AI-Based Breast Cancer Classification Using Ultrasound Images

Suriya P. Mohan

University of Wisconsin – La Crosse

DS785: Capstone

Dr. Tracy Bibelnieks

May 1, 2022

Abstract

Breast cancer is the most frequently diagnosed cancer in women, and it is the leading cause of cancer-related deaths in women. Identifying the tumor in the early stage is essential to increase the survival rate of the patients. The breast ultrasound images combined with biopsy are clinically proven to be an effective method of diagnosing tumors. This project utilizes Artificial Intelligence (AI), such as deep learning models, to segment and classify the tumor in the ultrasound images. A segmentation deep learning model is utilized to detect the tumor in the ultrasound image, identify the region of interest (ROI), and draw a bounding box around the tumor. The radiologist confirms the ROI. Later the ROI is sent to the classification model to categorize cancer as benign or malignant. The radiologist can use the results from these models to make informed decisions, thereby reducing the manual errors involved in the current process. The AI technology assistance and radiologist expertise can lead to a process that is less prone to errors.

The models were trained on augmented data set of 452 malignant images and 706 benign images. The accuracy of this model is 76.2%, precision 80%, and F1 score of 72.7%.

Keywords: Ultrasound, artificial intelligence, deep learning, machine learning

Table of Contents

List of Tables

List of Figures

List of Equations

Chapter 1 - Introduction

1.1 Background

Breast cancer is the leading cause of cancer-related death in women, and it is the most frequently diagnosed cancer in women (Siegel et al., 2017). It is estimated that there will be 281550 new invasive cases of breast cancer, leading to the death of approximately 43500 in the United States (Shen et al., 2021). Therefore, it is essential to identify the tumor in the early stage before metastasis to prevent deaths and increase the survival rate of the patients.

There are several imaging techniques to diagnose breast tumors. The ultrasound image combined with biopsy has been clinically proved to be an effective method for detecting and classifying breast tumors (Chiao et al., 2019). However, there is a high false-positive rate in this process.

1.2 Challenges in the Existing Process

To classify the breast tumor as benign or malignant, radiologists evaluate the tumors in the ultrasound image using various features such as shape, margin, orientation, etc. The benign tumors do not invade other parts of the body, while malignant tumors are cancerous, i.e., they invade other parts of the body. Identifying the malignant tumor earlier and starting the treatment is essential to increase the survival rate of the patients. Suppose the radiologist identifies the tumor as malignant. In that case, the patients are sent for a biopsy for further studies, and if the radiologist predicts the tumor to be benign, patients are requested for revisits.

However, there are false positives in the current process, i.e., radiologists predict the tumor to be malignant, but the biopsy result was benign. Also, there are false negatives, i.e., radiologists predict the tumor to be benign, but later diagnosis detects the tumor as malignant.

The biopsy is an invasive procedure and can lead to infections, so the patients with benign predictions are not sent for biopsy. Also, false negatives can be fatal because early detection of malignant tumors is essential for treating patients and avoiding complications.

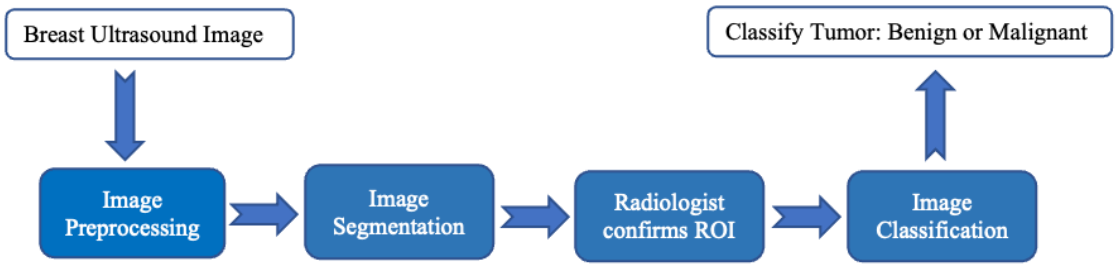
Further, new radiologists are entering the market who need guidance to make decisions. This project discusses the AI-based solution to provide information about the tumor to radiologists to mitigate these issues. The solution provides information about cancer so the radiologist can make an informed decision, thereby reducing the false-positive and negative involved in the existing process.

1.3 Proposed Approach

This project utilizes deep learning models such as convoluted neural networks (CNN) to detect and classify the tumors in the breast ultrasound images. The process involves image segmentation to detect the tumor in the ultrasound image, identify the region of interest (ROI) and draw the bounding box around the tumor in the ultrasound image. The image classification involves classifying the tumor as benign or malignant. Figure 1 displays the process flow.

Figure 1

Process flow



The radiologist feeds the breast ultrasound images into the system through a user interface. This ultrasound image is preprocessed to improve contrast and reduce noise before it is fed into the deep learning models. Later these images are center cropped to remove the text message on the

sides of the ultrasound image. Finally, this preprocessed image is fed to the image segmentation model, which detects the tumor, identifies the ROI, and draws the bounding box around the tumor.

The region of interest (ROI) is presented to the radiologist, who confirms if the model has identified the tumor in the ultrasound image correctly. For example, the radiologist could adjust the bounding box to encompass the entire tumor if the predicted bounding box missed any part of the tumor. Later the ROI confirmed by the radiologist is fed into a DL classification model to classify the tumor as benign or malignant. Finally, the classification results are presented to the radiologist, who makes informed decisions to reduce the errors.

Deep learning models are good at learning patterns from the previous trained images and apply the learning to new images to classify the tumor as benign or malignant. This guidance from the deep learning AI models, along with radiologist expertise, can reduce the error involved in the diagnosis process.

1.4 Organization of the Project

The project consists of eight discrete phases:

1. Explore the existing research performed in diagnosing breast cancer.
2. Data acquisition – Gather ultrasound images provided by Mayo clinic and public datasets.
3. Image preprocessing – To improve contrast and reduce noise in ultrasound images.
4. Image augmentation – Alter existing images to generate more images for training the model and reducing model overfitting.
5. Image segmentation – Train the segmentation model to identify the region of interest and draw a bounding box around the tumor in the ultrasound image.
6. Image classification – Train the classification model to classify the tumor as benign or malignant.

7. Model Evaluation – Identify the best model/approach.
8. UI development – Develop a user interface for the radiologist to interact with the system and view the results.

The following chapter discusses in detail each phase of the project.

Chapter 2 – Literature Review

There are numerous studies to segment and classify the breast tumors in the ultrasound sound images. Traditionally, CAD models were utilized to segment and classify the tumors. However, AI-based solutions have recently gained traction due to higher accuracy. This section discusses some of the existing studies conducted in the medical field to detect and classify breast tumors in ultrasound images.

