**Final Project Report for CS 184A/284A, Fall 2023**

**Project Title: Detect, Segment and Classify Breast Ultrasound Images**

**List of Team Members:**Sohyun Shin, 71686312, sohyuns1@uci.edu

Yujeong Yang, 39208092, [yujeony@uci.edu](mailto:yujeony@uci.edu)

Xinyu Zhang, 81641551, xzhang52@uci.edu

1. Introduction and Problem Statement

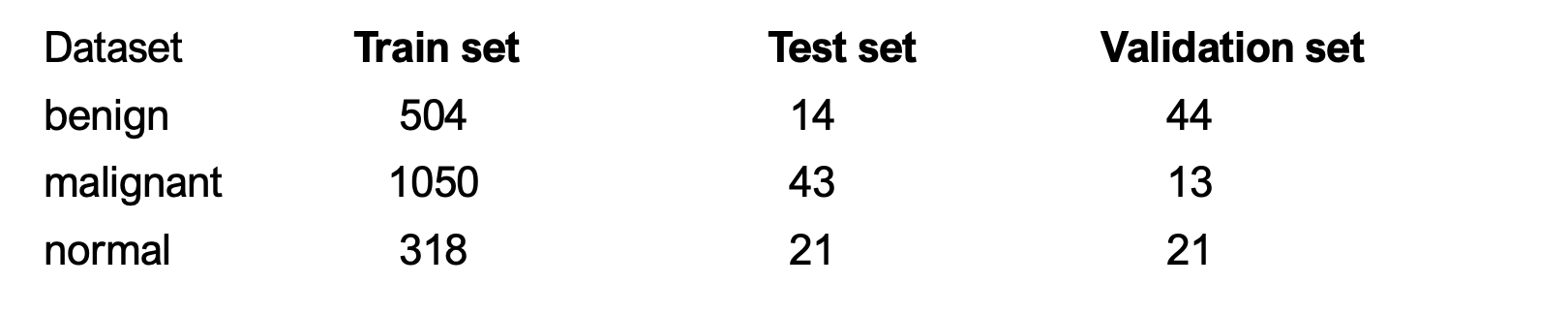
Breast cancer remains a significant health concern, being the most common cancer among women in the United States and the second leading cause of cancer-related deaths [1]. Early detection is pivotal in improving outcomes, and it is recommended to start screening every year for average women 45 or older [2]. This project aims to enhance breast cancer diagnosis through advanced technologies applied to breast ultrasound images. Specifically, we focus on detecting, segmenting, and classifying abnormal lesions within these images. The study utilized data from 600 patients, aged between 25 and 75 years, and employed a Faster R-CNN model as a baseline. We compared this baseline model with segmentation models specialized in medical images; U-Net, a convolutional neural network (CNN) tailored for medical images, and MedSAM, a vision transformer trained on medical data. Evaluation metrics used were precision, recall, F1 score for classification and segmentation, and IoU to gauge segmentation quality. By doing so, the project aims to answer pivotal questions: 1) Does segmentation outperform detection in classifying types of lumps in breast ultrasound? 2) If segmentation is superior, which segmentation model yields optimal results? 3) In the specific context of breast ultrasound images, does the transformer-based segmentation approach demonstrate superior accuracy and efficiency compared to other segmentation models?

2. Related Work on the Issue

The questions we ask align with the broader goal of refining breast cancer diagnosis by integrating cutting-edge technologies and methodologies, contributing to the growing body of knowledge in the field. As of December 2021, CNNs have been widely employed for their ability to extract valuable features from images [3]. While our project incorporates CNN methodologies, it distinguishes itself by adopting state-of-the-art segmentation models. Due to the novelty of these models, limited research has been conducted, with available studies like “BreastSAM: A Study of Segment Anything Model for Breast Tumor Detection in Ultrasound Images” [4], “Breast cancer detection from ultrasound images using attention U-nets model” [5], “Comparative Analysis of Segment Anything Model and U-Net for Breast Tumor Detection in Ultrasound and Mammography Images” [6], primarily addressing challenges associated with segmenting complex boundaries. Our study seeks to bridge this gap by assessing the performance of these advanced segmentation models in the context of breast ultrasound images, comparing them against established methods, and contributing valuable insights to the field.

3. Data Sets

The data [7] collected at baseline include breast ultrasound images among women in ages between 25 and 75 years old. This data was collected in 2018. The number of patients is 600 female patients. The dataset consists of 780 images with an average image size of 500\*500 pixels. The images are in PNG format. The ground truth images are presented with original images. The images are categorized into three classes, which are normal, benign, and malignant. Below is the example of the original input image and the image with mask.



