Detection of Tumor in Mammogram

A Dissertation submitted in partial fulfilment of the requirements for the award of degree of

Master of Engineering
in
Electronic Instrumentation and Control

Submitted By
Kanchan Sharma
8211510003

Under the Guidance of
Dr. Deepti Mittal
Assistant Professor,EIED
Thapar University,Patiala

Electrical and Instrumentation Engineering Department
Thapar University
(Established under the section 3 of UGC act, 1956)
Patiala, 147004, Punjab, India

July 2014

DECLARATION

I hereby certify that the work is being presented in this thesis work entitled **"Detection of Tumor in Mammogram"** in partial fulfilment of award of degree of Master of Engineering in Electronics Instrumentation & Control submitted in Electrical & Instrumentation Engineering Department,Thapar University, Patiala is an authentic record of my own work carried under the supervision of **Dr Deepti Mittal**, Assistant Professor, Department Of Electrical & Instrumentation Engineering, Thapar University, Patiala, Punjab.

Date: 18/07/2014

Kanchan Sharma
821151003

I certify that the above statement made by the student is correct to the best of my knowledge and belief.

Date: 18/07/2014

Dr.Deepti Mittal
Assistant Professor,EIED
Thapar University,Patiala
Punjab

Countersigned By

Dr.Ravinder Agarwal
Head Of Department
Department of Electrical &
Instrumentation Engineering
Thapar University, Patiala.
Punjab

Dr. S.K. Mohapatra
Dean of Academic Affair
Thapar University, Patiala
Punjab

ACKNOWLEDGEMENTS

Firstly I would like to express my sincere gratitude to **Dr. Prakash Golapan,** Director of Thapar University, Patiala, Punjab for providing the facilities to accomplish M.E. in Thapar University, Patiala.

I would like to thank **Dr. Ravinder Agarwal**, Head of the Department of Electrical and Instrumentation Engineering , Thapar University, Patiala who inspire and support me throughout the work.

I would like to express my sincere gratitude to my Guide, **Dr. Deepti Mittal,** Assistant Professor of Electrical and Instrumentation Engineering ,Thapar University, Patiala for her guidance and support throughout my work. Without her guidance I was unable to do my work with ease. During my work, the problems which I have faced, she has helped me at her best to solve my problems. I am very thankful to her to guide me and give her valuable time to me.

I would like to express my sincere gratitude to my staff members of Electronics Instrumentation and Control Department for helping me in my thesis work.

I am very thankful to God to give me a chance to pursue my academic goal and for giving me everything which I want.

Finally, I would like to express thanks to my parents for their sacrifice to make my life wonderful and full of knowledge.

DATE: 18/07/2014

Place: Thapar University, Patiala

Name: Kanchan Sharma

Roll No: 821151003

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

CLAHE	Contrast Limited Adaptive Histogram Equalization
PN	Patient
OIV	Original Image Values
CIV	Cancerous Image Values

ABSTRACT

Breast cancer is one of the common cancer today, mostly found in women. It is due to the abnormal growth of the breast cells. Research found in the year 2012 that Belgium had the highest rate of breast cancer and that is followed by Denmark and France. The highest incidence of breast cancer was found in Northern America and Oceania; and the lowest incidence was found in Asia and Africa [1].According to the Times of India ,the incident number of breast cancer rate in India in 2001 and 2016 respectively:89,914:140975[2].To decrease the death rate due to breast cancer, mammography is the process to diagnose and early detection of the breast cancer whereas the mammogram is the X-ray exam to reveal the abnormalities(benign and malignant).In mammogram, there are two plates in which breast is compressed and applying a radiation's small dose for producing the X-ray image. It is used for screening and diagnosing the breast cancer[3].

In the Mammography based research work, we have used Digital Image Processing with MAT LAB Vversion R2007b. In this proposed work of mammography, We have taken a database image of DDSM/MIAS and apply all the processes to find the Benign and Malignant in the breast. Firstly taken the DDSM/MIAS Breast image, then apply pre-processing and image enhancement technique .After that image segmentation (edge detection)technique is applied with thresholding and in the end GLCM texture features are extracted from the images.

Keywords: Database image DDSM/MIAS, Pre-processing, Image Enhancement, Segmentation, GLCM Features, Benign and Malignant.

CHAPTER 1

INTRODUCTION OF MAMMOGRAPHY

1.1 Breast Cancer

A cancer is a disease that is caused by the abnormal growth of the cells. Breast cancer is the breast disease that starts from the breast tissues and spread in the duct or lobes of the breast. When controlling in the breast cells are not proper, they divide and produce in the form of lumps or tumour, that is called breast tumor/cancer[4].Breast cancer is the second most common cancer and second leading cause of the death. Breast cancer is caused due to the genetic abnormalities with age that is 90% and cancer due to abnormalities received from father and mother is 5-10%.[5].Breast Cancer mostly occur in females as compared to males because of smaller and more muscular males chest as compared to females. The second most reason is that male chest has no glands and having no capability to produce milk for nursing babies.

1.2 Types of Breast Cancer

Breast cancer is one disease but there are many types of breast cancer. Mainly there are two types of Breast cancer :

1. Non invasive breast cancer
2. Invasive breast cancer

1.2.1 Non-invasive breast cancer

This is the cancer that has abnormal/cancerous cells inhabit in the breast lobules or ducts. It is also called Pre-invasive cancer

It is of two types:

1. LCIS(Lobular carcinoma in situ)
2. DCIS(Ductal carcinoma in situ)

In LCIS, cancerous cells inhabits in its place i.e. lobules and cannot roll out to the breast surroundings whereas in case of DCIS, cancerous cells inhabit in ducts itself and donot have ability to roll out outside its place.”In SITU” means” its place”. Benign is invasive type.

1.2.2 Invasive breast cancer

This is the cancer having cancerous cells that have ability to move outside the breast and roll out to the other parts of the body. It starts from original place i.e. milk ducts or lobules and move into surrounding breast tissues[6].

Mainly it can be of two types:

1. Lobular breast cancer
2. Ductal breast cancer

Both are invasive breast cancer but lobular breast cancer's inventive site is lobules, it starts from lobules and roll out other parts of the body.

In case of Ductal breast cancer, it originates from ducts and roll out to other parts of the body. Malignant is the invasive breast cancer[6].

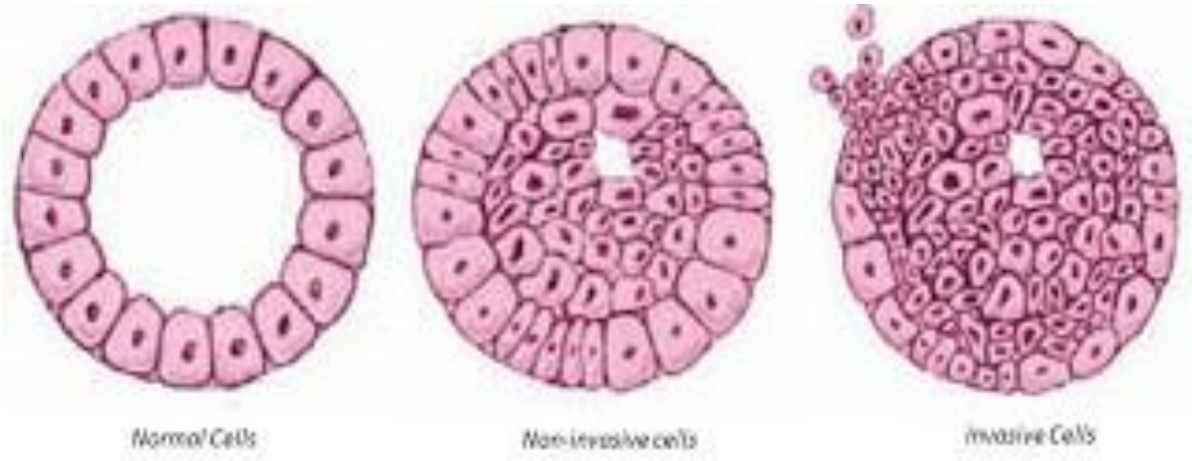


Figure 1: Normal, Non-invasive and invasive cells

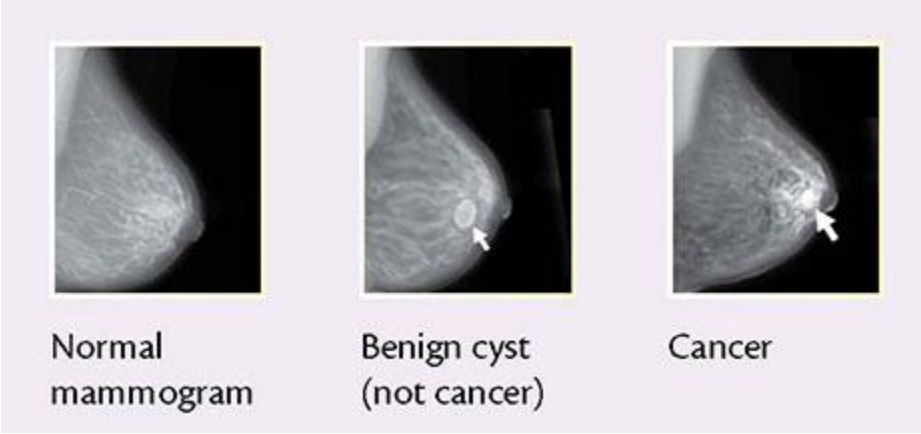


Figure 2: Normal,Benign Cyst and Cancer

1.3 Breast Cancer Diagnosis

There are many types of tests for determination of breast cancer:

1. Surgerical test
2. Imaging test

3. Blood test
4. Genetic test
5. Additional test

Surgerical tests require biopsy whereas image tests demand ultrasound, MRI and Mammography and blood test demand Complete Blood Count (CBC) and Serum[20]. Mammography is detailed below:

1.4 Mammography

Mammography is the screening tool/agenda for showing and identificating the breast cancer. This process is used for the prior detection of the breast cancer. Mammogram is the X-ray exam and it is used to detect and identify the prior breast cancer. It captures changes in the breast which are not normal. A mammogram can find the breast cancer when cancer is so small that we can't feel it. It identifies the breast cancer 2-3 years before we feel or know about it. The mammogram having good standard and it can find 80-90% breast cancer because there is no screening tool is 100% fruitful[8].

1.4.1 Mammography Types

There are main two types of mammography:

1. Film screen mammography
2. Digital mammography

The procedure how to work ,they are same. The object that differs them from each other is that images are captured from photography or from digital files recorded directly on computer[9].

1.4.1.1Filmscreenmammography

In this mammography, black and white images appear on large sheet of film[9].Here image is stored on the film.

1.4.1.2 Digital mammography

It is first approved in 2000 by Food and Drug administration. It is also called full-field Digital mammography. In this method ,electronic image is taken and storing that image by computer[7]. This process that has solid state detector that captures X-rays and converted X-Ray signals into digital image. These detectors are same as digital cameras. It offers a number of practical advantages:

1. Time Consumption: The digital machines are fast and patients remain uncomfort for small amount of time. So they take small amount of time to diagnose and early detect of breast cancer.
2. Transpotrtation: The digital images can be retrieved and stored easily. So transportation of these images is easy
3. Availability: Digital images are easily available because for development of film, there is no need to wait.
4. Transmision: Digital images are transmitted from one physician to other quickly and easily.

