Computer-Aided Diagnosis (CAD) of Breast Cancer from Ultrasound Images

Adam Silberfein

University of Wisconsin – La Crosse

DS 785: Capstone

August 11, 2021

Abstract

Breast ultrasound has long been one the most preferred modalities of breast cancer detection due to its relatively low cost, ease of access, and noninvasive nature. Advancements in imaging technology and improvements in technique and standardization have led to corresponding increases in accuracy, but the potential for an additional boost exists from the application of artificial intelligence and machine learning. This paper explores how the potential of a quantitative model and associated machine learning prediction can assist radiologists currently using what is primarily a qualitative model to assign a BI-RADS score and characterize breast lesions as benign or malignant. The steps taken involve having an expert radiologist annotate a publicly available data set of 135 breast ultrasound images according to an established rubric, using image processing techniques to convert certain highly predictive features of that rubric to a numerical representation, and then feeding that quantitative representation into a machine learning model to predict malignancy of a validation set of images. The resulting model confirms that not only were lesion shape and circumscription important features for classification, but so was internal echotexture, which is much more difficult to differentiate to the naked eye. The model will be used to classify future larger data sets provided by our client, either in its current manifestation or after retraining on a subset of those new images.

Keywords: Breast cancer, breast ultrasound, machine learning, radiomics, computer-aided diagnosis

Contents

List of Tables

List of Figures

Chapter 1: Introduction

1.1 Background

Breast cancer is the most diagnosed cancer among women in the United States, and is now the most common cancer worldwide, accounting for about 12% of all new cancer cases. Though death rates have been dropping, about 43,000 women in the U.S. are expected to die from breast cancer in 2021 (Breastcancer.org, n.d.). As with many cancers, early detection and diagnosis is vital in managing the disease. When the cancer is localized (no sign of spread outside the breast), the 5-year survival rate is as high as 99%.

Along with routine mammograms and physical exams, breast ultrasound (BUS) is one of the tools used in breast imaging to help detect and diagnose abnormalities in breast tissue. Often, the procedure is used as a safe and noninvasive next step when something suspicious is found during a physical exam, mammogram, or breast MRI. Sound waves are sent through the skin to produce in image of the internal tissue and structure, allowing the physician to determine whether an abnormality is solid (non-cancerous lump or cancerous tumor), filled with fluid (for example, a benign cyst), or both cystic and solid.

The project described below is part of a joint effort between The Mayo Clinic and researchers at the University of Wisconsin-La Crosse. The Mayo Clinic is a large nonprofit hospital system with major campuses in Minnesota, Arizona, and Florida. The Mayo Clinic Health System, founded in 1992, consists of hospitals, clinics, and medical service centers throughout Minnesota, Iowa, and Wisconsin. Their mission is to provide the highest quality care to every patient through integrated clinical practice, education, and research. They serve more than 1.3 million patients every year from all 50 states and 140 countries (Mayo Clinic, n.d.).

1.2 Terminology

BI-RADS is an acronym for Breast Imaging Reporting and Data System. It is a numerical scoring system running from 0 to 6 that helps radiologists communicate the level of abnormality in a breast image (mammogram, ultrasound, or breast MRI). In order to assign a BI-RADS score, the radiologist will consider various characteristics of the BUS image:

- Tissue Composition
- Shape, orientation, margin, echo pattern, and posterior features of any masses
- Calcifications
- Associated features, such as architectural distortion, duct changes, vascularity, and elasticity
- Special cases such as type of cyst, lymph nodes, mass location, fat necrosis, and postsurgical fluid collection
- Other factors such as a patient's medical history, age, and the size of any masses.

The assessment categories are as follows:

Table 1: BI-RADS Classification System (DocPanel, n.d.)

Score	Category	Recommended Action	Likelihood of Cancer
0	Incomplete	Need additional imaging to evaluate	N/A
1	Negative	Continue routine annual screening	0%
2	Benign	Continue routine annual screening	0%
3	Probably Benign	Short interval follow-up suggested (6 months)	<2% probability of malignancy
4	Suspicious for Malignancy	Biopsy	4A: Low suspicion (2-9%) 4B: Moderate suspicion (10-49%) 4C: High suspicion (50-94%)
5	Highly suggestive of malignancy	Biopsy	>95% probability of malignancy
6	Known biopsy proven malignancy	Confirmed biopsy and treatment planning	Proven malignancy

1.3 Statement of Problem

Our client contact has stated that the positive biopsy rate across the Mayo Clinic Enterprise for breast ultrasound lesions ranged from 31% to 51% between January 1, 2019, and November 1, 2020. This broad range indicates non-uniform patient management and an opportunity to introduce improved organizational standards. A low positive biopsy rate is an indication of low specificity, and our contact believes the introduction of computer assistance would both reduce the discrepancy across the enterprise and increase specificity and sensitivity.

While there is broad consensus among radiologists on BUS images having a BI-RADS score of 1, 2, or 5 (clearly benign or clearly malignant), there are widely differing interpretations between BI-RADS 3 and 4, particularly between BI-RADS 3 and 4A. The distinction is important, as it is the difference between a biopsy and short interval follow-up. Correctly classifying a lesion as a 3 instead of a 4A eliminates an unnecessary medical procedure, reduces cost, and spares the patient the psychological impact of a benign breast biopsy (Andrykowski, et al., 2002). Conversely, accurately identifying a malignant lesion as 4A instead of 3, with an associated positive biopsy, offers the patient more opportunity and time to fight the cancer.

1.4 Conceptual Framework

Computer-aided diagnosis (CAD) can be employed to aid radiologists as they attempt to characterize indeterminate breast lesions in ultrasound images. Researchers have taken two analytical approaches in designing models to assist with this task:

1. *Image processing* – In this approach, the system is trained to look for the same morphological and texture features as a radiologist would, such as lesion shape, orientation, margins, texture, and vascularity. Daoud et al. (2016) present a model which quantifies the morphology of a breast tumor and then analyzes the texture by breaking the image up into many small regions of

interest (ROI). They conclude that this hybrid ensemble is better able to predict malignancy than either feature alone, and at a very high accuracy overall.

2. *Deep learning* – Recent advances in deep learning and increased access to computing power have allowed improved object detection and classification of breast ultrasound images. Cao et al. (2019) evaluate various convolutional neural network (CNN) algorithms and concluded that SSD300 performed best for detecting lesions and ImageNet and DenseNet were best at lesion classification.

Each analytical approach offers benefits that the other does not. Due to the presence of noise, shadowing, and variation in echo patterns in surrounding tissue, a well-trained deep learning model is often better suited for isolating a ROI than traditional image processing. Because deep learning is like a black box, however, it does not necessarily leave a lot of room for manual overrides by the physician analyzing the image. Image processing can therefore provide better interpretability.

In the interactive system we envision, the radiologist will be presented with both an image and the various BI-RADS features identified by the hybrid model, with the opportunity to manually override any of them. The expectation is that synthesizing an image processing model and a deep learning model will improve the overall accuracy, particularly for the borderline cases that often prove difficult for radiologists.

1.5 Research Objective

The overarching objective of this research is to be able to develop one of the best breast ultrasound CAD systems by improving the accuracy of indeterminate lesion classification. If successful, it has the potential to raise the standard of care across the Mayo Clinic enterprise, improve patient management and outcomes, and reduce medical costs.

A secondary objective is trying to find whether there is a more optimal combination of artificial intelligence and human expertise than is currently in practice. Romeo, et al. (2021) recently published a similar study that exhibits the potential for improvement in this domain. For reasons of patient privacy and data ownership, much of the research in CAD for breast ultrasound has involved relatively small datasets consisting of a few hundred images or less, of varying quality. Training the CAD system by annotating images is also relatively labor intensive and requires a significant time commitment of expert radiologists. This project overcomes those limitations as much as possible potentially providing access to thousands of high-quality images from the Mayo system and an experienced radiologist who is one of the principal investigators with a strong motivation to see it succeed.

1.6 Organization of the Project

The project will consist of four discrete phases:

1. Acquisition and pre-processing of images
2. Model Generation
    a. Automated CAD using image processing techniques
    b. Automated CAD using deep learning techniques
3. Synthesis of models
4. Development of GUI

1.7 Practical Limitations

While the above sections detail the overall scope of the project, most of the rest of this paper will only deal with the part of the project for which I am most directly involved and responsible. Specifically, the research team has been divided into two groups, one focused on deep learning and the other on image processing. I am part of the latter group and so my analysis will be limited to the techniques, methodology, results, and discussion as it pertains to image processing.

Lastly, because the timeline of this project extends beyond the due date of this paper, my analysis will include only phases 1 and 2 of the project (acquisition/pre-processing of images and image processing model generation).

Chapter 2: Literature Review

2.1 Introduction

Breast ultrasound (BUS) has long been one of the most common and well-studied techniques in the early detection of breast cancer. Its noninvasive nature, low cost, and portability are major advantages over other screening technique. Sood, et al. (2019) performed one of the larger meta-analyses consisting of 26 studies and found that overall pooled sensitivity was 80.1% and specificity was 88.4%. These numbers suggest that BUS is an effective tool for detecting malignant lesions while simultaneously leaving some room for improvement.

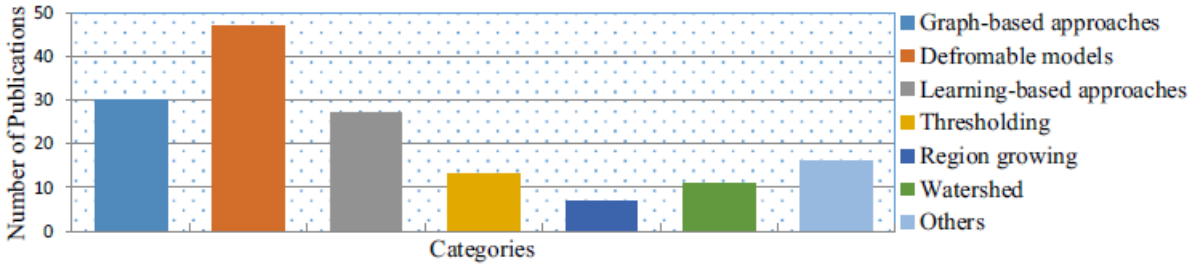
The literature review chapter will be divided into three sections: (1) Image Segmentation, (2) Image Classification, and (3) Radiomics. Though image segmentation is not the primary focus of our report, it is an important prerequisite that impacts the ability of the classification algorithm to accurately predict malignancy and thus it warrants a brief overview.

2.2 Image Segmentation

Advances in image processing technology have allowed researchers to apply techniques from other domains toward automatically identifying a region of interest (ROI) on a BUS image. Accurately establishing the boundaries of the region of interest is a crucial first step in further analysis, because the type of lesion margins are one of the key indicators of malignancy. Compared with other imaging modalities, BUS image segmentation is particularly challenging because (1) ultrasound images have very low quality due to speckle noise, low contrast, low sign noise ratio, and artifacts; (2) there are large

variations of breast structures between patients, making it difficult to standardize; and (3) strong priors based on tumor shape, size, and echo strength are important for segmentation in other imaging modalities, but are difficult to apply to BUS image segmentation (Xian, et al., 2018). Approaches to segmentation fall into roughly two categories: classic image processing techniques (thresholding, region growing, and watershed), and more computationally advanced techniques (graph-based, learning-based, and deformable models).

Figure 1: Distribution of automatic BUS image segmentation approaches (Xian, et al., 2018)



In their study, Xu, et al. (2019) discuss how the supervised or manual segmentation methods discussed above are mainly subject to limitations of the quality of the image. Techniques that are effective at segmenting some images will fail on others due to inherent artifacts and noise, attenuation, blurry boundaries, and intensity inhomogeneity. Consequently, the results of traditional methods must be "hand-crafted and experience-determined," requiring more intervention from the subject matter experts than is ideal. Using convolutional neural networks (deep learning), the team of researchers found that performance was roughly 10% better than traditional methods, with accuracy, precision, recall, and F1 score all exceeding 80%. In general, the vast majority of BUS image segmentation research in the last two or three years has focused on a deep learning approach. Our group has adopted this as well, but for the purposes of the classification sub-project, we will be using manually segmented and annotated BUS images.

2.3 Image Classification

As discussed earlier, the BI-RADS score was created to quantify the risk of a lesion for a given case. The first edition of BI-RADS was introduced in 1993 for mammography but it was not until the fourth edition was released in 2003 that advice for ultrasound was included. This version included language more specific to sonography meant to avoid the ambiguity that had existed up until that point. A team of researchers concluded that the newly updated BI-RADS for ultrasound was at least as accurate at lesion classification as BI-RADS for mammography (Heinig, Witteler, Schmitz, Kiesel, & Steinhard, 2008).

While a lesion is ultimately either malignant or benign, BI-RADS provides physicians with the ability to assign a likelihood of malignancy based on various predictive features. Hong, et al. (2005) studied 403 solid lesions to determine the positive and negative predictive values of different ultrasound features from the BI-RADS lexicon and the results are shown in Table 1. Their conclusions show the relative risk contributions of the three major features that make up the BI-RADS score: margins, shape, and orientation, but notably do not discuss echo pattern/echotexture, which is another important predictor.

Table 2: Negative and Positive Predictive Values of US Features

Predictor Group and US Descriptor Feature	Predictive Value (%)
Negative Predictors	
Oval Shape	NPV = 84
Circumscribed margins	NPV = 90
Parallel orientation	NPV = 78
Positive Predictors	
Irregular shape	PPV = 62
Spiculated margins	PPV = 86
Nonparallel orientation	PPV = 69

Different physicians will naturally have different opinions of these features. A 2006 study found substantial agreement for lesion orientation, shape, and boundary ($\kappa = 0.61$, 0.66, and 0.69,

respectively), moderate agreement for lesion margin and posterior acoustic features (κ = 0.40 for both), and fair agreement on lesion echo pattern (κ = 0.29) (Lazarus, Mainiero, Schepps, Koelliker, & Livingston, 2006).

Raza, et al. (2010) considered some of the reasons and ramifications of these differing interpretations. They posited that some of the variability was likely due to operator technique while much of it might be due to the level of training or experience of the radiologist or confusion about the BI-RADS lexicon. To address this, they presented several illustrative cases, providing the imagery and matching it up with the associated lexicon. Particular attention was paid to the most challenging circumstances, such as when age or previous medical history of the patient might lead the score to be different than it might otherwise be.

Daoud, et al. (2016) devised an approach which they termed "fused multiple-ROI texture analysis and morphological analysis" which seems very promising. Whereas BI-RADS puts particular importance on how well-circumscribed a lesion's margins are, this technique was able to achieve high accuracy by considering shape and texture in tandem. Because it is focused on classification rather than segmentation, they began by having a trained radiologist manually outline a region of interest. For the texture component, they divided the ROI into a grid of many smaller regions of interest, and then extracted twenty different texture features at varying distances and orientations. In terms of shape, they extracted ten different features directly from the tumor, six more from the best-fitting ellipse of the tumor, and two more from the normalized radial length (NRL) of the tumor. Though these researchers did not use the term directly, their overall technique dovetails very well into the next section which discusses radiomics. Table 3 highlights the additional accuracy gained by the fused approach of multiple ROI and morphology.

Table 3: Performance improvements from texture and morphological analysis

Performance metrics obtained using the (a) conventional classification approach using texture features, (b) conventional classification approach using morphological features, (c) conventional classification approach using both texture and morphological features, (d) proposed classification approach using multiple-ROI texture analysis, and (e) proposed classification approach using the fused multiple-ROI texture analysis and morphological analysis					
	(a)	(b)	(c)	(d)	(e)
Accuracy	85.5%	87.6%	90.9%	95.5%	98.2%
Specificity	84.4%	89.4%	90.6%	93.8%	98.4%
Sensitivity	87.0%	84.8%	91.3%	97.8%	97.8%
PPV	80.0%	84.8%	87.5%	91.8%	97.8%
NPV	90.0%	89.1%	93.6%	98.4%	98.4%

While the results seem very favorable, we cannot overlook the fact that this study, like many of the studies discussed here, employs only a very small data set of 110 tumor images (64 benign and 46 malignant). Such a limited data set raises questions of generalizability and overfitting.

2.4 Radiomics

An emerging field of research known as radiomics attempts to aid clinical decision-making and outcome prediction by making use of quantitative features mathematically extracted from clinical images (Rizzo, et al., 2018). As it pertains to BUS image processing, it has the potential to improve both image segmentation and image classification by removing some of the human bias and difficulty in quantifying or characterizing the pixels in an ultrasound image without the aid of software. For example, while a computer program can accurately and consistently measure the homogeneity of pixel intensity or echotexture within a lesion boundary, expert radiologists are often not able to reach a consensus. In addition to quantifying the shape features (morphology) of a region of interest, radiomics offers first-order statistics features (mean, median, asymmetry, flatness, uniformity, randomness, etc.), second order-order statistics features (inter-relationships between pixels, such as homogeneity metrics), and higher-order statistics features (e.g. transforms and fractal analysis).

A large team of researchers studied the use of radiomics applied to breast ultrasound and found promising results for differentiating benign and malignant lesions (Romeo, et al., 2021). The study was retrospective in nature and used 135 lesions from one institution to train a radiomics model and then tested the model on 66 lesions from a different institution. Among the goals of the study were to evaluate the accuracy of a radiomics approach paired with machine learning (ML), compare them with the diagnostic accuracy of an expert radiologist, and then determine if the performance of the radiologist could be improved by providing them with the ML learning output. The findings were that the radiomic-based ML model was able to achieve comparable accuracy to the radiologist (ML: 82%, Radiologist: 79.4%) with high sensitivity (93%) but low specificity (57%). When given the ML model, the performance of the radiologist improved slightly but not significantly, to 80.2% accuracy.

Despite the tremendous promise of radiomics for both image segmentation and image classification, it is important to note that it is a relatively new approach and comes with a few potential pitfalls. Reproducibility has proven to be challenging due to lack of standardization, insufficient reporting, or limited open source code or data. It also suffers from a difficulty of interpretability of results (correlation vs. causation), leading it to be less trusted in decision-support systems. Lastly, most of the studies have been retrospective rather than prospective, meaning that a radiomic approach has not always been able to prove its predictive ability (van Timmeren, Cester, Tanadini-Lang, Alkadhi, & Baessler, 2020).

2.5 Summary

Given the nature of the data available at the time of this writing, which consists of a manually segmented data set of 163 images with BI-RADS scores, annotations, and predicted malignancy, the logical next step involves applying radiomics and machine learning techniques to see if it is possible to

improve upon the predictions of the subject-matter expert. The methodology, results, and conclusions will thus all be geared toward this specific part of the project.

Chapter 3: Data Collection and Methodology

3.1 Overall Strategy

At the outset of the project, the goal was to produce one or more models that were capable of predicting whether a lesion was benign or malignant while conveying the model's confidence, in order to assist the radiologist in borderline or otherwise tricky cases. The vision for the final product allows the radiologist to manually override any of the features identified by the model and have the probability of malignancy and confidence intervals update accordingly. Therefore, it was important for our approach to mirror as much as possible the methodology of the radiologist. Early on, we received a lot of feedback from our subject matter expert as we learned to see what he is able to see in an ultrasound image. The middle phase consisted of building algorithms so our software could detect the sorts of features that a trained radiologist would use as clues to detect malignancy, and the last phase was the machine learning component -- taking those features, training a model, and using it to predict malignancy on new ultrasound images.

Initially, we attempted to build a machine learning model to predict either BI-RADS score or malignancy using only the qualitative descriptors provided by Dr. Ellis. However, it quickly became apparent that there was not enough variety or balance in these features for this particular data set. Many of the descriptors were not used at all, and even among those that were used, many only showed up on only a handful of images. Attempting to train on this distribution of data was very limiting in the type of classifier that was even possible; essentially any classifier that relied upon nearest neighbors would lead to a runtime error. Furthermore, the only BI-RADS scores that the model was predicting were 2, 3, and 5. It was evident that we needed to change strategies to one where, rather than

predicting descriptors, we would attempt to quantify the presence of a particular feature on a spectrum from 0 (not present) to 1 (perfect representation of feature).

3.2 Raw Data

Due primarily to privacy concerns, publicly available datasets of breast ultrasound images for research purposes are extremely difficult to find. The first dataset our we tried, commonly referred to as the BUSI dataset (Walid, Gomaa, Khaled, & Fahmy, 2020), contained too many inconsistencies, artifacts, and post-sonographic annotations to be considered very useful. Ultimately, our team intends to use images provided to us by The Mayo Clinic, although at the time of this analysis, those images were not yet made available.

In the interim, the images for this analysis come from a dataset collected in 2012 from the UDIAT Diagnostic Centre of the Parc Taul´ı Corporation, Sabadell (Spain) with a Siemens ACUSON Sequoia C512 system 17L5 HD linear array transducer (8.5 MHz). The dataset consists of 163 images from different women with a mean image size of 760 × 570 pixels, where each of the images presented one or more lesions. Within the 163 lesion images, 53 were images with cancerous masses and 110 with benign lesions. From the malignant images, 40 were invasive ductal carcinomas, 4 were ductal carcinomas in situ, 2 were invasive lobular carcinomas and 7 were other unspecified malignant lesions. From the benign images, 65 were unspecified cysts, 39 were fibroadenomas and 6 were of another type of benign lesion (Yap, et al., 2018).

3.3 Radiologic Evaluation

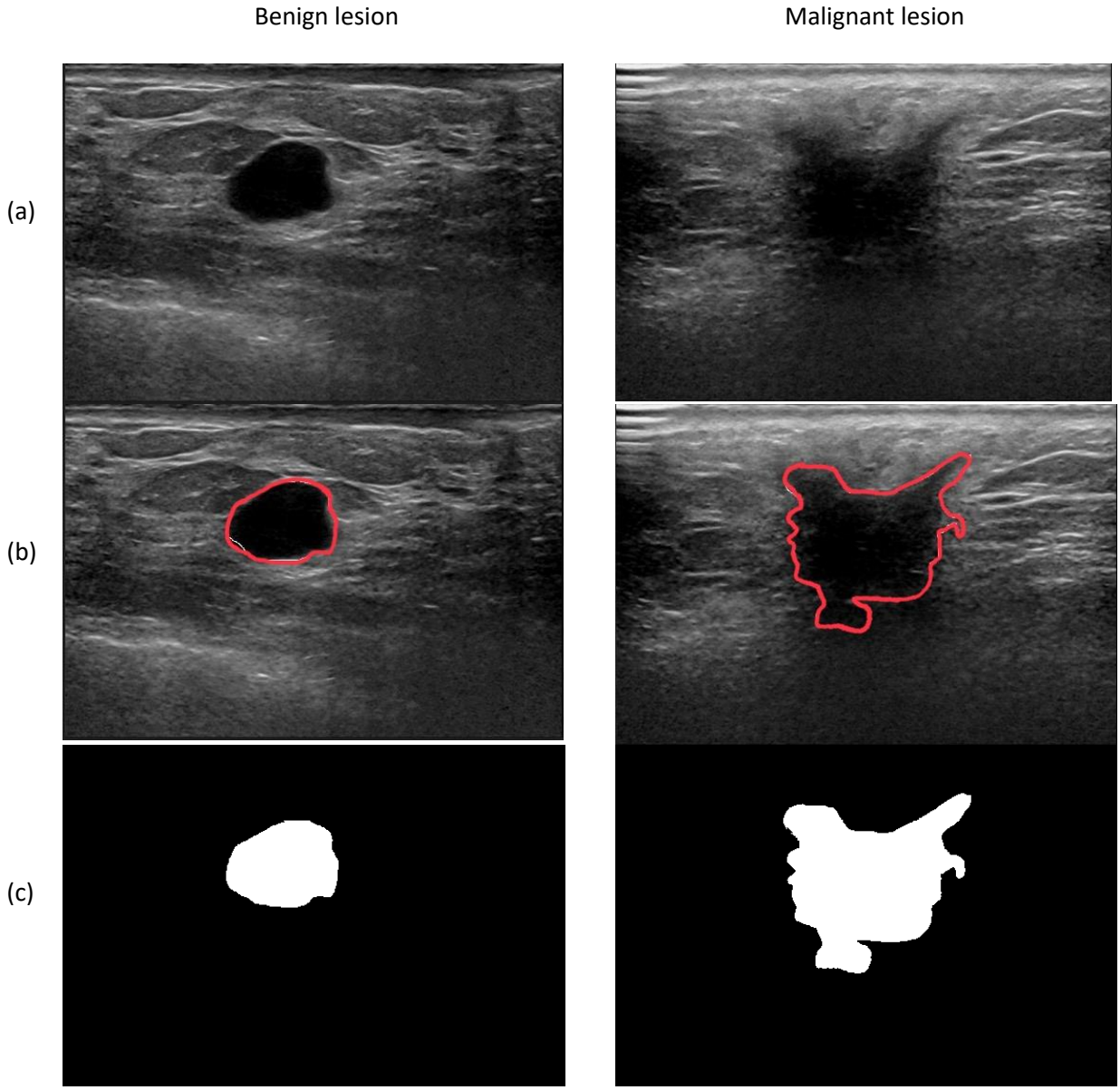
At the beginning of the project, our subject matter expert, Dr. Ellis, provided a rubric that his group uses to classify lesions. The rubric assigns qualitative descriptors to each of the three zones of a lesion (internal, marginal, and peripheral) as well as capturing the lesion's size, vascularity, and any posterior features. A summary of the rubric is shown in Table 1:

Table 4: Rubric for classifying lesions (R. Ellis, personal communication, March 9, 2021)

Feature	**Low Risk**	**Moderate Risk**	**High Risk**
Internal Zone	Oval, parallel, homogeneous hyperechoic or hyperechoic	Round, hypoechoic	Irregular shape, angles, not parallel
Marginal Zone	Well circumscribed, abrupt interface	Partially circumscribed, echogenic halo	Indistinct, angular, microlobulated, spiculated
Peripheral Zone	Normal tissue	Shadowing, enlarged ducts, edema	Spiculations, skin thickening, architectural distortion
Posterior Features	Enhancement	Partial shadowing	Full or majority shadowing
Vascularity	Absent	Internal or rim vascularity	

The first step in our process consisted of Dr. Ellis segmenting and annotating the 163 images in our dataset. Each of the images was given one or more descriptors for the first four features in Table 4 (it was not possible to accurately assess vascularity without doppler imaging, which was not available). In addition, Dr. Ellis also provided a BI-RADS score, assessment of malignancy, and most critically, a marked-up image with a border drawn around the lesion (segmentation). From this border, a ground truth mask was created that would be used for morphological analysis later on. Figure 2 shows examples of both a benign (left) and malignant (right) image, with the original image (a), the boundary of the lesion (b), the ground truth mask (c), and the classifications assigned by Dr. Ellis (d).

Figure 2: Samples of annotated images. (a) Original image, (b) lesion segmentation, (c) ground truth mask, and (d) feature descriptions for benign lesion (left) and malignant lesion (right)



3.4 Algorithm Development

The findings of Hong, et al. (2005) and the rubric provided by Dr. Ellis both indicated that the most predictive value was captured by examining the lesion margins (boundary), followed by the shape

(oval, round, or irregular), and the orientation (parallel or not). We decided to tackle these features in the same order, but rather than aim for discrete categorical variables, along the lines of the BI-RADS guidance and the rubric, we chose to quantify features wherever possible. The advantages in doing so were twofold. First, since our software would ultimately be generating a numeric value for these features (e.g., how well circumscribed is the boundary?), converting to a categorical value would reduce the amount of information available to the model. Secondly, there is no universal agreement among radiologists regarding where, for example, the distinction should be made between "round" and "oval" or at what degree of rotation a "parallel" lesion becomes "not parallel". We wanted our model to be able to handle all the in-between cases as well as it could handle the more obvious ones.

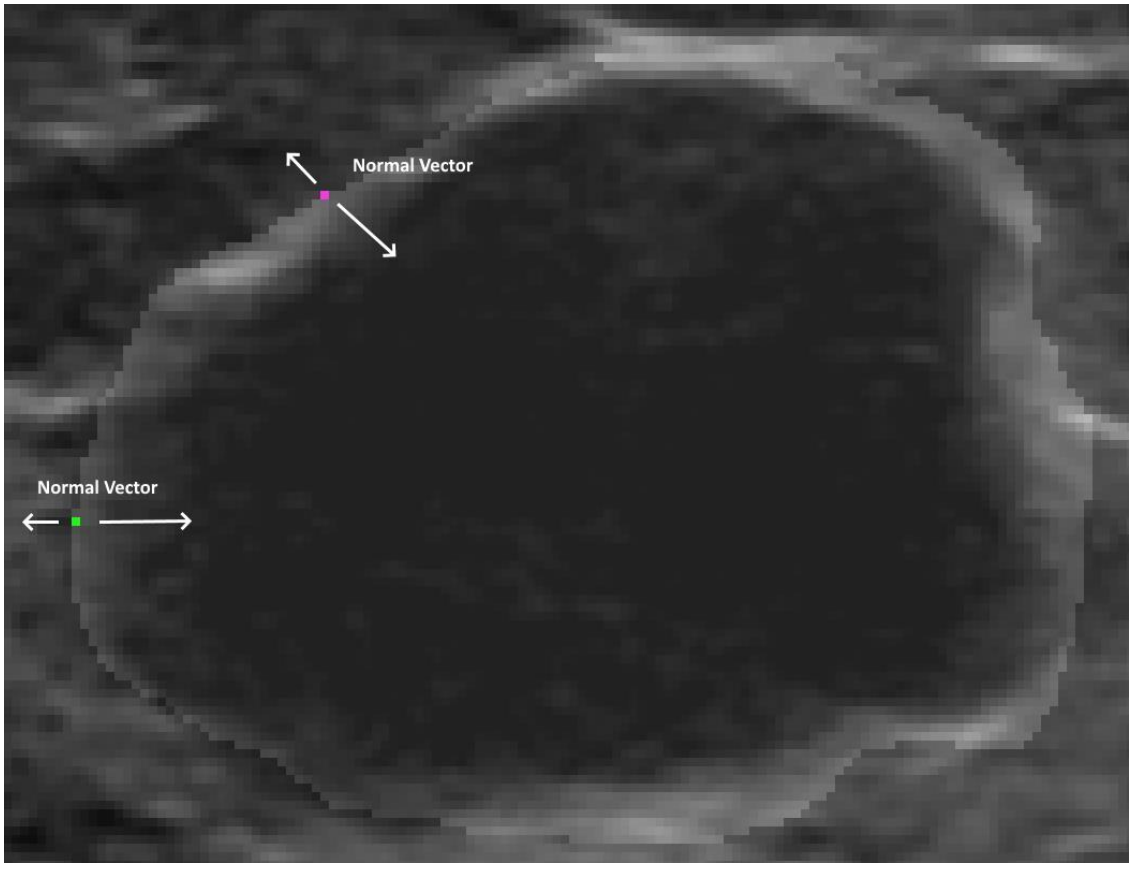
3.4.1 Margin Circumscription

Because margin circumscription, and its related concept, abruptness, carried so much predictive value, it was important to quantify this as carefully and accurately as possible. One dilemma that quickly became evident was the outlines provided by Dr. Ellis – the *ground truth* – were not absolutely precise down to the individual pixel level, and it would be unreasonable to expect that they could be. As we could see in the benign lesion in Figure 2, it is often the case that there is large pixel intensity difference between the internal zone of a lesion and the surrounding tissue.

In order to quantify the level of circumscription of a lesion's margin, we introduced two new metrics, which will be referred to as *abruptness* and *average maximum gradient*. Both involve tracing around each edge pixel of the ground truth overlayed on the original image and inspecting the gradient in the orthogonal direction. The gradient was calculated on the normal vector (orthogonal) looking mostly inward but also a few pixels outward from the lesion, in the event that the outline was not perfectly drawn. Figure 3 shows a zoomed in version of the benign lesion in Figure 2 along with normal vectors for two of the points on the perimeter.

Absent any sort of industry standard, we considered normal vectors of length 10 that looked inward 7 pixels and outward 3 pixels. For abruptness, an edge pixel is deemed abrupt if it the normal vector contains a gradient value greater than a certain threshold – a steep edge. For our model, we used 20 to be that threshold. The average max gradient of a lesion considers the maximum gradient of the normal vector of each pixel and then takes the mean value for the entire boundary.

Figure 3: Calculation of the abruptness, based on a zoomed in image of ground truth overlaid on original ultrasound image of benign lesion



3.4.2 Shape

Shape features were captured using a similar methodology to Romeo, et al. (2021). Pyradiomics is an open-source Python package capable of calculating hundreds of morphological and texture features of a masked image, including measures of roundness, major and minor axis length, measures of

image homogeneity in the region of interest, and numerous other potential predictors of malignancy. A description of all the features extracted by Pyradiomics can be found at https://pyradiomics.readthedocs.io/en/latest/features.html. There were three relevant shape features we included in our model that are defined below: *sphericity*, *elongation*, and *perimeter-to-surface ratio*.

- *Sphericity* refers to the ratio of the perimeter of the ROI to the perimeter of a circle with the same area as the ROI and is thus a measured of the roundness of the lesion. A perfectly circular lesion would have a sphericity of 1.
- *Elongation* shows the relationship between the major axis length and minor axis length of the ellipse that best approximates the shape of the ROI.
- *Perimeter-to-surface ratio* is simply the ratio of the perimeter of the ROI to its area. A lower value indicates a more compact shape (circle-like).

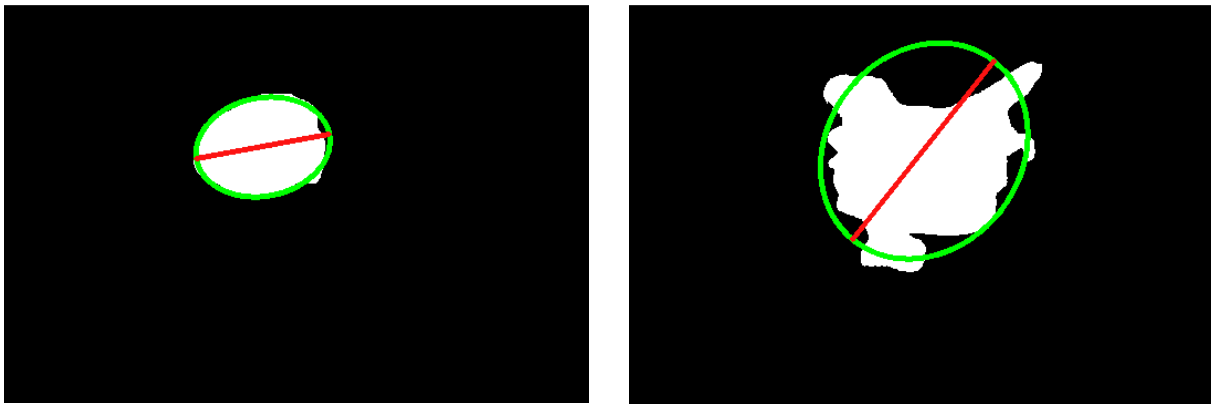
In addition to the shape features, one of the main reasons we chose to use this package is for its texture analysis. Characterizing the texture of a region, particularly of a relatively low-resolution image such as a breast ultrasound image, is difficult to do with the naked eye. The echo pattern, both inside the internal zone and in the peripheral zone, is used by radiologists as an indicator of potential malignancy, so it was important that our software might help quantify these features for use in a machine learning model.

3.4.3 Orientation

The last major feature we wanted to include in our model was the orientation, as the BI-RADS standard advises that lesions which are parallel to the tissue (x-axis) have a higher likelihood of being benign, while those which are not parallel have a higher likelihood of being malignant. Due to the highly irregular shapes of some of the lesions, defining parallel vs. not parallel might not be a straightforward exercise. For our analysis, we used the Python OpenCV package's `fitEllipse()` function to find an

ellipse that best approximates the shape of the ground truth in a least-squares sense, and then calculated the angle between the major axis of that ellipse and the x-axis, and the deviation of that angle from 90 degrees (90 degrees would represent a vertical ellipse). Figure 4 shows a visual representation of the algorithm for the benign and malignant images we have been looking at in this chapter.

Figure 4: Orientation of benign (left) and malignant (right) lesions by fitting an ellipse and calculating the angle of the major axis



Angle: 79.35°

Angle: 38.37°

3.6 Data Analysis

At this point, Dr. Ellis had reviewed and annotated all of the images in our dataset. He determined that 28 of the original 163 images were not fit for further analysis due to the fact that the image quality was poor, there was more than one lesion, or there was evidence of a breast implant. This left 135 images to be analyzed, for which we calculated quantitative values for each of the features described above. In addition to abruptness, average maximum gradient, and orientation, Pyradiomics was able to provide 102 more morphological and texture features.

Feature Selection

Due to the large number of features, we decided to reduce the number of features in our model. Certain features from Pyradiomics, such as the maximum, minimum, and range, seemed like

natural candidates to be cut. These statistics were based on individual pixel values within a region of interest -- they would have no predictive value as our model was based on the characteristics an entire boundary or two-dimensional area. Furthermore, it is sometimes the case that the brightest pixels in an image are just outside the internal zone, outside the area we are trying to isolate but possibly included because of an imprecisely drawn boundary.

Aside from these three features, though, there were no obvious candidates for removal, so we looked at a pairwise correlation matrix for the rest of the Pyradiomics variables. Using a correlation threshold of 0.8, we were able to reduce the number of Pyradiomics variables down to a more reasonable 23. After considering using PCA to reduce the number of variables further still, we decided to keep them for model selection, in hopes of keeping the model more explainable and the possibility of discovering new predictors which were not yet part of BI-RADS.

Model Selection

The data at this point consisted of 26 numeric variables (abruptness, maximum average gradient, orientation, and 23 morphological and texture variables from Pyradiomics), and we needed to create a model to classify each image as benign or malignant. For this task, we chose to use an open-source machine learning Python package named PyCaret that made it relatively simple to perform all of the necessary tasks: accounting for the imbalance in the data (2/3 of the remaining images were of benign lesions, 1/3 were malignant), creating a train/test split, and calculating the best cross-validated model. The results of this step can be found in the next chapter.

Chapter 4: Results

4. 1 Introduction

Since the goal of the project is to determine the extent to which machine learning could assist a radiologist assessing a breast ultrasound image, we will first establish a baseline by examining the accuracy of our subject matter expert, Dr. Ellis, in classifying the images in our sample data set, split out by the BI-RADS scores he assigned to each one. After that, we will look at how well our algorithm is able to numerically represent the three most predictive features of a BUS image: marginal zone, shape, and orientation. Lastly, we will evaluate the accuracy of a few different classification models based on these three features plus a collection of textural features provided by the Pyradiomics package.

4.2 Findings

4.2.1 Physician Performance

The curated data set contained 90 images of benign lesions and 45 images of malignant lesions. Dr. Ellis correctly classified 91.1% of them, with a precision of 97.1% and a recall of 75.6%, as shown in Table 5.

Table 5: Performance of subject matter expert for all images

	Predicted Benign	**Predicted Malignant**
True Benign	89	1
True Malignant	11	34

The most important distinction a radiologist needs to make is separating BI-RADS 3 from BI-RADS 4a, because the protocol for anything BI-RADS 4a and above is a biopsy, while a BI-RADS 3 results in continued monitoring. For this reason, let us consider the physician performance just among those

two BI-RADS scores. To review, BI-RADS 3 means that there is less than a 2% expected chance of being malignant and BI-RADS 4a means that there is a 2% to 9% expected probability of malignancy. Of these 32 images, Dr. Ellis correctly classified 78.1%, with a precision of 50% and a recall of 14.3%, as shown in Table 6.

Table 6: Performance of subject matter expert on BI-RADS 3 or 4a

	Predicted Benign	**Predicted Malignant**
True Benign	24	1
True Malignant	6	1

4.2.1 Feature accuracy

In the methodology section, we described in detailed our process for quantifying the three main features used to predict malignancy in a breast ultrasound image: margins, morphology, and orientation. The rubric that Dr. Ellis uses to help classify lesions consists of categorical variables, and is included in Appendix A. The following figures display the relationships between the descriptive annotations provided by Dr. Ellis and the numeric representations of them in our model.

4.2.1.1 Abruptness and Average Maximum Gradient

Figure 5: Abruptness vs. Marginal Zone ACR

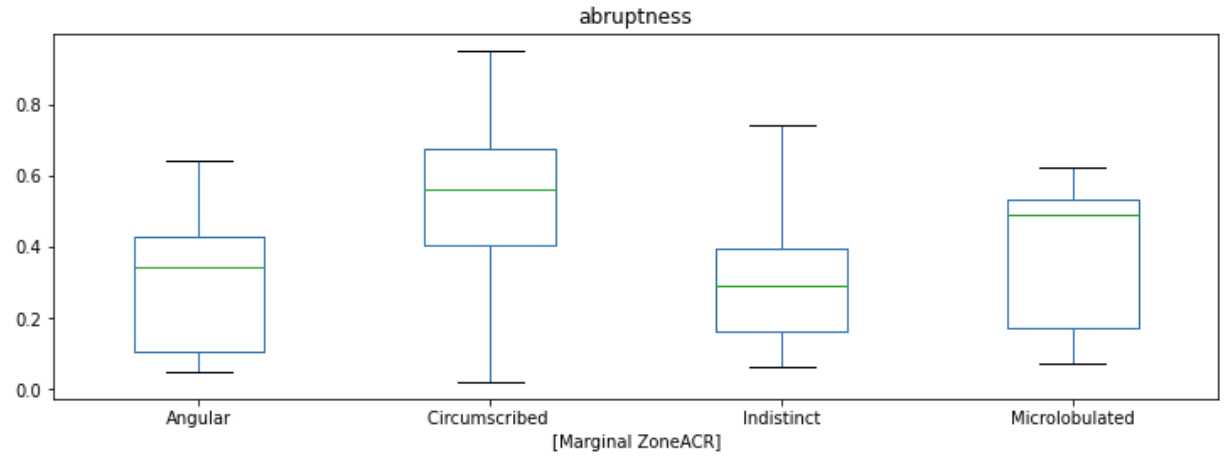
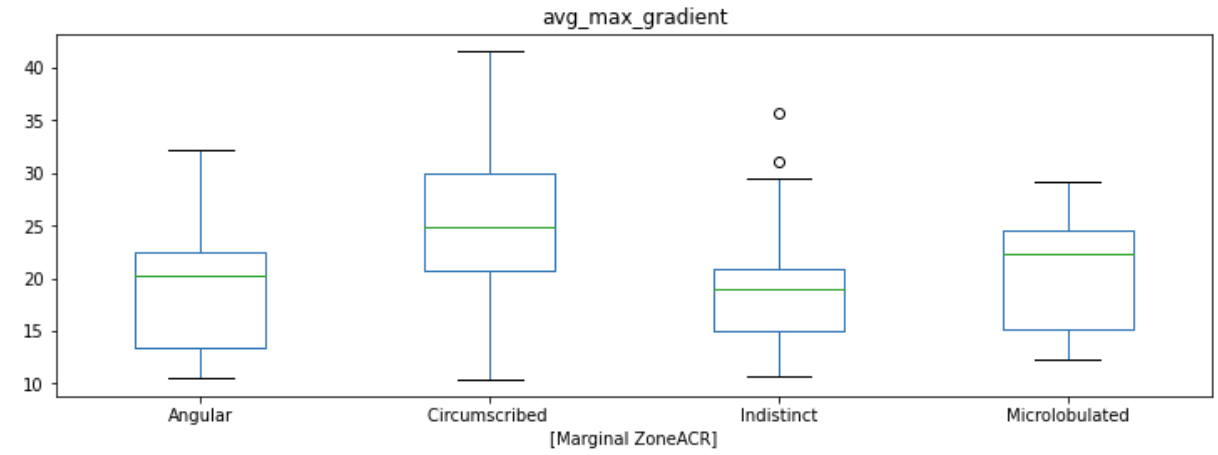


Figure 6: Average Maximum Gradient vs. Marginal Zone ACR



4.2.1.2 Morphology

There were two characterizations of shape in the rubric, “ShapeTS” and “ShapeACR”, with some overlapping values. Our model includes three morphological features: sphericity, elongation, and perimeter-surface ratio. Figures 3 and 4 show the comparison between sphericity and the categorical variables from the rubric.

Figure 7: Sphericity vs. ShapeTS

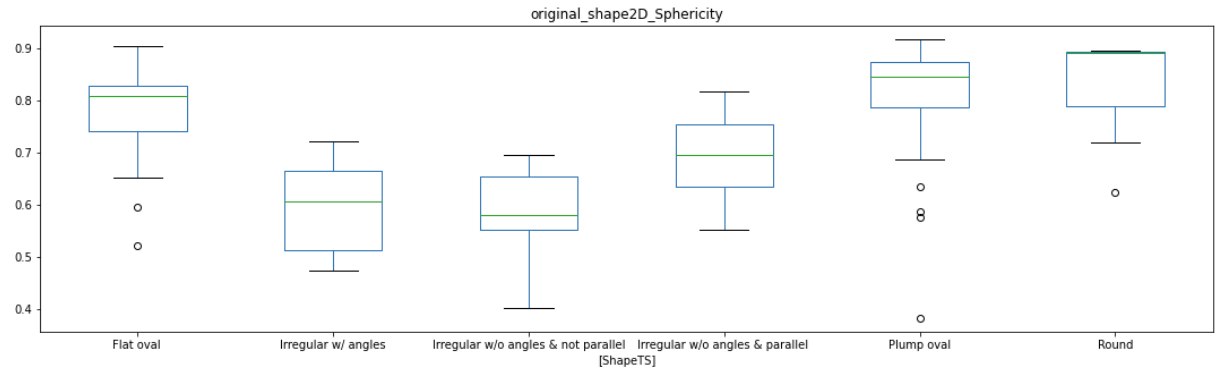
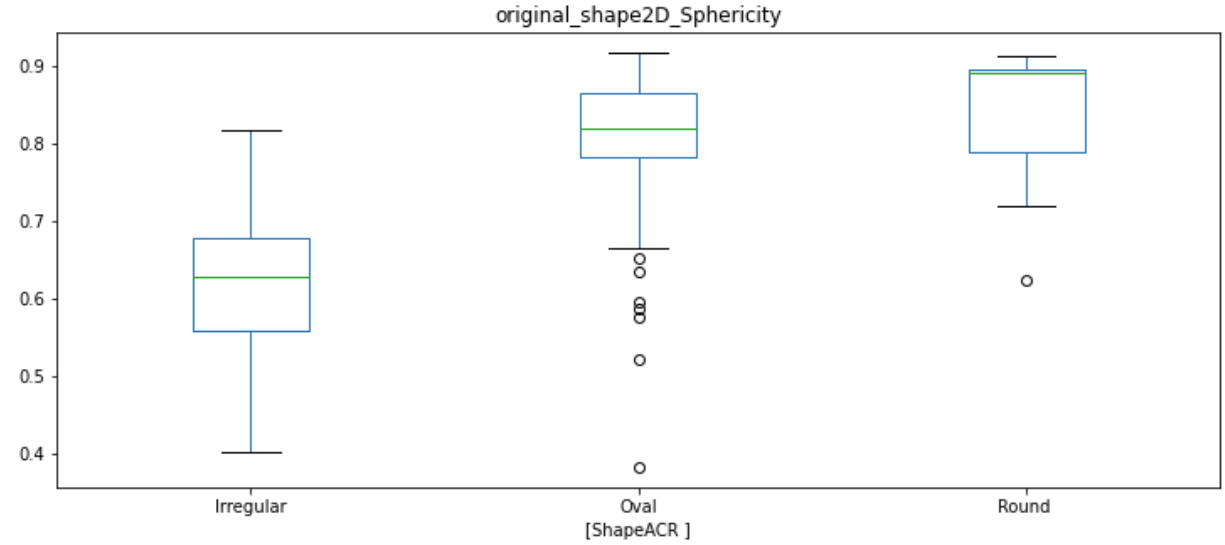


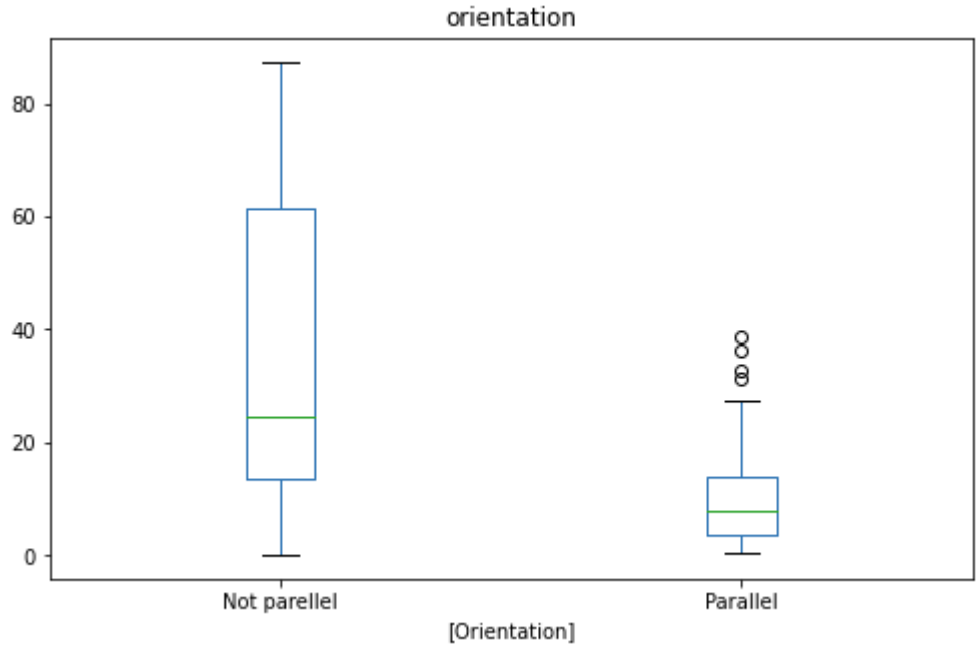
Figure 8: Sphericity vs. ShapeACR



4.2.1.3 Orientation

The rubric contains only one feature for orientation, with values of “parallel” or “not parallel”. Figure 9 plots the distribution of quantitative orientation values as compared to the label assigned by Dr. Ellis.

Figure 9: Orientation (model) vs. Orientation (rubric)



4.2.2 Classification Models

Pycaret generated statistics on 13 different types of models based on a training set size of 94 images, with the remaining 41 images being used for validation. The accuracy of the first several models on the list was very similar, but there seemed to be a small separation of the top three models from the rest. Table 7 shows the performance of each of the models:

Table 7: Classification model candidates

	Model	Accuracy	AUC	Recall	Prec.	F1	Kappa	MCC	TT (Sec)
ridge	Ridge Classifier	0.8633	0.0000	0.7333	0.8583	0.7338	0.6523	0.6937	0.024
et	Extra Trees Classifier	0.8500	0.9246	0.7333	0.8167	0.7229	0.6289	0.6658	0.474
rf	Random Forest Classifier	0.8400	0.8794	0.7000	0.7600	0.6731	0.5929	0.6353	0.488
lda	Linear Discriminant Analysis	0.8078	0.8413	0.7000	0.7917	0.6933	0.5618	0.5998	0.023
gbc	Gradient Boosting Classifier	0.7967	0.8825	0.7000	0.6350	0.6260	0.5061	0.5395	0.124
lightgbm	Light Gradient Boosting Machine	0.7889	0.8952	0.6667	0.6617	0.6062	0.4854	0.5297	0.092
ada	Ada Boost Classifier	0.7867	0.8333	0.6333	0.6933	0.6160	0.4805	0.5090	0.113
dt	Decision Tree Classifier	0.7544	0.7143	0.6000	0.5683	0.5674	0.4053	0.4156	0.027
lr	Logistic Regression	0.7033	0.7214	0.5000	0.3862	0.4226	0.2543	0.2702	0.335
qda	Quadratic Discriminant Analysis	0.6911	0.4929	0.0000	0.0000	0.0000	0.0000	0.0000	0.029
svm	SVM - Linear Kernel	0.6111	0.0000	0.2000	0.0600	0.0923	0.0000	0.0000	0.025
knn	K Neighbors Classifier	0.4278	0.4798	0.5000	0.2660	0.3369	-0.0874	-0.0997	0.122
nb	Naive Bayes	0.3811	0.6294	0.8500	0.3192	0.4622	0.0218	0.0089	0.025

4.2.2.1 Ridge Classifier

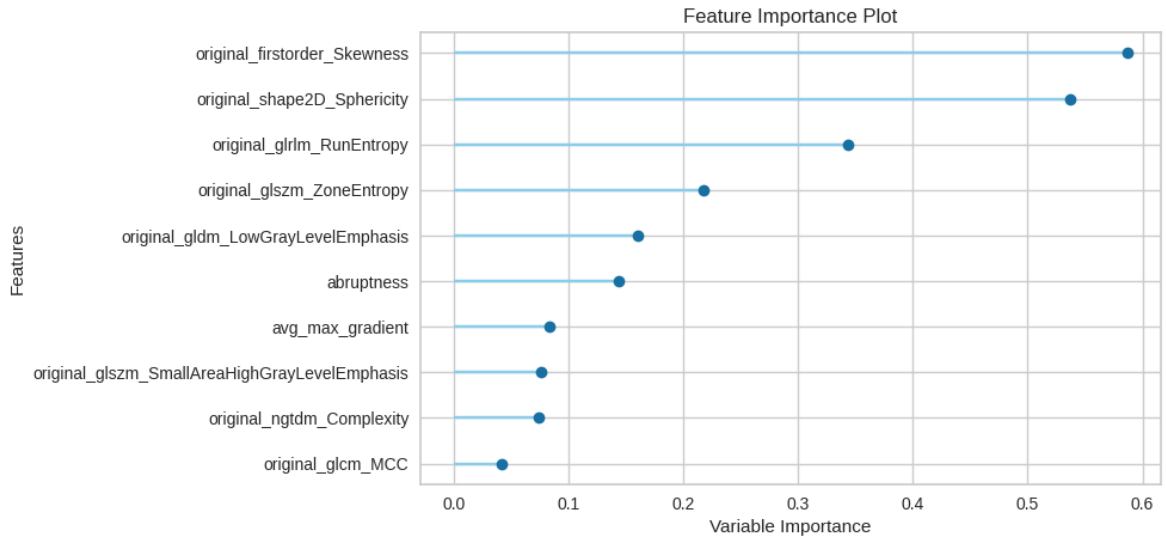
Based on accuracy, Pycaret identified the ridge classifier to be the best performing model. It was able to accurately predict 80.5% of the 41 images in the validation set, including 6 of the 7 images that were BI-RADS 3 or BI-RADS 4a.

Table 8: Ridge classifier on validation set

	Predicted Benign	**Predicted Malignant**
True Benign	19	6
True Malignant	2	14

Interestingly, the most important features in the ridge classifier were skewness (a measure of how skewed the distribution of pixel intensities are in the ROI), sphericity, and entropy (a representation of heterogeneity in the texture).

Figure 10: Ridge classifier feature importance



4.2.2.2 Extra Trees Classifier

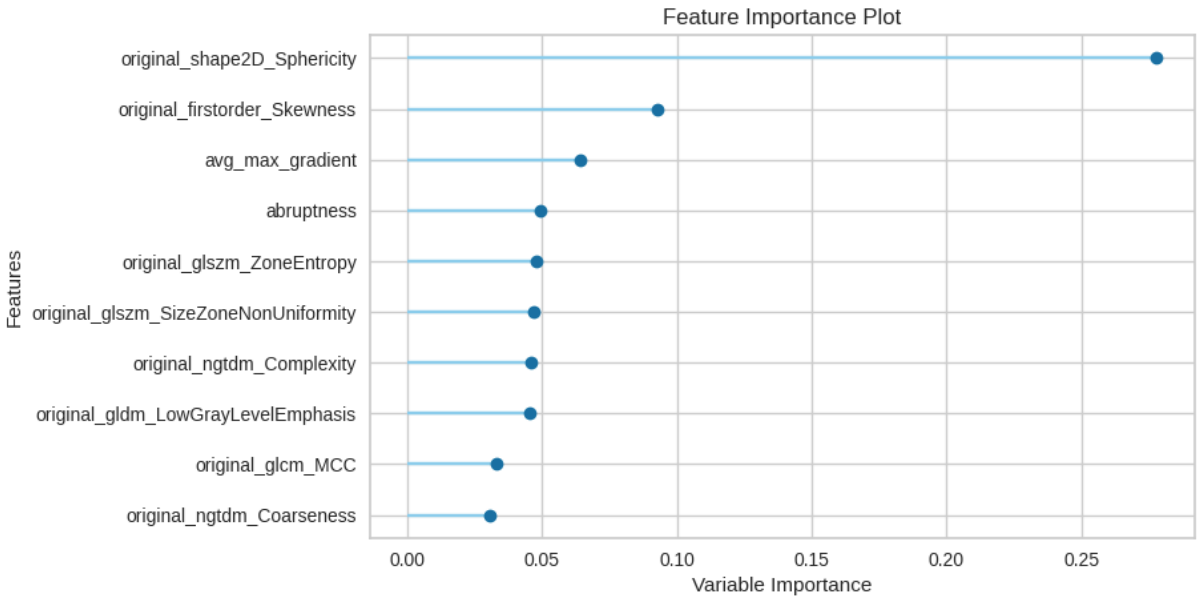
The extra trees classification model performed very similarly to the ridge classifier, accurately predicting 78.0% of the images in the test set, including 5 of the 7 BI-RADS 3 and BI-RADS 4a images.

Table 9: Extra trees classifier on validation set

	Predicted Benign	**Predicted Malignant**
True Benign	22	3
True Malignant	6	10

In terms of feature importance, this model more closely matches the BI-RADS guidance, with shape and margin characterization features hovering near the top of the list.

Figure 11: Extra trees classifier feature importance



4.2.2.3 Random Forest Classifier

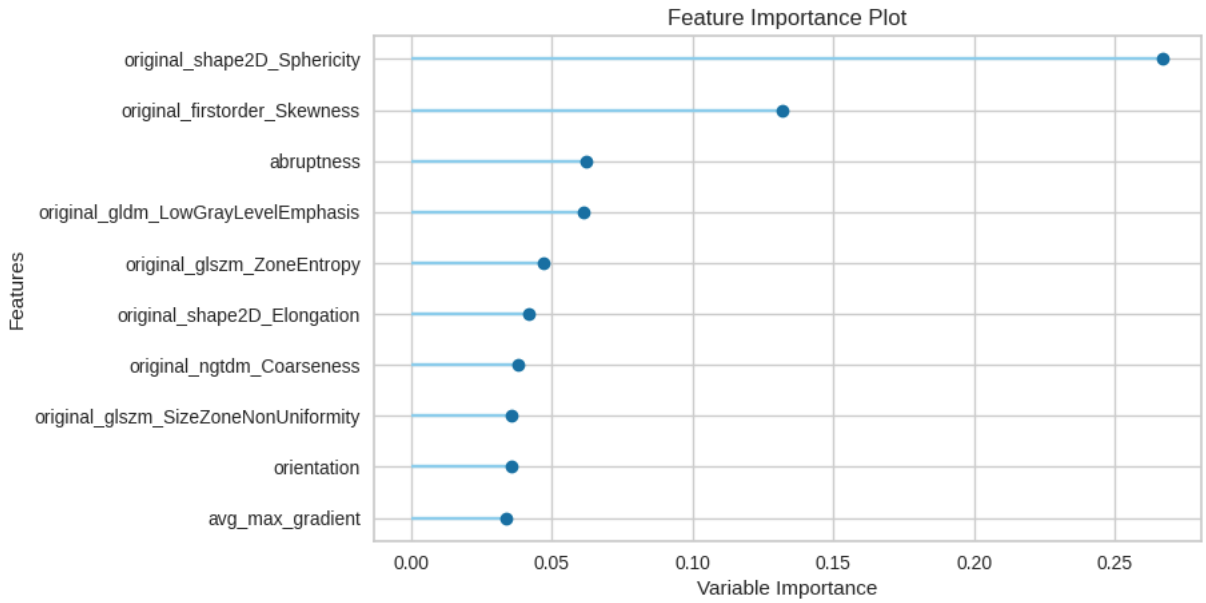
The last model we considered was the random forest classifier. It had a much lower accuracy on the validation set than either the ridge classifier or the extra trees classifier, predicting only 28 of the 41 images correctly (68.3%), including only 3 of the 7 BI-RADS 3 and BI-RADS 4a images.

Table 10: Random forest classifier on validation set

	Predicted Benign	**Predicted Malignant**
True Benign	20	5
True Malignant	8	8

As with the other two models, sphericity (shape) ranked very highly in feature importance, along with skewness. Figure 12 shows the top ten features and their relative importance.

Figure 12: Random forest classifier feature importance



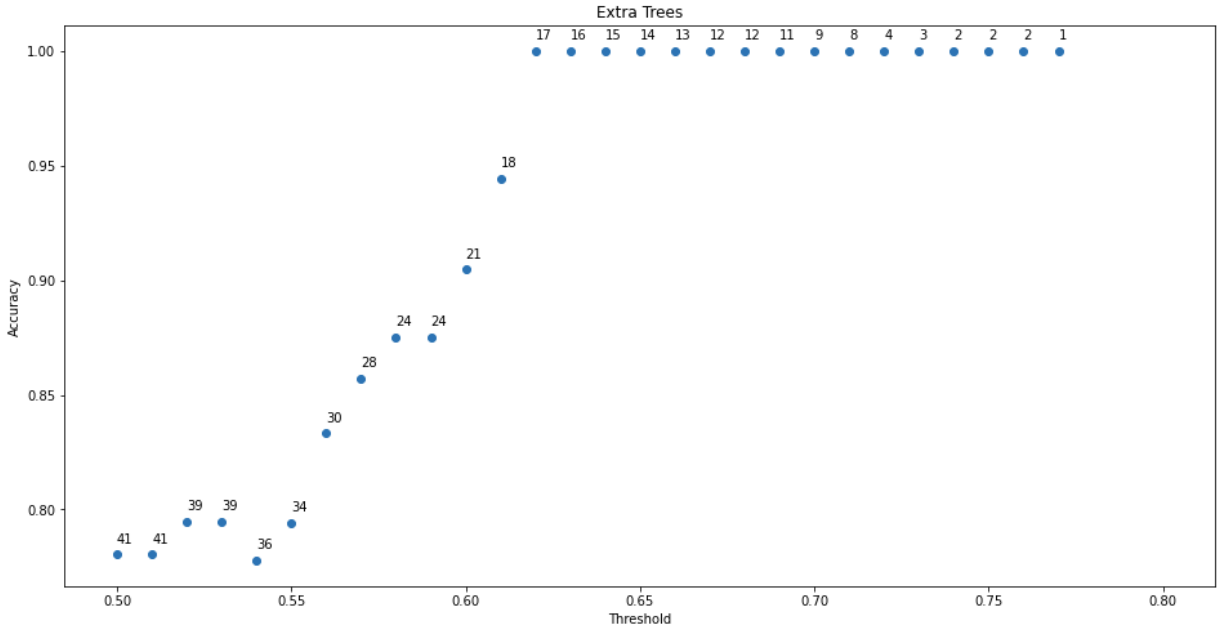
4.3 Conclusion

Dr Ellis's analysis of the data set showed that while a human subject matter expert is very good at predicting malignancy, there is room for improvement. As one would expect with borderline cases (BI-RADS 3 and BI-RADS 4a), the accuracy decreases considerably. This is true of the algorithmic approach as well, performing much better on the more clear-cut cases (BI-RADS 1, 2, 4b, 4c, and 5) that on the more difficult ones.

According to the literature and standard protocol, the best predictors of malignancy are lesion margins, shape, and orientation. Margin and shape both proved predictive in this dataset, depending on the algorithm, but orientation was much less so, at least by our quantification of this feature. Our analysis also suggests that skewness and entropy, or perhaps other measures of echotexture, also show potential as predictors of malignancy.

The ridge classifier and the extra trees classifier both performed well at predicting malignancy of the validation set, and despite the very small sample size, even showed hints that they could assist in separating the critical BI-RADS 3 and BI-RADS 4a lesions. The extra trees classifier also provides a probability score, allowing us to set a threshold above which we are confident in our prediction. The confusion matrix above is based on a threshold of 0.50, but if we increase that threshold even slightly, the accuracy rate increases considerably. Figure 13, below, plots the accuracy and number of samples that qualify for the extra trees classifier at different thresholds. The number above each point represents the number of images that reach a particular threshold. So, this algorithm correctly classified all 17 images (out of 41 total) in the validation set that met a minimum probably threshold of 0.62.

Figure 13: Extra trees classifier threshold vs. accuracy



Chapter 5: Discussion

5.1 Introduction

This chapter begins by providing some in depth discussion of the performance of both the radiologist and the machine learning model, including recommendations for how to best combine the

relative strengths of human evaluation and computer analysis. Next, we offer up some ideas for future research, depending on the availability of a richer data set that we expect to be receiving from Mayo shortly after this paper was written. We conclude by describing a forward-looking vision of more interactive ultrasound evaluation process, in which a radiologist is provided with the images, editable segmentations of those images, and the system's interpretation of the segmented images using the algorithms detailed in this paper.

5.2 Summary of Findings

Both Dr. Ellis and the machine learning models described above did a commendable job of classifying the images in the data set, though there was room for improvement in each case. Lesion shape and margins both proved to be important features for prediction, but orientation was not as important a predictor in this data set as echotexture. The best performing machine learning models were the ridge classifier and the extra trees classifier. The ridge classifier showed better recall, which is critical in cancer detection, but the extra trees classifier has the advantage of providing a probability threshold with its predictions, which end users can utilize to make better judgment calls as they see fit.

5.3 Discussion

Before getting too far into the discussion, it is worth reiterating that the number of images in this data set was far less than ideal and, according to Dr. Ellis, the quality of the images also was not up to current standards. The original plan was to have several hundred (or more) images provided by Mayo and annotated by Dr. Ellis, but the timeline of receiving these images did not match up with the timeline of this capstone. Nonetheless, we feel that the margin, shape, orientation, and texture algorithms we developed would work equally well and be easily transferable to a larger data set. A larger data set will also allow us to better train the model and potentially pick up more of the features from the rubric, many of which may be predictive.

It is very reasonable to look at the 91% of images that Dr. Ellis was able to correctly classify and conclude that an expert radiologist does a very acceptable job of separating out the malignant lesions from the benign ones. The real question we need to ask, though, is whether additional input from an AI system would improve that number even more. In particular, given the severe consequences of a false negative (a cancerous lesion that was not classified as such), could the computer model assist a human a human in detecting these misclassifications?

The best performing model in our analysis was the ridge classifier, and while the overall accuracy was lower than the radiologist (80.5%) the recall was higher (87.5% vs. 75.6%), suggesting that the model sacrificed some false positives for fewer false negatives. While there are downsides to both, as discussed above, the misclassification of a malignant lesion has a potentially life-threatening outcome. The fact that it accurately predicted 6 of the 7 BI-RADS 3 and BI-RADS 4a images – the line between definite biopsy and follow-up ultrasound -- provides further hope that the algorithm might help resolve the most difficult cases. The extra trees classifier performed about as well as the ridge classifier and comes with the added benefit of providing a probability. This information could be helpful to a radiologist by providing a numerical estimate of how much faith to put in the recommendations produced by the model, leaving it up to the radiologist to decide whether to switch their diagnosis in cases where there is disagreement.

Lastly, there is a major opportunity to update the rubric from one that is qualitative and descriptive to one that is quantitative. The most salient example is the shape feature, where the radiologist is asked to select whether a lesion is round, oval (flat or plump) or irregular. This is precisely the sort of subjective judgment that might vary between radiologists, or even for the same radiologist on different days. A quantitative metric like sphericity takes away a lot of judgment calls by converting the shape to an objective number, and perhaps getting rid of the descriptor altogether. For features that are difficult to ascertain to the human eye, like heterogeneity of pixel intensity, the computer is once

again at an advantage of being able to easily translating this into multiple numbers, which are not at the whims of what a human is focusing on.

5.4 Suggestions for Future Research

The timeline of a semester-long capstone project necessarily limits the overall scope of a project such as this one. For practical reasons, we were only able to take a somewhat streamlined path from inception (135 raw images) to conclusion (machine learning model). We recognize that there were several alternate routes which warrant consideration, described below.

- *Incorporate patient information* – All of the data in this set was deidentified, but radiologists routinely use contextual information (previous imagery, age, family history, etc.) to help interpret an ultrasound image. Progression of a lesion over time is certainly a key consideration that would make sense to add to a model if the data was available.
- *Multiple images per case* – Our research team only had access to a single image for each lesion, sometimes of dubious quality. In practice, radiologists have access to multiple transverse and longitudinal images for each case, along with doppler images that better capture vascularity (a key predictor of malignancy). We feel the accuracy of our model could potentially be improved by analyzing images as a group.
- *Larger data sets* – Machine learning models perform better when trained on as large a data set as possible. Our set consisted of only 135 images, with a training set of 94 images and a validation set of 41 images. Many of the features in the rubric were not even identified in any of the images by Dr. Ellis, making it impossible to train the software on them.
- *Automating the ground truth* – A subject matter expert segmenting the images was a critical prerequisite for our analysis, and many of our quantitative metrics were highly dependent on the precision of the lesion margins. It is impractical to have images manually segmented at scale,

so it would be interesting to see how a deep learning model performed at this task. Some nascent research in this area has already been done, but the limited size of data sets is a major impediment.

5.5 Conclusion

Computer-assisted diagnosis shows considerable promise, even from a limited data set such as the one in our analysis. While an expert radiologist is able to achieve a very high classification accuracy, a trained model is almost able to keep up and, more importantly, perform well on borderline images. We can very easily imagine an enhanced system where the software provides a malignancy prediction for a given ultrasound image, along with numerical representations of the margins, shape, orientation, echotexture. In this system, the radiologist would be able to update the boundaries of the lesion manually, and these metrics would automatically recalculate, along with the probability of malignancy. This interactive system will hopefully improve the lesion classification accuracy for both benign and malignant lesions for breast ultrasound patients, especially in the gray area between BI-RADS 3 and BI-RADS 4a.

5.6 Acknowledgements

I would like to thank my research supervisors, Dr. Song Chen and Dr. Jeff Baggett, for selecting me to be involved in this project, procuring funding, asking hard questions, providing excellent guidance, and handling all of the behind-the-scenes work that is necessary to make a project like this successful.

Also, I would like to thank my fellow students working on this project: David Halama, Simon Wagner, Justin Hall, Lucas Spellman, and Suriya Mohan. Your technical acumen, curiosity, creativity, optimism, and general friendliness have been incredibly helpful for getting through the difficult times, teaching me new techniques, and forcing me to try keeping up with you.

Lastly, I wanted to thank Dr. Ellis, who has been an invaluable part of our project. While it impossible to learn all there is to know about how to read a breast ultrasound in a few short months, you were somehow able to get us speaking the language and picking out details in the images we would otherwise have never been able to notice.

References

Andrykowski, M. A., Carpenter, J. S., Studts, J. L., Cordova, M. J., Cunningham, L. L., Beacham, A., . . . McGrath, P. (2002). Psychological Impact of Benign Breast Biopsy: A Longitudinal, Comparative Study. *Health Psychology, 21*, 485-494.

Breastcancer.org. (n.d.). *U.S. Breast Cancer Statistics*. Retrieved June 23, 2021, from Breastcanceer.org: https://www.breastcancer.org/symptoms/understand_bc/statistics

Cao, Z., Duan, L., Yang, G., & Chen, Q. (2019). An experimental study on breast lesion detection and classification from ultrasound images using deep learning architectures. *BMC Medical Imaging, 19*, 51.

Daoud, M. I., Bdair, T. M., Al-Najar, M., & Alazrai, R. (2016, November). A Fusion-Based Approach for Breast Ultrasound Image Classification Using Multiple-ROI Texture and Morphological Analyses. *Computational and Mathematical Methods in Medicine, 2016*, 1-12.

DocPanel. (n.d.). *A Practical Guide to Understanding BIRADS*. Retrieved June 22, 2021, from DocPanel: https://www.docpanel.com/blog/post/practical-guide-understanding-bi-rads

Gupta, D., & Anand, R. (2017). A hybrid edge-based segmentation approach for ultrasound medical images. *Biomedical Signal Processing and Control, 31*, 116-126.

Heinig, J., Witteler, R., Schmitz, R., Kiesel, L., & Steinhard, J. (2008). Accuracy of classification of breast ultrasound findings based on criteria used for BI-RADS. *Ultrasound in Obstetrics and Gynecology, 32*(4), 573-578.

Hong, A. S., Rosen, E. L., Soo, M. S., & Baker, J. A. (2005). BI-RADS for sonography: positive and negative predictive values of sonographic features. *American Journal of Roentgenology, 184*(4), 1260-1265.

Lazarus, E., Mainiero, M., Schepps, B., Koelliker, S., & Livingston, L. (2006). BI-RADS lexicon for US and mammography: interobserver variability and positive predictive value. *Radiology, 241*(2), 355-365.

Mayo Clinic. (n.d.). *Mayo Clnic Quick Facts*. Retrieved June 22, 2021, from Mayo Clinic Connect: https://connect.mayoclinic.org/champions/mayo-clinic-quick-facts/

Pfob, A., Barr, R. G., Duda, V., Busch, C., Bruckner, T., Spratte, J., . . . Golatta, M. (2021). A New Practical Decision Rule to Better Differentiate BI-RADS© 3 or 4 Breast Masses on Breast Ultrasound. *Journal of Ultrasound in Medicine, 9999*, 1-10.

Raza, S., Goldkamp, A. L., Chikarmane, S. A., & Birdwell, R. L. (2010). US of Breast Masses Categorized as BI-RADS 3, 4, and 5: Pictorial Review of Factors Influencing Clinical Management. *RadioGraphics, 30*, 1199-1213.

Rizzo, S., Botta, F., Raimondi, S., Origgi, D., Fanciullo, C., Morganti, A. G., & Bellomi, M. (2018). Radiomics: the facts and the challenges of image analysis. *European Radiology Experimental, 2*(36).

Romeo, V., Cuocolo, R., Apolito, R., Stanzione, A., Ventimiglia, A., Vitale, A., . . . 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*. Retrieved from https://doi.org/10.1007/s00330-021-08009-2

Sood, R., Rositch, A., Shakoor, D., Ambinder, E., Pool, K.-L., Pollack, E., . . . Harvey, S. C. (2019). Ultrasound for Breast Cancer Detection Globally: A Systematic Review and Meta-Analysis. *Journal of Global Oncololgy, 5*, 1-17.

van Timmeren, J. E., Cester, D., Tanadini-Lang, S., Alkadhi, H., & Baessler, B. (2020). Radiomics in medical imaging—"how-to" guide and critical reflection. *Insights into Imaging, 11*(91).

Walid, A.-D., Gomaa, M., Khaled, H., & Fahmy, A. (2020). Dataset of breast ultrasound images. *Data in Brief, 28*.

Xian, M., Yingtao, Z., Cheng, H., Xu, F., Zhang, B., & Ding, J. (2018). Automatic breast ultrasound image segmentation: A survey. *Pattern Recognition, 79*, 340-355.

Xu, Y., Wang, Y., Yuan, J., Cheng, Q., Wang, X., & Carson, P. L. (2019). Medical breast ultrasound image segmentation by machine learning. *Ultrasonics*(91), 1-9.

Yap, M. H., Pons, G., Marti, J., Ganau, S., Sentis, M., Zwiggelaar, R., . . . Marti, R. (2018). Automated Breast Ultrasound Lesions Detection Using Convolutional Neural Networks. *IEEE Journal of Biomedical and Health Informatics, 22*(4), 1218-1226.

Appendix A: Code

The code for this project can be found at the following location: https://github.com/adsilb/DS785

Appendix B: Ultrasound Rubric

The following table shows the rubric that Dr. Ellis and his team use to classify breast ultrasound lesions. Some features have a suffix of “ACR”, which stands for American College of Radiology, dating to 2012. Other features have a suffix of “TS”, which stands for Tom Stavros, a radiologist who developed refinements to the 2012 ACR rubric.

Feature Name	**Available Values**
Peripheral Zone ACR	Duct changes; Cooper ligament changes; Edema; Architectural Distortion; Skin Thickening; Skin retraction/irregularity
Peripheral Zone TS	Normal tissue; Shadowing; Enlarged ducts; Enlarged ducts with calcifications; Thin spiculations; Thick spiculations
Marginal Zone ACR	Circumscribed; Not circumscribed; Indistinct; Angular; Microlobulated; Spiculated
Boundary Zone ACR	Abrupt interface; Echogenic halo
Marginal Boundary Zone TS	Well circumscribed; Partial circumscribed; Well circ w/ uniform thick iso or hyperechoic capsule; Well circ w/o thin hyper rim; Indistinct; Thick echogenic rim; Short hyper or hypo echogenic spicules
Shape ACR	Oval; Round; Irregular
Shape TS	Flat oval; Plump oval; Round; Irregular w/o angles & parallel; Irregular w/o angles & not parallel; Irregular w/ angles
Orientation	Parallel; Not parallel
Echo Pattern ACR	Anechoic; Hyperechoic; Complex cystic/solid; Hypoechoic; Isoechoic; Heterogeneous
Echo Pattern TS	Homogeneous hyperechoic; Complex cystic/solid; Mild hypoechoic; Heterogeneous w/o calcs; Heterogeneous w/ calcs; Severe/Marked hypoechoic
Posterior Features ACR	No features; Enhancement; Shadowing; Combined pattern
Posterior Features TS	Enhancement; No features; Mixed, mostly enhanced & partially shadowing; Mixed, partially enhanced & mostly shadowing; Mixed, normal & shadowing; Shadowing
Vascularity ACR	Absent; Internal vascularity; Rim vascularity

Appendix C: Data Set Annotations

The following table shows the annotations provided by Dr. Ellis, his prediction of malignancy, the BI-RADS score he assigned, and the actual malignancy.

	Peripheral Zone		Marginal Zone			Internal Zone									
Sample name	Peripheral Zone ACR	Peripheral Zone TS	Marginal Zone ACR	Boundary Zone ACR	Marginal Boundary Zone TS	Shape ACR	Shape TS	Orientation	Echo Pattern ACR	Echo Pattern TS	Posterior Features ACR	Posterior Features TS	Histology Predicted	BI-RADS	Histology Actual
PA1		Normal tissue	Not circumscribed	Abrupt interface	Partial circumscribed	Oval	Plump oval	Parallel	Anechoic	Severe/Marked hypoechoic	Enhancement	Enhancement	Benign	2	Benign
PA2		Normal tissue	Circumscribed	Abrupt interface	Well circumscribed	Oval	Plump oval	Parallel	Anechoic	Severe/Marked hypoechoic	No features	No features	Benign	2	Benign
PA3		Normal tissue	Circumscribed	Abrupt interface	Well circumscribed	Round	Round	Parallel	Anechoic	Severe/Marked hypoechoic	Enhancement	Mixed, mostly enhanced & partially shadowing	Benign	2	Benign
PA4		Normal tissue	Circumscribed	Abrupt interface	Partial circumscribed	Oval	Flat oval	Parallel	Anechoic	Severe/Marked hypoechoic	No features	No features	Benign	2	Benign
PA5		Normal tissue	Circumscribed	Abrupt interface	Well circumscribed	Round	Round	Parallel	Anechoic	Severe/Marked hypoechoic	Enhancement	Mixed, mostly enhanced & partially shadowing	Benign	2	Benign
PA6		Normal tissue	Circumscribed	Abrupt interface	Well circumscribed	Oval	Flat oval	Parallel	Anechoic	Severe/Marked hypoechoic	No features	No features	Benign	2	Benign
PA7		Normal tissue	Circumscribed	Abrupt interface	Well circumscribed	Oval	Flat oval	Parallel	Hypoechoic	Mild hypoechoic	No features	No features	Benign	3	Benign
PA8		Normal tissue	Circumscribed	Abrupt interface	Well circumscribed	Oval	Flat oval	Parallel	Hyperechoic	Mild hypoechoic	No features	No features	Benign	3	Benign
PA9		Normal tissue	Circumscribed	Abrupt interface	Well circumscribed	Oval	Flat oval	Parallel	Anechoic	Severe/Marked hypoechoic	No features	No features	Benign	2	Benign
PA10		Normal tissue	Circumscribed	Abrupt interface	Well circumscribed	Oval	Flat oval	Parallel	Anechoic	Severe/Marked hypoechoic	No features	No features	Benign	2	Benign
PA12		Normal tissue	Circumscribed	Abrupt interface	Well circumscribed	Oval	Flat oval	Parallel	Anechoic	Severe/Marked hypoechoic	Enhancement	Mixed, partially enhanced & mostly shadowing	Benign	2	Benign
PA14		Normal tissue	Circumscribed	Abrupt interface	Well circumscribed	Oval	Plump oval	Parallel	Heterogeneous	Mild hypoechoic	No features	No features	Benign	3	Benign

PA15		Normal tissue	Circumscribed	Abrupt interface	Well circumscribed	Oval	Plump oval	Parallel	Anechoic	Severe/Marked hypoechoic	No features	No features	Benign	2	Benign
PA16		Normal tissue	Microlobulated	Abrupt interface	Short hyper or hypo echogenic spicuales	Round	Round	Parallel	Hypoechoic	Mild hypoechoic	No features	No features	Malignant	4a	Benign
PA17		Normal tissue	Circumscribed	Abrupt interface	Well circumscribed	Oval	Flat oval	Parallel	Anechoic	Severe/Marked hypoechoic	No features	No features	Benign	2	Benign
PA18		Thin spiculations	Indistinct	Echogenic halo	Indistinct	Irregular	Irregular w/o angles & not parallel	Not parallel	Hypoechoic	Mild hypoechoic	Combined pattern	Mixed, normal & shadowing	Malignant	5	Malignant
PA19		Normal tissue	Circumscribed	Abrupt interface	Well circumscribed	Oval	Flat oval	Parallel	Anechoic	Severe/Marked hypoechoic	No features	No features	Benign	2	Benign
PA20		Normal tissue	Circumscribed	Abrupt interface	Well circumscribed	Oval	Plump oval	Parallel	Hypoechoic	Mild hypoechoic	No features	No features	Benign	3 or 4a	Benign
PA21		Normal tissue	Circumscribed	Echogenic halo	Thick echogenic rim	Oval	Round	Not parallel	Hypoechoic	Mild hypoechoic	Enhancement	Enhancement	Benign	3	Benign
PA22		Normal tissue	Indistinct	Abrupt interface	Indistinct	Irregular	Irregular w/ angles	Not parallel	Complex cystic/solid	Complex cystic/solid	Enhancement	Enhancement	Benign	4b	Benign
PA23		Normal tissue	Circumscribed	Abrupt interface	Well circumscribed	Oval	Plump oval	Parallel	Hypoechoic	Severe/Marked hypoechoic	No features	No features	Benign	2	Malignant
PA24		Normal tissue	Circumscribed	Abrupt interface	Well circumscribed	Oval	Plump oval	Parallel	Hypoechoic	Mild hypoechoic	No features	No features	Benign	3	Malignant
PA25		Normal tissue	Angular	Echogenic halo	Partial circumscribed	Irregular	Irregular w/o angles & parallel		Hypoechoic	Mild hypoechoic	Shadowing	Shadowing	Benign	4b	Malignant
PA26		Normal tissue	Circumscribed	Abrupt interface	Well circumscribed	Oval	Flat oval	Parallel	Hypoechoic	Mild hypoechoic	No features	No features	Benign	3	Benign
PA27	Architectural distortion	Shadowing	Indistinct	Echogenic halo	Indistinct	Oval	Plump oval	Not parallel	Hyperechoic	Mild hypoechoic	Shadowing	Mixed, normal & shadowing	Malignant	4c	Malignant
PA28 Bad Image	Architectural distortion		Microlobulated		Indistinct	Oval	Plump oval	Parallel	Hypoechoic	Mild hypoechoic	No features	No features	Benign	0	Malignant
PA29	Architectural distortion	Thick spiculations	Indistinct	Echogenic halo	Indistinct	Irregular	Irregular w/ angles	Not parallel	Hypoechoic	Severe/Marked hypoechoic	Shadowing	Shadowing	Malignant	5	Malignant
PA30	Architectural distortion	Thick spiculations	Indistinct	Echogenic halo	Indistinct	Irregular	Irregular w/ angles	Not parallel	Hypoechoic	Mild hypoechoic	No features	No features	Malignant	5	Malignant
PA31	Architectural distortion	Thick spiculations	Indistinct		Indistinct	Irregular	Irregular w/ angles	Not parallel	Hyperechoic	Severe/Marked hypoechoic	No features	No features	Malignant	4b	Malignant
PA32		Normal tissue	Circumscribed	Abrupt interface	Well circumscribed	Oval	Plump oval	Parallel	Anechoic	Severe/Marked hypoechoic	No features	No features	Benign	2	Benign
PA33		Normal tissue	Circumscribed	Abrupt interface	Well circumscribed	Oval	Flat oval	Parallel	Anechoic	Severe/Marked hypoechoic	No features	No features	Benign	2	Benign
PA34 Bad Image		Normal tissue	Circumscribed	Abrupt interface	Well circumscribed	Oval	Plump oval	Not parallel	Anechoic	Severe/Marked hypoechoic	No features	No features	Benign	2	Benign
PA35		Normal tissue	Circumscribed	Abrupt interface	Well circumscribed	Oval	Plump oval	Parallel	Anechoic	Severe/Marked hypoechoic	Enhancement	Enhancement	Benign	2	Benign

PA36 Bad Image													Benign		Benign
PA37		Normal tissue	Circumscribed	Abrupt interface	Well circumscribed	Oval	Plump oval	Parallel	Anechoic	Severe/Marked hypoechoic	No features	No features	Benign	2	Benign
PA38		Normal tissue	Circumscribed	Abrupt interface	Well circumscribed	Oval	Flat oval	Parallel	Hypoechoic	Mild hypoechoic	No features	No features	Benign	3	Benign
PA39		Normal tissue	Circumscribed	Abrupt interface	Well circumscribed	Oval	Plump oval	Parallel	Hypoechoic	Mild hypoechoic	No features	No features	Benign	3	Benign
PA40		Normal tissue	Circumscribed	Abrupt interface	Well circumscribed	Oval	Plump oval	Parallel	Hypoechoic	Mild hypoechoic	No features	No features	Benign	2	Benign
PA41		Normal tissue	Microlobulated	Echogenic halo	Partial circumscribed	Oval	Flat oval	Parallel	Complex cystic/solid	Complex cystic/solid	Enhancement	Enhancement	Benign	4a	Malignant
PA42		Normal tissue	Circumscribed	Abrupt interface	Well circumscribed	Oval	Flat oval	Parallel	Anechoic	Severe/Marked hypoechoic	No features	No features	Benign	2	Benign
PA43		Normal tissue	Indistinct		Indistinct	Irregular	Irregular w/o angles & not parallel	Not parallel	Hypoechoic	Mild hypoechoic	Shadowing	Shadowing	Malignant	5	Malignant
PA44		Normal tissue	Circumscribed	Abrupt interface	Well circumscribed	Oval	Plump oval	Parallel	Hypoechoic	Mild hypoechoic	No features	No features	Benign	4a	Malignant
PA45		Normal tissue	Circumscribed	Abrupt interface	Partial circumscribed	Oval	Plump oval	Parallel	Hypoechoic	Mild hypoechoic	Enhancement	Enhancement	Benign	4a	Malignant
PA46	Architectural distortion	Thin spiculations	Indistinct	Echogenic halo	Indistinct	Round	Round	Not parallel	Hypoechoic	Mild hypoechoic	No features	No features	Malignant	4b	Malignant
PA47		Normal tissue	Microlobulated	Abrupt interface	Partial circumscribed	Oval	Plump oval	Parallel	Hypoechoic	Mild hypoechoic	Enhancement	Enhancement	Benign	4a	Malignant
PA49	Architectural distortion		Indistinct	Echogenic halo	Indistinct	Oval	Plump oval	Not parallel	Hypoechoic	Mild hypoechoic	Shadowing	Shadowing	Malignant	5	Malignant
PA50		Normal tissue	Microlobulated	Abrupt interface	Partial circumscribed	Oval	Plump oval	Parallel	Hypoechoic	Mild hypoechoic	Combined pattern	Mixed, mostly enhanced & partially shadowing	Benign	4a	Malignant
PA51		Normal tissue	Circumscribed	Echogenic halo	Well circumscribed	Oval	Plump oval	Parallel	Hypoechoic	Mild hypoechoic	Enhancement	Enhancement	Benign	3	Benign
PA52		Normal tissue	Circumscribed	Abrupt interface	Well circumscribed	Oval	Plump oval	Parallel	Hypoechoic	Mild hypoechoic	No features	No features	Benign	3	Benign
PA53		Normal tissue	Indistinct	Echogenic halo	Indistinct	Irregular	Irregular w/o angles & not parallel	Not parallel	Heterogeneous	Heterogeneous w/o calcs	Shadowing	Shadowing	Malignant	4b	Malignant
PA54		Normal tissue	Microlobulated		Indistinct	Round	Round	Not parallel	Heterogeneous	Heterogeneous w/o calcs	No features	No features	Malignant	4b	Malignant
PA55	Architectural distortion	Thick spiculations	Indistinct		Indistinct	Irregular	Irregular w/o angles & not parallel	Not parallel	Anechoic	Severe/Marked hypoechoic	Shadowing	Shadowing	Malignant	5	Malignant
PA56		Normal tissue	Circumscribed	Abrupt interface	Well circumscribed	Oval	Flat oval	Parallel	Hypoechoic	Mild hypoechoic	No features	No features	Benign	3	Benign
PA57	Architectural distortion	Thick spiculations	Indistinct	Echogenic halo	Indistinct	Irregular	Irregular w/o angles &	Not parallel	Anechoic	Severe/Marked hypoechoic	Shadowing	Shadowing	Malignant	4c	Malignant

							not parallel								
PA58		Normal tissue	Circumscribed	Abrupt interface	Well circumscribed	Oval	Plump oval	Parallel	Hypoechoic	Mild hypoechoic	Enhancement	Enhancement	Benign	2	Benign
PA59		Normal tissue	Circumscribed	Abrupt interface	Well circumscribed	Oval	Plump oval	Parallel	Anechoic	Severe/Marked hypoechoic	Enhancement	Enhancement	Benign	2	Benign
PA61		Normal tissue	Circumscribed	Abrupt interface	Well circumscribed	Oval	Plump oval	Parallel	Anechoic	Severe/Marked hypoechoic	Enhancement	Enhancement	Benign	2	Benign
PA62		Normal tissue	Circumscribed	Abrupt interface	Well circumscribed	Oval	Plump oval	Parallel	Anechoic	Severe/Marked hypoechoic	Enhancement	Enhancement	Benign	2	Benign
PA63		Normal tissue	Circumscribed	Abrupt interface	Well circumscribed	Oval	Plump oval	Parallel	Anechoic	Severe/Marked hypoechoic	Enhancement	Enhancement	Benign	2	Benign
PA64		Normal tissue	Circumscribed	Abrupt interface	Well circumscribed	Oval	Plump oval	Parallel	Anechoic	Severe/Marked hypoechoic	Enhancement	Enhancement	Benign	2	Benign
PA65		Normal tissue	Circumscribed	Abrupt interface	Well circumscribed	Oval	Plump oval	Parallel	Anechoic	Severe/Marked hypoechoic	Enhancement	Enhancement	Benign	2	Benign
PA66		Normal tissue	Circumscribed	Abrupt interface	Well circumscribed	Oval	Plump oval	Parallel	Anechoic	Severe/Marked hypoechoic	Enhancement	Enhancement	Benign	2	Benign
PA67		Normal tissue	Circumscribed	Abrupt interface	Well circumscribed	Oval	Plump oval	Parallel	Hypoechoic	Mild hypoechoic	No features	No features	Benign	2	Benign
PA68		Normal tissue	Circumscribed	Abrupt interface	Well circumscribed	Oval	Plump oval	Parallel	Anechoic	Severe/Marked hypoechoic	Enhancement	Enhancement	Benign	2	Benign
PA69		Normal tissue	Circumscribed	Abrupt interface	Well circumscribed	Oval	Plump oval	Parallel	Hypoechoic	Mild hypoechoic	No features	No features	Benign	3	Benign
PA70		Normal tissue	Circumscribed	Abrupt interface	Well circumscribed	Oval	Plump oval	Parallel	Anechoic	Severe/Marked hypoechoic	Enhancement	Enhancement	Benign	2	Benign
PA71		Normal tissue	Circumscribed	Abrupt interface	Well circumscribed	Oval	Plump oval	Parallel	Anechoic	Severe/Marked hypoechoic	Enhancement	Enhancement	Benign	2	Benign
PA72		Normal tissue	Circumscribed	Abrupt interface	Well circumscribed	Oval	Plump oval	Parallel	Anechoic	Severe/Marked hypoechoic	Enhancement	Enhancement	Benign	2	Benign
PA73		Normal tissue	Circumscribed	Abrupt interface	Well circumscribed	Oval	Plump oval	Parallel	Hypoechoic	Mild hypoechoic	No features	No features	Benign	3	Benign
PA74		Normal tissue	Circumscribed	Abrupt interface	Well circumscribed	Oval	Plump oval	Parallel	Hypoechoic	Mild hypoechoic	No features	No features	Benign	3	Benign
PA76		Normal tissue	Circumscribed	Abrupt interface	Well circumscribed	Oval	Plump oval	Parallel	Hypoechoic	Mild hypoechoic	No features	No features	Benign	3	Benign
PA77		Normal tissue	Circumscribed	Abrupt interface	Well circumscribed	Oval	Flat oval	Parallel	Hypoechoic	Mild hypoechoic	No features	No features	Benign	3	Benign
PA78		Normal tissue	Indistinct		Partial circumscribed	Irregular	Irregular w/o angles & parallel	Parallel	Hypoechoic	Mild hypoechoic	No features	No features	Benign	4a	Benign
PA79		Normal tissue	Circumscribed	Abrupt interface	Well circumscribed	Oval	Plump oval	Parallel	Hypoechoic	Mild hypoechoic	No features	No features	Benign	4a	Benign
PA80		Normal tissue	Circumscribed	Abrupt interface	Well circumscribed	Oval	Plump oval	Parallel	Hypoechoic	Mild hypoechoic	No features	No features	Benign	3	Benign
PA83		Normal tissue	Circumscribed	Abrupt interface	Well circumscribed	Oval	Plump oval	Parallel	Hypoechoic	Mild hypoechoic	No features	No features	Benign	3	Benign
PA84	Architectural distortion	Shadowing		Echogenic halo	Partial circumscribed	Oval	Plump oval	Parallel	Anechoic	Severe/Marked hypoechoic	Shadowing	Shadowing	Benign	4a	Benign