2.1 Imaging Techniques to Diagnose Breast Tumors

Imaging techniques such as magnetic resonance imaging (MRI), mammography, and breast ultrasound are utilized to diagnose breast tumors. The MRI technique is a costly procedure, requires a longer scan time, and has a high rate of false positives. Further, MRI is sensitive to soft tissue lesions. MRI is recommended for patients who are at higher risk of breast cancer. (Chiao et al., 2019).

Mammography is the process of using a low-energy x-ray to examine the human breast. This process is utilized for screening and diagnosis. However, mammography has limitations for patients with dense breast tissues (Chiao et al., 2019).

The breast ultrasound technique converts the electrical signals to ultrasound. This ultrasound generates an image through computer processing based on the magnitude of the reflected sound wave and echoes time. The ultrasound process does not involve ionizing radiation. Due to this, breast ultrasound and mammography are commonly used for screening (Chiao et al., 2019).

Radiologists utilize these imaging techniques to diagnose cancer. First, they study the tumor in the ultrasound image using various characteristics such as tissue composition, shape, orientation, margin, calcifications, etc. and assign a BIRADS score. The BIRADS stands for Breast Imaging Reporting and Data System. It is a numerical score from 1 to 6 to indicate the level of abnormality in the ultrasound Image, mammography, and MRI. Table 1 lists the BIRADS score, their likelihood of malignancy, and the next step in the evaluation.

Table 1

BIRADS - Breast Imaging Reporting and Data System

Category	Description	Likelihood of malignancy	Next step in the evaluation
0	Incomplete; need additional imaging evaluation	Unknown	Special mammographic views, ultrasonography, magnetic resonance imaging
1	Negative	No mammographic evidence of malignancy	Routine screening
2	Benign finding	No mammographic evidence of malignancy	Routine screening
3	Probably benign finding	Less than 2 percent	Follow-up imaging at six and 12 months
4	Suspicious abnormality	12 to 25 percent	Fine-needle aspiration, percutaneous or surgical biopsy
5	Highly suggestive of malignancy	Greater than 95 percent	Percutaneous or surgical biopsy
6	Known malignancy	100 percent	Definitive surgery, chemotherapy, radiation therapy

If the radiologist evaluates the tumor to be malignant, the patient is sent for a biopsy to further evaluate the tumor. However, if the radiologist finds the tumor benign, the patients are requested for revisits for routine screening. There are several techniques to segment and classify the tumors.

2.2 CAD Model

Traditionally, computer-aided diagnosis (CAD) models were used to classify breast tumors in breast ultrasound images. Wu et al. (2019) compare traditional ML models and deep learning models to detect and classify tumors in breast ultrasound images. It was found that the accuracy of the Data Learning software (DLS) was better as the learning was consistent as compared to a human.

2.3 Pyradiomics

Romeo et al. (2021) utilize the PyRadiomics tool to extract quantitative data from the medical images, such as texture, shape, orientation, etc. This data is used for building the predictive models. This model is based on the machine learning algorithm through 5-fold stratified cross-validation such that in each fold, data is split and thus proves to be more robust. In this setting, the algorithm outperformed both the baseline reference and radiologist. It is also suggestive of utilizing the improved performance of the radiologist with the aid of the ML algorithm. This model had an accuracy of 82%, which slightly improved the radiologist's performance from 79.4% to 80.2%.

2.4 Deep Learning Models

In recent times, AI is gaining more popularity in image segmentation and classification due to better accuracy than traditional methods. Chiao et al. (2019) utilize the deep learning model Mask RCNN model to segment the tumors, draw a bounding box around the tumor and classify the tumors. Using this model, the overall accuracy in classifying the benign and malignant tumors was 85% (Chiao et al., 2019).

Seokmin Han et al. (2017) utilize the GoogLeNet deep learning model to classify the tumors. In this model, the method used histogram equalization followed by image cropping and

margin augmentation. The data sets used were trained on augmentation and with no augmentation data. These data sets had an area under the curve (AUC) of over 0.9. Their model showed an accuracy of 0.9, a sensitivity of 0.86, and a specificity of 0.96. This model showed promising results by classifying the malignant lesions.

Table 2

Results of the GoogLeNet Deep Learning Model

Measures	Percentage
Accuracy	0.9
Sensitivity	0.86
Specificity	0.96

2.5 Research Objective

This project captures the best practices of traditional and the latest AI methods. The traditional approach has benefits because image preprocessing techniques such as histogram equalization improve the images' contract, which improves the model accuracy. Additionally, displaying the segmented tumor and bounding box to the radiologist will increase the radiologist's confidence in the model's predicted results. If the predicted bounding box is not accurate, the radiologist can modify the bounding box to capture the tumor in the ultrasound image. The classification model will utilize the image in the new bounding box to classify the tumor. Human interference along with model prediction can lead to better accuracy.

Chapter 3 – Methodologies

In medical terms, histology is the classification of Breast tumors into benign or malignant categories. The project uses deep learning models to detect tumors in ultrasound images and predict histology. The process consists of two steps. The first step is image segmentation to detect the tumor and region of interest (ROI) in the ultrasound image. The second step is to classify the identified tumor as begin or malignant.

Image preprocessing techniques such as histogram equalization and bilateral filters are applied to the ultrasound images to improve the contrast and reduce the noise. Additionally, the ultrasound image is center-cropped to remove text and numbers in the image boundaries. A Mask RCNN model identifies the region of interest (ROI) and draws the bounding box around the tumor in the ultrasound image. The bounding box and segmented tumor are displayed to the radiologist for confirmation. Later VGG-16 model is used to classify the tumor as benign or malignant.

Training a deep learning model requires good infrastructure, so Google Colab is utilized. The models were tested and trained using GPU runtime in Google Colab. The deep learning models were developed using PyTorch, and OpenCV was utilized for image preprocessing tasks. The images are preprocessed in the data preparation step before being used for model training purposes.

3.1 Data preparation

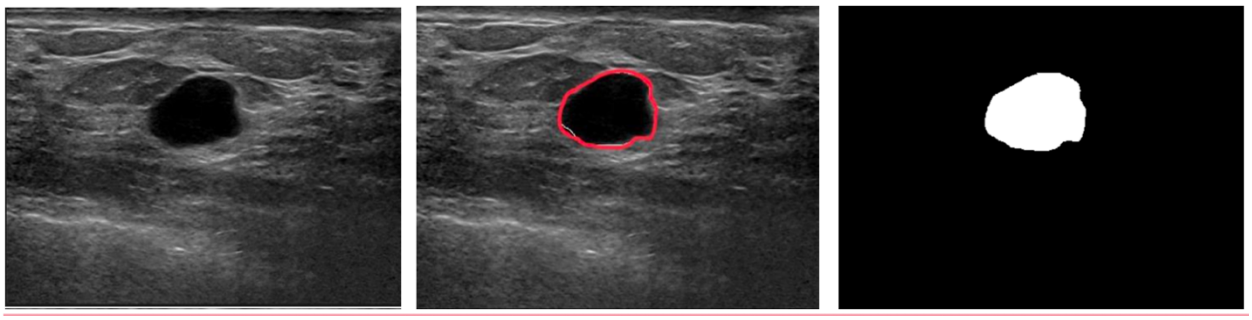
Mayo clinic provided the ultrasound Images for training and testing purposes. Publicly available Kaggle ultrasound images are used to create diversity in the dataset. Image data set from Mayo clinic consist of the ultrasound image and the outline of the tumor in the ultrasound image. This tumor outline in the ultrasound image is used to create a grayscale mask image

where the tumor is made white and the background is black. The ultrasound and mask images are essential for training segmentation models like Mask RCNN.

In figure 2, the image on the left is the original breast ultrasound image. The center image is an outline around the tumor created by the doctors. The ground truth mask is displayed on the right, which is created using the outline provided by the doctors.

Figure 2

Generating Ground Truth Mask

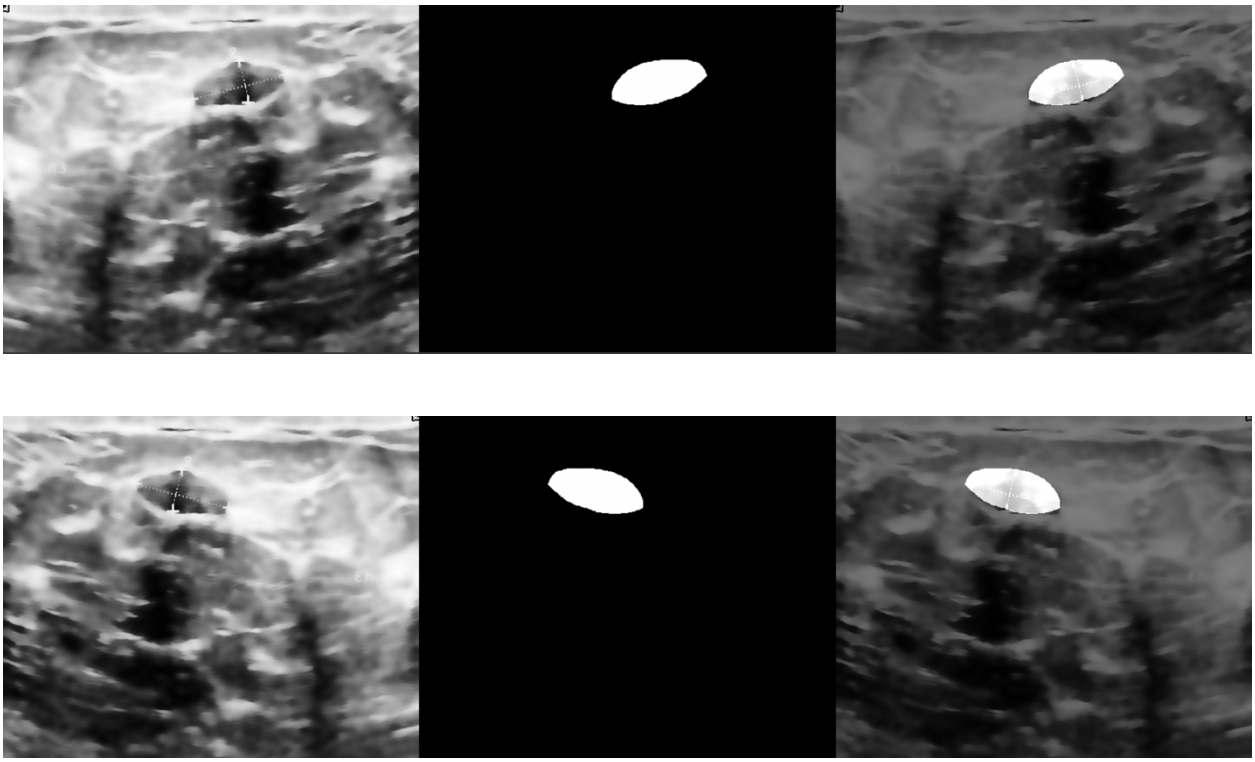


3.1.a Image Augmentation

Image augmentation is the process of altering the existing images to produce more images for model training purposes. Since the number of breast ultrasound images collected during the data collection process is limited, the image augmentation process generates more images for training purposes. The augmented data set will also reduce the model overfitting, i.e., the model performs well in the training dataset, but in the test data set, the model performs poorly. Figure 3 displays the image augmentation method of flipping the image to create mirror images. In the top row, the left image is the ultrasound image, the center is the ground truth mask, and the right image is the ground truth mask overlaid on the ultrasound image. The bottom row is the mirror image of the ultrasound image, ground truth mask, and ground truth mask overlaid on the ultrasound image.

Figure 3

Image augmentation - Mirror Image

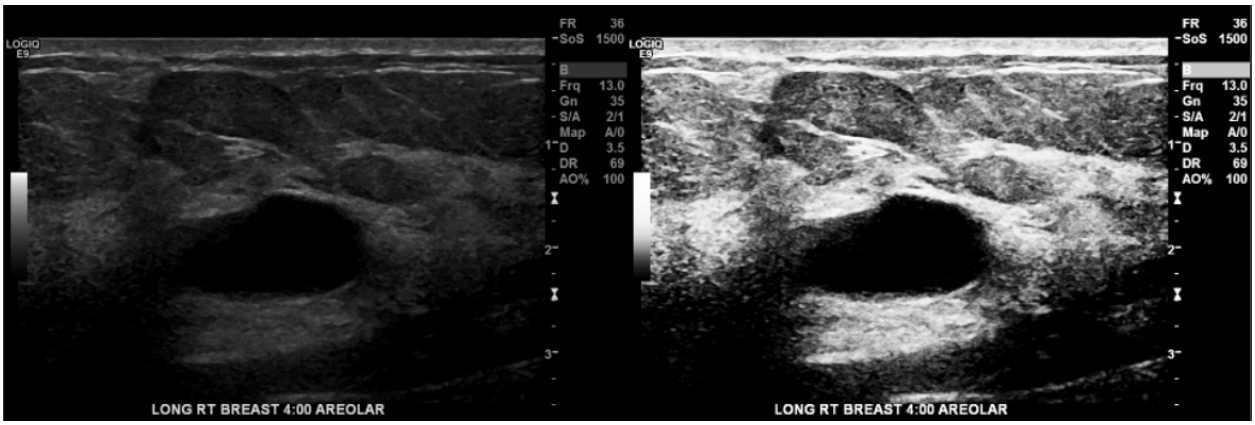


3.1.b Histogram equalization

Histogram equalization is an image processing technique to improve the contrast in ultrasound images. Histogram equalization is applied only on the ultrasound images and not the ground truth mask. In figure 4, the image on the left is the original breast ultrasound image, and the image on the right is after applying the histogram equalization.