   There is easier to see the subtle difference between tissues of mammogram images after completition of mammogram because after completition of mammogram brightness, contrast and darkness of images can be adjusted and sections of the images can be magnified[9]

1.5 Comparison between digital and film mammography

Similiarities

1. To take the images of the breast ,both types of mammography use compression and X-rays[11]

Differences:

1. Digital mammography uses less radiation as compared to film mammography.
2. Digital mammography is expensive as compared to film mammography.
3. In digital mammography images can be stored and shared more easily than film mammography[12]
4. Confirmation of image quality and patients satisfaction by Digital mammography is more as compared to film mammography[13]
5. Accuracy is more in digital mammography.
6. In film mammography, image is stored on the film while in digital mammography, image is stored on computer[14]

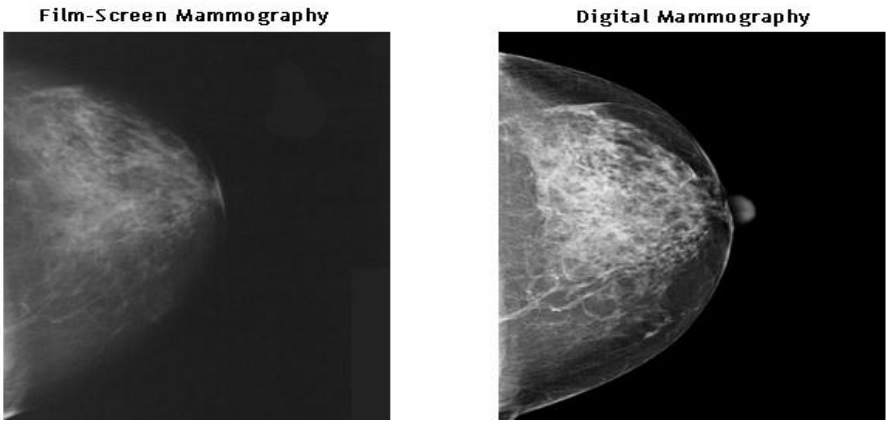


Figure 3: Film screen and Digital mammography

1.6 Diagnosis of Breast Cancer

Different types of tests are there for diagnosing of breast cancer, mainly two types of mammogram/mammography tests are:

1. Screening mammogram
2. Diagnostic mammogram

1.6.1 Screening Mammogram

This mammogram is done only then when women have no indication of cancer. It detects lumps and tumors and that lumps and tumors do not feel by women. It involves two x-rays of each breast[14].It takes two images of each breast.This mammogram can detect the micro-calcification [14].

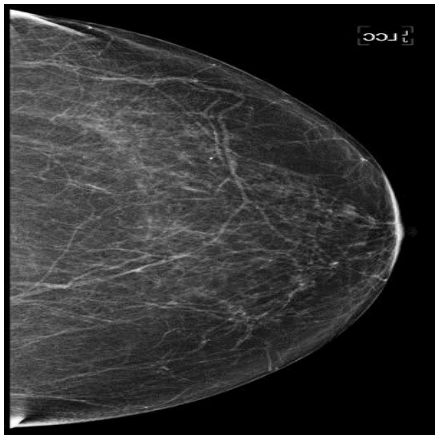


Figure 4: Screening Mammogram

1.6.2 Diagnostic Mammogram

When lump or other indication of mammogram found, then this mammogram is used. The symptoms of mammogram are like pain in breast, nipple discharge and change the shape or size of the breast. The breast tissues which are hard to find on screening mammogram ,diagnostic mammogram helps to see that tissues also. This mammogram takes many pictures of breast from different angles.

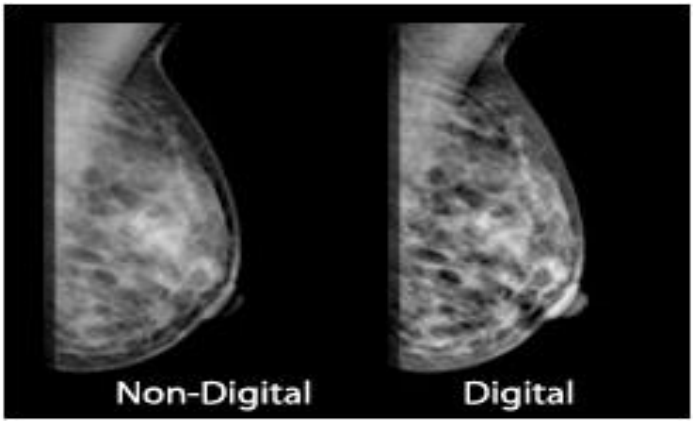


Figure 5: Diagnostic Mammogram

1.7 Mammogram Procedure

The person who takes X-rays is radiologist technician. In X-ray machine ,there are two plates X-ray plate and plastic plate. These plates are connected to X-ray machine. The radiologist technician adjusts the breast once in a time between those two plates and compress the breasts and make them flatten. The breast tissues move out for obtaining a clearer picture. Mostly two pictures are taken –one from above and one from the side. A screening mammogram occupies 20 minutes to finish the procedure

During mammogram, you will feel pressure but only for few seconds. You sense some discomfort and squashed or pinched and this feelings remain only for some minutes. More the breast is flattened ,more clearer the picture is[14]

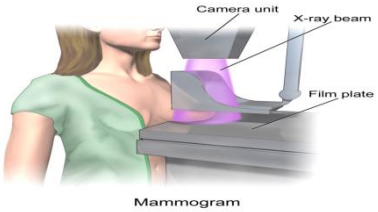


Figure 6: Mammogram

CHAPTER 2

LITERATURE REVIEW

R. Ramani *et a*l. worked on pre-processing and filtering. He has observed that doctors and radiologistics due to their inexperience in the cancer detection field can miss the abnormalities. He found that images are not of good quality and their capture quality are also very poor. He also found that there are noise present in the images. So to overcome the problem of image quality and to remove the noise in the images ,he researched on the pre-processing techniques. He worked on different type of filtering techniques like median, adaptive median , mean filtering etc.

Dr.Narain Ponraj *et a*l. has observed that cancer is due to abnormal growth of breast cells .It occur both in men and women and found that the abnormalities in the breast are either micro-calcifiaction and masses. He has observed that there there are two regions in mammogram. First is exposed breast region and second is unexposed breast region. He found that identification of breast region is necessary and removing of non-exposed breast region is also necessary. The problem of this solution is Pre-processing techniques.

Armen Sahakyan *et al.* found that for efficient and accurate breast region segmentation still is the problem in the mammography. He has proposed an automatic technique for mammogram segmentation. That technique uses morphological pre-processing algorithm .The use of this algorithm is to remove digitization noise and separate background region from breast profile region.

Monika Sharma *et al*. has observed the pattern recognition problem in the image processing. The solution of this problem is feature extraction because the accuracy of classification depends on feature extraction. She has extracted textural, structural and statistical features.

Pawar B.V. *et al* .has presented a method of detection tumor in breast. He has given a method to detect the breast tumor and that method is pixel based mass detection and has used the template matching procedure. He has also used median filtering to enhance the mammogram images before template matching procedure. He has observed that high pass filtering is used to enhance edge and edge detection method is used to detect edges and also found that only circular shape early stage tumor is detected.

Dr. Annopa Maria Sabu *et al*. has presented a paper on mass detection and microcalcification in mammography. He has found the result for the detection of these by using textural analysis that characterize spatial variation within the image by extracting information and also obtained the textural features obtained.

Aswini Kumar Mohanty *et al*. has presented a paper for the classification of Benign and malignant breast cancer. He has observed that only 15-30% of masses referred for surgical biopsy are actually malignant. The solution of this is to propose an approach to develop a computer-aided classification system for cancer detection from digital mammogram. This proposed system has three steps: first is ROI extraction, second is feature extraction in Which GLCM and GLRLM(Gray level Run Length Matrix) features are extracted which can distinguish benign and malignant cancer. The third step is classification process in which association Rule mining classifier is used.

Robert. M. Harallick *et al.* has described some easily computable textural features based on gray tone spatial dependencies and also illustrated their applications in category identification tasks for three different types of data photomicrograph of five kind of stand stones, panchromatic aerial photograph and ERTS multispectral images. He has given the solution that textural features are the solutions for wide variety of image classification applications.

Pradeep N *et al.* has found the problem of uncontrolled growth of breast cells and the most common abnormalities are masses and micro-calcification. The solution of this problem is the secondary tool, CAD(Computer Aided Detection) System to radiologists for segmentation and features extraction for the diagnosis of breast cancer.

P.Mohanaiah *et al.* has observed that primitive or low level image features can either be generic features or domain specific features. He has given an approach called GLCM approach to extract second order statistical textural features for motion estimation of images.

CHAPTER 3
METHODOLOGY

In this proposed method,we have to search the benign and malignant breast cancer. For detecting that , various methods are applied for that:

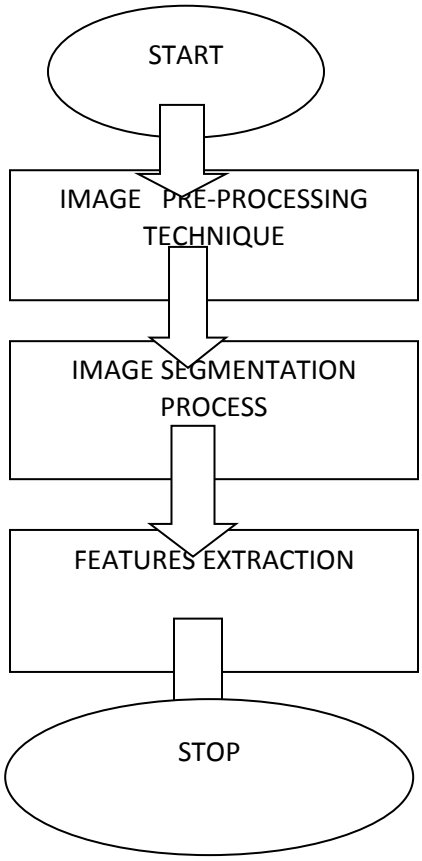
1. Pre-processing Method
2. Image Segmentation method
3. GLCM Features

In the Pre-processing method, noise reduction method is applied in which cropping and median filtering processes are done. Cropping eliminates the written labels. Median filtering isused to remove the noise like ‘salt and pepper’. After that Image Enhancement process is done by using histogram equalization and thresholding method. In histogram equalization method, CLAHE method is applied .CLAHE method improves the contrast of the image as compared to other histogram equalization methods. Global thresholding method is applied. After applying thresholding method, cancerous image is detected. Now Image segmentation process is applied by using edge detection method with Sobel operator. Other operators i.e Prewitt operator can also apply but Sobel gives good result as compared to Prewitt. Segmentation is used foe segmentation of the images. Edge Detection Method of the segmentation is used to detect the edges of the images.

GLCM features are extracted like entropy, energy, sum of variance etc in 4 directions i.e. 0,45,90 and 135 degree. GLCM stands for Gray Level Co-Occurrence Matrix. GLCM is also called Gray tone Spatial Ddependency Matrix. With the GLCM featuers benign and malignant cancer is detected. There is one more term GLRLM,Gray Level Run Length Matrices that is used to find the benign and malignant cancer. If combine both,GLCM and GLRLM the benign and malignant result would be most effective.

In this work, contrast, correlation, energy and homogeneity are extracted. All these feature values of different patients are shown in various tables. There are other features which can be extracted from GLCM i.e. sum variance, Difference Variance, Sum of Squares, sum average, Inverse Difference Moment (IDM) Entropy, Standard Deviation etc. These all features are helpful to determine Malignant and Benign Cancer.

3.0 Flow Chart Methodology



3.1 Pre-Processing

Pre-processing technique is very important technique before image segmentation process. This technique improves the quality of the image by reducing the noise and enhancing the image.It involves two processes:

1. Noise reduction
2. Image enhancement

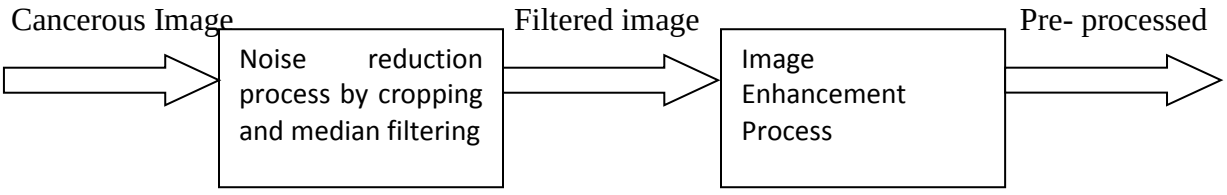


Figure 7: Pre-Processing Technique

3.1.1 Noise reduction process

This process is very important before the processing of the mammogram images because noise reduction process removes noise from the images. These noise can be artifacts like written labels, high intensity rectangular label, low intensity label etc[15]. To reduce these noise , different filtering processes are used.

3.1.1.1 Cropping

For eliminating the written labels in the images, cropping process has to be used[16].

