EJSIT
European Journal of Science, Innovation and Technology
ISSN: 2786-4936
www.ejsit-journal.com
Volume 4 | Number 3 | 2024

Breast Tumours and Cancer Detection and Identification Using MATLAB

Dr. Someyeh Kazemi, Rahman Hanoon Atiyah
Dr. Computer Engineering, Iraq
Sc. Biomedical Engineering, Iraq

ABSTRACT
In Iraq and across the world, cancer has become a deadly disease and more and more people are suffering from cancer. One study says that 1 in 30 women will suffer from this disease in their lifetime, that is why the project was conceived because of the increasing incidence of maternal cancer, and the most important thing is that we can detect cancer. In the afternoon there are many treatment possibilities. So, this project explores the foundations of automatic cancer detection so that everyone can be diagnosed early and more comprehensively, making breast cancer the most prevalent cancer worldwide (Parkin et al., 2005).

Keywords: breast cancer, medical image, Matlab, imaging data image, enhancement, mammograms

INTRODUCTION
Breast cancer occurs when healthy breast cells become abnormal, grow out of control, and form tumours. 1 in 8 women will be diagnosed with breast cancer in their life time. Breast cancer can develop in different parts of the breast. Mammography images consist of many artefacts and noises and make medical images too difficult to detect and understand the cancer at the primary stages. With a low dose of X Ray imaging using (Dhawan, 2003) mammography is most basic screening test for breast cancer and also records better visualised internal several methods have been reported for pre-processing mammography images because of its influences in detection of cancer. Moreover, the size of the significant details can be very small. Several research works have tried to develop computer aided diagnosis (CAD) tools that could help the radiologists in the interpretation of the mammograms and could be useful for an accurate diagnosis (Highnam & Brady, 1999; Kekre, Sarode, & Gharge, 2009).

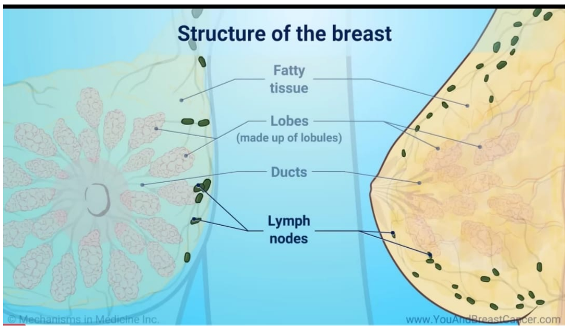

Figure 1: Structure of breast

WHAT IS BREAST CANCER
Breast cancer is major cause of death in women around the world. According to WHO (World Health Organisation), breast cancer accounted for maximum deaths (2.26 million cases) (Center for Disease Control and Prevention, n.d.), worldwide in 2020 out of the 10 million cases of cancer. Breast cancer starts when cells in the breast begin to grow out of

control. These accumulations of cells are called tumours and they can often be seen on an x-ray or felt as a lump. Breast cancer can spread when the cancer cells get into the blood or lymph system and are carried to other parts of the body making them prone to cancer. There are many different types of breast cancer and common ones include ductal carcinoma in situ (DCIS) and invasive carcinoma. The side effects of Breast Cancer are – Fatigue, Headaches, Pain and numbness (peripheral neuropathy), Bone loss and osteoporosis.

There are two types of tumours (IEEE Engineering in Medicine and Biology (IEEE EMBS), n.d.). One is benign which is non-cancerous and the other one is malignant which is cancerous. Benign breast tumours are abnormal growths in the breast, but they do not spread outside. So, this means that they are not life threatening, but some types of benign tumours can increase a woman's risk of getting breast cancer. Different imaging tests are used for detecting breast cancer. Some of them are mammograms, breast ultrasound and breast MRI. A **mammogram** is nothing but an x-ray of breast and it is used to look for any changes in the breast. A mammogram makes it easy to treat by finding and detecting breast cancer early, when the tumour is small and even before a lump can be felt.

Detection of breast cancer in its early stages using image processing techniques includes four parts. In the first part the digital images (mammograms) are pre-processed to remove any kind noise. Then in the second part the images undergo the segmentation process to enhance the tumour part. After this, in the third part, the important features in the segmented images are extracted. Finally, in the fourth part, with the help of the extracted features, the images are classified into normal, benign or malignant. Here, '**normal**' represents the breast with no tumour, '**benign**' represents the breast with non-cancerous tumour and '**malignant**' represents breast with cancerous tumour.

HOW BREAST CANCER SPREADS

Breast cancer can spread when the cancer cells get into the blood or lymph system and are carried to other parts of the body.