Figure 4

Image Preprocessing - Histogram equalization

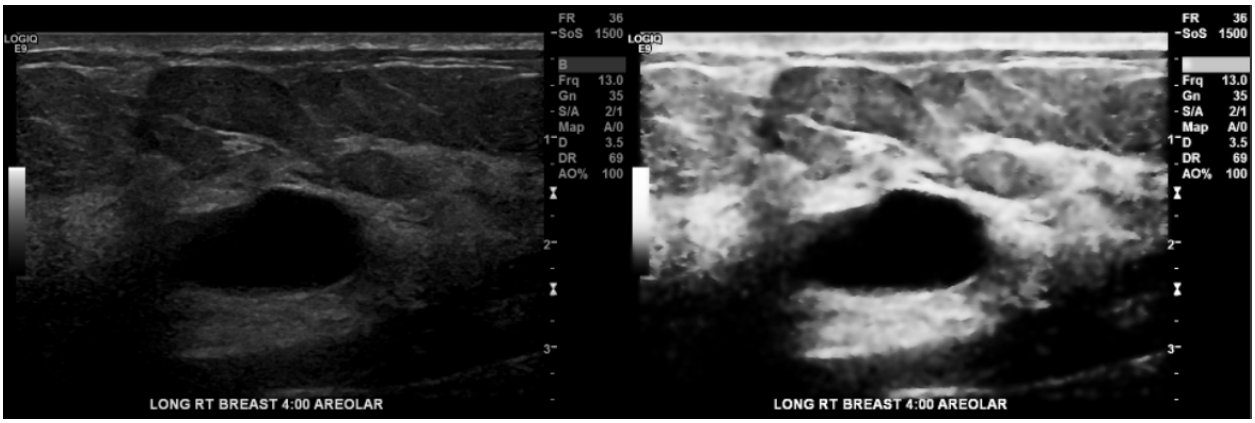


3.1.c Bilateral filter

Next, the image is preprocessed through a bilateral filter. This technique will reduce the noise and preserve the edge. In figure 5, the image on the left is the original breast ultrasound image, and the image on the right is after applying the histogram equalization and bilateral filters.

Figure 5

Image preprocessing - Bilateral filter



3.1.d Center crop

The image is center-cropped to remove the text messages around the ultrasound image. The center crop technique needs to be applied to both the ultrasound image and the generated mask so that the tumor position will remain the same in the ultrasound image and the mask. The image

on the left is the original ultrasound image after center cropping; the middle image is the center cropped mask image. In figure 6, the image on the right is the cropped mask overlaid on top of the ultrasound image to verify the tumor position is not shifted.

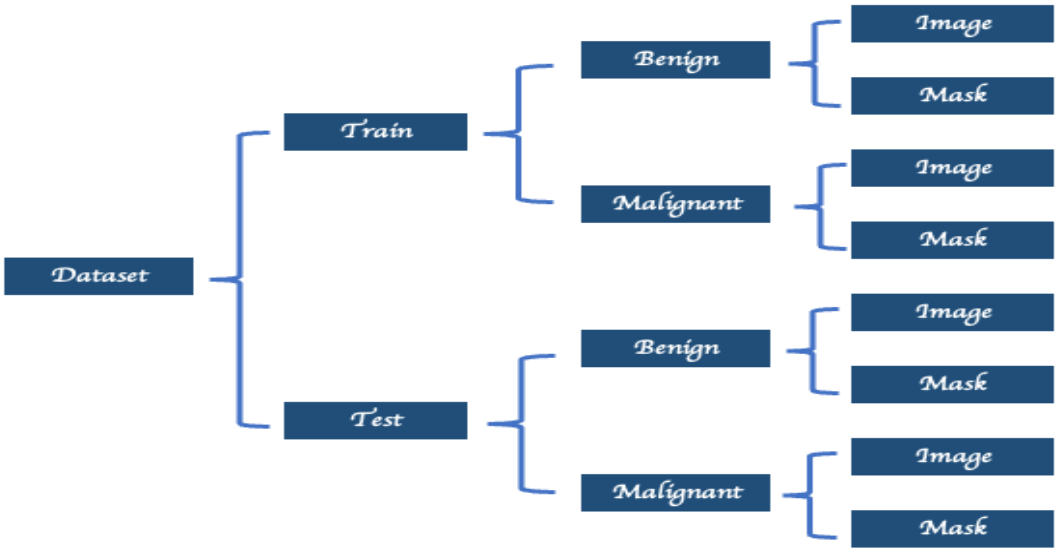
Figure 6

Center crop ultrasound image and mask

The data is split into train and test datasets. 90% of the images are used for training purposes, and 10% of the images are used for validation. The data is further organized into benign and malignant. Figure 7 is the tree diagram of how the data is organized for training and testing purposes.

Figure 7

Organize images into folders for model training and testing



3.2 Model Training

In this project, a few models were experimented with for image segmentation and classification. For image segmentation, UNET and Mask RCNN models were tried. However, since only a few hundred images are available for training, the UNET model did not perform well in detecting the tumor in the ultrasound images. However, Mask RCNN performed well even in detecting cancer in the ultrasound images with the limited training dataset. Therefore, the Mask RCNN model is utilized for image segmentation for this project.

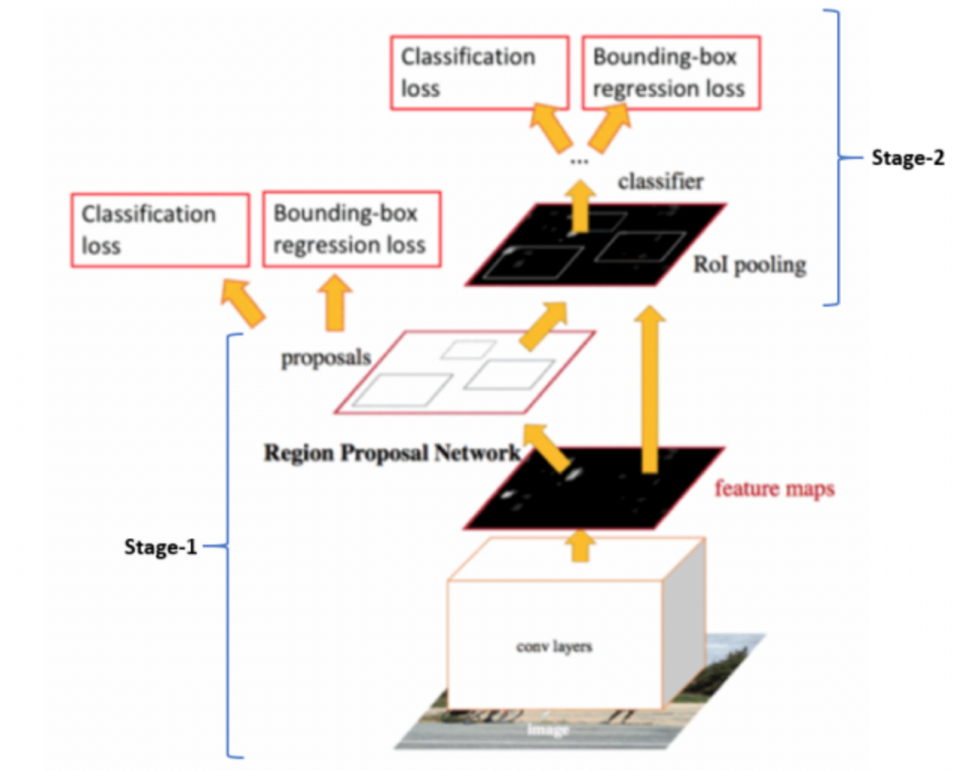
VGG16, VGG19, and Resnet50 models were experimented with for image classification. The VGG16 model outperformed other models. Therefore, the VGG16 model is used for image classification for this project.

3.2.a Image segmentation

The Mask RCNN model is utilized to segment the tumor and draw a bounding box around cancer in the ultrasound image. Mask RCNN is a region-based convoluted neural network, and it is the state-of-the-art model for image segmentation. Figure 7 displays the Mask RCNN architecture. The Mask RCNN works in two stages. In the first stage, the tumor in the image is identified. These regions are called region proposals. In the second stage, the network identifies the bounding box and the class for each region.

Figure 8

Mask RCNN architecture



The ultrasound image and ground truth mask are required for training the Mask RCNN model. The ground truth masks are generated from the doctor's tumor outline. A pre-trained Mask RCNN model trained on the COCO dataset is utilized for this project. A pre-trained model is already trained on images, using the existing model weights as a starting point for greater accuracy right from the first epoch.

The image dataset received from Mayo Clinic is limited, so the image augmentation technique is used to alter the existing images to produce more images. Also, image preprocessing techniques such as histogram equalization and bilateral filters are applied to ultrasound images to improve the contrast and reduce noise, so the deep learning models perform better.

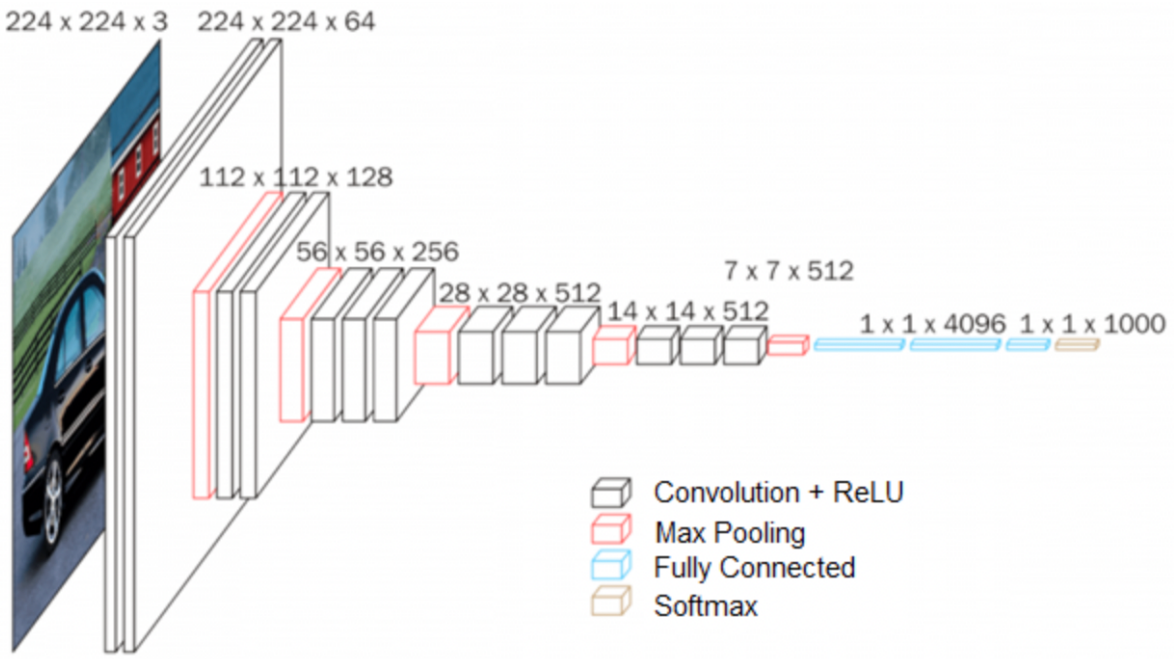
After training the Mask RCNN model for 100 epochs on a few hundred images, the model is validated on the test images. The predicted mask generated by the Mask RCNN model on the test images is stored in google drive, so these generated masks can be utilized in the classification model to classify the tumor.

3.2.b Image Classification

The predicted mask from the Mask RCNN model is fed to the VGG 16 model for tumor classification. VGG-16 is a 16-layer deep convolutional neural network commonly used for image classification. Figure 8 displays the architecture of the VGG-16 model.

Figure 9

VGG-16 architecture



A pre-trained VGG-16 model is used for training purposes. A pre-trained model is already trained on the images, so the training process can utilize the existing model weights to have better accuracy right from the first epoch. The final classification layer is modified to match

the number of target classes. For tumor classification, there will be two classes, benign and malignant. During the training, only the final layer of the model is trained. The weights of the intermediate layers of the model are frozen. The accuracy of the model is tested in each epoch. If the model accuracy is better than the accuracy from the previous epoch, the model weights are stored in a checkpoint directory. The model is trained for 25 epochs. The loss and accuracy in each epoch are stored for comparison.

The VGG-16 model is trained on both ultrasound images, and the ground truth mask and two classification models are prepared. The model trained on the ground truth mask is utilized to classify the predicted mask from the Mask RCNN model. The model trained on the ultrasound image classifies the ultrasound image directly. Developing two models will provide a benchmark compare to compare the performance metrics.

The test dataset is prepared by randomly choosing an equal number of benign and malignant cases. These models are not trained on the test dataset. The following chapter covers the performance metrics for these trained models using the test dataset.

Chapter 4 – Findings / Results

The project's objective is to detect and segment the tumor in the ultrasound images and present the results to the radiologist. The predicted result contains the segmented tumor and the bounding box. The radiologist confirms the bounding box indicated by the model encompasses the entire tumor. If the bounding box does not cover the whole tumor, then the radiologist adjusts the bounding box. This effort will help identify the region of interest in the ultrasound image. The mask is generated using a deep learning model, later fed to another deep learning model to classify the tumor as benign or malignant. Finally, the classification results are presented to the radiologist, who makes informed decisions.

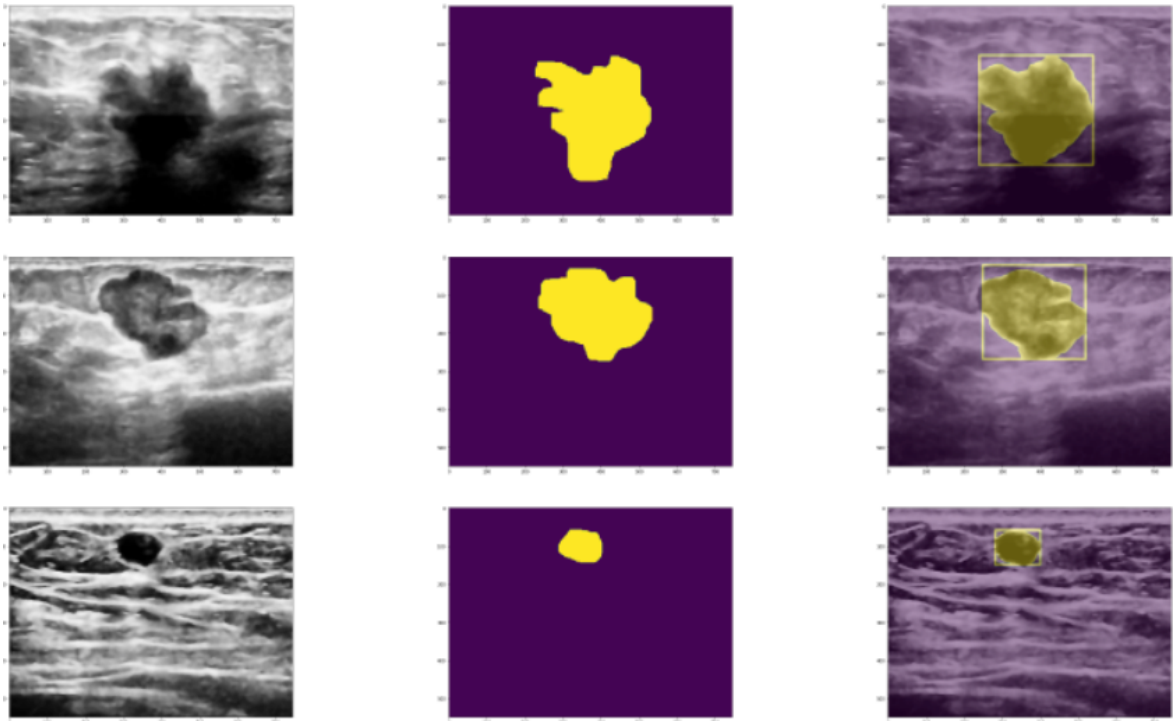
There are two ways to classify the tumor. First, the classification can be performed on the ultrasound image using a deep learning model. The second approach is to run the classification on the predicted mask from the Mask RCNN model. The classification is also performed on the ground truth mask provided by the doctors to calculate the baseline validation metrics. The confusion matrix and validation metrics are utilized to compare the classification results on the ground truth mask, predicted mask, and ultrasound image in the next section.

4.1 Image segmentation

The Mask RCNN is trained for 100 epochs on benign and malignant ultrasound images. Epoch is the number of times the algorithm will work on the entire training dataset. After training the model on the training images, the model is validated using the test images. Figure 10 displays the predicted results from the Mask RCNN model on the test images. The left image is the preprocessed image; the center image is the mask generated by the doctors; the right image is the mask, and the bounding box is generated by the model overlaid on top of the ultrasound image for comparison.

Figure 10

Predicted mask and bounding box



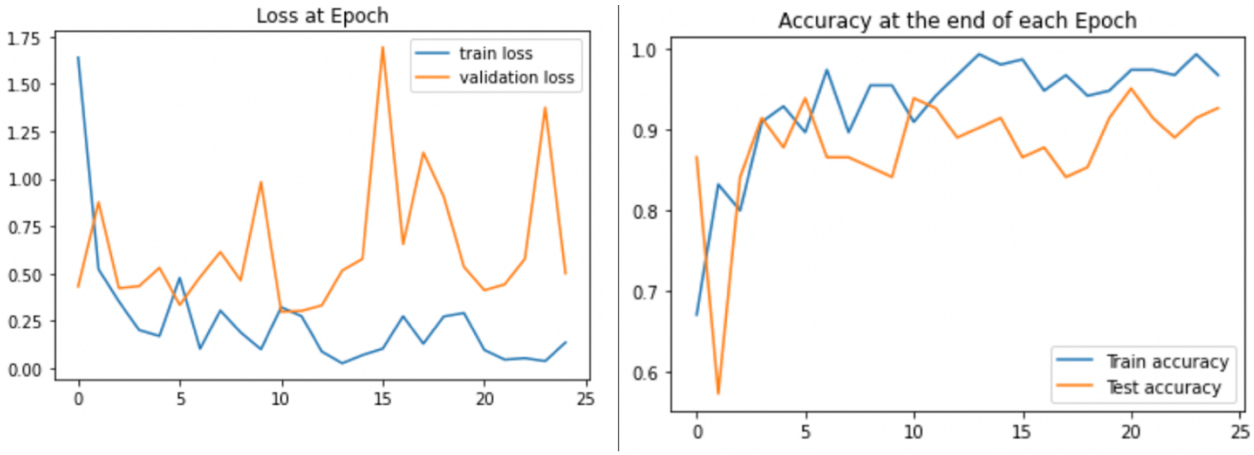
The predicted mask from the mask RCNN is saved to a google drive location and later utilized by the classification model – VGG 16 to classify the tumor as benign or malignant

4.2 Accuracy and Loss plot:

The classification VGG-16 model is trained on ultrasound images and the ground truth masks to produce two classification models. The classification model is trained for 25 epochs, and each epoch's accuracy and loss are stored. Loss is a penalty for bad prediction, and it is zero if the model correctly classifies all the ultrasound images as benign or malignant. Figure 9 displays the loss and the accuracy for each epoch. After ten epochs, the model accuracy fluctuated between 86% to 91%, and training losses were around 0.25.