3.1.1.2 Adaptive median filtering

This filter is used where rectangular region Sxy occurs. During filtering operation, it changes the size of rectangular sections and that depends on certain conditions. There is window surrounding each pixel that is changeable and the variation depends upon the median of pixels in present window. The size of window becomes larger if median value is impulse. In other respects within the window specification, processing is done.On the image, processing is done in the following way:

If center pixel of the image is not impulse then the value of the centre pixel occupies in filtered image and if center pixel value is impulse, then the new value of that pixel in the filtered image will be the median value of the pixels in that window. If centre pixel is impulse then gray scale pixel of the filtered image is identical of the input image[17] .Zeros takes the place of edge of the images. The filter's value is single value and this value replaces current pixel value at (x, y) and S is centered on this mark. The system of symbols which are used are

1. Zmin = minimum pixel value in Sxy
2. Zmax = maximum pixel value in Sxy
3. Zmed = median pixel value in Sxy
4. Zxy= pixel value at coordinates (x, y)
5. Smax = maximum allowed size of Sxy[14]

Adaptive filtering does two tasks

1. Separate impulse noise from the image
2. Separate distortion from the image[17]

3.1.1.3 Median filter

Median filtering is used to separate impulse noise such as 'salt &pepper noise' from the image. In salt & pepper noise, at two ends of image pixel intensities, impulse values exists. The impulses with high pixel intensities called salt noise and the impulses with low pixel intensities are called pepper noise[18].Median filtering is also used to remove all background noise[30].

3.1.1.4 Mean Filter

This filter is the spatial filter that exchanges the center value of the pixel with average values of the pixels in the same window. It is not used in the case of mammography.

Adaptive mean filtering and mean filtering are not used in this images[15].

3.2 Image Enhancement

Image enhancement process is used to make the quality of the image better. This technique has two categories:

1. Spatial domain method
2. Frequency domain method

Spatial domain method has characteristics to work on pixels directly whereas in frequency domain method , fourier transform of an image is captured then frequency domain method is applied on fourier transformed image.

3.2.1 Mammography Image enhancement techniques

Image enhancement technique is very significant before image segmentation. Scanner creates artifacts, excessive background noise, scratches and dust artifacts ,these were the troubles with image acquisition and they could impact the reliability of the algorithm. The mammogram figure has highly non uniform background and the area higher than the breast tissue area having small-scale contrast. That's why image enhancement is necessary before image segmentation.

3.2.1.1 Histogram Equalization

Due to summation operation, the normal and adaptive histogram equalization may over intensify the noises and needle-like region in noises. In the histogram of nearly steady areas in the original image, it gives the large values in the enhanced image, an image enhancement technique was selected, CLAHE(Contrast Limited Adaptive Histogram Equalization)[19].

3.2.1.2 Eliminate background noise

In the mammogram images, there are radio non-transport artifacts such as wedges and label, in order to reduce them, thresholding technique and morphological operations are used. Image segmentation's simplest method is thresholding. Thresholding is used to create binary images from gray scale and coloured images[20]

The algorithm for finding artifacts and labels and to separate breast profile is as follows:

1. Gray scale image is transformed into binary image by thresholding.
2. Binary image objects are tagged and number of pixels in the image are calculated.
3. The breast profile is the tremendous one other than that all binary image objects are removed from noise. After that morphological operation is applied, removing isolated pixels.
4. The binary image is then multiplied by original image and earn the final image without artifacts[20,19]

3.2.1.3 Thresholding

There are two types of thresholding basically

1. Global thresholding
2. Local thresholding

Global thrsholding methods can fail when the background light is uneven. A common action in such situations is to preprocess the image to repay for lightening problems and then register global threshold to preprocessed image.

G(x,y)=1 if f(x,y)>=T(x,y)

G(x,y)=0 if f(x,y)<T(x,y)

Where T(x,y) is varying thresholding function

T(x,y)=f0(x,y)+T0 where f0(x,y) is the morphological opening of f and constant T0 is the result of function graythresh applied to f0.

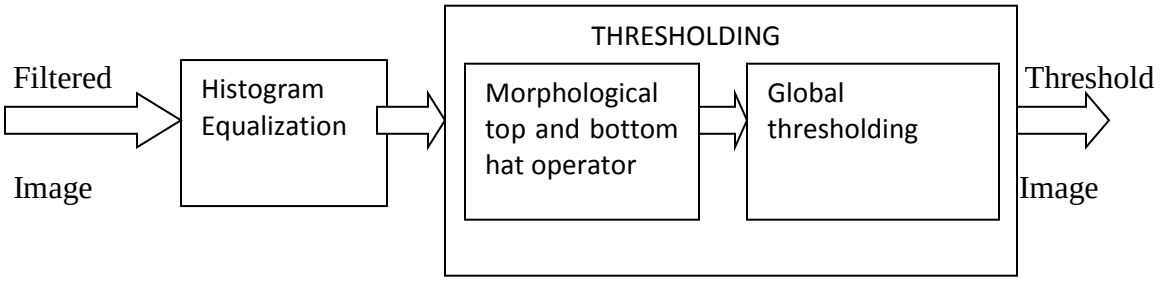


Figure 8: Image Enhancement Block Diagram

3.3 Image Segmentation

Image segmentation is the operation which subdivides an image into integral areas. Mammogram segmentation algorithm can be categorized according to regions to be segmented:

1. Breast region segmentation
2. Region of interest segmentation(ROI)

In the **breast region segmentation** process, the mammogram image is fragmented into breast regions and background for focusing and restricted the search for irregularities on the breast region and having no reaction on background result for superior detection[18]

Region of interest segmentation:In this process, segmentation of suscpious region to be examine for irregularities[18]

3.3.1 Types of Image Segmentation

Basically there are three types of image segmentation:

1. Point detection
2. Line detection
3. Edge detection

   In the mammography, Edge detection method is used to find the edges of the mammogram image.

3.3.1.1 Edge Detection Method

Boundary between two areas with relationally well-defined gray level properties[20].So to detect tumor ,extraction of image edges are necessary.

Different operators are used for edge detection and the operators are

1 Prewitt operator

2 Sobel edge operator

3 Kirch operator

But Sobel operator is used to detect the edges in mammography because it gives superior (more sharp and clear) edges as compared to other operators[20].

Sobel Edge Detector

Sobel operator has same purpose as prewitt operator but sobel operator finds the edge detection wider. The images with lot of noises and having gray gradient well, only these images can process with sobel operation.

Gx=(z7+2z8+z9)-(z1+2z2+z3)

Gy=(z3+2z6+z9)-(z1+2z4+z7)

g=[Gx^2+Gy^2]^1/2

The direction of this point is

b=1/tan(Gx/Gy)

Gx

-1	-2	-1
0	0	0
1	2	1

Gy

-1	0	1
-2	0	2
-1	0	1

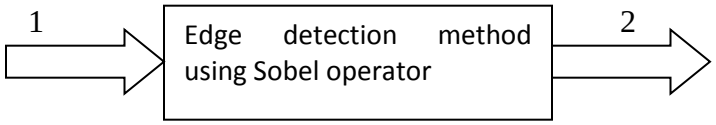


Figure 9: Image Segmentation Process Block Diagram

1 - Pre-processed Image

2 - Segmented Image

3.3.1.2 Filtering

After image segmentation, still there are noise in the images which will affect the feature extraction stage.This noise is detached by any type of filtering method. The apt method can be applied. These noises can be easily seen by sectioned output. If there are small white spots in the segmented output that means noises are there. These noises are filling by their intensity values by 0.There are different types of filtering method

1. Median filtering
2. Adaptive filtering
3. Mean filtering
4. Adaptive mean filtering

3.4 Feature Extraction

Features extraction gives the information of the image by observing patterns. It is relevant for achievement of relevant task related to correlate application. After the segmented output , image feature extraction method is very important to extract the features[21].Mainly there are three types of feature extraction methods:

1. Geometric feature method

2. Texture feature extraction method
3. Gradient feature extraction method

3.4.1 Geometric Features

Geometric features describe the geometric properties of ROI. It is the collection of pixels in the images. The basic characteristics of Geometric Features are:

1 area

2 perimeter

3 compactness

In medical diagnose, these features are essential to recognize any object regardless of breast.

3.4.2 Gradient Features

The gradient image is the value of the local image.an edge in the image corresponds to the higher pixel value in the gradient image and sobel operation is used in this[21].

3.4.3 Texture Features

Texture features are used to classify the image, detection of an image and segmentation of the image. These feature can be extracted into:

1 Statistical features

2 Structural features

The texture feature's extraction method plays a very important role to detect the mammogram abnormalities. Texture features have been proven to be useful to isolate normal and abnormal lesions with masses and micro calcification[21]

Statistical texture features are used for early detection of masses in mammogram. GLCM is the statistical texture measure.

GLRLM is the another statistical texture measure fully named Gray Run Length Matrix.

3.4.3.1 GLCM(Gray Level Co-occurrence Matrix)

It is of 2^{nd} order statistics, so information about pixels of pairs are collected by GLCM.GLCM shows how the pixel brightness in an image occurs. So it is called the tabulation of frequencies. At a distance of d=1 and for direction of theta=0,45,90,135(all in degree),a matrix is constructed[35].GLCM contains rows and columns which are equal

to number of Gray levels G, in the image. P(I,j/del x del y) is the pixel element which shows the relative frequency with which two pixels separated by del x del y which occur within the neighbourhood , one with intensity I and other with intensity j.

The matrix element P(I,j/d,theta) has the second order statistical probabilities values for changing between gray levels ‘i’ and ‘j’ at particular distance displacement d and particular angle ‘theta’.[22]

GLCM Is also called Gray tone spatial dependency matrix.A reliable texture information might not be given by single direction. So four directions are used to extract information. To classify ROI as masses or non masses, texture feature is used. Textural features are

3.4.3.2 GLRLM(Gray Level Run –Length Matrix)

GLRLM is in the form of matrix which can be used to extract the the features.In this run length means the numer of adjacent pixels in the particular direction that have same length. It is 2D matrix in which every matrix is of 2D having each element p(i,j\theta) is the number of elements j with intensity i,in the direction.

Galloway suggest 5 features from GLRLM and these are

1. SRE(Short Run Emphasis)
2. LRE(Long Run Emphasis)
3. GLN(Gray Level Non uniformity)
4. RLN(Run Length Non –uniformity)
5. Run Percentage(RP)[21]

Table 1: GLCM Texture features with Formulas

Sr.No	Texture Features of GLCM	Formulas
1	Correlation	$\frac{\sum_{i,j=1}^{N}(i-\mu i)(j-\mu i)}{\sigma i \sigma j}$
2	Energy	$\sum_{i,j=1}^{N} p(i,j)^2$
3	Homogeneity	$\frac{\sum_{i,j=1}^{N)} p(i,j)}{1+\lvert i-j \rvert}$
4	Contrast	$\sum_{i,j} \lvert i-j \rvert p(i,j)$

This table shows the four features of GLCM which are applied to detect the Benign and Malignant cancer. Correlation means similiarity. It shows the linear dependency of grey levels on those of neighbouring pixels.

In Contrast case, if I and j are equal, the weight is zero and having no contrast. If i and j values are differ by 1,the weight 1 and having little contrast. The weight continues to increase exponentially as (i-j) increases.

Homogeneity is also called "inverse difference moment". The weight in the homogeneity formula is 1\[1+(i-j)^2]. It increases with less contrast.

Energy is the square root of angular second moment (ASM).Energy is opposite of entropy. It measures order in an image,it is applied in texture.

CHAPTER 4

RESULTS

Malignant and Benign are the tumors of mammogram.There are two standard databases for mammogram DDSM and MIAS .For detecting the malignant and benign tumor of mammogram,some steps are followed and the steps are:

STEP 1: Firstly any database breast image has taken whether it is DDSM based or MIAS based.

STEP 2: For removing written labels and taking the exact size of image, Crop the image as you want. Crop the needed part of the image and remove the unnecessary part.

S**TEP 3**: For removing the noise, Median Filtering is used. It is the best-known order statistics filter in digital image processing, which corresponds to 50^{th} percentile.

STEP 4: Histogram Equalization is needed to improve the contrast of the image. Here CLAHE histogram equalization is used because it gives better contrast ratio as compared to other equalization techniques.