The lymph system is a network of lymph (or lymphatic) vessels found throughout the body that connects lymph nodes (small bean-shaped collections of immune system cells). The clear fluid inside the lymph vessels, called lymph, contains tissue byproducts and waste material, as well as immune system cells. The lymph vessels carry lymph fluid away from the breast. In the case of breast cancer, cancer cells can enter those lymph vessels and start to grow in lymph nodes. Most of the lymph vessels of the breast drain into:

- Lymph nodes under the arm (auxiliary nodes);
- Lymph nodes around the collar bone (supraclavicular [above the collar bone] and infraclavicular [below the collar bone] lymph nodes);
- Lymph nodes inside the chest near the breast bone (internal mammary lymph nodes).

As proposed by the World Health Organization, early detection and screening, especially when combined with adequate therapy, offer the most immediate hope for a reduction in breast cancer mortality (National Cancer Control Programs, 2002).

If cancer cells have spread to your lymph nodes, there is a higher chance that the cells could have travelled through the lymph system and spread (metastasized) to other parts of your body. The more lymph nodes with breast cancer cells, the more likely it is that the cancer may be found in other organs. Because of this, finding cancer in one or more lymph nodes often affects your treatment plan. Usually, you will need surgery to remove one or more lymph nodes to know whether the cancer has spread.

Still, not all women with cancer cells in their lymph nodes develop metastases, and some women with no cancer cells in their lymph nodes develop metastases later.

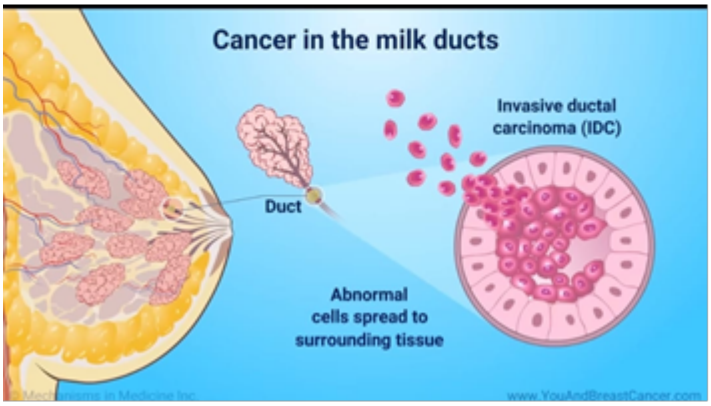

Figure 2: Cancer in the milk ducts

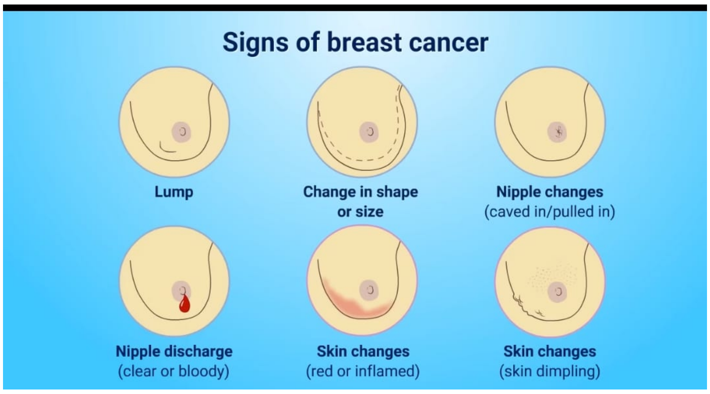

Figure 3: Signs of breast cancer

PROBLEM STATEMENT

Mammography images consist of many artefacts and noises and make medical images too difficult to detect and understand the cancer at the primary stages. Standardization of images quality and extraction of ROI is essential to limit the hunt for normalities required to store the compression image.

OBJECTIVES

- To identify abnormal areas that cannot be experienced physically or visualized.
- To diagnose and identify the abnormality quickly and easily.

TYPES OF BREAST CANCER

There are many different types of breast cancer and common ones include ductal carcinoma in situ (DCIS) and invasive carcinoma. Others, like phyllodes tumours and angiosarcoma are less common.

Once a biopsy is done, breast cancer cells are tested for proteins called estrogen receptors, progesterone receptors and HER2. The tumour cells are also closely looked at in the lab to find out what grade it is. The specific proteins found and the tumour grade can help decide treatment options.

Type of breast cancer

*Ductal carcinoma in situ (DCIS)
-starts in milk duct
-it is non invasive
-detected through mammogram
-can be treated successfully

*Triple - Negative breast cancer (INBC)
1-negative for:
Estrogen receptors (ER)
Progesterone receptor (PR)
HER2 protein

2- it is more aggressive other types

*Invasive ductal carcinoma (IDC)
-accounts for 80% of all breast cancer cases
-it is invasive in nature
*Invasive lobular carcinoma (ILC)
-starts in lobules
-it is also invasive

*HER2-positive breast cancer
1- it is +ve for HER- protein, which promotes cancer cell growth
2- can be treated with targeted therapy drugs
*Inflammatory breast cancer (IBC)
1-a counts for<5% of all breast cancer
2-Require immediate treatment.