Figure 11

Accuracy and Loss plot for training and test datasets



4.3 Confusion Matrix

The confusion matrix is a table. It defines the performance of a classification model. The predicted results from the model and actual results from the biopsy are needed to create a confusion matrix. The dataset is prepared for validation by randomly selecting an equal number of benign and malignant cases that are not used for model training purposes. The confusion matrix is prepared by comparing the predicted results with the actuals. The confusion matrix consists of the following cells –

Figure 12

Confusion Matrix

		Predicted Class: Positive	Predicted Class: Negative	
Actual Class	**Positive**	True Positive (TP)	False Negative (FN) **Type II Error**	**Sensitivity** $\frac{TP}{(TP+FN)}$
	Negative	False Positive (FP) **Type I Error**	True Negative (TN)	**Specificity** $\frac{TN}{(TN+FP)}$
		Precision $\frac{TP}{(TP+FP)}$	**Negative Predictive Value** $\frac{TN}{(TN+FN)}$	**Accuracy** $\frac{TP+TN}{(TP+TN+FP+FN)}$

The malignant tumors are cancerous, so they are considered positive, and the benign tumors are non-invasive, so they are deemed negative. The true positive, false positive, true negative, and false negative in the confusion matrix are calculated as below –True positive (TP): The ultrasound image is predicted as malignant, and the actual biopsy result is malignant.

- False positive (FN): The ultrasound image is predicted as malignant; however, the actual biopsy result is benign.
- True negative (TN): The ultrasound image is predicted as benign, and the actual biopsy result is benign.
- False negative (FN): The ultrasound image is predicted as benign; however, the actual biopsy result is malignant.

Table 3 displays the confusion matrix for the ground truth mask, predicted mask, and ultrasound image. In the table, yellow cells represent the confusion matrix for the ground truth mask. The ground truth mask is ideal since doctors review the ultrasound image and outline the tumor, and the mask is generated, so ground truth mask results provide the benchmark numbers. The cells in green are the confusion matrix for classifying the predicted mask from the Mask RCNN model. The cells in blue are the confusion matrix for classifying the ultrasound image directly.

Table 3

Classification Results

Total Malignant Test cases: 30

Total Benign Test cases: 32

		Predicted					
		Ground Truth Mask		**Predicted Mask**		**Ultrasound Image**	
		Malignant	**Benign**	**Malignant**	**Benign**	**Malignant**	**Benign**
Actual	**Malignant**	24	6	20	10	17	13
	Benign	2	31	5	28	2	31

The validation metrics such as accuracy, precision, recall, and F1 score are utilized for validating the model results. These metrics can be calculated using a confusion matrix.

4.4 Accuracy

The accuracy metric describes how well the model performs across all the classes. It is the ratio of the correct predictions by total test cases, i.e., the ratio sum of true positives and true negatives by total test cases. Equation 1 displays the formula for calculating the accuracy.

Equation 1

Accuracy Formula

$$\frac{TP + TN}{TP + FP + TN + FN}$$

Table 4 displays the accuracy of the models trained on the ground truth mask, the predicted mask, and the ultrasound images.

Table 4

Model Accuracy

	Ground truth mask	**Predicted Mask**	**Ultrasound images**
Accuracy	0.87	0.762	0.762

4.5 Precision

The precision is the ratio between the number of correctly classified positive samples to the total number of samples classified as positives. The precision metrics indicate how reliable the model is in classifying the positive samples. The metric is calculated by dividing the true positive (TP) by the sum of true positive (TP) and false positives (FP). Equation 2 displays the formula for calculating the precision.

Equation 2

Precision formula

$$\frac{TP}{TP + FP}$$

Table 5 displays the precision of the models trained on the ground truth mask, the predicted mask, and the ultrasound images.

Table 5

Precision metric of the models

	Ground truth mask	Predicted Mask	Ultrasound images
Precision	0.92	0.8	0.895

4.6 Recall

The recall is the ratio between correctly classified positive cases and the total number of actual positive cases. It is the model's ability to detect positive cases. The metric is calculated by dividing the true positive (TP) by the sum of true positive (TP) by false negative (FP). Equation 3 displays the formula for calculating the recall.

Equation 3 *Recall formula*

$$\frac{TP}{TP + FN}$$

Table 6 displays the recall of the models trained on the ground truth mask, the predicted mask, and the ultrasound images.

Table 6

Recall metric of the models

	Ground truth mask	Predicted Mask	Ultrasound images
Recall	0.8	0.667	0.567

4.7 F1 Score

The F1 score metric is the harmonic mean of precision and recall. It takes the contribution from both the precision and the recall, so the higher the F1 score better the model. Equation 4 displays the formula for calculating the F1 score.

Equation 4

F1 score formula

$$\frac{2 * precision * recall}{precision + recall}$$

Table 7 displays the F1 score of the models trained on the ground truth mask, the predicted mask, and the ultrasound images.

Table 7

F1 score of the models

	Ground truth mask	Predicted Mask	Ultrasound images
F1 score	0.8	0.667	0.567

Table 8 displays the summary of the validation metrics for the ground truth mask, predicted mask, and ultrasound image.

Table 8

Summary of validation metrics

		Ground Truth Mask	Predicted Mask	UltraSound Image
Metrics	**Accuracy**	0.873015873	0.761904762	0.761904762
	Precision	0.923076923	0.8	0.894736842
	Recall	0.8	0.666666667	0.566666667
	F1 Score	0.857142857	0.727272727	0.693877551

The model accuracy for the predicted mask and ultrasound image is 76.2%. However, the F1 score for the predicted mask model is slightly better than the model trained on the ultrasound image. Furthermore, a model trained on the ultrasound image is a black box because the users are unaware of the Region of Interest (ROI) used for classifying the tumor. In the model trained on the predicted mask, the mask is already reviewed by the radiologist. This approach can lead to better confidence in the end-users. As a result of this study, the model trained on classifying the predicted mask is selected as the preferred method for classifying tumors.

Chapter 5 – Next Steps

A semester-long timeline is certainly a limitation for this large-scale project. While every effort was made to complete the project within the timeline, few activities can be experimented with to enhance the model and the process further. They are discussed as follows -

5.1 Larger dataset

The number of breast ultrasound images gathered from the Mayo clinic and online is limited. Therefore, image augmentation was utilized to modify the collected images to generate more images. However, even after the image augmentation, the number of malignant images was 450, and the number of benign images was around 700. Some deep learning models did not produce expected results because of this limited data set. Therefore, gathering more data points is essential to improve the existing model tried in this project and to experiment few other models.

5.2 Model overfitting

Training a deep learning model on insufficient data set results in overfitting (Al-Dhabyani et al., 2019). There are two ways to overcome this problem. The first way is to utilize image augmentation to generate more images. The second approach is the regularization process by adding an additional penalty in the error function. Image augmentation is utilized in this project to avoid overfitting. However, regularization techniques such as drop-out can be experimented with to enhance the model further.

