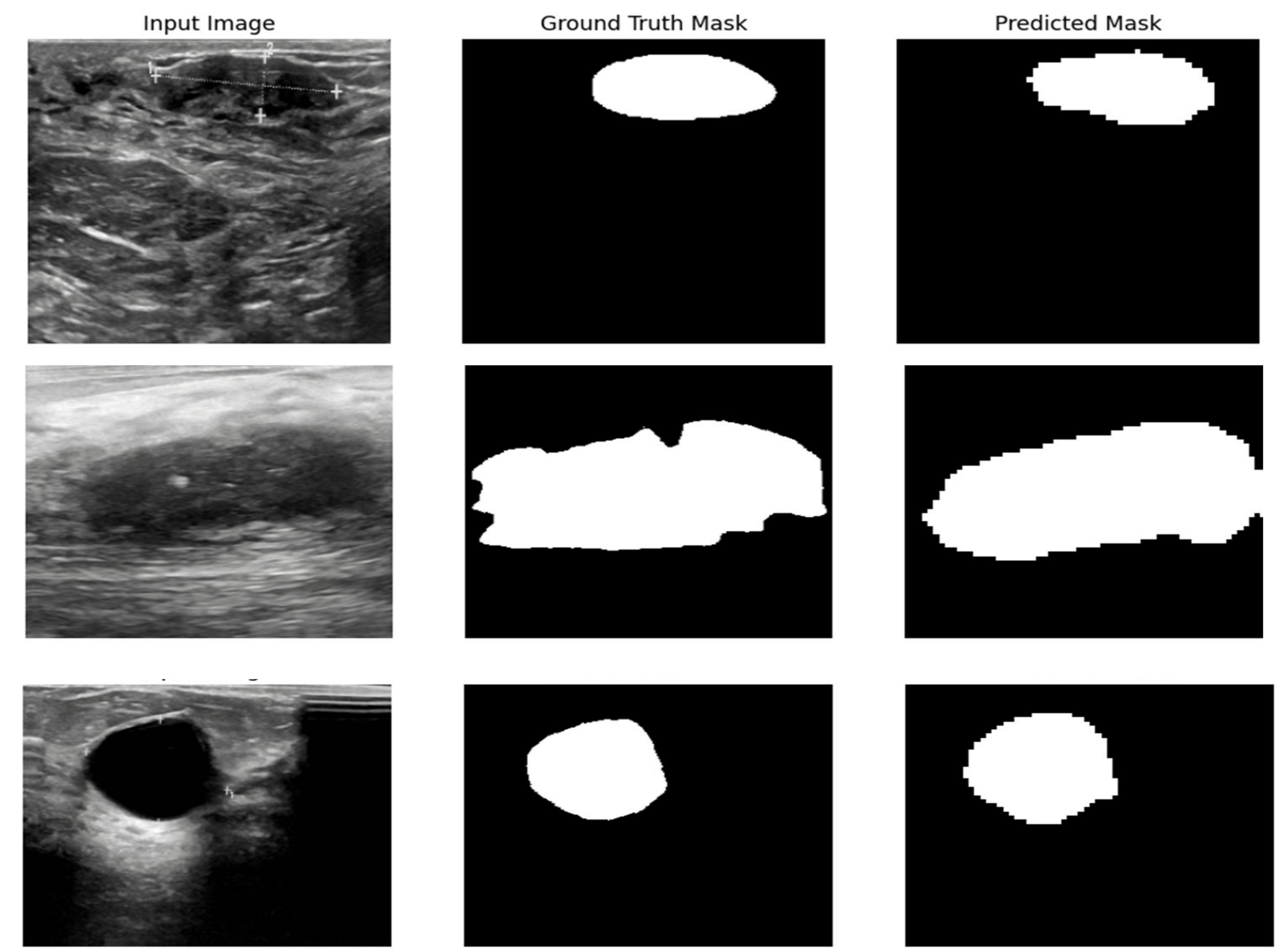

Figure 3: Segmentation Results of UResnet MTL Model

Conclusion

Although the predictions made by our multitask learning (MTL) model for breast cancer detection and segmentation are not yet perfect, they show significant promise for improving the efficiency and accuracy of diagnostic processes. The MTL model achieves comparable performance to single-task models, with slightly lower precision and recall for classification compared to ResNet and VGG. However, it demonstrates superior segmentation performance compared to UResNet and FCN, achieving higher IoU and Dice scores. This is most likely due to the larger weight assigned to segmentation, encouraging the model to focus more on that task. By using both classification and segmentation tasks simultaneously, the MTL model has the potential to reduce the time and cost of breast cancer detection, making it more accessible and timely for individuals. While further refinements are necessary, these results highlight the future potential of deep learning in transforming cancer diagnostics. With continued development, these techniques could significantly enhance clinical decision-making, leading to earlier detection and better outcomes for patients. This approach represents a more efficient diagnostic process for the healthcare system.

References

Acknowledgment

This material is based upon work supported by the National Science Foundation under Grants numbers CNS-2137791, and HRD-1834620. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation.

This material is based upon work supported by the National Science Foundation under Grants numbers CNS-2137791, and HRD-1834620. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation.