Automated Segmentation and Classification of Lesion on Breast Ultrasound

Submitted By:
ZAHRA FAZAL
NUST201260464MSMME62412F

Supervisor:
DR. SYED OMER GILANI

Co-Supervisor:
DR. SYED IRTIZA ALI SHAH

Department of Biomedical Engineering and Sciences
School of Mechanical and Manufacturing Engineering
National University of Sciences and Technology
Islamabad, Pakistan
2014

Declaration

It is hereby declared that this research study has been done for the partial fulfillment of requirements for the degree of Master of Science in Biomedical sciences. I hereby declared that no portion for this work has submitted in support of an application for another degree, to any other university.

Zahra Fazal

MASTER THESIS WORK

We hereby recommend that the dissertation prepared under our supervision by:

Zahra Fazal (NUST201260464MSMME62412F)

Titled**:** **Automated segmentation and classification of lesion on breast ultrasound** be accepted in partial fulfillment of the requirements for the award of **MS** degree with ___ Grade.

Examination Committee Members

1. Name:_____ Signature:_____

2. Name:_____ Signature:_____

3. Name:_____ Signature:_____

Supervisor's name: _____ Signature:_____
Date:_____

Co-Supervisor's name: _____ Signature:_____
Date:_____

_____ _____
Head of Department Date

COUNTERSINGED

_____ _____
Dean/Principal Date

Dedicated to my parents ……...for their love, support and encouragement

And

Dedicated to my brothers Muhammad Ather,Faisal Fazal and Zayed Fazal… for their care and endless happiness.

Acknowledgement

First and foremost thanks to Almighty ALLAH for his countless blessing upon me and strength provided to be by him to face and be patient with all the challenges.

I would like to acknowledge various people for their contribution towards the completion of this project. I would like to express my deep gratitude to Dr. Syed Irtiza Ali Shah and Dr. Syed Omer Gilani for their patience, guidance, encouragement and useful critiques on this project.

I would also like to thanks Dr. Muhammad Nabeel Anwar for his advice and assistance in keeping my progress on schedule. I would like to acknowledge to Dr. Khubaib for providing me with data.

I would like to thanks Muhammad Samran Navid and Rabeil Sakina for always being there for me to guide and support me. I would like to appreciate for the support and care they have given me throughout this project.

Zahra Fazal

Table of Contents

1 INTRODUCTION 1

1.1 HISTORICAL BACKGROUND 1

1.2 BREAST IMAGING-REPORTING AND DATA SYSTEMS (BI_RAD) 2

1.3 MAMMOGRAPHY 2

1.4 ULTRASOUND 4

1.5 FILTRATION 5

1.5.1 Non-Linear Anisotropic Diffusion Filter 6

1.5.2 Median Filter 6

1.5.3 Sigmoid Filter 7

1.6 AUTOMATED SEGMENTATION 7

1.7 FEATURE EXTRACTION 8

1.8 MACHINE LEARNING 8

1.8.1 Unsupervised learning 8

1.8.2 Supervised learning 9

1.9 MOTIVATION FOR CURRENT STUDY 12

1.9.1 Operational Definition 13

1.9.2 Objective 13

2 LITERATURE REVIEW 14

2.1 HISTORICAL PERSPECTIVE 14

2.2 STAVROS (1995) 14

2.3 GIGER *ET*.AL (1990) 15

2.4 GIGER *ET*.AL (2002) 15

2.5 CHEN *ET*.AL (2002) 15

2.6 SEGYEONG JOO *ET*.AL (2004) 16

2.7 DRUKKER AND GIGER *ET*.AL (2004) 16

2.8 CHEN ET.AL (2005) 16

2.9 CHANG ET.AL (2005) 17

2.10KARLA HORSCH *ET*.AL (2006) 17

2.11 HUANG *ET*.AL (2007) 17

2.12WOO KYUNG MOON*ET*.AL (2013) 17

3 METHODOLOGY 19

3.1 Experimental protocol 19
3.1.1 Data Acquisition *19*
3.1.2 Sample Size *19*
3.1.3 Study population *20*
3.1.4 Study type *20*
3.1.5 Inclusive criteria *20*
3.1.6 Exclusive criteria *20*
3.2 CAD System Design 21
3.2.1 Preprocessing *22*
3.2.2 Non-linear Anisotropic Diffusion *22*
3.2.3 Median Filter *23*
3.2.4 Sigmoid Filter *24*
3.3 Segmentation 24
3.3.1 Active Contours *24*
3.4 Morphological Features 27
3.4.1 Shape *27*
3.4.2 Orientation *29*
3.4.3 Margin *29*
3.4.4 Lesion Boundary *30*
3.4.5 Posterior Acoustic enhancement *31*
3.5 Textural Features 31
3.5.1 Echogenity *31*
3.6 Evaluation 34
3.6.1 Receiver operating characteristics (ROC) *34*
3.6.2 Accuracy *35*
3.6.3 Sensitivity *36*
3.6.4 Specificity *36*
3.7 Support vector Machine (SVM) Classifier 37
3.7.1 One vs. All *37*
4 RESULTS 39
4.1 Filtration results 39
4.2 Segmentation results 40
4.3 Evaluation metrics 41
4.4 Features analysis for CAD model 44
4.5 Morphological features analysis 44

4.6 TEXTURAL FEATURES ANALYSIS 47

5 DISCUSSION 51

5.1 BENIGN-MALIGNANT GROUP 51

5.2 SIGNIFICANT FEATURE FOR BENIGN-MALIGNANT GROUP 51

5.3 BENIGN-MALIGNANT-NORMAL GROUP 52

5.4 SIGNIFICANT FEATURE FOR BENIGN-MALIGNANT-NORMAL GROUP 52

5.5 LIMITATIONS 52

6 CONCLUSION &.FUTURE WORK 53

REFERENCES 55

List of Figures

Figure 1-1: CAD models assessment in Mammography. This figure describes the order in which CAD model works to help radiologist at reporting table 3

Figure 1-2: CAD Model Construction. This figure describes the key components in CAD Model. 4

Figure 1-3: Speckle noise removal sequence. This figure describe the procedure followed in developing proposed CAD model 5

Figure 1-4: Median filters functioning. This figure describes the procedure in which median filter work by sorting al pixels and replacing central pixel with the median value[8]. 6

Figure 1-5: Sigmoid Function adopted from [9] 7

Figure 1-6: SVM with separating hyperplane separating the two classes, defined by the equation 10

Figure 1-7:SVM with optimal hyperplane and optimization of decision boundary[13] 11

Figure 3-1: Segmentation using Active contour model 27

Figure 3-2: ROC Curve adopted form (http://www.ebm.ugent.be/ROC-curve.jpg) 35

Figure 4-1: This image shows the results provide on applying various results in developed CAD model 39

Figure 4-2:(a) Image filtered with Anisotropic diffusion filter for removal of speckle noise.(b) anisotropic image filtered with median filter to enhance edge and blur the background (c) Image filtered through sigmoid filter for contrast enhancement of lesion .(d)After processing through all filters, image segmented by active contour method. 41

Figure 4-3: (a) Sensitivity of accuracy to change in σ values for radial kernel. Az (area under the curve) for various γ values used in radial kernel for SVM in benign-malignant group and locating the highest possible accuracy of classifier with γ to be used in C 41

Figure 4-4: ROC curve indicating accuracy of proposed CAD model as area under the curve. (a)Receiver operating characteristics (ROC) curve of benign-malignant-normal group with textural features and area under the ROC curve is (Az = 0.92) for benign, (Az value= 0.84) for malignant and (0.86 for normal) 42

Figure 4-5: ROC curve indicating accuracy of proposed CAD model as area under the curve. Roc curve of benign-malignant group with morphological features and area under the ROC curve is (Az=0.97) 43

List of Tables

Table 1.1: BI-RADS stages and findings..... 2

Table 3.1: Sonographic Features and their description according to the Extracted Features..... 33

Table 3.2: Confusion matrix 36

Table 4.1: Performance of the benign malignant group with morphological features. This table indicates the performance of the proposed CAD model for benign and malignant group 44

Table 4.2: Performance of the benign-malignant-normal group with textural features. This table indicates the performance of the proposed CAD model for benign-malignant and normal group44

Table 4.3: Az (area under the curve) for thirteen morphological features in classifying benign and malignant lesion on breast ultrasound images. This table indicates the accuracy of features used in the CAD model for classifying benign and malignant tumor. Thirteen features are arranged in the order of (Spiculation, branch-pattern, lobulation, relative brightness of nodule, ENC, LS ratio, roundness, form factor, solidity, convexity, extent, aspect ratio, posterior enhancement). 45

Table 4.4: The Sensitivity of thirteen morphological features for classifying benign and malignant lesions on breast ultrasound images. This table represents the sensitivity of features used in CAD model for classifying benign and malignant tumor. Thirteen features are arranged in the order of (Spiculation, branch-pattern, lobulation, relative brightness of nodule, ENC, LS ratio, roundness, form factor, solidity, convexity, extent, aspect ratio, posterior enhancement). 46

Table 4.5: The Specificity of thirteen morphological features in classifying benign and malignant lesion on breast ultrasound images. This table represents the specificity of features used in CAD model for classifying benign and malignant tumor. Thirteen features are arranged in the order of (Spiculation, branch-pattern, lobulation, relative brightness of nodule, ENC, LS ratio, roundness, form factor, solidity, convexity, extent, aspect ratio, posterior enhancement) . 47

Table 4.6: Az (area under the curve) for textural features in classifying benign, malignant and normal images. This table indicates the accuracy of features used in CAD model for classifying benign, malignant and normal images. Sixteen features are arranged in the order of (contrast, correlation, energy and homogeneity) 48

Table 4.7: Sensitivity of Textural features in classifying benign and malignant and normal image. This table indicates the sensitivity of features used in CAD model for classifying benign, malignant and normal images. Sixteen features are arranged in the order of (contrast, correlation, energy and homogeneity)..... 49

Table 4.8: Specificity of Textural features in classifying benign and malignant and normal image. This table indicates the specificity of features used in CAD model for classifying benign, malignant and normal images. Sixteen features are arranged in the order of (contrast, correlation, energy and homogeneity)..... 50

List of Abbreviations

USG	……	Ultrasound
BI-RADS	……	Breast imaging reporting system
CAD	……	Computer Aided Diagnosis
SVM	……	Support vector Machine
ROC	……	Receiver operating characteristic
Az	……	Area under the curve
$\|\nabla u\|^2$	……	Describing the like hood of edges
$G_\sigma * \mu_0$	……	Convolution of image with Gaussian filters
α	……	Smoothness of the contour
β	……	Smoothness of the contour
λ	……	Attracts the contours toward the boundary
D	……	Diffusion tensor, symmetric matrix
C1	……	Average intensity inside the contour +1
C2	……	Average intensity outside the contour -1
ACM	……	Active Contour Model

Abstract

In this research we have developed and evaluated a computer aided diagnosis (CAD) model that is based on the automated segmentation of breast lesion on ultrasound images. Active contour model has been used for segmentation after removal of speckle noise followed by morphological and textural features computation. The CAD model is based on the breast imaging reporting and Data system (BI-RAD) for major feature selection to classify benign-malignant group and benign-malignant-normal group. Ultrasound images were collected from multiple hospitals with their consent. The data set consisted of 163 actual ultrasound images of benign, malignant and normal images used for the analysis. In the proposed method, after segmentation and feature extraction test image is placed into either of two groups of benign-malignant group and benign-malignant-normal group. Binary support vector machine classifier has been used to lesion based identification of breast tumor as benign or malignant in the first group, whereas multiclass SVM using one vs. all method has been used for similar identification in the second group. All cases were samples with k-fold cross validation method, performance of the classifier for classifying were evaluated by receiver operating characteristics in both groups. The area under the curve for benign-malignant group using morphological features were 0.97, classifier accuracy to be 94% with sensitivity and specificity of 97% and 88%.In benign-malignant and normal group areas under the curve for benign, malignant and normal group was 0.94%, 0.84 and 0.86 % with a sensitivity and specificity of 94% and 83%. It was concluded that proposed CAD model was able to differentiate, with acceptable accuracy, benign from malignant breast tumor using morphological features and normal from benign and malignant using textural features. Hence this model can be used to detect tumor and can provide a reliable second opinion to the radiologist.

***Key words**: breast imaging reporting and Data system (BI-RAD), support vector machine (SVM), ultrasound (US), computer aided diagnosis (CAD)*

1
INTRODUCTION

Breast cancer is the leading cause of death in developing countries. According to pink ribbon Pakistan last year 30,000 women died from breast cancer in Pakistan. Mortality rate depends upon the earlier diagnosis of breast cancer as early the diagnosis is better chance of treatment it provides as there will be less chance of cancer spreading to the distinct parts of the body. Most common equipment used for diagnosis in breast cancer is mammography and ultrasound, mammography is usually recommended in women above 40 years as a regular screening procedure to reduce mortality rate.Mammography is used in adjunct with ultrasound due to its superior diagnostic ability in women with dense breasts and also occult lesion not clearly visible on mammography can be seen in ultrasound.

Sonography is also utilized in needle guided biopsies to confirm if the tumor is benign or malignant and is also used for preoperative evaluation. However, ultrasound is operator dependent, lesions are evaluated on the basis of six BI-RADS descriptive and CAD model based on features extracted on the basis of BI-RADS descriptors can proves useful for eliminating operator dependency. A CAD model has potential to improve diagnosis in less experienced reader, generate accurate reports and reduce unnecessary biopsies

1.1 Historical Background

Computer aided diagnosis (CAD) in ultrasound are discovered by various researchers.Most of the CAD models are focused on differentiating benign from malignant lesion. Giger *et.*al used various methods of differentiating benign from malignant lesion as well as malignant lesions from various benign classes[1]. Chen *et.* al, utilized the textural features to classify lesion using neural network[2],Woo *et*.al, uses BI-RADS for finding and categorizing lesion[3] .In the benchmark study by Stavros *et*.al,they described all features that were associated with breast cancer[4]. Most of the CAD models use these features to differentiate benign and malignant lesion.

1.2 Breast Imaging-Reporting and Data systems (BI_RAD)

The BI_RADS stands for breast imaging reporting and data systems which is widely accepted standardize assessment criteria developed by American College of Radiology for evaluating breast cancer on mammography, ultrasound and MRI (Magnetic Resonance imaging)[5]. Table 1.1lists the stages of BI-RAD and findings associated with them.

Table 1.1: BI-RADS stages and findings

Stages	Findings
BIRADS0	Incomplete (further imaging or information is required)
BIRADS I	Negative
BIRADS II	Benign Findings
BIRADS III	Probably Benign
BIRADS IV	Suspicious abnormality (needle biopsy recommended)
BIRADS V	Mammographic appearance highly suggestive of malignancy
BIRADS VI	Proven malignant (biopsy confirmed)

1.3 Mammography

Mammography is the most common diagnostic method used for breast cancer, women above 40 years are usually recommended for regular screening. Mammography uses low dose X-rays to examine breast and is used in the detection and diagnosis of breast disease in women. Mammography was printed on X-ray film now it has been printing on digital film that converts x-rays after falling onto them in electric signals.

Mammography mostly involves two views craniocaudal view and mediolateral oblique views with sometime additional views true lateral and spot compression. In females having dense breast, lesions cannot be clearly visible in mammograms and can increase false positive rate. Commercially available CAD(computer aided diagnosis) and CADe (computer aided detection) models are used in mammography to reduce unnecessary biopsies, as well as prove useful in diagnosing breast tumor for inexperienced radiologist, also provides second set of eyes to support and enhance the radiologist's judgment in making accurate reports. A typical CAD session works in the following order described below in Figure 1-1.

Some of most common commercially available CAD models in mammography are "Image Checker" (R2 Technology, Sunnyvale, CA) and "SecondLook", these systems are under investigation for their benefits in diagnostic capabilities [6].

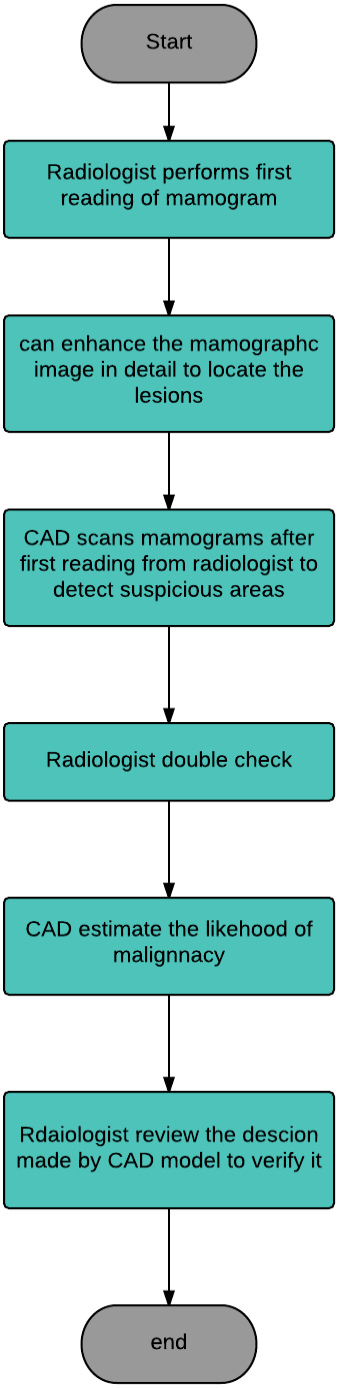


Figure 1-1: CAD models assessment in Mammography. This figure describes the order in which CAD model works to help radiologist at reporting table

1.4 Ultrasound

Ultrasound uses high frequency sound waves to picture inside the body and is also used for diagnosis of breast cancer. Ultrasound is used in combination with mammography due to its superior diagnostic capabilities. Ultrasound and can visualize occult lesions obstruct lesion not clearly visible on mammogram of young females having high fibrocystic tissue.

Ultrasound is useful in both palpable and non-palpable nodules and provides useful clinical information about the location of the lesion. Ultrasound is used in needle guided biopsy to determine if the lesion is benign or malignant and is also used for preoperative evaluation. Ultrasound also proves to be useful in locating satellite tumors that is sometime not visible in mammography. CAD models have been developed by various researchers[1], in past to decrease the false positive rate, provide reliable opinion to the radiologist reporting and to improve the sensitivity and specificity of ultrasound machine in discriminating benign from malignant lesion. Some researcher were also focused on differentiating various classes of benign from malignant to help increase the overall diagnostic capability of the system and can provide the better and accurate second opinion to the radiologist.

Ultrasound CAD models have been developed with various feature, i.e. morphological and textural depending upon the use of CAD models. CAD models are constructed by the following method as described in Figure 1-2. CAD models in ultrasound can decrease the unnecessary biopsies by reducing the false alarm rate.

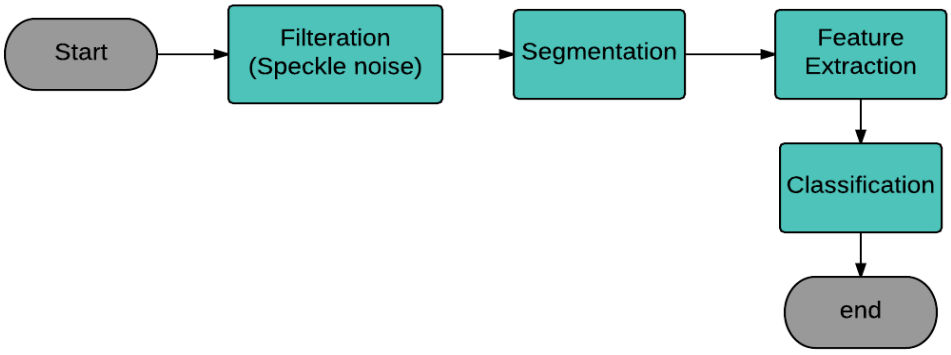


Figure 1-2: CAD Model Construction. This figure describes the key components in CAD Model.

1.5 Filtration

Ultrasound images have inherited speckle noise in them that degrades the resolution of ultrasound images, In order to increase contrast resolution as well as for segmentation of breast tumor from ultrasound image, speckle noise has to be reduced.

Speckle noise is reduced by applying speckle denoising filters. Speckle noise is a multiplicative noise and in order to remove speckle noise major information of image should not lost (i.e. not cause blurring of image, preserving edges) in ultrasound image.

The most commonly used filter for this purpose is lee, kaun, frost and Gamma MAP. These filters depends upon the filter size, increasing the size of filter can result in blurring of edges, over smoothing and small size will leave the speckle behind.The CAD model that has been developed removes speckle noise by applyingsequence of filters to achieve the desirable result. Figure 1-3 mentions the sequence of filters applied

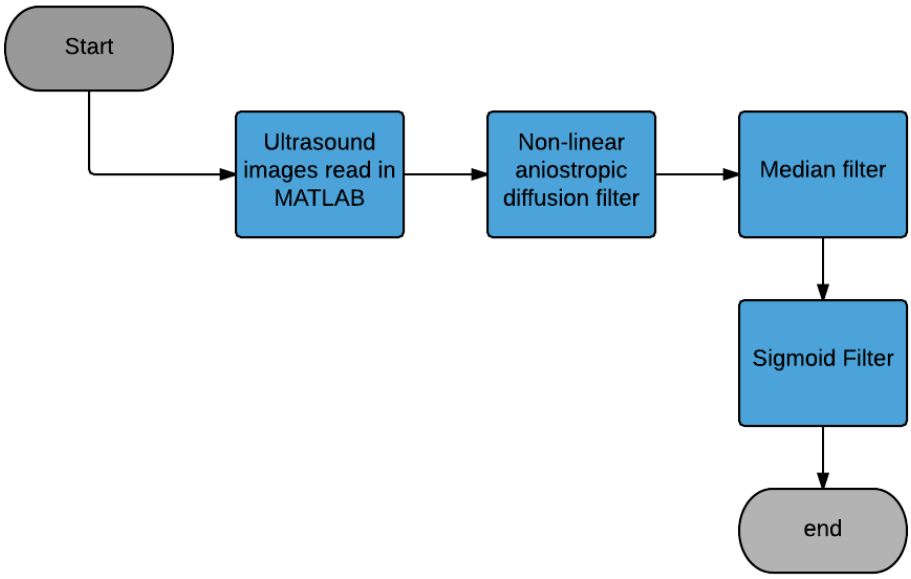


Figure 1-3: Speckle noise removal sequence. This figure describe the procedure followed in developing proposed CAD model

1.5.1 Non-Linear Anisotropic Diffusion Filter

Anisotropic filters works well on ultrasound image in reducing speckle noise and preserving edges.Ultrasound images suffer from speckle noise due to interference from the back-scattered signal. This noise degrades the visual quality of images, resulting in decrease in contrast and affecting diagnostic accuracy. Non-linear anisotropic filter based on Perona and Malik method was used for removal of speckle noise and preserving edges in ultrasound images.

Perona and Malik method was based on a non-linear diffusion method for avoiding the blurringand localization problems of linear diffusion filtering[7].

1.5.2 Median Filter

Median filter is non-linear filter and is useful in reducing noise and preserving edges. After preprocessing from anisotropic filter, the images will be further processed with median filter [10*10].Median filter will cause further reduction of speckle noise that will result in blurring of image but preservation of edge information that will be used for automated segmentation of lesion. Filters dimensions were empirically determined. Images filteredwith median filter can be seen below in Figure 1-4.

Median filter works when applied on image as by sorting all pixels in the neighborhood depending upon the size of filer and replacing the middle value with the median value of the sorted order.

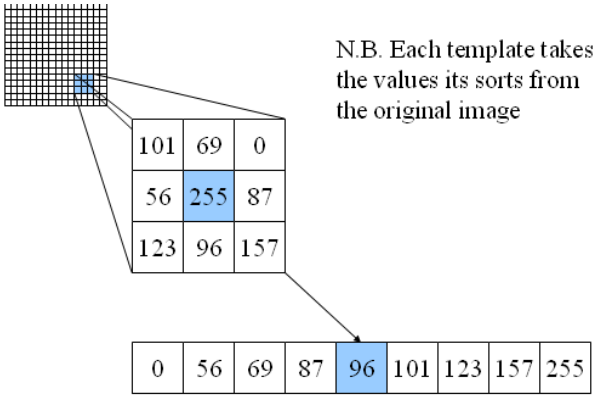


Figure 1-4: Medianfilters functioning. This figure describes the procedure in which median filter work by sorting al pixels and replacing central pixel with the median value[8].

1.5.3 Sigmoid Filter

Sigmoid filter is a spatial domain filter. Sigmoid filter will be applied on median filter processed image to enhance edges of breast nodule from the background. Sigmoid function is a smooth continuous function and its output varies from -1 to 1.The output pixel is determined by the following equation, sigmoid function is a continuous non-linear function, this function is also known as logistic function.

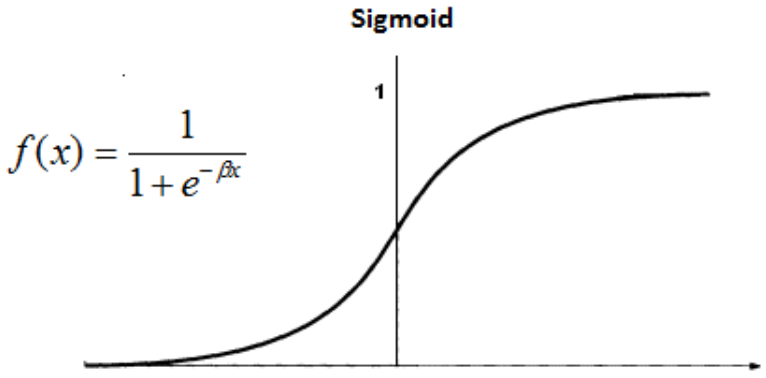


Figure 1-5: Sigmoid Function adopted from [9]

1.6 Automated Segmentation

Segmentation in ultrasound is a difficult task due to many reasons, first it has been inheritably degraded with speckle noise and in order to remove noise, edges can blur and weak edges making segmentation even tougher. Various segmentation techniques have been used in ultrasound images such as watershed segmentation that is a region based segmentation. Segmentation in ultrasound images requires proper preprocessing for removal of speckle noise and to overcome this problem segmentation step is followed after preprocessing.Active contours model proves useful in ultrasound segmentation due to its capability to evolve along the boundaries of the segmented object.

1.7 Feature Extraction

Morphological and textural features were extracted from the segmented lesion as well as from the images and will be based on BI-RAD descriptors to achieve higher accuracy in classifying benign-malignant group and benign-malignant-normal group. Benign and malignant lesions vary from each other in shape, size, echogenity, margin, lobulation, orientation and posterior acoustic enhancement. These features were also described in BI-RAD lexicon method developed by American College of Radiology to standardize the assessment criteria for a radiologist to diagnose a lesion[5].

Thirteen morphological features were used to describe the shape, margin, lobulation, orientation, lesion boundary and posterior acoustic enhancement for benign-malignant group also listed inTable 3.1,sixteen textural features (contrast, correlation, homogeneity and energy) on 4 different angles (0, 45, 90, 135) were used to represent the echogenity of lesion for benign-malignant-normal group.

1.8 Machine Learning

Machine learning is used to optimize computer performance using exampled data or past experience. Machine learn patterns in the training data using input features, pattern learned to unseen data to ensure generalization, If generalization fails features are modified and more training data are fed into algorithm. Predication is made on the basis of trained data. There are two main categories for machinelearning supervised and unsupervised learning[10].

1.8.1 Unsupervised learning

It is a type of machine learning algorithm that is used to draw inferences from dataset without having any labeled response[1]. Dataset is explored to find out some intrinsic structures in them[11].

[1] Data is unlabeled i.e. data set belongs to which class is not mention.

1.8.2 Supervised learning

It is a type of machine learning in which patterns are discovered that relates data attributes with target (class) attribute. These patterns are then utilized to predict the values of target attribute in future data. Thegroups that were formed in our CAD model,is benign-malignant group and benign-malignant-normal and they were classify with binary SVM for first group and multiclass SVM using one vs. all methodfor benign-malignant-normal group[11].

1.8.2.1 Support Vector Machine

Support vector machine is a type of supervised learning method based on decision plane that define decision boundaries. Decision plane separates classes having different class membership.

SVM goal is to find a hyperplane that can separate training data with maximum margin, SVM not only classifies the data but also optimize the decision boundary.

Given a set of training data, input vectors are given in the form of $x_{i,}$ with each input vectors having number of features and they are labeled with corresponding labels, represented as y_i,and there are m such pairs $(i=1,\ldots..m)$.

The training data can be viewed as labeled data point as can be seen below inFigure 1-6: SVM with separating hyperplane separating the two classes, defined by the equation, for binary class classification in well separated data, learning depends upon finding an optimal hyperplane that data points on one side of the plane can be labeled y_i = +1, data points on the other side of the plane can be labeledy_i= -1, optimize the decision boundary and can separate the data with maximal margin[12]. Support vector machine finds an optimal hyperplane that can separate the data points on both side of the margin with maximal distance. Such points closest to the margin are known as support vectors.Hyperplane isgiven as also shown inFigure 1-6.

$$\boldsymbol{w}.\boldsymbol{x}+\boldsymbol{b}=\mathbf{0} \qquad \textbf{1.1}$$

Where dot $(\cdot)$ represents the scalar product.

b is the bias or offset of the hyperplane from the origin in input space.

X are points located within the hyperplane.

The weights w, determines the orientation of the hyperplane

Hyperplane separates the two classes to data points with one side labeled as

$$\boldsymbol{yi = w.xi + b \geq 1} \quad \textbf{1.2}$$

Points on the other side can be labeled as

$$\boldsymbol{yi = wx1 + b \leq -1} \quad \textbf{1.3}$$

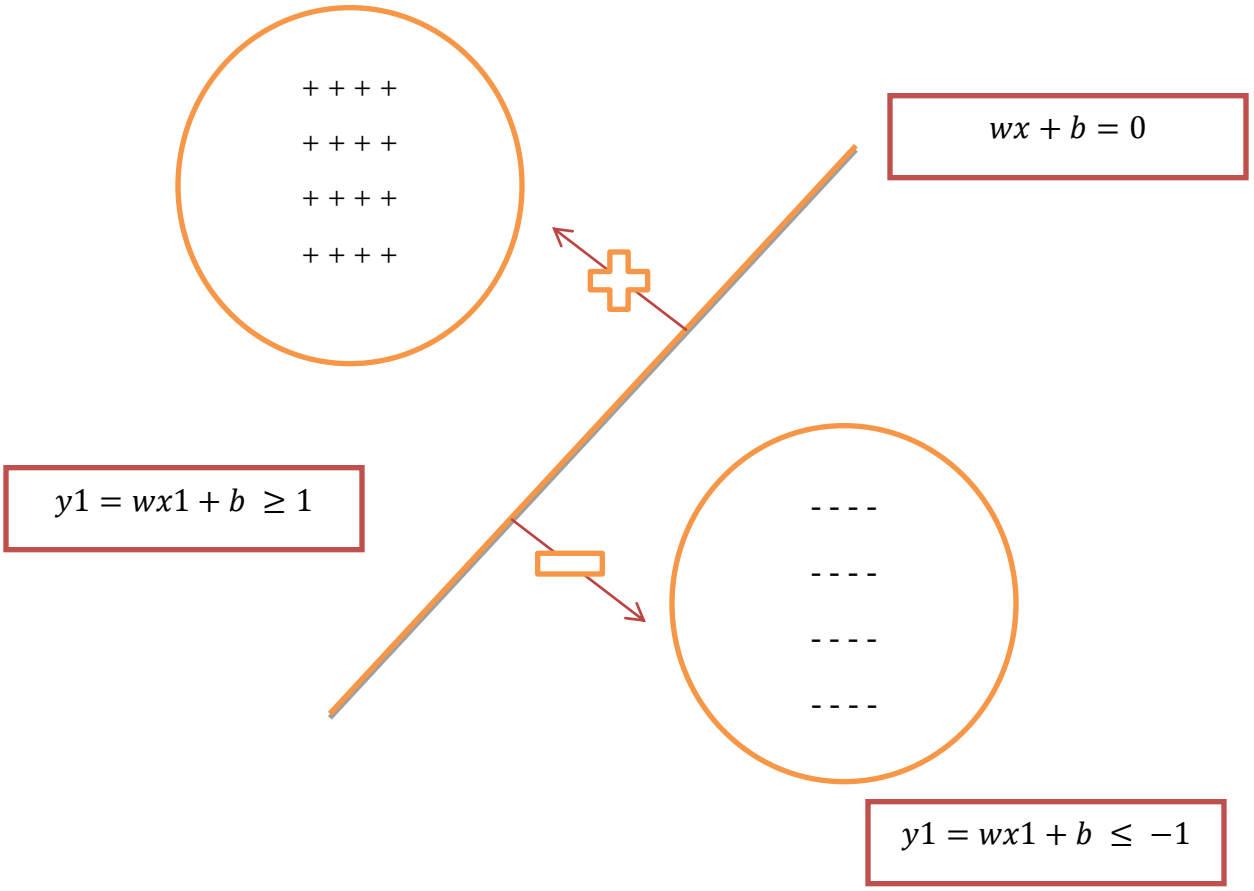


Figure 1-6: SVM with separating hyperplane separating the two classes, defined by the equation wx+b=0.

For binary classification, the decision function be

$$\boldsymbol{f(x) = sign(\sum_{i=1}^{l} \alpha i y i k(s, .x) + b)} \qquad \textbf{1.4}$$

αi= positive Lagrange multiplier

si= Support vectors

k(si, x)=function for convolution of kernel of decision function

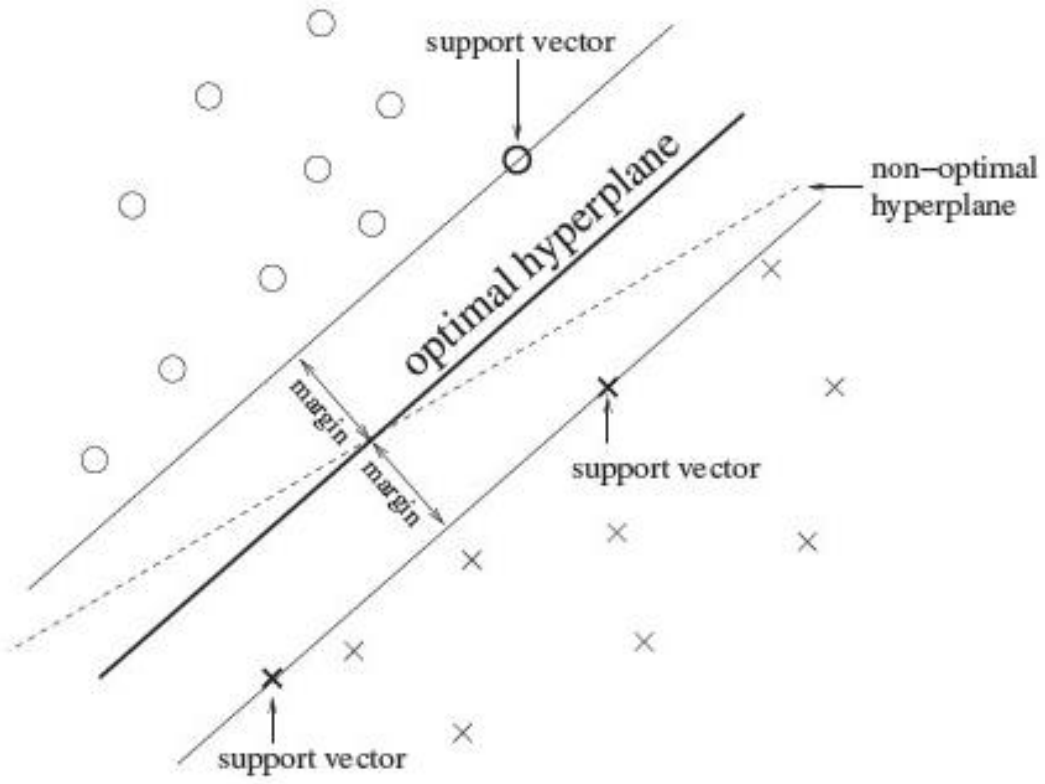


Figure 1-7:SVM with optimal hyperplane and optimization of decision boundary[13]

SVM identify the hyperplane the can maximize the margin between hyperplane by maximizing the distance between classes, i.e. distance from the nearest point also shown in

Figure **1-7**. SVM enables decision on the basis of linear hyperplane also known as linear classifier. SVM can be used to make non-linear decision boundaries depending upon the distribution of the data points in the input space. This introduces the concept of kernels in SVM that can be used according to the need of the data. This involves mapping the data into another space of much higher dimensionality using kernel functions[4, 13].

Kernels most commonly used in SVM are polynomial and RBF, SVM have good generalization properties, some parameters need to be set by hand, namely regularization parameter and RBF kernel width σ. The radial kernel is defined as

$$\boldsymbol{k(x,y) = exp(-\gamma(x-y)^2)} \qquad \textbf{1.5}$$

Where $\boldsymbol{\gamma \in R}$ is a non-zero parameter.

1.8.2.2 Multiclass SVM

Support vector machine is actually a binary classifier to use it for multiclass purpose it has to decompose into multiple binary classes and the most common way for using SVM in multiclass is one vs. all or one vs. one method. The most widely used method for multiclass classification is one vs. all

1.9 Motivation for Current Study

CAD model has an active area of research for many years, and many new developments have been made in this area.

- Main focus of this area is to develop an efficient method to differentiate benign from malignant with higher accuracy to provide reliable second opinion to the radiologists.
- CAD model based on the BI-RAD descriptors can help radiologist to diagnose a lesion on the standard basis of characterization.

- CAD models can help inexperienced reader to provide efficient report CAD model development is also focused on using ultrasound as an screening method with mammography

1.9.1 Operational Definition

Computer-aided diagnosis (CAD) is a diagnosis made by a clinician who uses the output from a computerized analysis of medical images as a second opinion in detecting lesions and in making diagnostic decisions. The final diagnosis is rendered by the clinician, e.g., the radiologist.

1.9.2 Objective

1. *To develop a CAD model that can provide second opinion to the radiologist.*
2. *Computer aided diagnosis (CAD) system that is based on the automatic segmentation of breast lesion on ultrasound using active contour model after removal of speckle noise from ultrasound images*
3. *Feature extraction based on morphological and textural features extraction by the breast imaging reporting and Data system (BI-RAD) to classify breast tumor in benign or malignant*
4. *To develop a CAD model that can be able to differentiate normal high fibrocystic image from the benign and malignant image using multiclass SVM to decrease false positive rate*

2
LITERATURE REVIEW

Breast cancer is the leading cause of death in developing countries. Early and accurate diagnosis is the key role towards the treatment of the disease and stop spreading it to the distant part of the body. Computer aided diagnosis was introduced to reduce the human error and provide a reliable second opinion to the radiologist in interpretation of breast cancer. CAD models are now plays an important role for the detection of breast tumor in mammography in United States. Several commercial CAD models are available, such as "Image Checker" (R2 Technology, Sunnyvale, CA) and "SecondLook", these systems are under investigation for their benefits in diagnostic capabilities

2.1 Historical perspective

Early attempts for computerized analysis of medical images were made in 1960, systematic advancement begin in 1980s at Kurt Rossmann Laboratories for Radiologic Image Research in the Department of Radiology at the University of Chicago. Mammography and ultrasound is used in combination for the detection of breast cancer but mammography cannot visualize lesion in dense breasts. Ultrasound is useful in localizing the lesions not easily visible on the mammography, safer for the patients but image interpretation in ultrasound is operator dependent.

CAD models were introduced to reduce the operator dependency and to increase the diagnostic accuracy, decrease the false positive rate and can provide a reliable second opinion.

2.2 Stavros (1995)

Early investigation on finding the reliability of ultrasound in differentiating benign from malignant was done by Stavros in (1995). This study provides many useful aspects that help the

researchers in developing CAD model for ultrasound images because at that time ultrasound reliability in differentiating benign from malignant lesions were was not known[4]. Stavros also gives detailed description on features related to the benign and malignant lesion; these features were later used by many researchers for feature extraction in their CAD model. These features describe the shape, orientation, margins, lobulation common among both groups.

2.3 Giger *et*.al (1990)

In the late 1990s Giger *et*.al with her team in Chicago provides many CAD models on breast ultrasound images that were based on automated segmentation, morphological features extraction and classification of lesions. They also developed their own method of automated segmentation method that can provide higher accuracy in segmentation. Features were morphological features related to the shape and contour of tumor and these features were also significant for diagnosis by radiologist.

2.4 Giger *et*.al (2002)

In 2002 Giger *et*.al investigated the used of radial gradient index technique for the automated detection of breast cancer, and used round robin analysis for the classification. Results showed Az value of 0.84, overall performance for CAD model by case is 94% sensitivity and 0.48 false positives per case. They concluded that computerized analysis of breast ultrasound can help in breast cancer screening program on sonography[1]. Ultrasound was used at that time for identifying the lesion not visualizes in mammography but these studies provided evidence on the efficacy of ultrasound machine in identifying as well as differentiating the lesion.

2.5 Chen *et*.al (2002)

In 2002 Chen et.al developed a novel CAD model based on neural network and textural features used for the first time in CAD model. Textural features used were autocorrelation, correlation, Results showed that accuracy Az of 0.93, sensitivity and specificity of 98% and 81%. There purposed model provide potential for using textural features for the feature extraction in CAD model[14], as tumors have significant textural variation. Textural features have several benefits over morphological features and among that the most significant is that

textural features do not require an efficient segmentation method and works without segmentation.

2.6 Segyeong Joo *et*.al (2004)

In 2004 Segyeong Joo et.al presents CAD algorithm to identify solid breast nodule malignancy using multiple sonographic features and artificial neural network classifier was developed from the data set. In their CAD model they used five morphological features representing shapes, edge characteristics and darkness of a nodule. Results showed that area under the ROC curve to be 0.95 and 99.3% sensitivity[15].This research provides contribution in using some important features not used before at that time and using ANN as a classifier. This model computed an important set of features that were most significant for radiologist to diagnose. They provided encouraging results with small set of features and relatively simple segmentation technique than used by various researchers.

2.7 Drukker and Giger *et*.al (2004)

In 2004 Drukker and Giger *et*. al develop and evaluated a two stage computerized method that first identify the suspicious regions on ultrasound images and then subsequently distinguishes among different lesion types, Bayesian neural network was used a classifier[16]. Results showed that Az of 0.94 and 0.91 with training and testing data, sensitivity of 90%. The results signify that computerized lesion detection and classification methods provide promising results and indicate the potential of such system in clinical breast ultrasound. This study was mostly focused on identifying false positives and then detecting the type of lesion.

2.8 Chen et.al (2005)

In 2005 Chen et.al developed a CAD model on which he proposed novel morphological features. Data set consist of two sets and the results showed Az of 0.95 ,proposed CAD model proves useful to differentiate benign from malignant lesion and the significant contribution was morphological features (NSPD, LI, ENC and ENS) that was used by various researcher and provided significant results[2]. Chen provided an important contribution to the CAD researchers

by allowing these features to provide significant results, the main drawback of these features were that they were computationally intensive and need long processing time for computation.

2.9 Chang et.al (2005)

In 2005 Chang et.al proposed a method based on the automatic segmentation and morphology based diagnosis of solid breast tumors. SVM was used for classification, sensitivity and specificity was 88% and 92.5%[17]. This introduces SVM as potential classifier, SVM has many benefits as classifier but that was not discussed at that time and after this many researchers proves the robustness of SVM in CAD models.

2.10 Karla horsch *et.*al (2006)

Karla et.al in 2006 evaluated a computer aided diagnosis model for multimodality intelligent workstation as an aid to radiologists in the interpretation of mammograms and breast sonograms at same time. Use of computer aided results in average performance of Az of 0.82-0.92and results showed that use of multimodality workstation can improves the task of differentiating benign from malignant lesions. This proposed model provides an efficient ideology based on providing radiologists with both modalities preview for providing better results.

2.11 Huang *et.*al (2007)

Huang et.al in 2007 evaluated computer aided diagnosis system with automatic contouring and morphological analysis to aid in the classification of breast tumors using ultrasound. Az value of 0.91 and 0.90 was achieved respectively. The system differentiates benign from malignant breast tumor and provides clinically useful second opinion.

2.12 Woo kyung Moon*et.*al (2013)

Woo *et.*al developed CAD model on the breast masses using BI-RADS findings, six BI-RAD features were extracted and used for the CAD model. Proposed model have Az value of 0.96 this system with BI-RADS findings also provide promising results to use BI-RADS descriptors for the CAD model. BI-RADS provide an efficient diagnosis system to provide systematic results. CAD models in past were focused on features significant in benign and malignant but never uses

a standard assessment criteria for features extraction, this study gives a useful insight into the proposed CAD system based on BI-RAD system can proves to be useful for follow a standard assessment criteria.

.

3
METHODOLOGY

CAD model based on the BI-RAD descriptors for feature extraction were used for classifying benign-malignant and benign-malignant-normal group.

3.1 Experimental protocol

The proposed CAD model was prepared in following way. First Ultrasound images were collectively preprocessed for removal of speckle noise, edge enhancement and automated segmentation. Morphological features were extracted from the segmented lesion in classifying benign-malignant group and textural features were extracted from all the images to classify benign-malignant-normal group with multiclass SVM using one vs. all method

3.1.1 Data Acquisition

This study used an actual 163 breast ultrasound images from various hospitals with their consent, having benign lesions (65 images), malignant lesions (53 images) and normal (45 images).

3.1.2 Sample Size

The number of images obtained from each patient depends on the number of lesions and varied from one to ten. Classification will be lesion based not image based and data set contains total 222 lesions of which benign lesions were (94) and malignant lesions were (128). The average tumor size of benign lesion was 53 mm (size range 1.15-367 mm) and average size for malignant lesion was 93 mm (size range 2-414 mm) .The patient's age range from 18 to 40 (mean,25 years).

3.1.3 Study population

Young females and women coming to the hospitals for ultrasound for screening purpose or feeling lump/nodule in breast, feeling pain around the areolar area and discharge of fluid from areolar area.

3.1.4 Study type

Study was an experimental study. Data set was collected prospectively from December 2013-May 2014 and this data set was obtained from (Xario, SSA-660A).

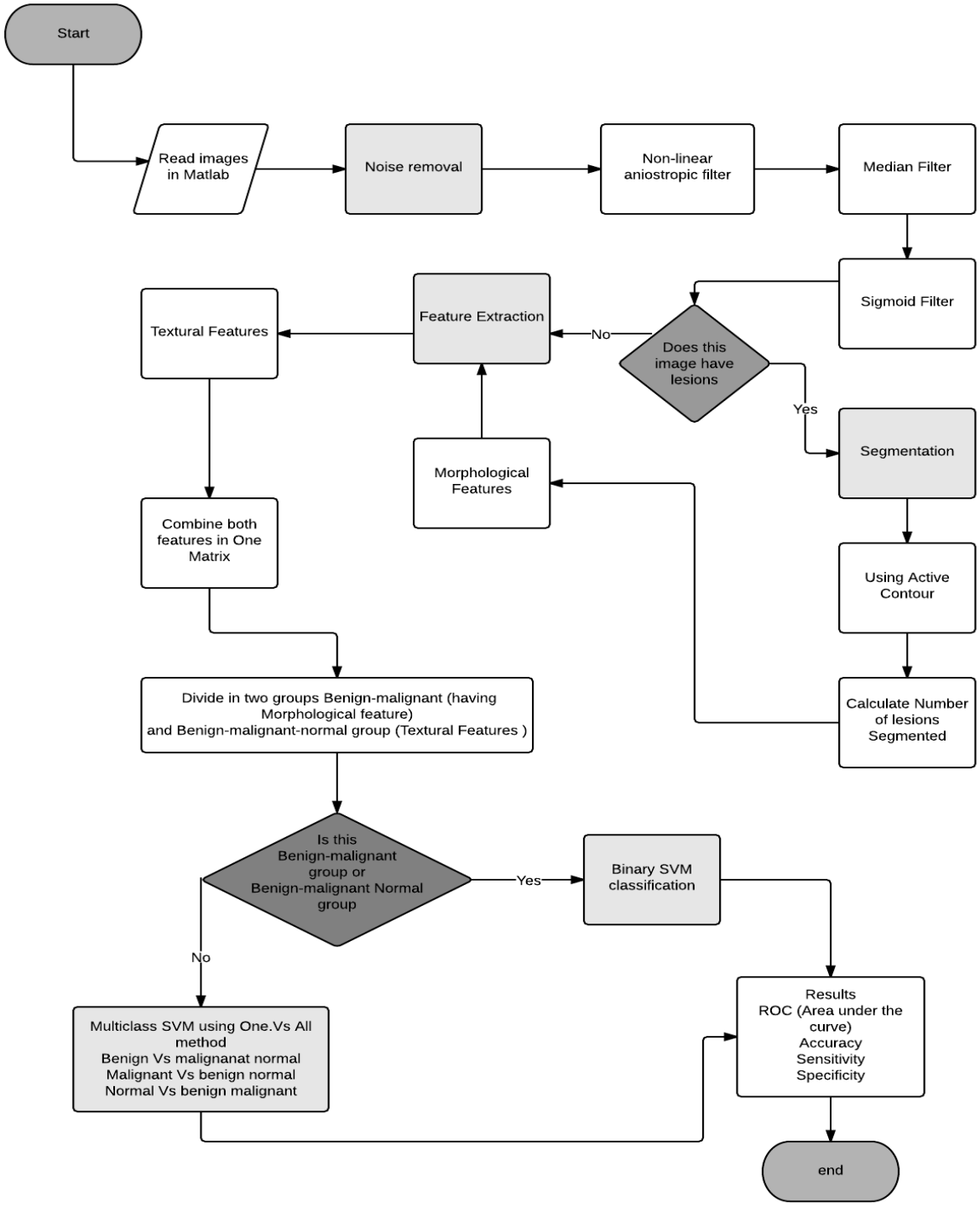
3.1.5 Inclusive criteria

Images were stored in the machine and were taken from the machine, for analysis using MATLAB (2013).

3.1.6 Exclusive criteria

Cystic images were not included in the study.

3.2 CAD System Design



3.2.1 Preprocessing

First step for CAD model development is preprocessing for removal of noise. In ultrasound images speckle noise is removed for accurate segmentation and classification. Speckle noise is a multiplicative noise and requires various filters for complete removal. In our CAD model various filters were applied for despeckling.

3.2.2 Non-linear Anisotropic Diffusion

Ultrasound images suffer from speckle noise due to interference from the back-scattered signal. This noise degrades the visual quality of images, resulting in decrease in contrast and affecting diagnostic accuracy. Non-linear anisotropic filter based on Perona and Malik method was used for removal of speckle noise and preserving edges in ultrasound images.

Perona and Malik method was based on a non-linear diffusion method for avoiding the blurringand localization problems of linear diffusion filtering[7]. Non-linear anisotropic diffusion filters overcome the problem of diffusion at edges and provides smoothing between edges.
They apply an inhomogeneous process that reduces the diffusivity at those locations which have a larger likelihood to be edges. This likelihood is measured by $|\nabla u|^2$.The Perona–Malik filter is based on the equation:

$$\partial_t u = div(g(|\nabla u|^2)\nabla u) \quad \textbf{3.1}$$

And it uses diffusivities such as, filtered image can been seen below in figure 1(a).

$$g(|\nabla u|^2) = \frac{1}{1 + \frac{|\nabla u|^2}{\lambda^2}} \quad \textbf{3.2}$$

Diffusion should be parallel to the edges D is define as diffusion tensor, positive symmetric matrix given as[18].

$$D=\begin{bmatrix} v1 \\ v2 \end{bmatrix}\begin{bmatrix} \lambda 1 & 0 \\ 0 & \lambda 2 \end{bmatrix}\begin{bmatrix} v1 \\ v2 \end{bmatrix} \quad \textbf{3.3}$$

Where v1 and v2 is the Eigen vector one is parallel to the gradient and other is tangential to the edge.

$$v1 = \frac{\nabla u}{|\nabla u|} \quad \textbf{3.4}$$

$$v2 = \begin{pmatrix} [v1]y \\ [v2]x \end{pmatrix} \quad \textbf{3.5}$$

To stop the diffusion over the edges λ is define as

$$\lambda 1 = g(|\nabla u|^2) \quad \textbf{3.6}$$

Perpendicular to the edges diffusion should not be stooped

$$\lambda 2 = 1 \quad \textbf{3.7}$$

3.2.3 Median Filter

Median filter is non-linear filter and is useful in reducing noise and preserving edges. After preprocessing from anisotropic filter, the images will be further processed with median filter [10*10].

Median filter will cause further reduction of speckle noise that will result in blurring of image but preservation of edge information that will be used for automated segmentation of lesion. Filters dimensions were empirically determined.

3.2.4 Sigmoid Filter

Sigmoid filter is a spatial domain filter. Sigmoid filter will be applied on median filter processed image to enhance edges of breast nodule from the background. Sigmoid function is a smooth continuous function and its output varies from -1 to 1.

$$f(x) = min + \frac{max - min}{1 + e^{\frac{b-x}{a}}} \quad \textbf{3.8}$$

The values **a** is (cut off) and **b is** (gain) can define the value range to be enhanced and the interval [min, max] determines the target range of the new value[9]. Sigmoid filters gain and cutoff vary for benign and malignant images. Contrast enhancement of the filtered image can be seen below in Figure 1-5.

3.3 Segmentation

Active contours were used for automated segmentation of lesions in breast ultrasound. Results for segmentations are shown in Figure 4-2.

3.3.1 Active Contours

In active contour models or snakes curve will be evolved, subjected to constraint from a given imageμ_0, in order to detect objects in an image.

The snake model is:

$$Inf_c J_1\ (C) \quad \textbf{3.9}$$

Where:

$$J_1(C) = \alpha \int_0^1 |C''(s)|^2 \, ds + \beta \int_0^1 |C''(s)| \, ds - \lambda \int_0^1 \left|\nabla \mu_0\big(C(s)\big)\right|^2 ds \quad \textbf{3.10}$$

Here α, β,λ are positive parameters. The first two terms control the smoothness of contour, while the third term attracts the contour toward the object. Observing the above equation curve is located at the point of maxima$|\nabla \mu_0|$, acting as an edge- detector.

A general edge detector is positive decreasing function g, depending on the gradient of the image μ_0

$$g(|\nabla \mu_0(x,y)|) = \frac{1}{1 + |\nabla G_\sigma(x,y) * \mu_0(x,y)|^p}, \quad p \geq 1 \quad \textbf{3.11}$$

Where $G_\sigma * \mu_0$is convolution of image with the Gaussian $G_\sigma(x,y) = \sigma - \frac{1}{2}\, e - |x^2 + y^2|^{4\sigma}$. The function $g|\nabla \mu_o|$ is positive in homogenous regions and zero at the edges.

Classical approaches of active contour uses the image gradient to stop the evolution of curve on boundary. However, when image is noisy or hasa weak boundary, active contour can completely miss the object.After preprocessing from anisotropic diffusion filter, median filter and contrast enhancement from sigmoid filter, lesion is segmented using Chan and Vese active contour without edges.

This model is based on trying to separate the image into region on based on intensities.

$$\begin{aligned} &F(c_1, c_2, C) = \mu L(C) + \nu A(in(C)) \\ &+ \lambda_1 \int_{in(C)} |u_0(x,y) - c_1|^2 \, dxdy + \lambda_2 \int_{out(C)} |u_0(x,y) - c_2|^2 \, dxdy \end{aligned} \quad \textbf{3.12}$$

Where$\mu \geq 0, v \geq 0, \lambda_1, \lambda_2 0$ $are\ fixed\ parameters$, C_1, C_2 are the average intensities inside and outside of the contour.Therefore, the minimization problem is

$$\inf_{c_1, c_2, C} F(c_1, c_2, C)$$

The Chan and Vese is a special case of Mumford Shah model:

$$F^{MS}(u, C) = \mu L(C) + \lambda \int_{\Omega} |u_0 - u|^2 d + \int_{\Omega \setminus C} |\nabla u|^2 d \qquad \textbf{3.13}$$

The above model is solved using level set formulation:

$$C = \{(x, y) \in \Omega : \phi(x, y) = 0\}$$
$$in(C) = \{(x, y) \in \Omega : \phi(x, y) > 0\}$$
$$out(C) = \{(x, y) \in \Omega : \phi(x, y) < 0\}$$

The Chan Vese model minimizes the energy for more detail on active contour models readers can refer to [1]

$$\frac{\partial \phi}{\partial t} = \delta(\phi)\left[\mu \kappa(\phi) |\nabla \phi| - \nu - \lambda_1 (u_0 - c_1)^2 + \lambda_2 (u_0 - c_2)^2\right] \qquad \textbf{3.14}$$

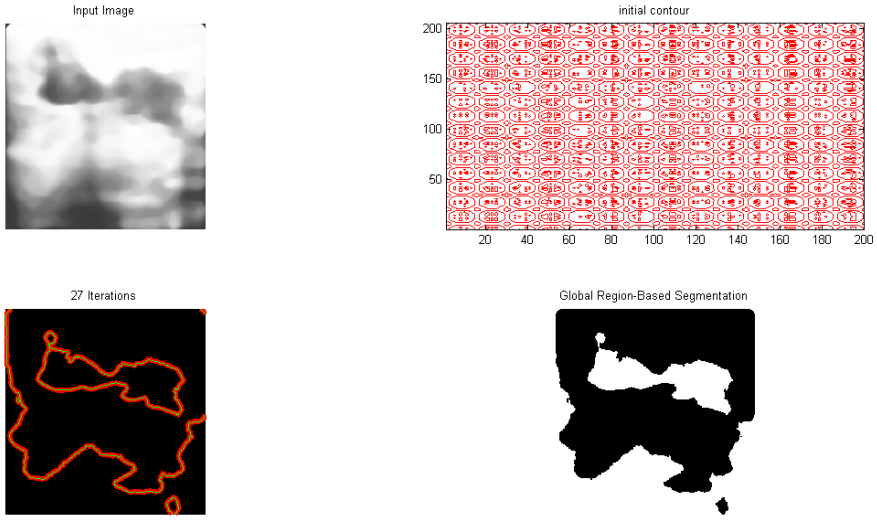


Figure 3-1: Segmentation using Active contour model

Where μ_0 is a given image and δ is delta function, and μ is a constant. Breast ultrasound images suffer from noise as well as weak edges and in high fibrocystic tissues, lesions have same intensity as the surrounding tissue making it difficult for segmentation.

This method combined with the preprocessing approach results in automated segmentation of lesions on breast ultrasound. Segmentation results are also shown in Figure 3-1.

3.4 Morphological Features

3.4.1 Shape

Benign and Malignant lesions differ in their shapes. Benign images are usually round,oval and malignant images are irregular[2, 4] . Morphological features that compute the shapes of lesions are listed in Table 3.1.

3.4.1.1 (A). Form factor

Form factor close to 1 mean, shape is round and lesion is benign [14].

$$Form\_Factor = \frac{4\pi * Area}{Perimeter^2} \quad \textbf{3.15}$$

3.4.1.2 (B). Roundness

$$Roundness = \frac{4\pi * Area}{\pi * MaxDiamter^2} \quad \textbf{3.16}$$

Maximum diameter is the length of major axis from the equivalent ellipse of tumor [14].

3.4.1.3 (C).Solidity

$$Solidity = \frac{Area}{Convex\_Area} \quad \textbf{3.17}$$

Convex area is convex hull of tumor. Solidity close to 0 means lesion is malignant[14].

3.4.1.4 (D).Convexity

$$Convexity = \frac{Convex\ perimeter}{Perimeter} \quad \textbf{3.18}$$

Convex perimeter is the convex hull of a tumor [14].

3.4.1.5 (E).Extent

$$Extent = \frac{Area}{Bounding\ \_Rectangle} \quad \textbf{3.19}$$

Bounding rectangle is the rectangle containing the tumor [14].

3.4.2 Orientation

According to BI-RAD category orientation describes if lesions are parallel, not-parallel. Malignant lesions are usually taller less wider, they are larger in anterior-posterior direction as compared to the sagittal or coronal section[4, 19] and is vice versa in benign lesion.

3.4.2.1 (A). Aspect Ratio

Aspect ratio is the length ratio of the tumor depth and width. If tumor depth is greater than its width then there is a high chance of lesion being malignant and will be closer to 1[14].

3.4.2.2 (B). LS Ratio (Long Axis to Short Axis Ratio)

$$LS_{RATIO} = \frac{MajorAxis}{Minor\ Axis} \tag{3.20}$$

Major and Minor axis is of equivalent ellipse defined below[2].

3.4.3 Margin

Benign and malignant lesion varies in margin, benign lesions have microlobulation and malignant lesion have spiculation[4].

3.4.3.1 (A). Number of Lobulations

In solid breast nodule, gentle lobulation defined as fewer than four have been regarded as a sign of benignancy[20].

$$F_x = Number\ of\ local\ extrema\ in\ curve\ fitted\ r(\theta) \tag{3.21}$$

3.4.3.2 (B). Spiculation

In Stavros' study Spiculation was the most important feature in identifying malignant lesion.

$$F_x = \frac{\sum_0^{\pi/4} |R(w)|}{\sum_{\pi/4}^{\pi} |R(w)|} \quad \textbf{3.22}$$

Malignant lesion has small spiculation[20].

3.4.4 Lesion Boundary

Benign lesion usually have abrupt interface at boundary and malignant lesion have echogenic halo [4].

3.4.4.1 (A). ENC(Elliptical Normalized Circumference)

$$ENC = \frac{Equivalent\_ellipse\_perimeter}{Perimeter} \quad \textbf{3.23}$$

ENC close to 1 means boundary is smooth and there is a high chance of lesion being benign[2].

3.4.4.2 (B). Branch Pattern

Branch pattern is defined as multiple projections from the nodule or around ducts extending away from the nipple.

$$F_x = Number\ of\ local\ extrema\ in\ low\ pass\ filtered\ r(\theta) \quad \textbf{3.24}$$

Malignant lesions have more Branch pattern[2].

3.4.4.3 (C). Relative Brightness of Nodule

$$F_x = {}^{Average\ darkness\ of\ nodule}/_{Avergae\ darkness\ of\ surroundings} \quad \textbf{3.25}$$

Malignant nodules are darker when compare with their surroundings[2].

3.4.5 Posterior Acoustic enhancement

Malignant lesions have posterior shadow behind them and are most prominent feature in ultrasound for malignant lesions.

3.4.5.1 (A). Posterior Acoustic enhancement

Malignant lesions usually have posterior shadow behind them and that can be measured by comparing the gray level values posterior to the lesion to the gray level values adjacent to them at same depth and height. The posterior acoustic behavior is defined as

$$MSD = \min(A_{post} - A_{left}, A_{post} - A_{right}) \quad \textbf{3.26}$$

A_{post}, A_{left}, A_{right} is average gray level value [21].

3.5 Textural Features

Textural feature were used to describe the echogenity of lesion. Benign, malignant and normal images have different echogenic pattern having various intensity values.In this study, sixteen textural features were computed from the breast ultrasound to discriminate between benign, malignant and normal images[22].

3.5.1 Echogenity

Benign, malignant and normal image have various echogenity that is they are hyperechoic, hypoechoic and isoechoic[4].Young females and women above 40 years can be misdiagnosed on ultrasound image due to high fibrocystic tissue, CAD model developed for differentiating normal image from benign and malignant image can be useful in eliminating this error. Table 3.1all the BI-RADS descriptors with their characteristics in benign and malignant lesions.

3.5.1.1 (A). Gray Level Co-occurrence Matrix (GLCM)

Gray level co-occurrence matrix provides statistical information about (contrast, correlation, homogeneity and energy) between the pixels. These four features were calculated in four angles (0, 90, 145 and 30). GCLM gives total of sixteen features. These features will help in classifying benign, malignant and normal image[3].

$$CON = \sum_{i,j} |i-j|^2 \, P(i,j) \tag{3.27}$$

$$Corr = \sum (i,j) \frac{(I-\mu i)(j-\mu j)p(i,j)}{\sigma_i \sigma_j} \tag{3.28}$$

$$Energy = \sum_{(i,j)} p(i,j)^2 \tag{3.29}$$

$$Hom = \sum_{i,j} \frac{p(i,j)}{1} + |i-j| \tag{3.30}$$

Table 3.1: Sonographic Features and their description according to the extracted features

Sonographic Significant features	Feature	Benign	Malignant
Shape	Form Factor	Round	Irregular
	Roundness		
	Solidity		
	Convexity		
	Extent		
Margin	Number of lobulation	Circumscribed	Not-circumscribed
	Elliptical Normalized circumference		
	Spiculation	Lobulated	Spiculated
Orientation	Long Axis to short axis	Parallel	Not Parallel
	Aspect Ratio		
Lesion Boundary	Branch Pattern	Abrupt Interface	Echogenic halo
	Relative brightness of nodule		
Posterior acoustic enhancement	Posterior Acoustic enhancement	No posterior acoustic feature	Posterior Enhancement
Echo Pattern	Contrast	Hyperechoic	Isoechoic
	Correlation		
	Homogeneity		Hypoechoic
	Energy		

3.6 Evaluation

The K-fold cross–validation method has been used with one vs. all method as well as with binary SVM classification for classifying two groups. Data was randomly divided into k fold (k=5). The first group was treated as testing data and remaining (k-1) as training data, (k =5) and this process was repeated k times until all groups have been used in turn for testing.

The performance of each classifier (i.e. how well lesions were categorize according to their classes) was evaluated on the basis of ROC (receiver operating characteristics) analysis with an index of area Az(area under the curve).This provides quantitative information about the performance of the system. The performance of the classifier was also measured for both groups by computing there Az value, Sensitivity and Specificity. These parameters were also calculated for all the features (morphological and textural) for identifying the best feature in classifying the data.

3.6.1 Receiver operating characteristics (ROC)

Roc was described by lusted in 1971 to assess the accuracy of the diagnostic test. An ROC curve is a plot of sensitivity (true positive rate) plotted on y-axis and 1-specificty (false positive rate) on x-axis. Each point on a graph is generated by using various thresholds. The data generated from various thresholds arepresents points on the graph and is connected by the line and is known as empirical ROC and then after joining the points one can see how both variables vary with each other.

Area under the curve of ROC curve measures the accuracy of the test. ROC curve can take any value between 0 and 1. Area under the ROC curve of 1.00 means perfect accuracy because it will have 1.0 sensitivity and 0 false positive rate and Az (area under the curve) of 0 has inaccurate results because all patients having disease is diagnosed as not having disease and patient having disease is diagnosed as not having a disease[23].

The lower bound for area under the ROC curve is 0.5, line segment from 0,0 to 1,1 has an area of 0.5 which means chance it has probability of being in either class of 50%. The closer the ROC curve is to 1.0 better is the diagnostic abilities of test. The ROC curve is a good measure of measuring the accuracy of the system can also be seen in Figure 3-2.

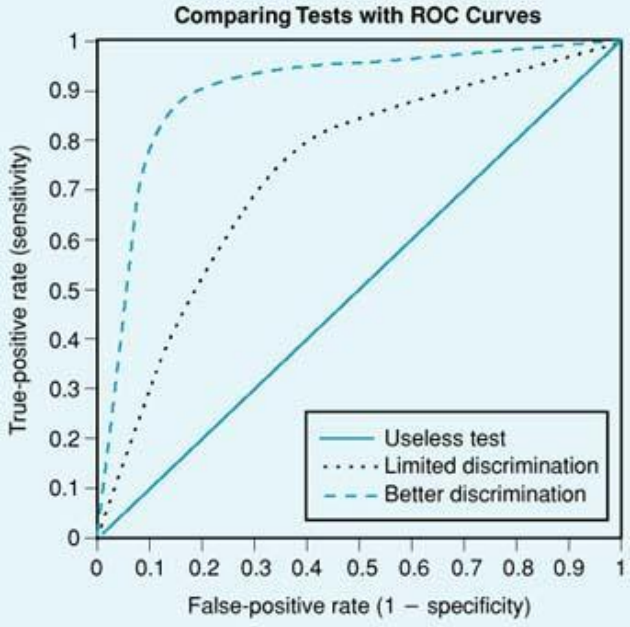


Figure 3-2: ROC Curve adopted form (http://www.ebm.ugent.be/ROC-curve.jpg)

3.6.2 Accuracy

Classifiers accuracy in classifying is measures through computing accuracy that is calculated with the help of confusion matrix as described below in Table 3.2. Accuracy is proportion of true resulti.e. how many benign lesions were actually classifies as benign and how many actual malignant lesions were classified as being malignant.

Table 3.2: Confusion matrix		
Actual	**Predicted**	
	Yes (having disease)	**No(not having disease)**
Yes	**True Positive** (correctly identify as having disease)	**False Positive** (wrongly identifies as having disease,)
No	**False Negative** (Wrongly identify as not having disease)	**True Negative** (Correctly identifies as not having a disease)

$$Accuracy = \frac{Numberoftruepositives + NumberoftrueNegatives}{Numberoftruepositives + falsenegative + falsepositive + truenegative} \quad \textbf{3.31}$$

3.6.3 Sensitivity

Sensitivity (true positive) measures classifiers abilityto correctly identify the proportion of patients actually having that disease

$$Sensitivity = \frac{Number\ of\ true\ positives}{Number\ of\ true\ positives + Number\ of\ false\ negatives} \quad \textbf{3.32}$$

3.6.4 Specificity

Specificity measures the classifiers ability in correctly excluding the conditions or disease in patients not having that disease.

$$Specificity = \frac{Number\ of\ true\ negatives}{Number\ of\ true\ negatives + Number\ of\ false\ positives} \quad \textbf{3.33}$$

3.7 Support Vector Machine (SVM) Classifier

The K-fold cross–validation method was used to evaluate the performance of the CAD model with multiclass classification for benign-malignant-normal group as well as with binary SVM for benign-malignant group. Data was randomly divided into k fold (k=5), first group was treated as testing data and remaining (k-1) as training data this process was repeated k times until all groups have been used in turn for testing.

The performance of each classifier (i.e. how well lesions were categorize according to their classes) was evaluated on the basis of ROC (receiver operating characteristics) analysis with an index of area Az(area under the curve).This provides quantitative information about the performance of the system[23]. Performance of each classifier was also measured by computing there accuracy, sensitivity and specificity. These parameters were also calculated for all the features (morphological and textural) for identifying the best feature in classifying the data.

Support vector machine is the most common tool used for classification in data mining, pattern recognition and image processing. SVM has been used before in many studies for classifying benign and malignant lesion. SVM is primarily used for binary classification. In this study we used binary as well as multiclass SVM using one vs. all method for classification. SVM is used to separate the training data by finding a hyperplane with maximal margin. SVM is an appropriate tool for classification because of its high generalization performance and of its robust nature and it works well on small data[24].

There are currently two approaches for multiclass SVM, one vs. all and one vs. one. The earliest method for solving multiclass problem is one vs. all method. This method builds M different binary classifier between each class and all other classes, Suppose there is $\boldsymbol{n}$ classes to classify than there will be $\boldsymbol{n}$ classification functions. For example, if the ith classifier has to separate the samples of theith class from all other classes then for doing this class labels are modified i.e. the ith class will be labeled as +1 and other classes as -1.In the end class having the maximum votes will be selected as he prediction class[25].

3.7.1 One vs. All

There are currently two approaches for multiclass SVM, one vs. all and one vs. one. The earliest method for solving multiclass problem is one vs. all method. This method builds M

different binary classifier between each class and all other classes, Suppose there is *n* classes to classify than there will be n classification functions.

For example, the ith classifier to separate the samples of the ithclass from all other classes. For doing this class labels are modified i.e. the ith class will be labeled as +1 and other classes as -1.In the end class having the maximum votes will be selected as he prediction class[26].

$$\boldsymbol{F(x) = Argmax(i) fi(x)} \qquad \textbf{3.34}$$

4
RESULTS

The Ultrasound images used for the analysis consists of 163 images in total having benign (65 images), malignant (53 images)and normal (45 images) of breast ultrasound.This data set contains total 222 lesions of which benign lesions were (94) and malignant lesions were (128).

The proposed CAD model was tested with k-fold cross validation method to identify tumor as benign or malignant. The diagnostic performance of the system was analyzed with the ROC curve and area under the curve Az. There were thirteen morphological features to classify benign-malignant group listed in Table 3.1and sixteen textural features for discriminating normal images from benign and malignant images.

4.1 Filtration results

Ultrasound images had speckle noise and for removal of speckle noise sequence of filters were applied in order to achieve the desired results. Filtration results can also be seen below

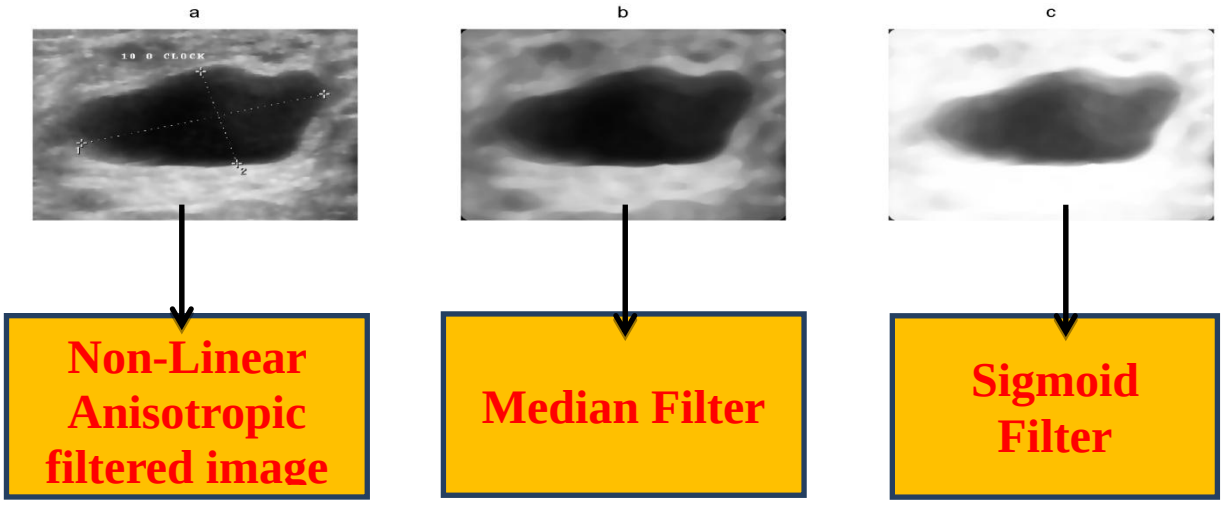
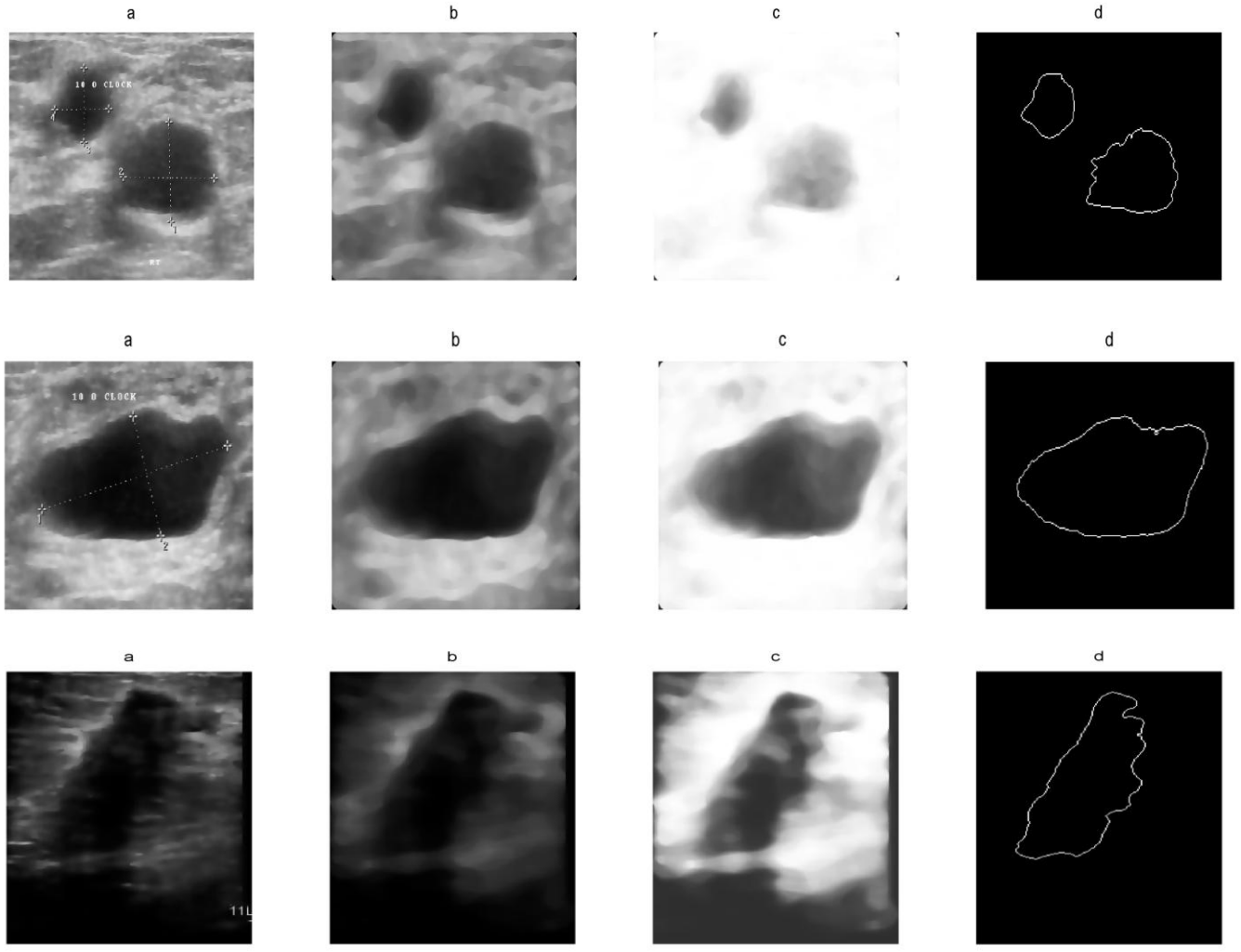


Figure 4-1: This image shows the results provide on applying various results in developed CAD model

4.2 Segmentation results

Active contours were applied in order to remove the lesion from the background for feature extraction and classification. Active contours proves to be useful segmentation technique as it is a region based technique and depends upon the intensity variation between image and contour. Lesion is segmented when there is minimal difference of intensity between contour and image. Segmentation results can be seen below in Figure 4-2.



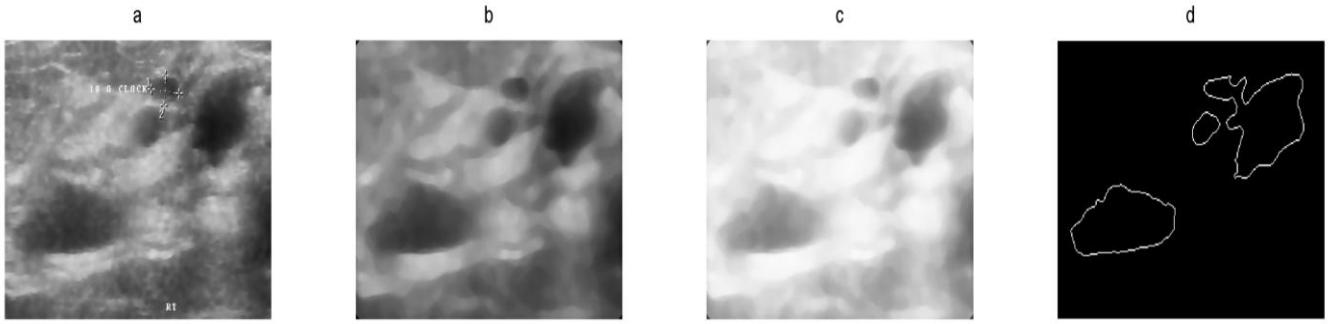
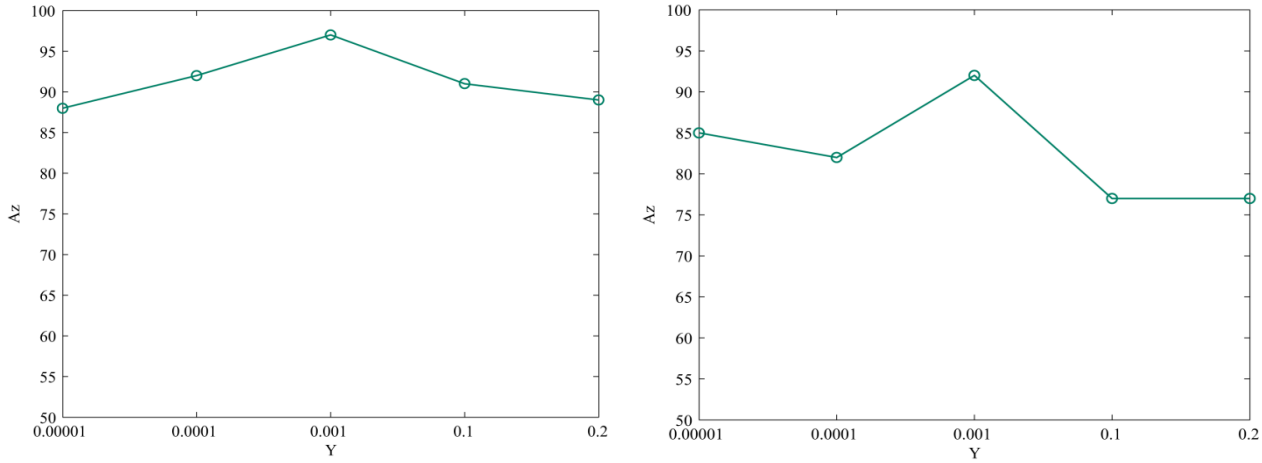


Figure 4-2:(a) Image filtered with Anisotropic diffusion filter for removal of speckle noise.(b) anisotropic image filtered with median filter to enhance edge and blur the background (c) Image filtered through sigmoid filter for contrast enhancement of lesion .(d)After processing through all filters, image segmented by active contour method.

4.3 Evaluation metrics

Radial kernel was used in SVM for classification, various γ values were used and the highest accuracy for our CAD model was achieved from $\gamma = 0.0001$,these values was determined empirically. γ Values were optimized for both groups (benign-Malignant) group and (benign-Malignant-Normal)group and was used in CAD model for classification and can be seen below in Figure(4-1).



(a) (b)

Figure 4-3: (a) Sensitivity of accuracy to change in σ values for radial kernel. Az (area under the curve) for various γ values used in radial kernel for SVM in benign-malignant group and locating the highest possible accuracy of classifier with γ to be used in C

SVM results showed 94 % accuracy and Az value of 0.97 for benign-malignant group, Sensitivity and specificity for benign-malignant group were 97% and 88%. Figure (4-2) demonstrates the ROC curve and **Table 4.1**listed the performance for benign malignant group.

SVM result showed 82% accuracy,Az value of 0.91 for benign, Az value of 0.81 for malignant and Az value of 0.83 for normal image,sensitivity and specificity of 94% and 83 % in benign-malignant-normal group. Figure (4-3) demonstrates the ROC curve for benign-malignant-normal group.**Table 4.2**listed the perfomance for benign-malignant-normal group.

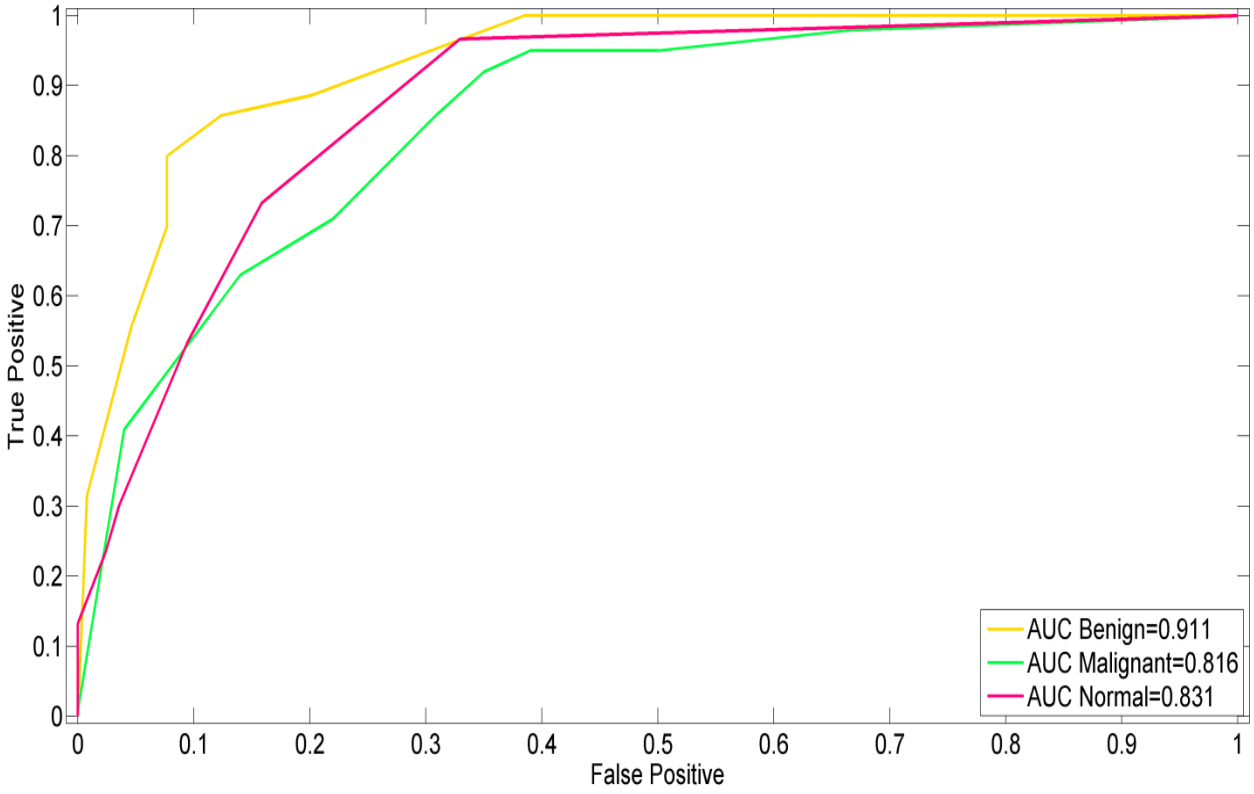


Figure 4-4: ROC curve indicating accuracy of proposed CAD model as area under the curve. (a)Receiver operating characteristics (ROC) curve of benign-malignant-normal group with textural features and area under the ROC curve is (Az = 0.92) for benign, (Az value= 0.84) for malignant and (0.86 for normal)

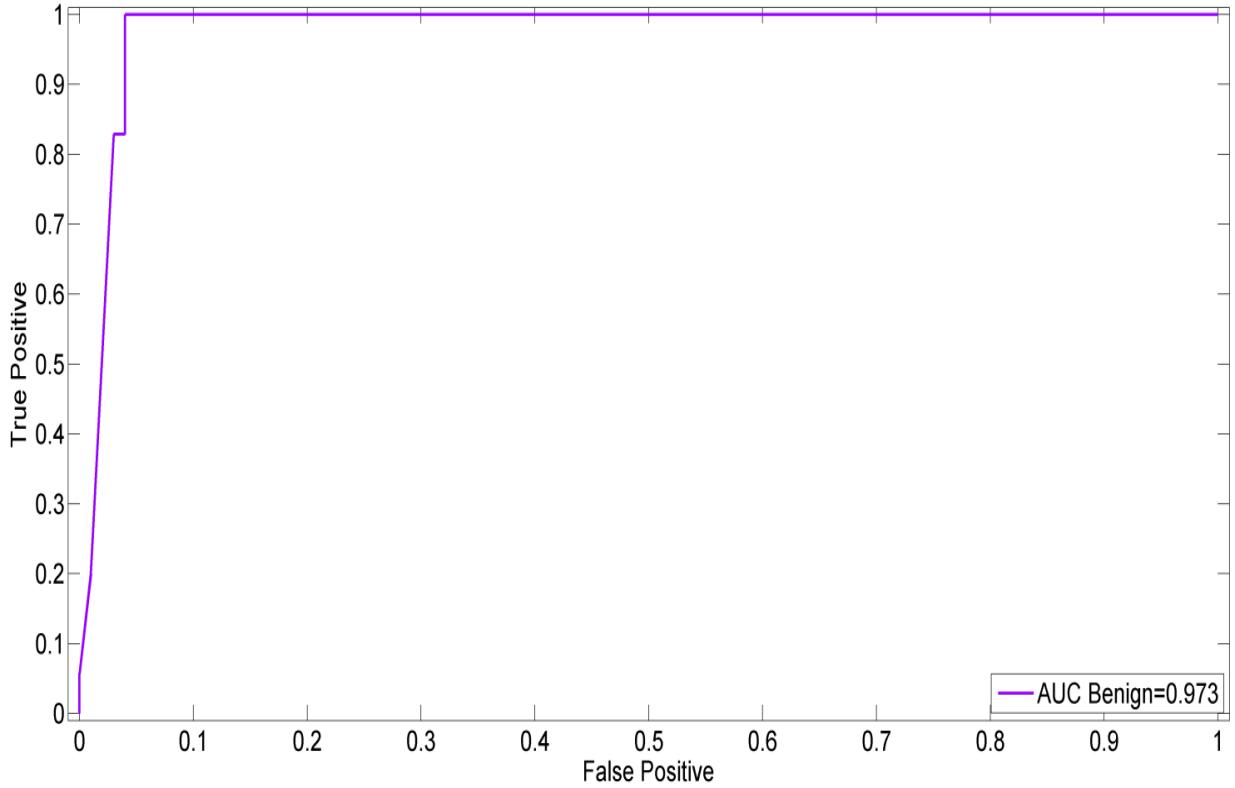


Figure 4-5: ROC curve indicating accuracy of proposed CAD model as area under the curve. Roc curve of benign-malignant group with morphological features and area under the ROC curve is (Az=0.97)

Table 4.1: Performance of the benign malignant group with morphological features. This table indicates the performance of the proposed CAD model for benign and malignant group	
Performance Metric	Benign-Malignant group
Accuracy	94%
Az	0.97
Sensitivity	97%
Specificity	88%

Accuracy= (TP+TN)/(TP+TN+FP+FN);Sensitivity=TP/(TP+FN);Specificity=TN/(TN+FP)

Table 4.2: Performance of the benign-malignant-normal group with textural features. This table indicates the performance of the proposed CAD model for benign-malignant and normal group	
Performance Metric	Benign-Malignant-Normal group
Accuracy	82%
Az(Benign)	0.92
Az(Malignant)	0.84
Az(Normal)	0.86
Sensitivity	94%
Specificity	83%

Accuracy= (TP+TN)/ (TP+TN+FP+FN); Sensitivity=TP/ (TP+FN); Specificity=TN/ (TN+FP)

4.4 Features analysis for CAD model

Morphological and textural features were analyzed in order to find the best combination of features in classifying benign malignant and benign-malignant and normal group.

4.5 Morphological features analysis

There were thirteen morphological features used for classifying benign-malignant group, **Table 4.3**provides performance analysis of all features with Az (area under the curve), sensitivity and specificity to find the most significant feature in classifying benign and malignant lesions. Results showed that features related to shape (i.e. roundness, form factor, solidity,

convexity, extent) provide higher Az value of about (0.85-0.87%) and posterior acoustic enhancement have Az of (0.85 %).

Table 4.3: Az (area under the curve) for thirteen morphological features in classifying benign and malignant lesion on breast ultrasound images. This table indicates the accuracy of features used in the CAD model for classifying benign and malignant tumor. Thirteen features are arranged in the order of (Spiculation, branch-pattern, lobulation, relative brightness of nodule, ENC, LS ratio, roundness, form factor, solidity, convexity, extent, aspect ratio, posterior enhancement).

	1	2	3	4	5	6	7	8	9	10	11	12	13
1	0.35	0.57	0.58	0.55	0.70	0.78	0.83	0.80	0.78	0.77	0.76	0.75	0.73
2	0.73	0.73	0.73	0.72	0.72	0.72	0.74	0.74	0.74	0.74	0.73	0.73	0.73
3	0.73	0.72	0.72	0.72	0.72	0.73	0.74	0.74	0.74	0.74	0.73	0.73	0.73
4	0.73	0.72	0.72	0.72	0.73	0.74	0.75	0.75	0.74	0.74	0.74	0.74	0.73
5	0.74	0.74	0.74	0.75	0.76	0.76	0.77	0.78	0.78	0.79	0.80	0.80	0.81
6	0.81	0.81	0.81	0.82	0.82	0.82	0.83	0.83	0.83	0.84	0.84	0.84	0.85
7	0.85	0.85	0.86	0.86	0.86	0.86	0.87	0.87	0.87	0.87	0.87	0.88	0.88
8	0.88	0.87	0.87	0.87	0.87	0.87	0.88	0.87	0.87	0.87	0.87	0.87	0.87
9	0.86	0.86	0.86	0.86	0.86	0.86	0.86	0.86	0.86	0.86	0.86	0.86	0.86
10	0.85	0.85	0.85	0.85	0.85	0.85	0.86	0.85	0.85	0.85	0.85	0.85	0.85
11	0.85	0.85	0.85	0.84	0.85	0.85	0.85	0.85	0.85	0.85	0.85	0.84	0.84
12	0.84	0.84	0.84	0.84	0.84	0.84	0.84	0.84	0.84	0.84	0.84	0.84	0.84
13	0.83	0.83	0.83	0.83	0.83	0.83	0.84	0.83	0.83	0.83	0.83	0.83	0.83

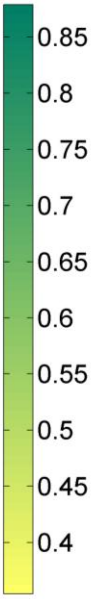


Table 4.4: The Sensitivity of thirteen morphological features for classifying benign and malignant lesions on breast ultrasound images. This table represents the sensitivity of features used in CAD model for classifying benign and malignant tumor. Thirteen features are arranged in the order of (Spiculation, branch-pattern, lobulation, relative brightness of nodule, ENC, LS ratio, roundness, form factor, solidity, convexity, extent, aspect ratio, posterior enhancement).

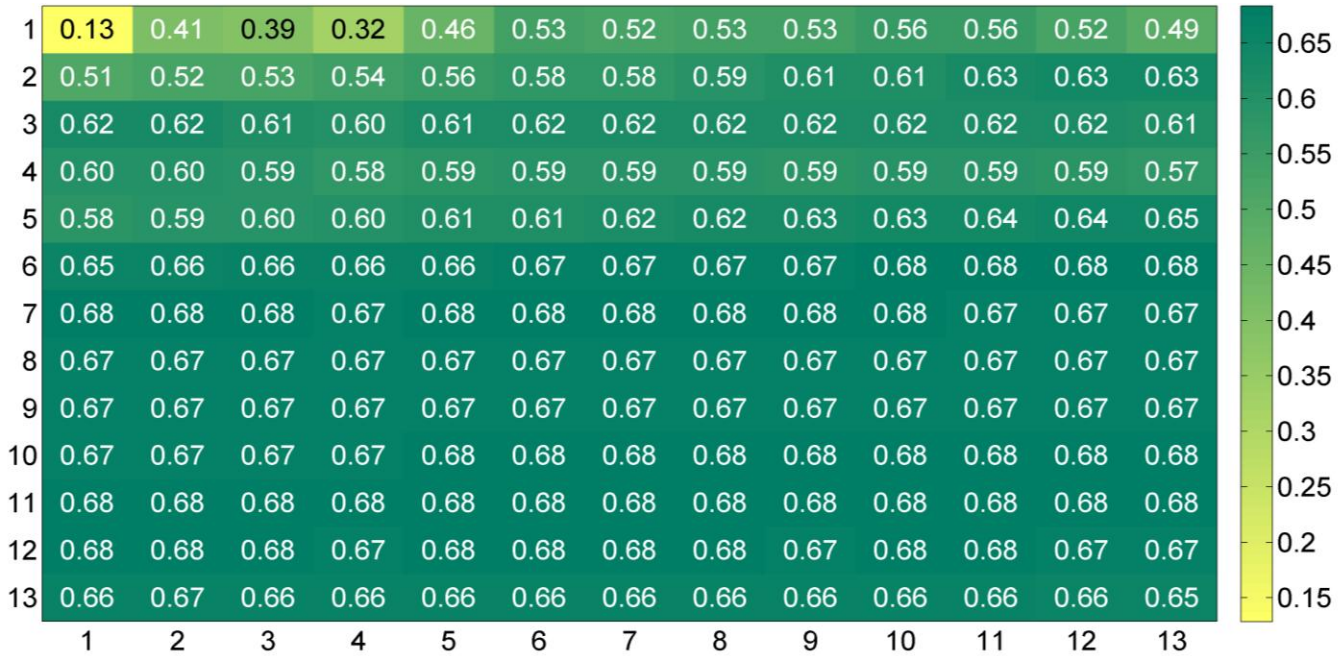
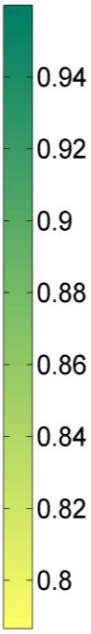


Table4.5: The Specificity of thirteen morphological features in classifying benign and malignant lesion on breast ultrasound images. This table represents the specificity of features used in CAD model for classifying benign and malignant tumor. Thirteen features are arranged in the order of (Spiculation, branch-pattern, lobulation, relative brightness of nodule, ENC, LS ratio, roundness, form factor, solidity, convexity, extent, aspect ratio, posterior enhancement)

	1	2	3	4	5	6	7	8	9	10	11	12	13
1	0.96	0.81	0.79	0.83	0.87	0.89	0.91	0.89	0.86	0.86	0.83	0.84	0.85
2	0.83	0.82	0.81	0.80	0.81	0.82	0.82	0.81	0.81	0.82	0.82	0.81	0.80
3	0.80	0.79	0.79	0.79	0.80	0.80	0.81	0.81	0.80	0.80	0.80	0.80	0.80
4	0.80	0.80	0.80	0.80	0.81	0.81	0.81	0.81	0.81	0.81	0.80	0.80	0.81
5	0.81	0.81	0.82	0.82	0.82	0.83	0.83	0.83	0.84	0.84	0.84	0.84	0.84
6	0.85	0.85	0.85	0.85	0.86	0.86	0.86	0.86	0.86	0.87	0.87	0.87	0.87
7	0.87	0.87	0.87	0.87	0.88	0.88	0.88	0.88	0.88	0.88	0.88	0.88	0.88
8	0.88	0.88	0.88	0.88	0.88	0.88	0.88	0.88	0.88	0.88	0.88	0.87	0.87
9	0.87	0.87	0.87	0.86	0.87	0.87	0.87	0.87	0.86	0.86	0.86	0.86	0.86
10	0.86	0.86	0.86	0.86	0.86	0.86	0.86	0.86	0.86	0.86	0.86	0.86	0.86
11	0.86	0.86	0.85	0.85	0.85	0.85	0.86	0.85	0.85	0.85	0.85	0.85	0.85
12	0.85	0.85	0.84	0.84	0.85	0.85	0.85	0.85	0.85	0.85	0.84	0.84	0.84
13	0.84	0.84	0.84	0.84	0.84	0.84	0.85	0.84	0.84	0.84	0.84	0.84	0.84



4.6 Textural features analysis

There were sixteen Textural features used for differentiating benign-malignant-normal images. A performance analysis of all textural features in identifying the most significant feature is listed in **Table 4.6**. The result signifies that among textural features contrast and homogeneity provides higher accuracy of 0.90, sensitivity and specificity of 90% and 82%.

Table 4.6: Az (area under the curve) for textural features in classifying benign, malignant and normal images. This table indicates the accuracy of features used in CAD model for classifying benign, malignant and normal images. Sixteen features are arranged in the order of (contrast, correlation, energy and homogeneity)

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
1	0.75	0.75	0.83	0.82	0.80	0.79	0.82	0.81	0.81	0.80	0.82	0.81	0.81	0.80	0.81	0.81
2	0.81	0.79	0.80	0.80	0.80	0.79	0.80	0.80	0.79	0.78	0.79	0.79	0.79	0.78	0.78	0.78
3	0.79	0.80	0.80	0.81	0.81	0.82	0.82	0.82	0.83	0.83	0.83	0.84	0.84	0.84	0.84	0.85
4	0.85	0.84	0.85	0.85	0.84	0.84	0.84	0.84	0.84	0.84	0.84	0.84	0.84	0.84	0.84	0.84
5	0.84	0.84	0.84	0.84	0.84	0.84	0.84	0.84	0.84	0.83	0.84	0.83	0.83	0.83	0.83	0.83
6	0.83	0.83	0.83	0.83	0.83	0.83	0.83	0.83	0.83	0.82	0.82	0.82	0.82	0.82	0.82	0.82
7	0.82	0.82	0.83	0.83	0.83	0.83	0.83	0.83	0.83	0.83	0.84	0.84	0.84	0.84	0.84	0.84
8	0.84	0.84	0.84	0.84	0.84	0.84	0.84	0.84	0.84	0.84	0.84	0.84	0.84	0.84	0.84	0.84
9	0.84	0.84	0.84	0.84	0.84	0.84	0.84	0.84	0.84	0.84	0.84	0.84	0.83	0.83	0.83	0.83
10	0.83	0.83	0.83	0.83	0.83	0.83	0.83	0.83	0.83	0.83	0.83	0.83	0.83	0.83	0.83	0.83
11	0.83	0.83	0.83	0.83	0.83	0.83	0.83	0.83	0.84	0.84	0.84	0.84	0.84	0.84	0.84	0.84
12	0.84	0.84	0.84	0.84	0.84	0.84	0.84	0.84	0.84	0.84	0.84	0.84	0.84	0.84	0.84	0.84
13	0.84	0.84	0.84	0.84	0.84	0.84	0.84	0.84	0.84	0.84	0.84	0.84	0.84	0.83	0.84	0.83
14	0.83	0.83	0.83	0.83	0.83	0.83	0.83	0.83	0.83	0.83	0.83	0.83	0.83	0.83	0.83	0.83
15	0.83	0.83	0.83	0.83	0.83	0.83	0.83	0.83	0.83	0.84	0.84	0.84	0.84	0.84	0.84	0.84
16	0.84	0.84	0.84	0.84	0.84	0.84	0.84	0.84	0.84	0.84	0.84	0.84	0.84	0.84	0.84	0.84

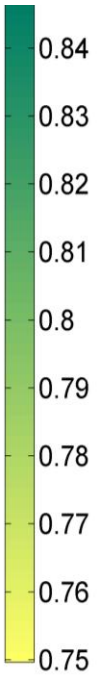


Table4.7: Sensitivity of Textural features in classifying benign and malignant and normal image. This table indicates the sensitivity of features used in CAD model for classifying benign, malignant and normal images. Sixteen features are arranged in the order of (contrast, correlation, energy and homogeneity)

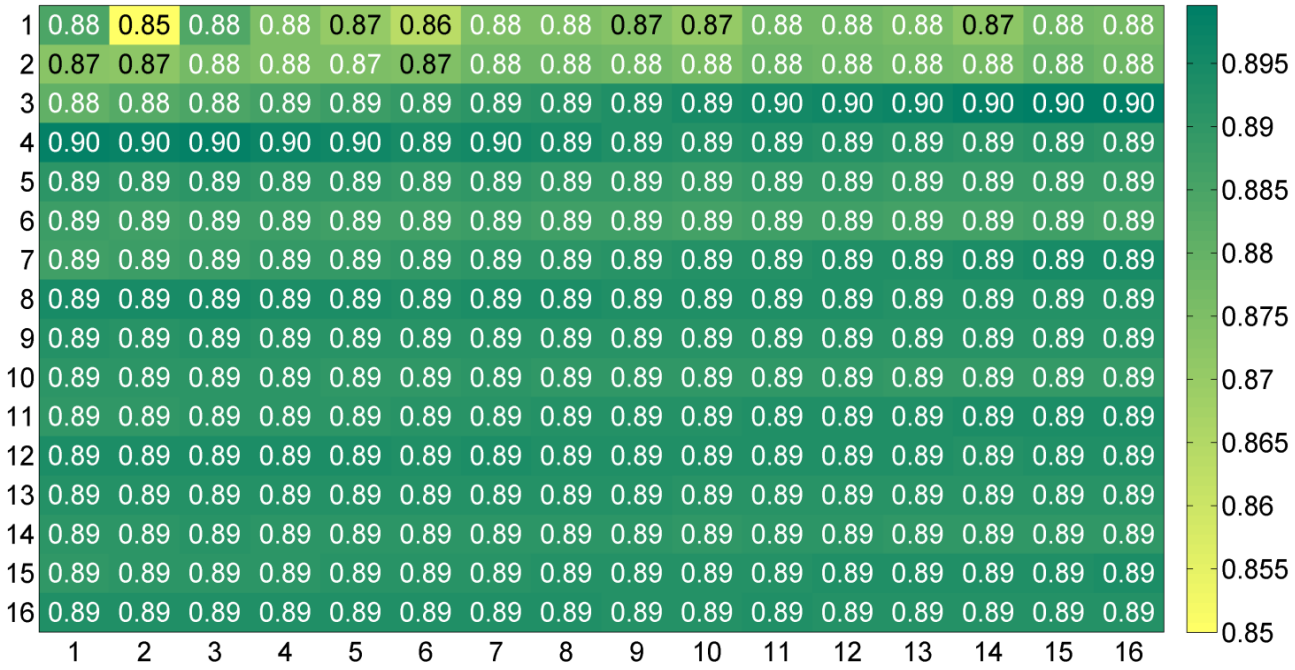
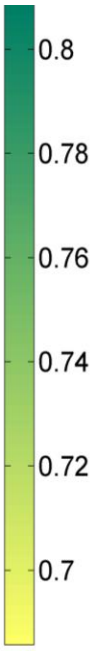


Table 4.8: Specificity of Textural features in classifying benign and malignant and normal image. This table indicates the specificity of features used in CAD model for classifying benign, malignant and normal images. Sixteen features are arranged in the order of (contrast, correlation, energy and homogeneity).

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
1	0.69	0.71	0.76	0.79	0.77	0.76	0.78	0.78	0.77	0.77	0.78	0.77	0.76	0.76	0.77	0.77
2	0.77	0.76	0.77	0.77	0.77	0.76	0.77	0.77	0.77	0.76	0.76	0.76	0.76	0.76	0.76	0.76
3	0.76	0.77	0.77	0.77	0.78	0.78	0.78	0.78	0.78	0.79	0.79	0.79	0.79	0.79	0.80	0.80
4	0.80	0.80	0.80	0.80	0.80	0.80	0.80	0.80	0.81	0.81	0.81	0.81	0.81	0.81	0.81	0.81
5	0.81	0.81	0.81	0.81	0.81	0.80	0.81	0.81	0.80	0.80	0.80	0.80	0.80	0.80	0.80	0.80
6	0.80	0.80	0.80	0.80	0.80	0.79	0.80	0.80	0.79	0.79	0.79	0.79	0.79	0.79	0.79	0.79
7	0.79	0.79	0.79	0.79	0.79	0.79	0.79	0.79	0.79	0.80	0.80	0.80	0.80	0.80	0.80	0.80
8	0.80	0.80	0.80	0.80	0.80	0.80	0.80	0.80	0.80	0.80	0.80	0.80	0.80	0.80	0.80	0.80
9	0.80	0.80	0.80	0.80	0.80	0.80	0.80	0.80	0.80	0.80	0.80	0.80	0.80	0.80	0.80	0.80
10	0.80	0.80	0.80	0.80	0.80	0.80	0.80	0.80	0.80	0.79	0.79	0.79	0.79	0.79	0.79	0.79
11	0.79	0.79	0.79	0.79	0.79	0.79	0.80	0.80	0.80	0.80	0.80	0.80	0.80	0.80	0.80	0.80
12	0.80	0.80	0.80	0.80	0.80	0.80	0.80	0.80	0.80	0.80	0.80	0.80	0.80	0.79	0.80	0.80
13	0.79	0.79	0.79	0.79	0.79	0.79	0.79	0.79	0.79	0.79	0.79	0.79	0.79	0.79	0.79	0.79
14	0.79	0.79	0.79	0.79	0.79	0.79	0.79	0.79	0.79	0.79	0.79	0.79	0.79	0.79	0.79	0.79
15	0.79	0.79	0.79	0.79	0.79	0.79	0.79	0.79	0.79	0.79	0.79	0.79	0.79	0.79	0.79	0.79
16	0.79	0.79	0.79	0.79	0.79	0.79	0.79	0.79	0.79	0.79	0.80	0.80	0.80	0.80	0.80	0.80



5
DISCUSSION

Breast cancer is the leading cause of death in developing countries. Breast cancer has a high incidence rate and the cause of breast cancer is still unknown. Early diagnosis plays an important role toward the diagnosis and treatment of cancer because early detection can prevent cancer from being spread to distant parts of the body. Breast ultrasound is one of the useful diagnostic tools used today for the breast cancer detection. However, ultrasound image interpretations are operator dependent and depend upon the radiologist experience and fatigue level. The role of CAD models in ultrasound image is to provide second opinion for the interpretation of tumors to the radiologist and reduces inter and intra-observer variability.

5.1 Benign-malignant group

In this study a CAD model had been developed and successfully implemented for the classification among two groups. First group was benign-malignant group for which morphological features were extracted and SVM were used for the classification. Classifiers accuracy for first group was 94% and area under the ROC curve was 0.97, sensitivity and specificity for the first group came out to be 97% and 88%. . Results showed that morphological features play an important role in differentiating benign from malignant tumor as these features are associated with the shape and contours of the tumor that significantly varies between benign and malignant and is not dependent upon the USG settings.

5.2 Significant feature for benign-malignant group

The most significant morphological feature, in classifying first group was determined with area under the curve (Az), sensitivity and specificity as listed in Table 3. The highest accuracy for classification was in between roundness and posterior acoustic enhancement having Az value of 88%, sensitivity of 80%, specificity of 88%. Roundness and posterior enhancement is also the most discriminating feature for the radiologist in diagnosing a tumor, benign tumor tends to be in round or oval shape having no posterior enhancement.

5.3 Benign-malignant-normal group

In the developed CAD model, second group was among benign-malignant and normal images for which textural features were used and multiclass SVM using one vs. all method for classification. Classifiers results showed area under the curve for benign, malignant and normal was 0.92, 0.84, and 0.86. Textural features are associated with intensity variations in each image and this variation is different among three classes, as they tend to be hyperechoic, hypoechoic and isoechoic.

5.4 Significant feature for benign-malignant-normal group

The most significant textural feature for classifying benign, malignant and normal group was determined with area under the curve, sensitivity and specificity. The highest accuracy was achieved among contrast and homogeneity having Az of 0.90, sensitivity and specificity of 90% and 82%. Contrast defines the intensity variation in the image and can be different among different categories of breast tumor on ultrasound images, resulting in providing the better accuracy for classification.

CAD model can be improved further by increasing the data set, calculating the more features related to the echogenity.

5.5 Limitations

- CAD model is focused on differentiating benign from malignant lesion.
- Cystic images are not included in CAD model.
- Images used in CAD model have clearly visible lesion.

6
CONCLUSION & FUTURE WORK

Early diagnosis of breast cancer provides the best chances of survival to the patients. Breast ultrasound is an important diagnostic tool used for the diagnosis of lesion as well as for the identification of lesions not easily visualized on mammography. CAD(Computer aided diagnosis) model are focused on developing an intelligent system that can provide reliable second opinion to the radiologist for the diagnosis and can help decrease the observer variability. CAD system includes preprocessing, segmentation and classification of lesion. Segmentation and feature extraction plays an important part in developing an efficient CAD model. Features develop on the basis of BI-RAD descriptors provides standard criteria for the assessment of lesion among various observers.

CAD models that can be used for the diagnosis of breast tumor should also be able to identify a normal breast ultrasound image because some young females have high fibrocystic tissues. Images having high fibrocystic tissues can lead to the false alarm, requiring hospital visits and unnecessary biopsy. In our proposed CAD model normal images were part of the second group that was classified on the basis of textural features using multiclass SVM. The proposed CAD model can be useful in eliminating such false alarm and can provide reliable second opinion to the radiologist.

In conclusion, the developed CAD model is a multipurpose system that can be used for the differentiation of benign from malignant as well as normal ultrasound image from the benign and malignant depending upon the use of the system. This model can be extended further for multiclass classification between different classes of benign and malignant using multiclass SVM.

A CAD model depends upon developing an efficient system that can provide reliable decision in diagnosis. CAD models for breast cancer are focused on early detection to provide better treatment to the patient. Breast cancer is diagnosed using mammography and ultrasound.

CADmodel for breast ultrasound images need to be focused on the following issues in order to be feasible for clinical application.

- Ultrasound displays a real time video and most of the CAD models are focused on capturing still images.CAD model that can able to work on real time video of ultrasound can be a practical solution to the problem.
- Having CAD working on video it can be implemented on machines that can provide the operator doing the scan with valuable diagnostic information and preventing the false positive.

REFERENCES

[1] M. L. Giger, H. Al-Hallaq, Z. Huo, C. Moran, D. E. Wolverton, C. W. Chan, and W. Zhong, "Computerized analysis of lesions in US images of the breast," *Academic radiology*, vol. 6, pp. 665-674, 1999.

[2] C.-M. Chen, Y.-H. Chou, K.-C. Han, G.-S. Hung, C.-M. Tiu, H.-J. Chiou, and S.-Y. Chiou, "Breast Lesions on Sonograms: Computer-aided Diagnosis with Nearly Setting-Independent Features and Artificial Neural Networks," *Radiology*, vol. 226, pp. 504-514, 2003.

[3] M. Woo Kyung, L. Chung-Ming, C. Nariya, C. Jung Min, H. Chiun-Sheng, C. Jeon-Hor, and C. Ruey-Feng, "Computer-aided diagnosis of breast masses using quantified BI-RADS findings," *Comput. Methods Prog. Biomed.*, vol. 111, pp. 84-92.

[4] A. T. Stavros, D. Thickman, C. L. Rapp, M. A. Dennis, S. H. Parker, and G. A. Sisney, "Solid breast nodules: use of sonography to distinguish between benign and malignant lesions," *Radiology*, vol. 196, pp. 123-134, 1995.

[5] H.-J. Lee, E.-K. Kim, M. J. Kim, J. H. Youk, J. Y. Lee, D. R. Kang, and K. K. Oh, "Observer variability of Breast Imaging Reporting and Data System (BI-RADS) for breast ultrasound," *European journal of radiology*, vol. 65, pp. 293-298, 2008.

[6] R. M. Rangayyan, F. b. J. Ayres, and J. E. Leo Desautels, "A review of computer-aided diagnosis of breast cancer: Toward the detection of subtle signs," *Journal of the Franklin Institute*, vol. 344, pp. 312-348, 2007.

[7] P. Perona, T. Shiota, and J. Malik, "Anisotropic diffusion," in *Geometry-driven diffusion in computer vision*: Springer, 1994, pp. 73-92.

[8] M. Nixon, *Feature extraction & image processing*: Academic Press, 2008.

[9] Hassan, Naglaa, and N. Akamatsu, "A new approach for contrast enhancement using sigmoid function," 2004.

[10] E. Alpaydin, *Introduction to machine learning*: MIT press, 2004.

[11] A. G. Barto, *Reinforcement learning: An introduction*: MIT press, 1998.

[12] C. Campbell and Y. Ying, "Learning with support vector machines," *Synthesis Lectures on Artificial Intelligence and Machine Learning*, vol. 5, pp. 1-95.

[13] F. Lotte, M. Congedo, A. LÃ©cuyer, F. Lamarche, and B. Arnaldi, "A review of classification algorithms for EEG-based brainâ€“computer interfaces," *Journal of neural engineering*, vol. 4, 2007.

[14] D.-R. Chen, R.-F. Chang, W.-J. Kuo, M.-C. Chen, and Y.-L. Huang, "Diagnosis of breast tumors with sonographic texture analysis using wavelet transform and neural networks," *Ultrasound in medicine & biology*, vol. 28, pp. 1301-1310, 2002.

[15] D.-R. Chen, R.-F. Chang, and Y.-L. Huang, "Computer-aided diagnosis applied to US of solid breast nodules by using neural networks 1," *Radiology*, vol. 213, pp. 407-412, 1999.

[16] K. Drukker, M. L. Giger, C. J. Vyborny, and E. B. Mendelson, "Computerized detection and classification of cancer on breast ultrasound< sup> 1</sup>," *Academic radiology*, vol. 11, pp. 526-535, 2004.

[17] D.-R. Chen, Y.-L. Huang, and S.-H. Lin, "Computer-aided diagnosis with textural features for breast lesions in sonograms," *Computerized Medical Imaging and Graphics*, vol. 35, pp. 220-226.

[18] K. Z. Abd-Elmoniem, A. Youssef, and Y. M. Kadah, "Real-time speckle reduction and coherence enhancement in ultrasound imaging via nonlinear anisotropic diffusion," *Biomedical Engineering, IEEE Transactions on*, vol. 49, pp. 997-1014, 2002.

[19] G. Rahbar, A. C. Sie, G. C. Hansen, J. S. Prince, M. L. Melany, H. E. Reynolds, V. P. Jackson, J. W. Sayre, and L. W. Bassett, "Benign versus Malignant Solid Breast Masses: US Differentiation 1," *Radiology*, vol. 213, pp. 889-894, 1999.

[20] S. Joo, Y. S. Yang, W. K. Moon, and H. C. Kim, "Computer-aided diagnosis of solid breast nodules: use of an artificial neural network based on multiple sonographic features," *Medical Imaging, IEEE Transactions on*, vol. 23, pp. 1292-1300, 2004.

[21] K. Horsch, M. L. Giger, L. A. Venta, and C. J. Vyborny, "Computerized diagnosis of breast lesions on ultrasound," *Medical Physics*, vol. 29, pp. 157-164, 2002.

[22] R. Sivaramakrishna, K. A. Powell, M. L. Lieber, W. A. Chilcote, and R. Shekhar, "Texture analysis of lesions in breast ultrasound images," *Computerized medical imaging and graphics*, vol. 26, pp. 303-307, 2002.

[23] N. A. Obuchowski, "Receiver Operating Characteristic Curves and Their Use in Radiology 1," *Radiology*, vol. 229, pp. 3-8, 2003.

[24] S.-T. Chen, Y.-H. Hsiao, Y.-L. Huang, S.-J. Kuo, H.-S. Tseng, H.-K. Wu, and D.-R. Chen, "Comparative analysis of logistic regression, support vector machine and artificial neural network for the differential diagnosis of benign and malignant solid breast tumors by the use of three-dimensional power Doppler imaging," *Korean Journal of Radiology*, vol. 10, pp. 464-471, 2009.

[25] Z. Lihong, S. Ying, Z. Yushi, Z. Cheng, and Z. Yi, "Face recognition based on multi-class SVM," presented at Control and Decision Conference, 2009. CCDC'09. Chinese, 2009.

[26] R. Ryan and K. Aldebaro, "In Defense of One-Vs-All Classification," *J. Mach. Learn. Res.*, vol. 5, pp. 101-141, 2004.