STEP 5: Thresholding is very important part of this procedure. It is used to extract the image from the background. Global Thresholding is used for the detection of exact cancerous area.

STEP 7: Use the segmentation process if you need. Edge Detection Process is used using Sobel operator because Sobel operator gives more sharp and clear images as compared to other operator.

STEP 8: For Features Extraction ,GLCM texture features are used for images of different patients at four different angles(0,45,90,135 degree).

4.1: A Database Mammogram Image(PN1)

STEP 1: A MIAS breast image is taken

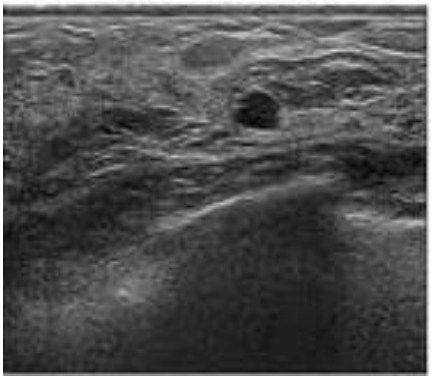


Figure 10(a)

STEP 2: Convert this image into gray scale image

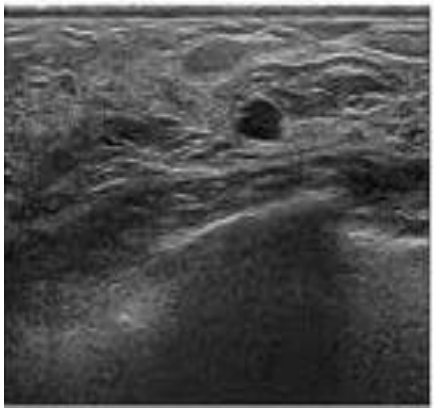


Figure 10(b)

STEP 3: To remove the noise like written labels,cropping is done

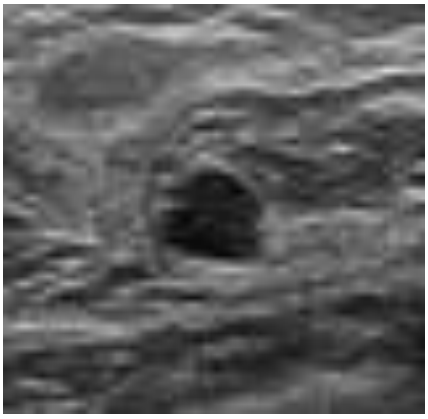


Figure 10(c)

STEP 4: To remove the noise from the breast image,Median filtering is used.

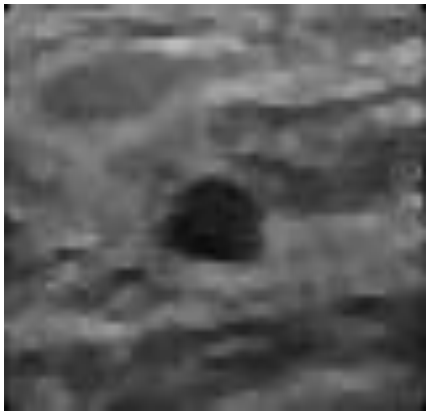


Figure 10(d)

STEP 5: Histogram Equalization is done by using CLAHE method.

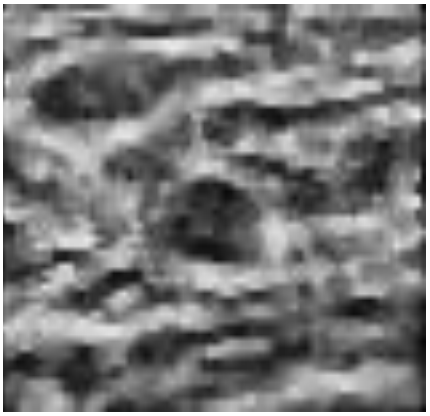


Figure 10(e)

STEP 7: Global thresholding is used to detect the cancerous part.

Figure 10(f)

STEP 8: Edge detection method with Sobel operator is used for Segmentation.

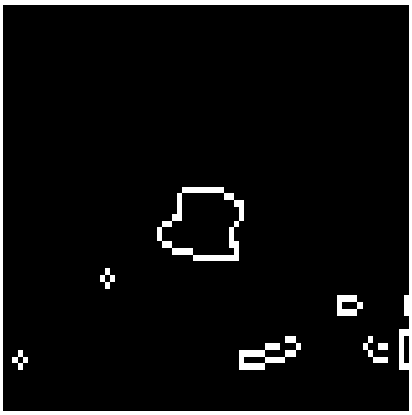
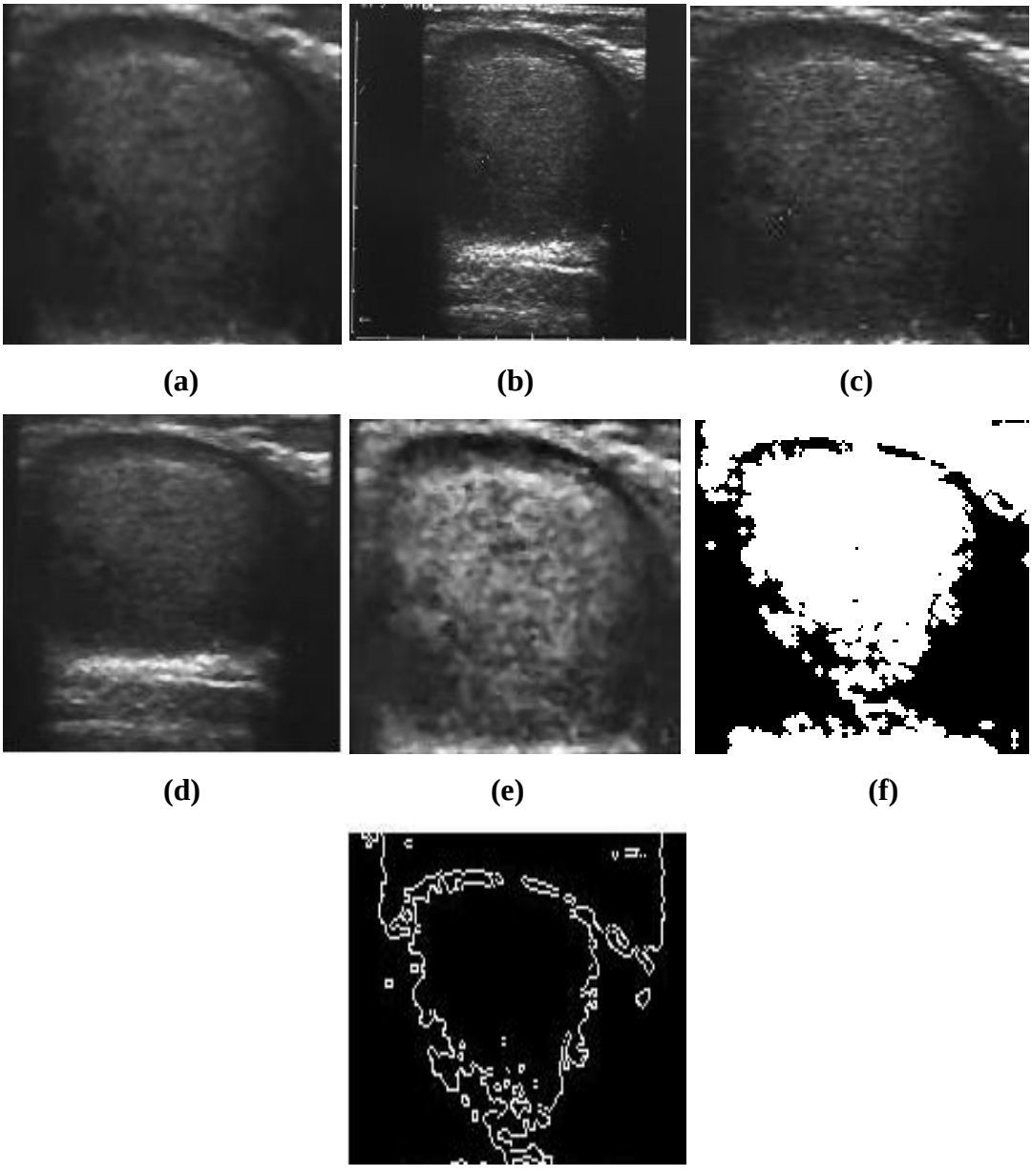


Figure 10(g)

Here Fig 10(a) shows the original image and Fig 10(b) shows the gray scale image.If there is any noise like artifacts it is removed by cropping.Fig 10(c) shows the cropped image,Fig 10(d) and Fig 10(e) shows the contrast and cancerous image (black colour) respectively.Here for better contrast CLAHE is used.Thresholding is used to detect the malignant and benign cancerous area.Fig 10(f) shows the segmented image which got by using Edge Detection Segmentation process.

4.2 A Database Breast Cancer Image(PN2)



(a) **(b)** **(c)**

(d) **(e)** **(f)**

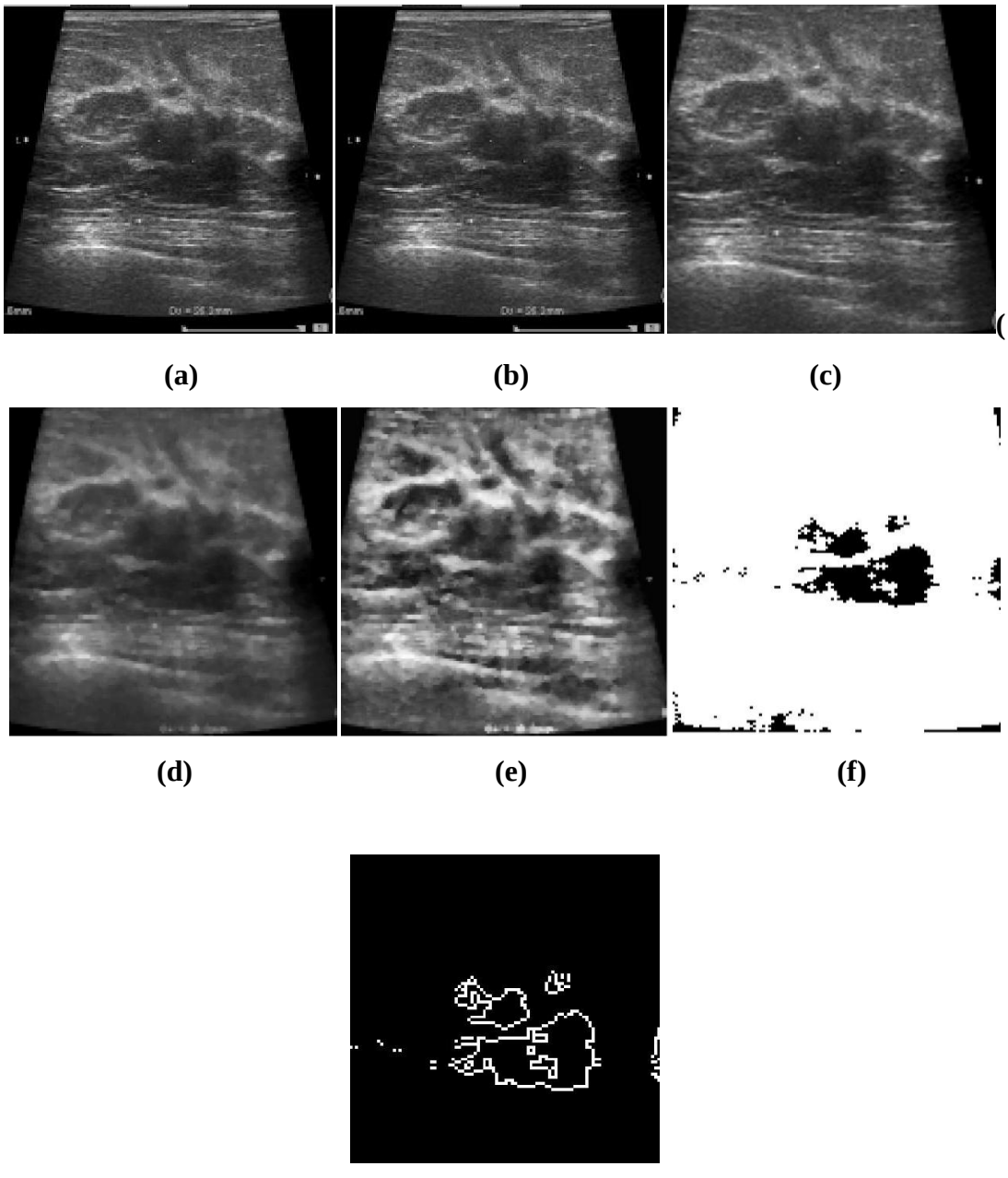
(g)

Figure 11 (a,b,c,d,e,f)

Above figures show the original, gray scale, cropped, filtered, contrast and cancerous image respectively and Original and cancerous image's GLCM features values at angles(0,45,90,135) degrees.

Cancer part of the image shown in white colour in Fig 11(f).

4.3 A Database Breast Cancer image(PN3)



(a) **(b)** **(c)**

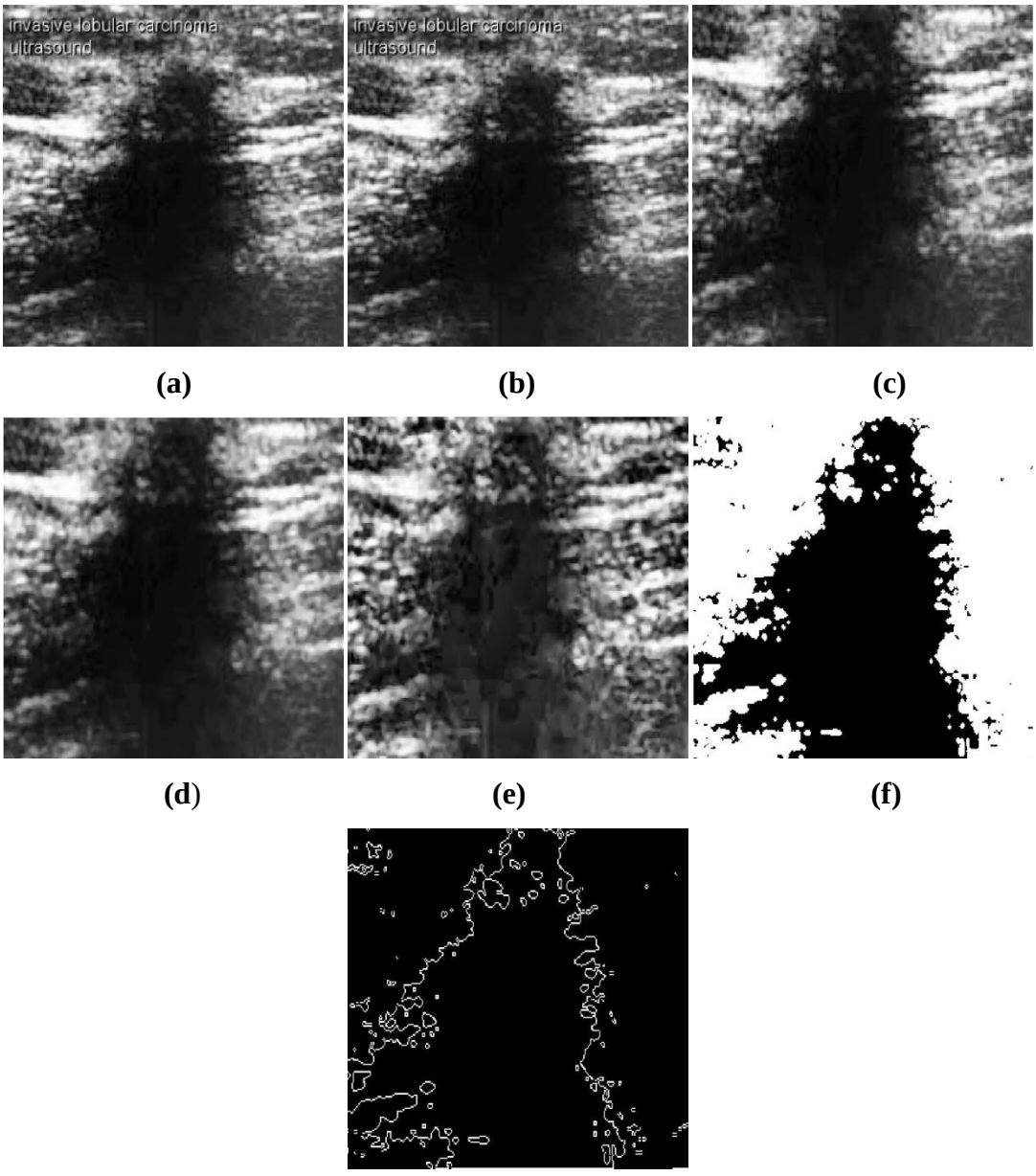
(d) **(e)** **(f)**

(g)

Figure 12 (a,b,c,d,e,f)

The above figure shows the original image, gray scale image, cropped image ,filtered image, contrast, cancerous and segmented image respectively and GLCM features of Original and Cancerous image at various angles(0,45,90,135) degrees respectively in Table 2.

4.4 A Database Mammogram image(PN4)



(a) **(b)** **(c)**

(d) **(e)** **(f)**

(g)

Fig 13 (a,b,c,d,e,f)

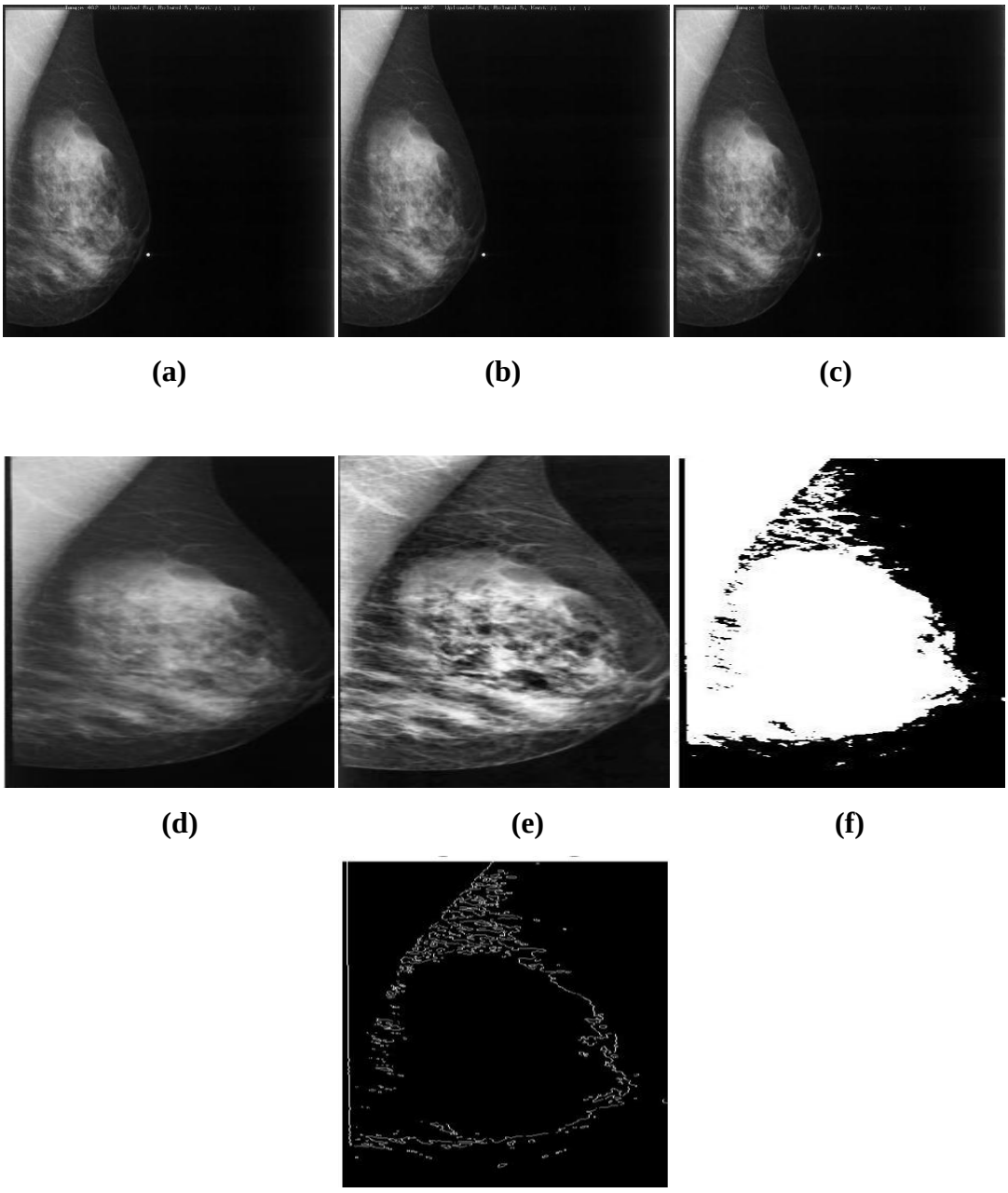
The above figure shows the original,gray scale,croppes,filtered , contrast, cancerous and segmented image respectively and GLCM features of Original and Cancerous image at various angles(0,45,90,135) degrees respectively in Table2.

Table 2:GLCM Texture Feature Values

Texture Features	PN1 OIV	PN1 CIV	PN2 OIV	PN2 CIV	PN3 OIV	PN3 CIV	PN4 OIV	PN4 CIV
Offset[0 1]								
Contrast	0.4112	0.4859	.4112	.5129	.2079	1.8167	.5984	1.9665
Correlation	0.9494	0.842	.9494	.8748	.9796	.9253	.8429	0.8192
Energy	0.2473	0.9275	.2473	.9060	.2074	.4682	.2002	.4647
Homogeneity	0.9423	0.9913	.9423	.9908	.9505	.9676	.8760	.9649
Offset[-1 1]								
Contrast	0.983	0.9856	.9830	1.1053	.2421	2.7378	1.0200	2.1514
Correlation	0.8727	0.7004	.8727	0.7347	.9762	.8874	.8788	.8322
Energy	0.2318	0.9176	.2318	.8829	.1977	.4511	.1713	.4466
Homogeneity	0.8724	0.9833	.8724	.9803	.9336	.9511	.8044	.9473
Offset[-1 0]								
Contrast	0.6646	0.4983	.6646	1.0859	.0707	2.3431	.7470	2.5710
Correlation	0.9155	0.8781	.9155	.7349	.9931	.9036	.8944	.8040
Energy	0.2353	0.4065	.2353	.8947	.2136	.4584	.1795	.4535
Homogeneity	0.886	0.9911	.8860	.9506	.9647	.9582	.8204	.9541
Offset[-1-1]								
Contrast	0.9919	1.0906	.9919	1.1974	.2232	2.7491	1.0110	2.9852
Correlation	0.8716	0.7375	.8716	.7418	.9782	.8869	.7350	.7274
Energy	0.2316	0.8935	.2316	.8911	.2025	.4509	.1709	.4460
Homogeneity	0.8716	0.9805	.8716	.9786	.9429	.9509	.8038	.9467

This Table shows that the GLCM features (Contrast,Energy,Homogeneity) except Correlation feature at different 4 angles(0.45,90,135 degrees) have more values in case of Original Images as compared to their Cancerous Image feature values.

4.5 A Databse Mammogram Image(PN5)



(a) **(b)** **(c)**

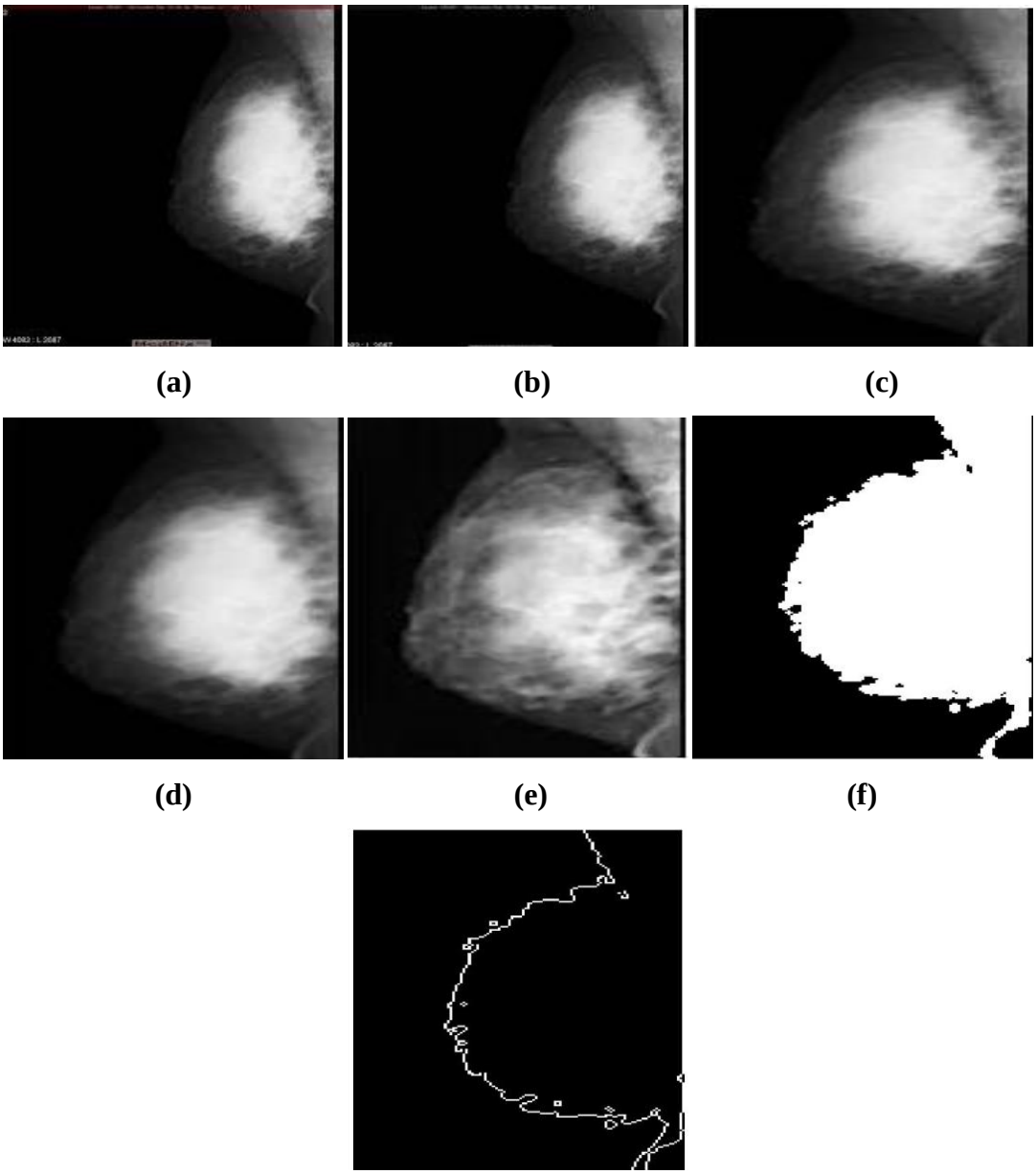
(d) **(e)** **(f)**

(g)

Figure 14 (a,b,c,d,e,f,g)

The above figure shows the original, gray scale, cropped, filtered , contrast, cancerous and segmented image respectively and GLCM features of Original and Cancerous image at various angles(0,45,90,135) degrees respectively in Table 3.

4.6 A Databse Mammogram Image(PN6)



(a) **(b)** **(c)**

(d) **(e)** **(f)**

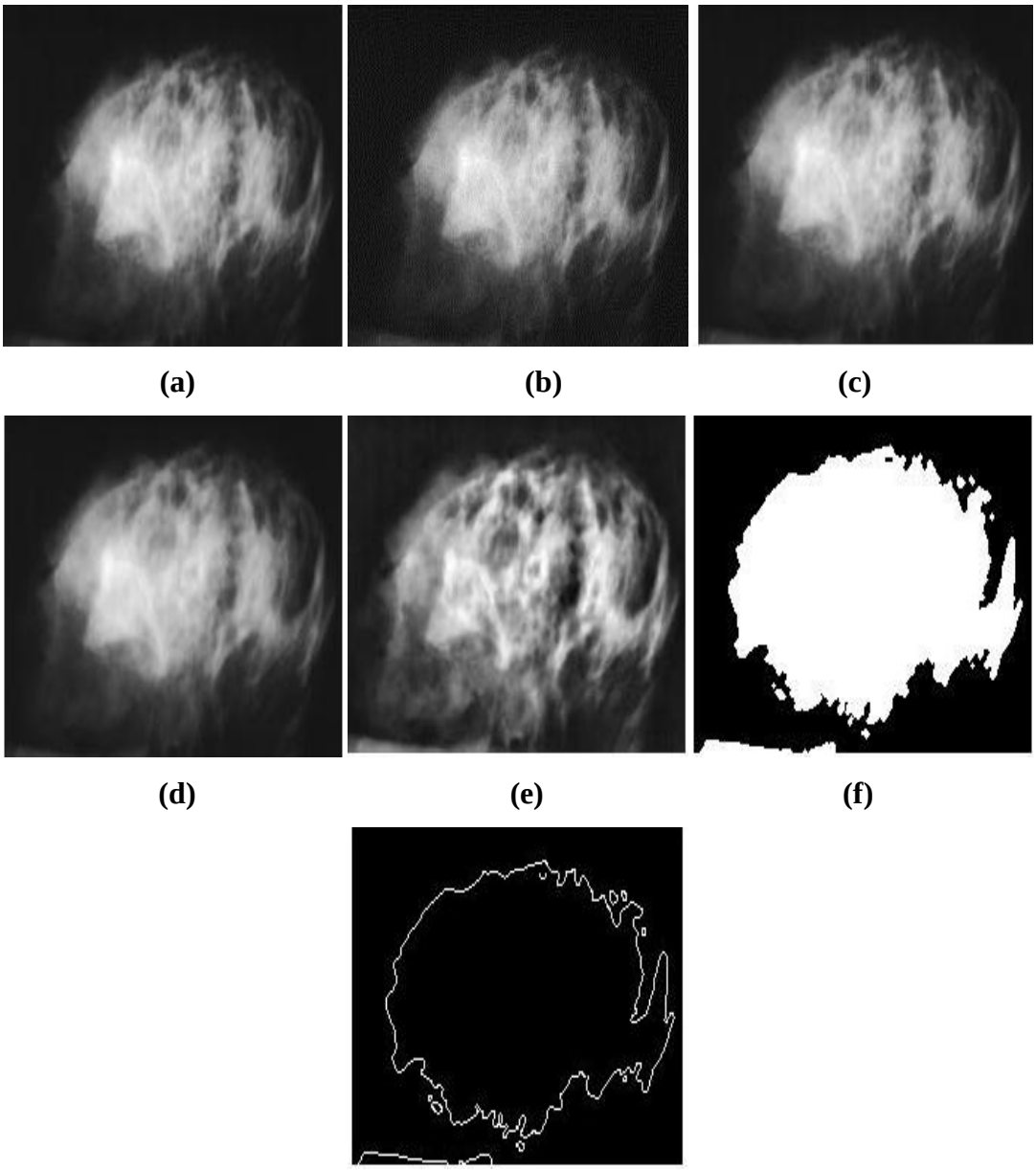
(g)

Figure 15 (a,b,c,d,e,f,g)

The above figure shows the original,gray scale,croppes,filtered , contrast, cancerous and segmented image respectively and GLCM features of Original and Cancerous image at various angles(0,45,90,135) degrees respectively in Table 3.

In Histogram Equalization image Fig15(e), the good contrast of the image shows,the Cancerous part shown in white colour in Fig 15(f) and its segmented part shown in Fig 15(g).

4.7 A Databse Mammogram Image(PN7)



(a) **(b)** **(c)**

(d) **(e)** **(f)**

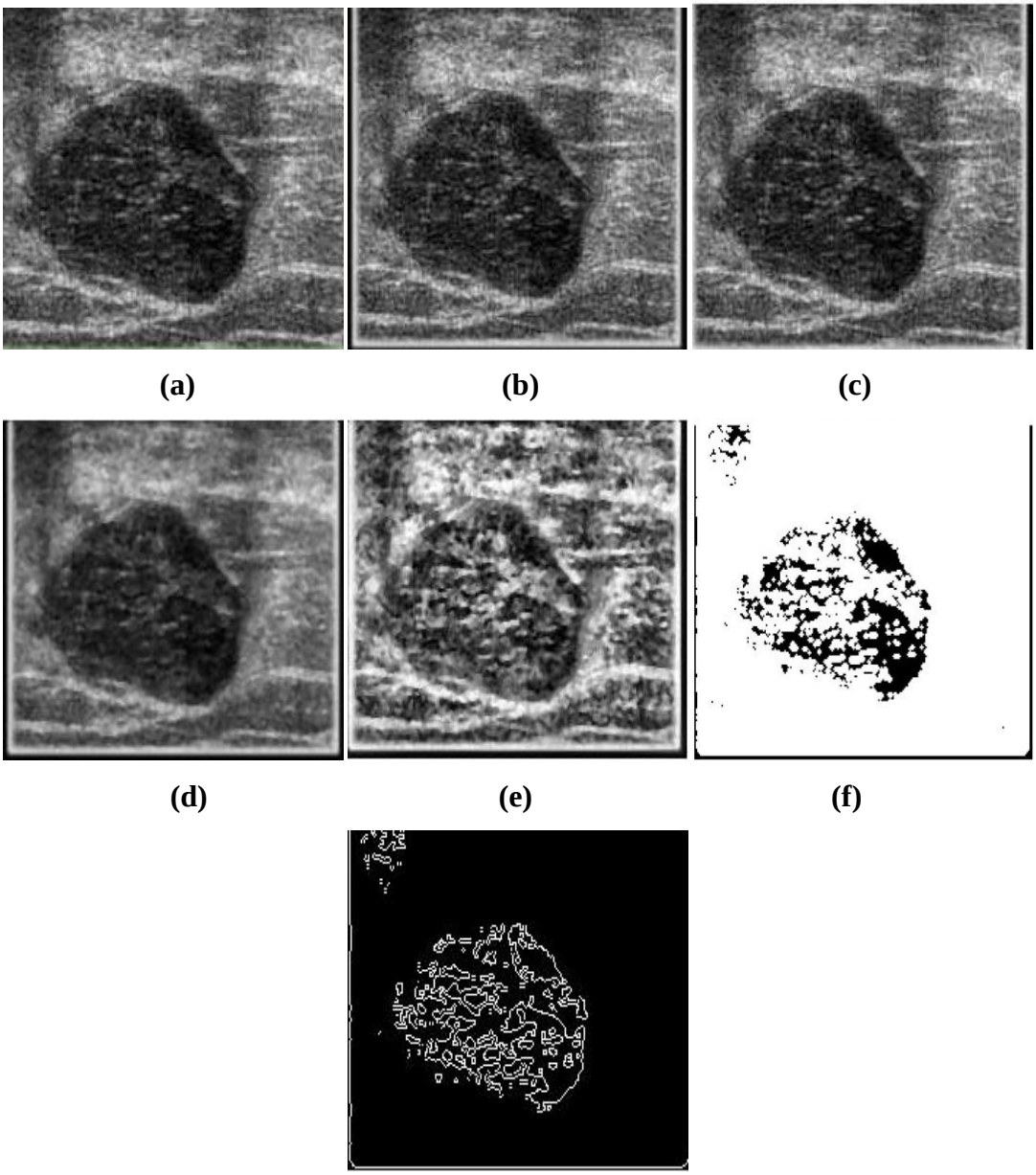
(g)

Figure 16 (a,b,c,d,e,f,g)

The above figure shows the original,gray scale,cropped,filtering,contrast,cancerous image and segmented image respectively.Here Cancerous image shown in Fig 16 (f).GLCM features of cancerous and original images are shown in Table 3.

In Fig 16(e) contrast image is shown which clearifies the cancerous image part and by using threshold, the cancerous part is detected which is shown in white colour in Fig 14(f) and its segmented image is shown in Fig 16(g).