5.3 Mask RCNN to perform both segmentation and classification

The Mask RCNN model can perform both image segmentation and classification. This project utilized a pre-trained RCNN model, which was already trained on the MS COCO dataset. The MS COCO (Microsoft – Common Objects in Context) is multiclass object detection, segmentation, and captioning dataset. Since the RCNN model trained on the multiclass data

COCO dataset, the ultrasound image prediction yielded more classes than two classes - benign or malignant. If thousands of ultrasound images are gathered, the Mask RCNN model can also be utilized for classification. Instead of using two models, one for classification and another for segmentation, one model, Mask RCNN, can be utilized for both segmentation and classification.

5.4 Segment and classify normal ultrasound images

The project focuses on classifying benign and malignant tumors in ultrasound images. However, models are not trained to identify the normal ultrasound images without tumors. If a sufficient number of normal ultrasound images are available, then the model can be trained to detect and classify the normal, benign, and malignant ultrasound images.

Conclusion

Early detection and identification the breast tumor is essential to provide treatment to the patients, thereby increasing their survival rate. In the existing process, a radiologist classifies the breast tumor by studying the tumor size, orientation, shape, etc., in the ultrasound image. However, the current process has inaccuracies, and any failure to detect the malignant tumor can be fatal. This project outlines a method utilizing a deep learning model to detect, segment, and draw a bounding box around cancer in the ultrasound image. Later, the radiologist confirms the predicted mask covers the entire tumor. Finally, another deep learning model utilizes this mask to classify the tumor as benign or malignant. The final results from the classification model are presented to the radiologist to make an informed decision to reduce the false-negative and positive.

Therefore, the proposed method can work in tandem with radiologists to reduce the human errors involved in the current process.

Acknowledgments

Firstly, I thank my research supervisors, Dr. Song Chen and Dr. Jeffery Baggett, for providing me the opportunity to work on this project. With their guidance, funding and prompt responses, I was able to make this project successful.

Also, I thank my fellow students working on this project: Adam Silberfein, David Halama, Justin Hall, Lucas Spellman and Simon Wagner. With your support I was able to advance in this new area, learn new things and execute this project at given timelines.

I want to thank, my wife Raksha Kadam for the constant support and encouragement throughout this project. Without your help it is not feasible to complete this part-time project on time.

Lastly, I wanted to thank Dr. Ellis. It is hard to learn a new area with little or no background, but you made it look simple providing your guidance.

Glossary

Term	Expansion
AI	Artificial Intelligence
AUC	Area under the curve
BIRADS	Breast Imaging Reporting and Data System
CAD	Computer-Aided Diagnosis
CNN	Convoluted Neural Network
DL	Deep Learning
DLS	Data Learning software
ML	Machine Learning
MRI	Magnetic Resonance Imaging
RCNN	Region-based Convoluted Neural Network
ROI	Region of Interest
VGG-16	A 16-layer deep learning model used for image classification

Appendix A: Code

The code for this project can be found at the following location

Google Colab

https://colab.research.google.com/drive/1GZ3og3IlTbvxzOuIslYiCt9Zpr2s1t6Z?usp=sharing

Github link

https://github.com/Suriya0404/Deep-Learning/blob/master/Final_AI_based_breast_tumor_classification_using_ultrasound_images.ipynb

References

Al-Dhabyani, W., Gomaa, M., Khaled, H., & Fahmy, A. (2019). Deep learning approaches for data augmentation and classification of breast masses using ultrasound images. *International Journal of Advanced Computer Science and Applications (IJACSA)*, *10*(5). https://doi.org/10.14569/IJACSA.2019.0100579

Chiao, J.-Y., Chen, K.-Y., Liao, K. Y.-K., Hsieh, P.-H., Zhang, G., & Huang, T.-C. (2019). Detection and classification the breast tumors using mask R-CNN on sonograms. *Medicine*, *98*(19). https://doi.org/10.1097/md.0000000000015200

Han, S., Kang, H.-K., Jeong, J.-Y., Park, M.-H., Kim, W., Bang, W.-C., & Seong, Y.-K. (2017). A deep learning framework for supporting the classification of breast lesions in ultrasound images. *Physics in Medicine & Biology*, *62*(19), 7714–7728. https://doi.org/10.1088/1361-6560/aa82ec

Hou, D., Hou, R., & Hou, J. (2020). On-Device training for breast ultrasound image classification. *2020 10th Annual Computing and Communication Workshop and Conference (CCWC)*. https://doi.org/10.1109/ccwc47524.2020.9031146

Liao, W.-X., He, P., Hao, J., Wang, X.-Y., Yang, R.-L., An, D., & Cui, L.-G. (2020). Automatic identification of breast ultrasound image based on supervised block-based region segmentation algorithm and features combination migration deep learning model. *IEEE Journal of Biomedical and Health Informatics*, *24*(4), 984–993. https://doi.org/10.1109/jbhi.2019.2960821

Shen, Y., Shamout, F. E., Oliver, J. R., Witowski, J., Kannan, K., Park, J., Wu, N., Huddleston, C., Wolfson, S., Millet, A., Ehrenpreis, R., Awal, D., Tyma, C., Samreen, N., Gao, Y.,

Chhor, C., Gandhi, S., Lee, C., Kumari-Subaiya, S., & Leonard, C. (2021). Artificial intelligence system reduces false-positive findings in the interpretation of breast ultrasound exams. *Nature Communications*, *12*(1). https://doi.org/10.1038/s41467-021-26023-2

Romeo, V., Cuocolo, R., Apolito, R., Stanzione, A., Ventimiglia, A., Vitale, A., Verde, F., Accurso, A., Amitrano, M., Insabato, L., Gencarelli, A., Buonocore, R., Argenzio, M. R., Cascone, A. M., Imbriaco, M., Maurea, S., & Brunetti, A. (2021). Clinical value of radiomics and machine learning in breast ultrasound: A multicenter study for differential diagnosis of benign and malignant lesions. *European Radiology*, *31*(12), 9511–9519. https://doi.org/10.1007/s00330-021-08009-2

Siegel, R. L., Miller, K. D., & Jemal, A. (2017). Cancer statistics, 2017. *CA: A Cancer Journal for Clinicians*, *67*(1), 7–30. https://doi.org/10.3322/caac.21387

Wu, G. G., Zhou, L. Q., Xu, J. W., Wang, J. Y., Wei, Q., Deng, Y. B., Cui, X. W., & Dietrich, C. F. (2019). Artificial intelligence in breast ultrasound. *World Journal of Radiology*, *11*(2), 19–26. https://doi.org/10.4329/wjr.v11.i2.19

Xian, M., Zhang, Y., Cheng, H. D., Xu, F., Huang, K., Zhang, B., Ding, J., Ning, C., & Wang, Y. (2021). BUSIS: A benchmark for breast ultrasound image segmentation. *Arxiv*, *2*. https://arxiv.org/abs/1801.03182