4-1. Description of Technical Approach

In this project, we have implemented three different segmentation models and a neural network classification model on top of these segmentation models in order to examine if segmentation improves the result of classification of legions. We then compared these results with the regular detection model using faster R-CNN as a baseline model. The three segmentation models we utilized were Mask R-CNN, U-Net and MedSAM.

In the realm of segmentation and classification, we opted for Mask R-CNN, serving as a benchmark against which other segmentation models were evaluated. This model extends Faster R-CNN by predicting segmentation masks on each Region of Interest (RoI), enabling precise object detection and pixel-level instance segmentation.

U-Net, recognized as a biomedical image segmentation algorithm, played a pivotal role. It assigned class labels to individual pixels and delineated object boundaries. Through convolution and symmetrical deconvolution paths, U-Net extracted feature maps, shedding light on the distribution of lesions within the image and their categorization.

MedSAM [8], a specialized medical image segmentation model derived from SAM (a vision transformer; ViT), addressed the unique challenges posed by medical images. Trained on an extensive dataset comprising over a million medical image-mask pairs, MedSAM exhibited proficiency in segmenting lesions and distinguishing between benign and malignant cases. Unlike traditional CNNs, which rely on convolutional layers to process image patches, ViT employs a transformer architecture and is already pretrained on large-scale datasets. This means that it is versatile and reduces the need for extensive training on task-specific datasets, but it is also computationally expensive since the model is trained on extensive and high-resolution images.

In the domain of classification, following the stages of lesion detection or segmentation, we developed a neural network classifier. This classifier aimed to discern between benign and malignant lumps, exploring how additional information derived from segmentation could enhance precision in categorizing identified objects. The integration of segmentation and classification models culminated in a comprehensive approach designed to elevate breast cancer diagnosis through advanced image analysis.

4-2 Software

For U-Net, we used the original model introduced by Ronneberger et al [10]. On the basis of the original structure, the model was then modified by Yujeong Yang in terms of activation function of the final layer, loss function and parameters for fine-tuning.

For MedSAM pretrained model developed by BoWang's Lab was used [8]. The code used to image preprocessing to fit the model’s requirement, the fine-tuning of the model and also create and prompt bounding boxes were written by Sohyun Shin. Sohyun Shin also wrote the code for examining the results of MedSAM. The classification model to feed the segmentation results were written by all the team members.

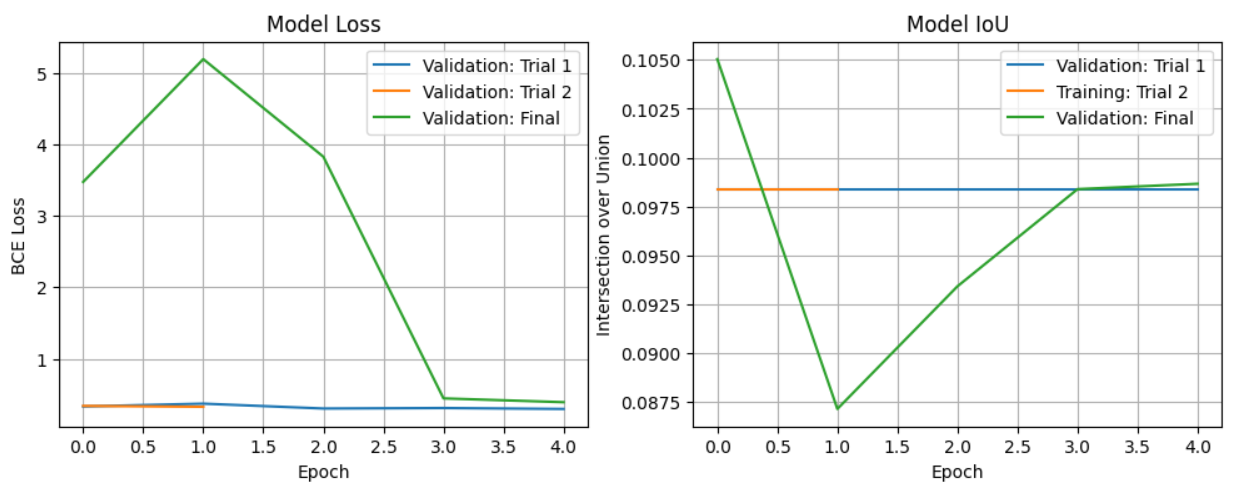
5. Experiments and Evaluation

5.1 Faster R-CNN / Mask R-CNN

5.2 U-Net

3 different sets of parameters was implemented for U-Net as below:

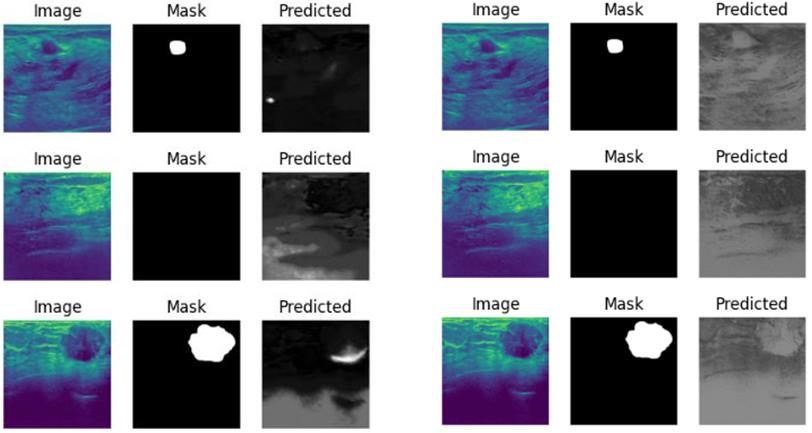
| Model | Dropout | Activation | Optimizer | Batch Size | Learning Rate |
| --- | --- | --- | --- | --- | --- |
| Trial 1 | 0.2 | ReLU/Sigmoid | Adam | 8 | 0.005 |
| Trial 2 | 0.2 | ReLU/Sigmoid | Adam | 8 | 0.01 |
| Final | 0.3 | ReLU/Sigmoid | Adam | 8 | 0.01 |



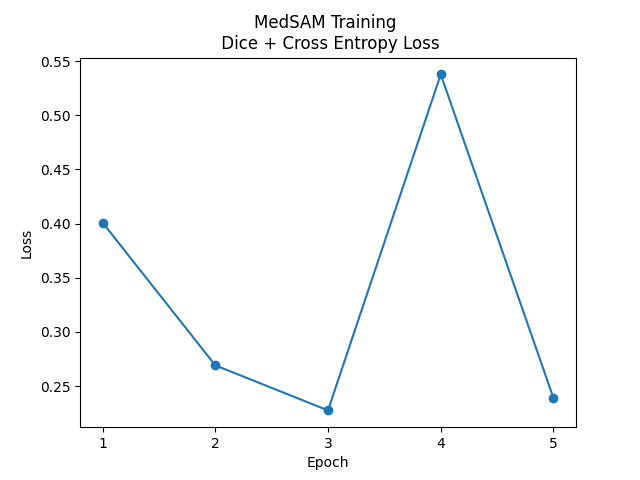
As you can see on the graph, the values of BCE loss and IoU are both low and almost fixed on the trial 1 and 2. For the low BCE loss, we adjusted the dropout rate to prevent overfitting. And For the fixed IoU, the learning rate was increased to enhance the performance of the model faster.

Since we evaluate the performance of the model based on IoU, the number of epochs for training and the model selection was also decided on IoU of the validation data. Training the model was stopped after 5 epochs, and we set that model as our best one.

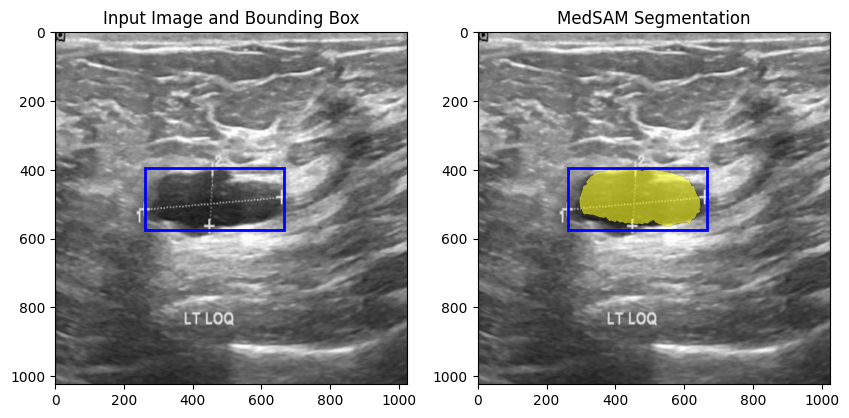
For the performance of the segmentation on test data, the average BCE loss was 0.7097, and the average IoU 0.0925. We can make several assumptions on the poor performance: lack of training, the structural characteristic of U-net, etc.

Some examples of prediction on the test data are as below (left: after epoch 1 of trial 1; right: after epoch 5 of final model):

5.3 MedSAM

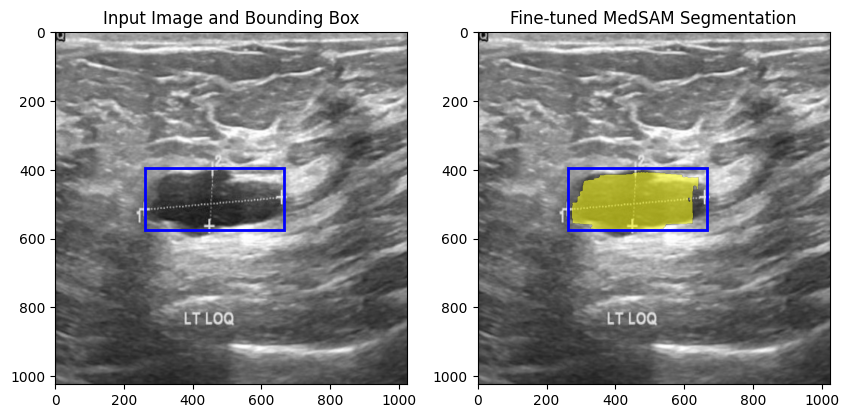


Pre-trained MedSAM without fine-tuning



Test fine-tuned IoU:

0.6578



Pre-trained mean IoU on validation: 0.8761

Fine-tuned mean IoU on validation: 0.6578

6. Discussion and Conclusion

7. Individual Contributions

Citation

[1] American Cancer Society, “Breast Cancer Statistics | How Common Is Breast Cancer?,” www.cancer.org, Jan. 12, 2023. https://www.cancer.org/cancer/types/breast-cancer/about/how-common-is-breast-cancer.html

[2] American Cancer Society, “ACS breast cancer screening guidelines,” www.cancer.org, Jan. 14, 2022. https://www.cancer.org/cancer/types/breast-cancer/screening-tests-and-early-detection/american-cancer-society-recommendations-for-the-early-detection-of-breast-cancer.html

[3] M. F. Mridha et al., “A Comprehensive Survey on Deep-Learning-Based Breast Cancer Diagnosis,” Cancers, vol. 13, no. 23, p. 6116, Dec. 2021, doi: https://doi.org/10.3390/cancers13236116.

[4] M. Hu, Y. Li, and X. Yang, “BreastSAM: A Study of Segment Anything Model for Breast Tumor Detection in Ultrasound Images” arXiv.org, May 21, 2023. https://arxiv.org/abs/2305.12447 (accessed Oct. 23, 2023).

[5] Muhammad Azaz Farooq, Z. Gong, Y. Liu, M. Zubair, A. Manzoor, and G. Zhang, “Breast cancer detection from ultrasound images using attention U-nets model” Fourteenth International Conference on Digital Image Processing (ICDIP 2022), Oct. 2022, doi: https://doi.org/10.1117/12.2643599.

[6] M. Ahmadi et al., “Comparative Analysis of Segment Anything Model and U-Net for Breast Tumor Detection in Ultrasound and Mammography Images” arXiv.org, Jun. 21, 2023. https://arxiv.org/abs/2306.12510 (accessed Sep. 07, 2023).

[7] W. Al-Dhabyani, M. Gomaa, H. Khaled, and A. Fahmy, “Dataset of breast ultrasound images,” Data in Brief, vol. 28, p. 104863, Feb. 2020, doi: https://doi.org/10.1016/j.dib.2019.104863.

[8] J. Ma and B. Wang, “Segment Anything in Medical Images,” arXiv.org, Apr. 24, 2023. https://arxiv.org/abs/2304.12306

[9] S. Gokhale, “Ultrasound characterization of breast masses,” Indian Journal of Radiology and Imaging, vol. 19, no. 3, p. 242, 2009, doi: <https://doi.org/10.4103/0971-3026.54878>.

[10] Ronneberger, O., Fischer, P., Brox, T. (2015). U-Net: Convolutional Networks for Biomedical Image Segmentation. In: Navab, N., Hornegger, J., Wells, W., Frangi, A. (eds) Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015. MICCAI 2015. Lecture Notes in Computer Science(), vol 9351. Springer, Cham. https://doi.org/10.1007/978-3-319-24574-4\_28