4.8 A Databse Mammogram Image(PN8)



(a) **(b)** **(c)**

(d) **(e)** **(f)**

(g)

Figure 17 (a,b,c,d,e,f,g)

The above figure shows the original, gray scale, cropped, filtering,contrast,cancerous image and segmented image respectively.Here Cancerous image shown in Fig 20(f).GLCM features of cancerous and original images are shown in Table 3

Table 3: GLCM Texture Features Values

Texture Features	PN5 OIV	PN5 CIV	PN6 OIV	PN6 CIV	PN7 OIV	PN7 CIV	PN8 OIV	PN8 CIV
Offset[0 1]								
Contrast	.5195	2.2578	.1200	.7626	.1093	1.0467	.0692	.9315
Correlation	.8803	.7738	.9685	.9588	.9829	.9569	.9867	.9620
Energy	.0973	.7523	.4099	.4865	.4499	.4835	.3747	.4816
Homogeneity	.8374	.9597	.9756	.9864	.9717	.9813	.9654	.9834
Offset[-1 1]								
Contrast	.9435	3.2000	.1832	1.0501	.1422	1.1304	.0845	1.1065
Correlation	.7806	.6741	.9519	.9570	.9778	.9535	.9838	.9548
Energy	.0674	.7388	.4056	.4809	.4452	.4816	.3690	.4781
Homogeneity	.7343	.9429	.9665	.9812	.9627	.9798	.9580	.9802
Offset[-1 0]								
Contrast	.6730	2.7003	.1111	.7403	.0730	.5996	.0631	.9404
Correlation	.8448	.7327	.9709	.9697	.9886	.9753	.9879	.9616
Energy	.0769	.7420	.4092	.4868	.4547	.4927	.3756	.4815
Homogeneity	.7787	.9518	.9763	.9868	.9752	.9893	.9684	.9832
Offset[-1-1]	.							
Contrast	.9450	3.1794	.1794	.9676	.1428	1.2481	.0995	1.3849
Correlation	.7803	.6761	.9529	.9604	.9778	.9486	.9809	.9435
Energy	.0675	.7391	.4059	.4825	.4455	.4793	.3652	.4727
Homogeneity	.7358	.9432	.9680	.9827	.9626	.9777	.9507	.9753

This Table shows that the GLCM features(Contrast,Energy,Homogeneity) except Correlation feature,at different 4 angles(0.45,90,135 degrees),having values more in case of Original images as compared to their Cancerous Image feature values.

4.9 :A Database Mammogram Image(PN9)

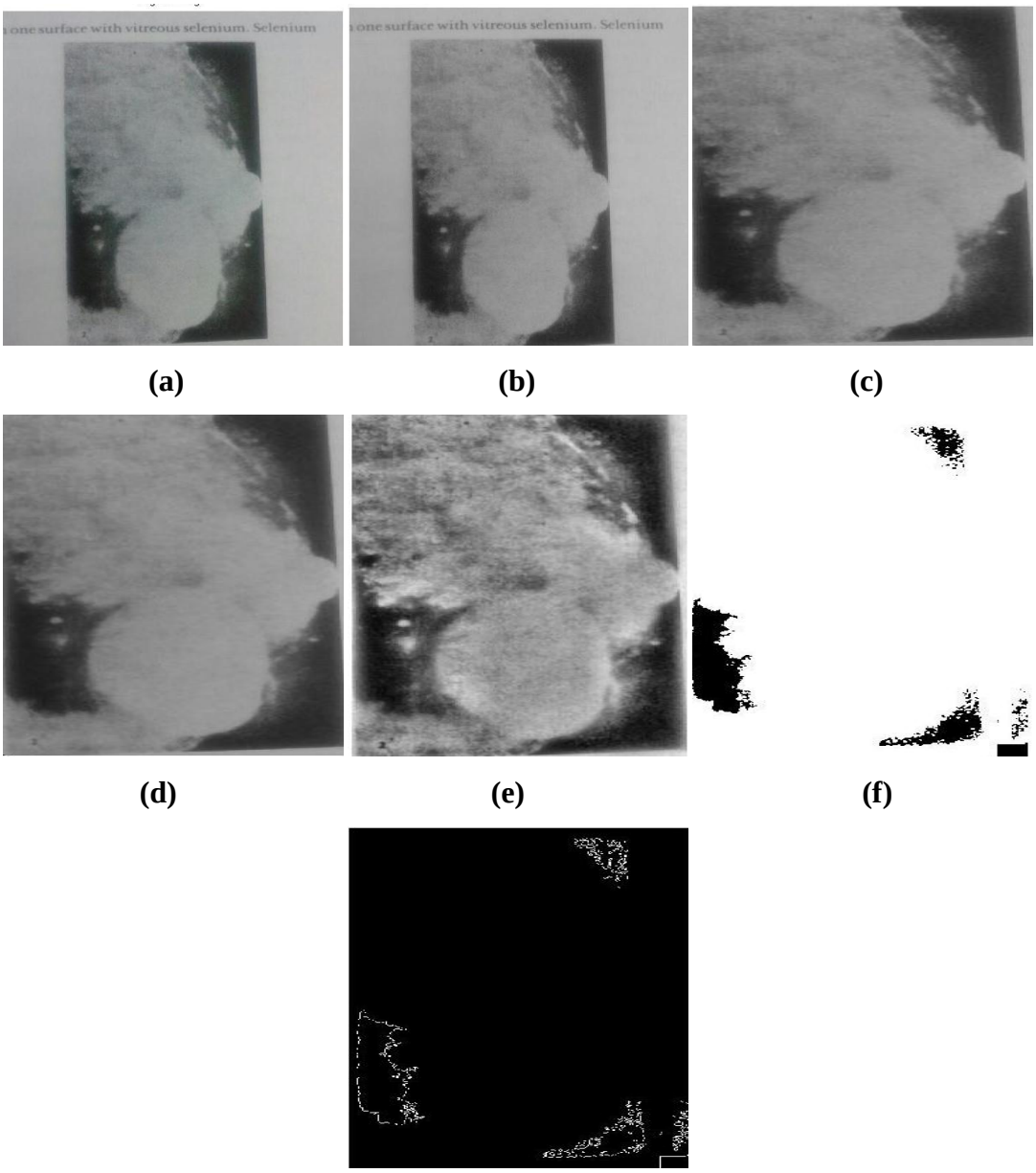
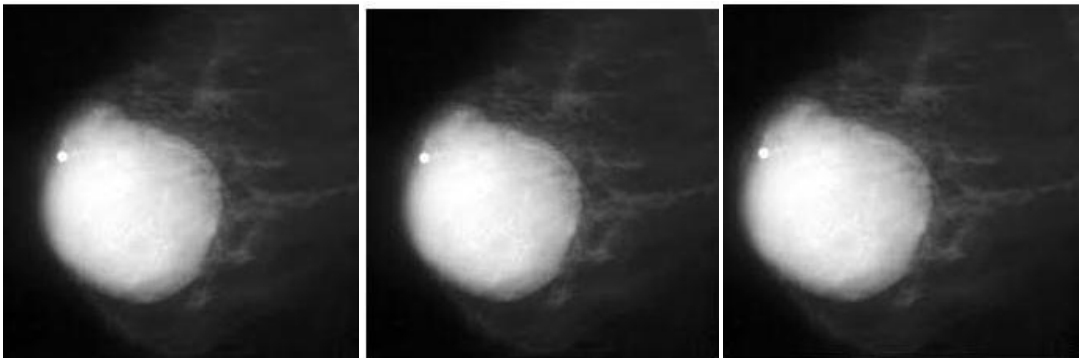


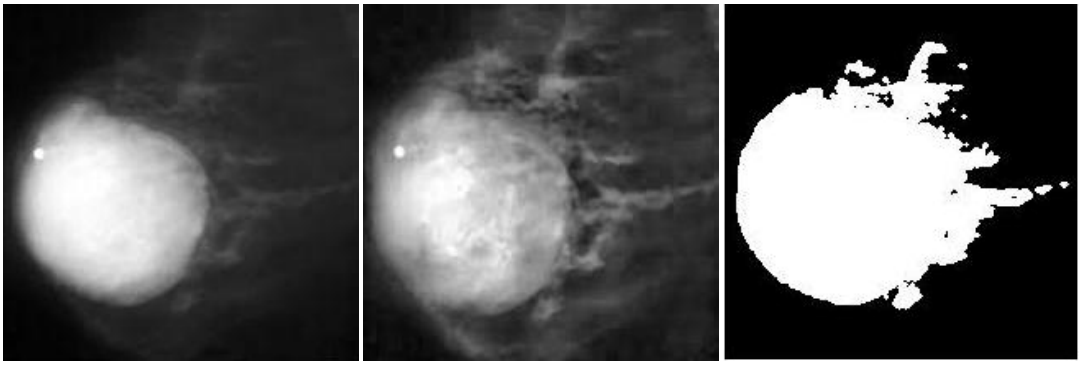
Figure 18 (a,b,c,d,e,f,g)

The above figure shows the original,gray scale,cropped,filtering,contrast,cancerous image and segmented image respectively.Here in Fig 18(f) shows the cancerous image which is white in colour and its segmented image is shown in Fig 18(g).

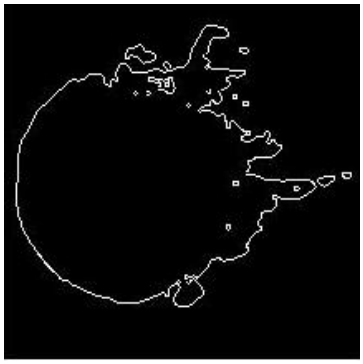
4.10: A Database Mammogram Image(PN 10)



(a) **(b)** **(c)**



(d) **(e)** **(f)**

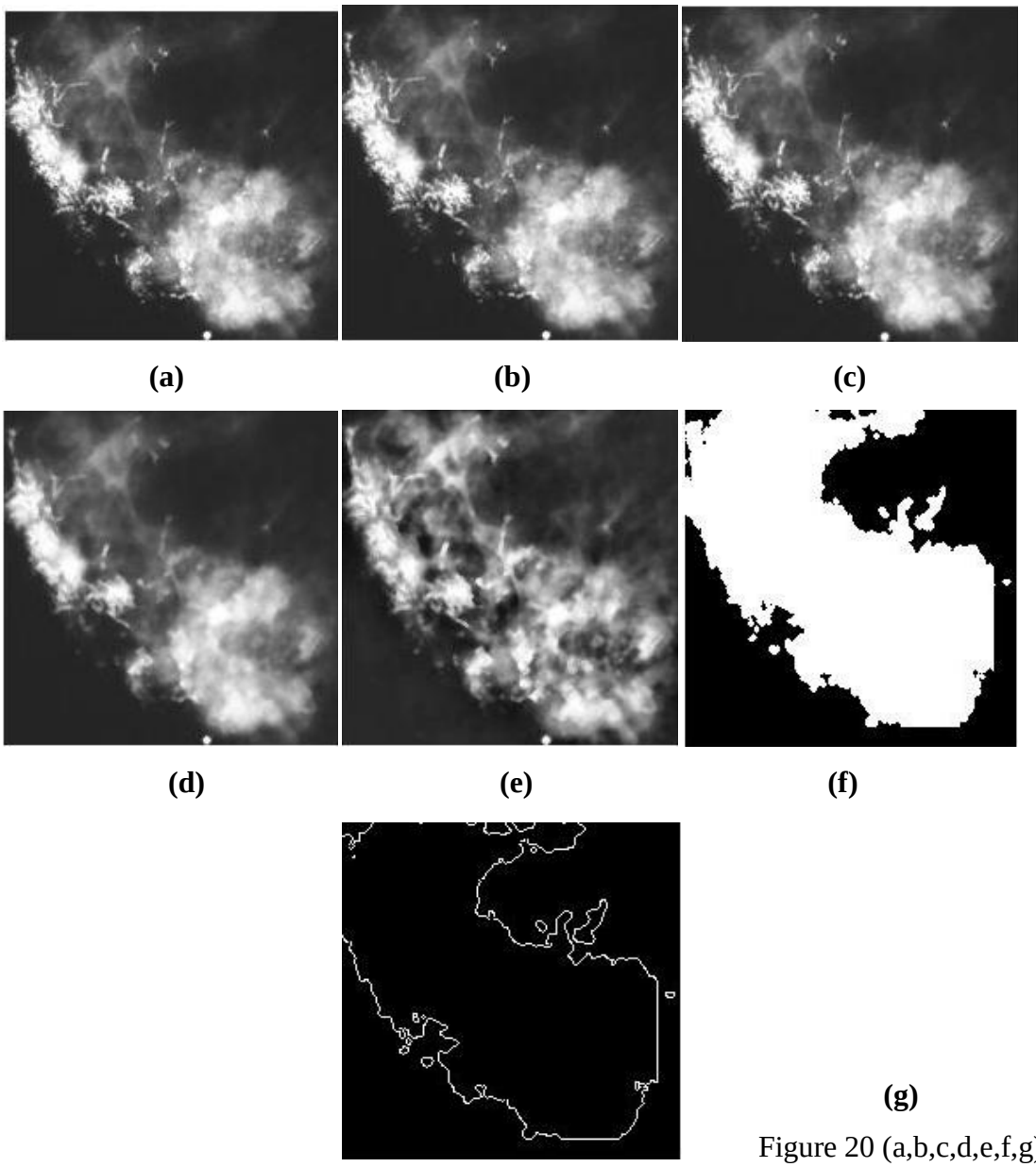


(g)