PA85		Normal tissue	Circumscribed	Abrupt interface	Well circumscribed	Oval	Plump oval	Parallel	Hypoechoic	Mild hypoechoic	No features	No features	Benign	3	Benign
PA86		Normal tissue	Circumscribed	Abrupt interface	Well circumscribed	Oval	Plump oval	Parallel	Hypoechoic	Mild hypoechoic	Enhancement	Enhancement	Benign	3	Benign
PA87		Normal tissue	Circumscribed	Abrupt interface	Well circumscribed	Oval	Flat oval	Parallel	Hypoechoic	Mild hypoechoic	No features	No features	Benign	2	Benign
PA89		Normal tissue	Circumscribed	Abrupt interface	Well circumscribed	Oval	Plump oval	Parallel	Hypoechoic	Mild hypoechoic	Enhancement	Enhancement	Benign	2	Benign
PA91		Normal tissue	Circumscribed	Abrupt interface	Well circumscribed	Oval	Flat oval	Parallel	Hypoechoic	Mild hypoechoic	No features	No features	Benign	2	Benign
PA92		Normal tissue	Circumscribed	Abrupt interface	Well circumscribed	Oval	Plump oval	Parallel	Anechoic	Severe/Marked hypoechoic	Enhancement	Enhancement	Benign	2	Benign
PA93		Normal tissue	Circumscribed	Abrupt interface	Well circumscribed	Oval	Plump oval	Parallel	Anechoic	Severe/Marked hypoechoic	No features	No features	Benign	3	Benign
PA94			Microlobulated		Indistinct	Oval	Plump oval	Parallel	Hypoechoic	Mild hypoechoic	No features	No features	Benign	4a	Benign
PA95		Normal tissue	Circumscribed	Abrupt interface	Well circumscribed	Oval	Plump oval	Parallel	Hypoechoic	Mild hypoechoic	No features	No features	Benign	2	Benign
PA96		Normal tissue	Circumscribed	Abrupt interface	Well circumscribed	Oval	Plump oval	Parallel	Hypoechoic	Mild hypoechoic	Enhancement	Enhancement	Benign	2	Benign
PA97		Normal tissue	Circumscribed	Abrupt interface	Well circumscribed	Oval	Plump oval	Parallel	Hypoechoic	Mild hypoechoic	No features	No features	Benign	2	Benign
PA98		Normal tissue	Circumscribed	Abrupt interface	Well circumscribed	Round	Plump oval	Parallel	Hypoechoic	Mild hypoechoic	Enhancement	Enhancement	Benign	2	Benign
PA102		Normal tissue	Circumscribed	Abrupt interface	Well circumscribed	Oval	Plump oval	Parallel	Anechoic	Severe/Marked hypoechoic	No features	No features	Benign	2	Benign
PA103		Normal tissue	Circumscribed	Abrupt interface	Well circumscribed	Round	Round	Parallel	Anechoic	Severe/Marked hypoechoic	Enhancement	Enhancement	Benign	2	Benign
PA104		Normal tissue	Circumscribed	Abrupt interface	Well circumscribed	Oval	Plump oval	Parallel	Anechoic	Severe/Marked hypoechoic	Enhancement	Enhancement	Benign	2	Benign
PA105		Normal tissue	Circumscribed	Abrupt interface	Well circumscribed	Oval	Plump oval	Parallel	Anechoic	Severe/Marked hypoechoic	No features	No features	Benign	2	Benign
PA106		Normal tissue	Circumscribed	Abrupt interface	Well circumscribed	Oval	Plump oval	Parallel	Hypoechoic	Mild hypoechoic	No features	No features	Benign	2	Benign
PA109		Normal tissue	Circumscribed	Abrupt interface	Well circumscribed	Oval	Plump oval	Parallel	Anechoic	Severe/Marked hypoechoic	No features	No features	Benign	2	Benign
PA110		Normal tissue	Circumscribed	Abrupt interface	Well circumscribed	Oval	Plump oval	Parallel	Anechoic	Severe/Marked hypoechoic	No features	No features	Benign	2	Benign
PA112		Normal tissue	Circumscribed	Abrupt interface	Well circumscribed	Oval	Plump oval	Parallel	Anechoic	Severe/Marked hypoechoic	No features	No features	Benign	2	Benign
PA113		Normal tissue	Circumscribed	Abrupt interface	Well circumscribed	Oval	Flat oval	Parallel	Anechoic	Severe/Marked hypoechoic	No features	No features	Benign	2	Benign
PA114		Normal tissue	Circumscribed	Abrupt interface	Well circumscribed	Oval	Flat oval	Parallel	Anechoic	Severe/Marked hypoechoic	No features	No features	Benign	2	Benign
PA115		Normal tissue	Circumscribed	Abrupt interface	Well circumscribed	Oval	Plump oval	Parallel	Anechoic	Severe/Marked hypoechoic	Enhancement	Enhancement	Benign	2	Benign
PA116		Normal tissue	Circumscribed	Abrupt interface	Well circumscribed	Oval	Flat oval	Parallel	Anechoic	Severe/Marked hypoechoic	No features	No features	Benign	2	Benign

PA117		Normal tissue	Circumscribed	Abrupt interface	Well circumscribed	Oval	Plump oval	Parallel	Anechoic	Severe/Marked hypoechoic	No features	No features	Benign	2	Benign
PA118		Normal tissue	Circumscribed	Abrupt interface	Well circumscribed	Oval	Plump oval	Parallel	Anechoic	Severe/Marked hypoechoic	No features	No features	Benign	2	Benign
PA119		Normal tissue	Circumscribed	Abrupt interface	Well circumscribed	Oval	Plump oval	Parallel	Anechoic	Severe/Marked hypoechoic	Enhancement	Enhancement	Benign	2	Benign
PA120		Normal tissue	Circumscribed	Abrupt interface	Well circumscribed	Oval	Plump oval	Parallel	Anechoic	Severe/Marked hypoechoic	Enhancement	Enhancement	Benign	2	Benign
PA121		Normal tissue	Circumscribed	Abrupt interface	Well circumscribed	Oval	Plump oval	Parallel	Anechoic	Severe/Marked hypoechoic	No features	No features	Benign	2	Benign
PA122		Normal tissue	Circumscribed	Abrupt interface	Well circumscribed	Oval	Plump oval	Parallel	Anechoic	Severe/Marked hypoechoic	Enhancement	Enhancement	Benign	2	Benign
PA124		Normal tissue	Circumscribed	Abrupt interface	Well circumscribed	Round	Round	Parallel	Anechoic	Severe/Marked hypoechoic	Enhancement	Enhancement	Benign	2	Benign
PA125		Normal tissue	Circumscribed	Abrupt interface	Well circumscribed	Round	Round	Parallel	Anechoic	Severe/Marked hypoechoic	Enhancement	Enhancement	Benign	2	Benign
PA126		Normal tissue	Circumscribed	Abrupt interface	Well circumscribed	Oval	Flat oval	Parallel	Anechoic	Severe/Marked hypoechoic	Enhancement	Enhancement	Benign	2	Benign
PA127		Normal tissue	Circumscribed	Abrupt interface	Well circumscribed	Oval	Flat oval	Parallel	Anechoic	Severe/Marked hypoechoic	No features	No features	Benign	2	Benign
PA128		Normal tissue	Circumscribed	Abrupt interface	Well circumscribed	Oval	Flat oval	Parallel	Anechoic	Severe/Marked hypoechoic	Enhancement	Enhancement	Benign	2	Benign
PA129		Normal tissue	Circumscribed	Abrupt interface	Well circumscribed	Oval	Flat oval	Parallel	Anechoic	Severe/Marked hypoechoic	Enhancement	Enhancement	Benign	2	Benign
PA130		Normal tissue	Circumscribed	Abrupt interface	Well circumscribed	Oval	Plump oval	Parallel	Anechoic	Severe/Marked hypoechoic	No features	No features	Benign	2	Benign
PA131		Normal tissue	Circumscribed	Abrupt interface	Well circumscribed	Oval	Plump oval	Not parallel	Anechoic	Severe/Marked hypoechoic	Enhancement	Enhancement	Benign	2	Benign
PA132	Architectural distortion	Shadowing	Indistinct		Indistinct	Irregular	Irregular w/o angles & parallel	Not parallel	Hypoechoic	Mild hypoechoic	Shadowing	Shadowing	Malignant	5	Malignant
PA133			Microlobulated	Echogenic halo	Indistinct	Irregular	Irregular w/o angles & parallel	Parallel	Hypoechoic	Mild hypoechoic	No features	No features	Malignant	4b	Malignant
PA134		Normal tissue	Circumscribed	Abrupt interface	Well circumscribed	Oval	Flat oval	Parallel	Anechoic	Severe/Marked hypoechoic	Enhancement	Enhancement	Benign	2	Malignant
PA136		Thin spiculations	Angular	Echogenic halo	Indistinct	Oval	Plump oval	Parallel	Hypoechoic	Mild hypoechoic	Shadowing	Shadowing	Malignant	4b	Malignant
PA137		Normal tissue											Benign	1	Malignant
PA140	Architectural distortion	Thin spiculations	Microlobulated	Echogenic halo	Indistinct	Oval	Plump oval	Not parallel	Anechoic	Severe/Marked hypoechoic	Shadowing	Shadowing	Malignant	4c	Malignant

PA141	Architectural distortion	Thick spiculations	Angular		Indistinct	Irregular	Irregular w/o angles & not parallel	Not parallel	Hypoechoic	Mild hypoechoic	Shadowing	Shadowing	Malignant	5	Malignant
PA142	Architectural distortion	Thin spiculations	Angular		Indistinct	Irregular	Irregular w/o angles & not parallel	Not parallel	Anechoic	Severe/Marked hypoechoic	Shadowing	Shadowing	Malignant	5	Malignant
PA143	Architectural distortion	Thick spiculations	Angular		Indistinct	Irregular	Irregular w/o angles & not parallel	Not parallel	Hypoechoic	Mild hypoechoic	Shadowing	Shadowing	Malignant	5	Malignant
PA144	Architectural distortion	Thin spiculations	Angular		Indistinct	Irregular	Irregular w/o angles & not parallel	Not parallel	Anechoic	Severe/Marked hypoechoic	Shadowing	Shadowing	Malignant	5	Malignant
PA145	Architectural distortion	Thin spiculations	Angular		Indistinct	Irregular	Irregular w/o angles & not parallel	Not parallel	Hypoechoic	Mild hypoechoic	Shadowing	Shadowing	Malignant	5	Malignant
PA147		Normal tissue	Microlobulated	Echogenic halo	Partial circumscribed	Round	Round		Hypoechoic	Mild hypoechoic		No features	Malignant	4c	Malignant
PA295	Architectural distortion		Indistinct	Echogenic halo	Indistinct	Irregular	Irregular w/o angles & not parallel	Not parallel	Anechoic	Severe/Marked hypoechoic	Shadowing	Shadowing	Malignant	5	Malignant
PA296		Thin spiculations	Indistinct		Indistinct	Irregular	Irregular w/o angles & parallel	Not parallel	Hypoechoic	Mild hypoechoic	Shadowing	Shadowing	Malignant	4b	Malignant
PA297	Architectural distortion	Thick spiculations	Angular	Echogenic halo	Indistinct	Irregular	Irregular w/o angles & not parallel	Not parallel	Hypoechoic	Mild hypoechoic	Shadowing	Shadowing	Malignant	5	Malignant
PA298	Architectural distortion	Thin spiculations	Indistinct		Indistinct	Irregular	Irregular w/o angles & parallel	Parallel	Hypoechoic	Mild hypoechoic	No features	No features	Malignant	5	Malignant
PA299	Architectural distortion	Thin spiculations	Indistinct	Echogenic halo	Indistinct	Irregular	Irregular w/o angles & not parallel	Not parallel	Hypoechoic	Mild hypoechoic	Shadowing	Shadowing	Malignant	5	Malignant
PA300			Angular		Indistinct	Irregular	Irregular w/o angles & parallel	Parallel	Hypoechoic	Mild hypoechoic	Shadowing	Shadowing	Malignant	4b	Malignant
PA301		Shadowing	Indistinct		Indistinct	Irregular	Irregular w/o angles & parallel	Parallel	Hypoechoic	Mild hypoechoic	Shadowing	Shadowing	Malignant	4b	Malignant
PA302		Normal tissue	Microlobulated	Echogenic halo	Well circ w/ uniform thick iso or hyperechoic capsule	Irregular	Irregular w/o angles & parallel	Parallel	Hypoechoic	Mild hypoechoic	No features	No features	Malignant	5	Malignant

PA304		Normal tissue	Circumscribed	Abrupt interface	Well circumscribed	Oval	Flat oval	Parallel	Hypoechoic	Mild hypoechoic	No features	No features	Benign	2	Malignant
PA305		Normal tissue	Microlobulated		Partial circumscribed	Oval	Plump oval	Parallel	Complex cystic/solid	Complex cystic/solid	Enhancement	Enhancement	Malignant	4a	Malignant
PA306	Architectural distortion		Angular	Echogenic halo	Indistinct	Irregular	Irregular w/ angles	Not parallel	Hypoechoic	Mild hypoechoic	Shadowing	Shadowing	Malignant	5	Malignant
PA307		Normal tissue	Circumscribed	Abrupt interface	Well circumscribed	Oval	Flat oval	Parallel	Hypoechoic	Mild hypoechoic	No features	No features	Benign	2	Malignant
PA308	Architectural distortion	Thick spiculations	Indistinct	Echogenic halo	Indistinct	Irregular	Irregular w/o angles & not parallel	Not parallel	Anechoic	Severe/Marked hypoechoic	Shadowing	Shadowing	Malignant	5	Malignant
PA309			Indistinct	Echogenic halo	Indistinct	Irregular	Irregular w/o angles & not parallel	Not parallel	Hypoechoic	Mild hypoechoic	Shadowing	Shadowing	Malignant	4c	Malignant
PA310		Normal tissue	Microlobulated	Abrupt interface	Indistinct	Oval	Plump oval	Parallel	Hypoechoic	Mild hypoechoic	No features	No features	Malignant	4b	Malignant