Fig 19 (a,b,c,d,e,f,g)

The above figure shows the original,gray scale,cropped,filtering,contrast,cancerous image and segmented image respectively.Here Cancerous image shown in Fig 19(f) with white area.GLCM features of cancerous and original images are shown in Table 4

4.11: A Database Mammogram Image(PN 11)



(a) **(b)** **(c)**

(d) **(e)** **(f)**

(g)

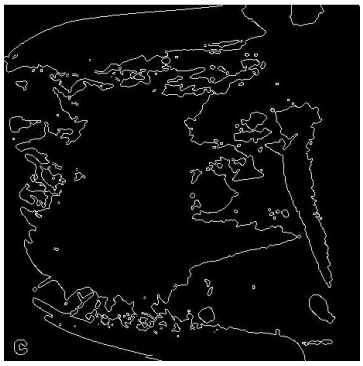
Figure 20 (a,b,c,d,e,f,g)

The above figure shows the original,gray scale,cropped,filtering,contrast,cancerous image and segmented image respectively.Here Cancerous image shown in Fig 20 (f) with white colour. GLCM features of cancerous and original images are shown in Table 4

4.12: A Database Mammogram Image(PN 12)

(a) **(b)** **(c)**

(d) **(e)** **(f)**



(g)

Fig 21 (a,b,c,d,e,f,g)

The above figure shows the original, gray scale, cropped, filtering,contrast,cancerous image and segmented image respectively.Here Cancerous image shown in Fig 21(f).GLCM features of cancerous and original images are shown in Table 3

Table 4:GLCM Texture Features Values

Texture Features	PN9 OIV	PN9 CIV	PN10 OIV	PN10 CIV	PN11 OIV	PN11 CIV	PN12 OIV	PN12 OIV
Offset[0 1]								
Contrast	.0454	.4277	.0425	.6826	.1876	.9710	.2512	1.2198
Correlation	.9801	.9222	.9929	.9709	.9674	.9600	.9634	.9479
Energy	.5064	.8792	.2790	.5080	.2460	.4846	.1604	.4979
Homogeneity	.9773	.9924	.9789	.9878	.9256	.9827	.9233	.4977
Offset[-1 1]								
Contrast	.0607	.5104	.0665	1.0185	.2658	1.2378	.4450	.9782
Correlation	.7734	.9071	.9920	.9566	.9540	.9490	.9348	.9318
Energy	.2991	.8475	.2709	.5006	.2364	.4797	.1453	.9174
Homogeneity	.9697	.9909	.9670	.9816	.9664	.9979	.8854	.4840
Offset[-1 0]								
Contrast	.0392	.4064	.0496	.8917	.1618	.8607	.2872	.9654
Correlation	.9828	.9260	.9940	.9620	.9719	.9646	.9529	.9297
Energy	.0392	.8797	.2755	.5038	.2480	.4867	.1545	.4892
Homogeneity	.9804	.9927	.9753	.9841	.9313	.9846	.4104	.9706
Offset[-1-1]								
Contrast	.0591	.5194	.0671	1.1131	.2456	1.1755	.4472	1.8822
Correlation	.9741	.9055	.9919	.9526	.9575	.9516	.9345	.9186
Energy	.2998	.8773	.2698	.4987	.2375	.4810	.1454	.4851
Homogeneity	.9705	.9907	.9670	.9801	.9079	.9790	.8860	.9664

This Table shows that the GLCM features(Contrast, Energy, Homogeneity) except Correlation feature at different 4 angles(0.45,90,135 degrees),having values greater in case of Original images as compared to their Cancerous Image feature values.

Summary: Original and their Cancerous Images Feature values are extracted from GLCM feature extraction. If comparing the both Original and its Cancerous image values, then it is shown from the table that all the three feature values Contrast, Energy, Homogeneity at different 4 angles(0.45,90,135 degrees),in case of Original Images have large values as compared to their detected Cancerous Images. Only Correlation feature values of Original Images are small as compared to its detected Cancerous images.

If size of Original and cancerous images increase, all the feature values of the images increase at 4 different angles(0,45,90,135) degrees.

CHAPTER 5

CONCLUSIONS AND FUTURE PERSPECTIVES

5.1 Conclusion

It is very necessary to detect the breast cancer as soon as possible. For early breast cancer detection, mammography is the best screening tool. There are other methods also to early detect the breast cancer like surgerical tests, biopsy and etc. It reduces the death rate of women because it detects the breast cancer at early stage.

By verifying the results of this research, we concluded that there are two cancers of Mammogram that is Benign and Malignant breast cancers. Benign cancer is non cancerous but malignant cancer is cancerous one.Mammography is required to detect the breast cancer early as soon as possible. Mostly problems occur in the women having age 40 and above.To detect the breast cancer,women should have the tests time to time.

To detect the breast cancer, database images whether DDSM or MIAS are taken and apply some MATLAB Version R200b programming using digital image processing.It includes pre-processing technique for noise reduction and image enhancement. Image segmentation used to detect the segmented part and then extract the features of GLCM.

By extracting GLCM features for Mammogram, we concluded that except correlation, the other features Contrast, Energy and Homogeneity all have larger values for Original Images as compared to their Cancerous images.

If the size of images increase, all the feature values increase whether it is the case of original images or cancerous images. Image size can be of 64x64, 128x128, 256x256 etc. Out of these the optimal value of image is taken i.e.128x128.

5.2 The Future Work

In future work on various classifiers have to be done and find the best classifier for mammography. Classifiers are used to classify the data. They show that how much accuracy is achieved to detect the benign and malignant cancer . Haralick has also proposed the methods of 3D images and try to work on 3D mammogram images for the detection of malignant and benign breast cancer.

[1] "Breast Cancer", available at http://www.wcrf.org/cancer_statistics/data_specific_cancers/breast_cancer_statistics.php

[2] "Breast Cancer in India", available at www.slideshare.net/NarjisSayeed/breast-cancer-in-india

[3] "Mammography for Breast Cancer Screening: Harm/Benefit Analysis", available at http://www.breastcancerdeadline2020.org/breast-cancer-information/breast-cancer-information-and-positions/mammography-for-breast-cancer.pdf

[4] "What is Breast Cancer",available at https://www.google.co.in/url?sa=t&rct=j&q=&esrc=s&source=web&cd=6&cad=rja&uact=8&ved=0CEAQFjAF&url=http%3A%2F%2Fwww.cancercare.mb.ca%2Fresource%2FFile%2Fabhr%2Fcommunity_groups%2F8-powerpoint_presentations%2Flesson_1_what_is_breast_cancer.pps&ei=yJuqU-DhCNWWuATnvYGQBA&usg=AFQjCNG4jFOAel-YUWPAwNvAeDghT5a_dQ&bvm=bv.69620078,d.c2E

[5] "Overview of Breast Cancer", available at www.slideshare.net/drizsyed/breast-cancer-ppt

[6] "Types of Breast Cancer", available at http://www.dhhs.tas.gov.au/__data/assets/pdf_file/0014/25520/breastscreen_resource_sect3.pdf[

7] "What Is the Difference between Digital and Film Mammography",available at http://cameron.agrilife.org/files/2014/05/Differences-in-Mammography-Screenings-Revisions.pdf

[8] "Mammography", available at https://www.google.co.in/url?sa=t&rct=j&q=&esrc=s&source=web&cd=4&cad=rja&uact=8&ved=0CEwQFjAD&url=http%3A%2F%2Fwww.elcamino.edu%2Ffaculty%2Fdcharman%2FRT%2520255%2520MAMMO%2520A%2520%26%2520P%2520200810.ppt&ei=cDTCU6zNPMXHuASWw4DwBg&usg=AFQjCNEspEY8ZNtsBgDeYWyMK_eBcxeWMg&sig2=0-HDN-NhCILPxPY-z4o40w

[9] "Mammography Techniques and Types", available at http://www.breastcancer.org/symptoms/testing/types/mammograms/types

[10] "Advantages of Digital Mammography",available at http://www.urmc.rochester.edu/highland/departments-centers/breast-imaging/tests-procedures/advantages-digital-mammography.aspx

[11] "Digital Mammography", available at http://www.udiwp.com/services-digital-mammography.html

[12] "What Is the Difference between Digital and Film Mammography",available at http://cameron.agrilife.org/files/2014/05/Differences-in-Mammography-Screenings-Revisions.pdf

[13] "Realizing the Benefits of Full-Field Digital Mammography,The Role of the Film Digitizer" ,available at http://www.vidar.com/film/images/stories/PDFs/products/diagnosticpro/edge/pdf/ffdm_whitepaper.pdf

[14] "Mammograms", available at http://womenshealth.gov/publications/our-publications/fact-sheet/mammograms.pdf

[15] D.NarainPonraj, M.K.Evangelin Jennfier,P Poongodi,J Samuel Manoharan, " A Survey on the Preprocessing Techniques of Mammogram for the Detection of Breast Cancer*" Journal of Eme rging Trends in Computing and Information Sciences,* vol. 2, NO. 12, December 2011.

[16] R. Ramani, Dr. N. Suthanthira Vanitha, S. Valarmathy," The Pre-Processing Techniques for Breast Cancer Detection in Mammography Images", *I.J. Image, Graphics and Signal Processing,*vol 5,pp. 47-54,2013.

[17] Blaine Martinez, Haidi Ibrahim, Nicholas SiaPik Kong, Theam Foo Ng,"Simple Adaptive Median Filter for the Removal of Impulse Noise from Highly Corrupted Images*",IEEE Transactions on Consumer Electronics,* Vol 54, pp .4, Novermber 2008.

[18] Blaine Martinez," Digital Image Processing Final Project" and Nicholas Sia Pik Kong, Theam Foo Ng ,"Simple Adaptive Median Filter for the Removal of Impulse Noise from Highly Corrupted Images HaidiIbrahim*", IEEE Transactions on Consumer Electronics*, Vol 54, No.4 ,Novermber 2008

[19] Robert M. Haralick ,K. Shanmugam and Its'hak Dinstein,"Textural Features for Image Classification", *IEEE Transcations on systems,man and cybernetics,v*ol.SMC-3 pp. 610-621,6-Nov 1973.

[20] Pawar B.V., Patil Sharma,"Early Stage Detection ofTumors in Mammograms",Electronic Dept.,W.C.E. Sangli.

[22] Aswini Kumar Mohanty,

Swapnastikta Beberta,S aroj Kumar Lenka," Classifying Benign and Malignant Mass using GLCM and GLRLM based Features from Mammogram", *International Journal of Engineering Research and Applications*,col.1,pp.6876-693

[23] P. Mohanaiah, P. Sathyanarayana, L. Guru Kumar," Image Texture Feature Extraction Using GLCM Approach", *International Journal of Scientific and Research Publications,* Volume 3, Issue 5, May 2013.