TREATMENT OF BREAST CANCER

*Surgery
-First line of treatment
-type depends on size and metastasis

*Radiation therapy
-used after surgery

*Chemotherapy
-used in combination with surgery and radiation therapy

*Hormone therapy
-used for hormone receptor positive breast cancer

*Targeted therapy
-used for patients with HER-2 positive breast cancer

Ductal cells: line the milk ducts that carry milk from the lobules to the nipple.
Lobular cells: make up the lobules, which produce milk.
Connective tissue cells: provide structure and support to the breast tissue as in figure below.

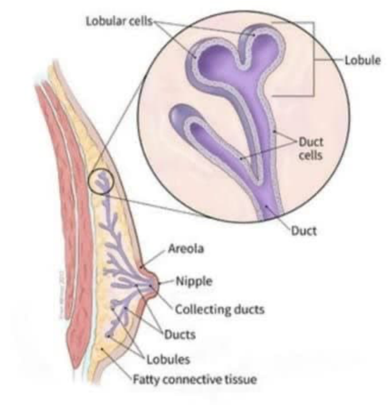

Figure 4: Breast tissue and ductal milk

METHODOLOGY

Background removal

Initially image was binarized with threshold value 0.1 then the connected component organized in descending order to extract the largest blob which is the breast profile but consists of pectoral muscle.

Removal of pectoral muscle

*To reduce the pectoral muscle part by using modified region growing technique;

*The seeded region growing is one of the image segmentation methods, it works in two ways based on selected pixel locational value and other is selection of seed point;

*Seed point is selected automatically by considering the orientation of the mammography;

*This approach determines the neighbouring pixel of seed point and examine whether the next pixel should be added to the region or not.

Image enhancement

Used to enhance the quality of the image using wiener filter and CLAHE.

Background removal

Initially image was binarized with threshold value 0.1 then the connected component organized in descending order to extract the largest blob which is the breast profile but consists of pectoral muscle.

Removal of pectoral muscle

*To reduce the pectoral muscle part by using modified region growing technique;

*The seeded region growing is one of the image segmentation methods, it works in two ways based on selected pixel locational value and other is selection of seed point;

*Seed point is selected automatically by considering the orientation of the mammography;

*This approach determines the neighbouring pixel of seed point and examine whether the next pixel should be added to the region or not.

Image enhancement

Used to enhance the quality of the image using wiener filter and CLAHE.

The figure below presents the detection of steps to detect breast cancer.

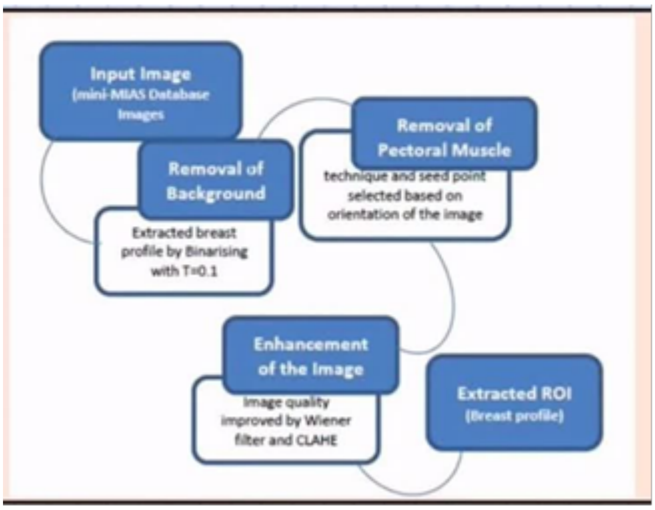

Figure 5: Detection of steps to detect breast cancer

IMAGE PRE-PROCESSING

In this process, the image is converted into the accessible form. Image smoothing act as the pre-processing step for image segmentation, as almost all of the images suffer from the problem of noise effect, uneven illumination and local irregularities. In this present work, the noise and uneven illumination are filtered out from the mammogram image using Gaussian filter which is a type of a linear filter. The contrast of the smoothened image is enhanced using the image processing toolbox functions available in Matlab 7.9. This improves the visualization effect of the original image. The imadjust function is used for this purpose. It has the following syntax:

J = imadjust(I, [low_in; high_in],[low_out; high_out]);

This maps the values in intensity image I to new values in J such that values between low_in and high_in maps to values between low out and high out. After the intensity adjustment the image format is changed from „indexed‟ to „intensity‟ image into the „binary‟ image. Im2bw function is used to do this process. It converts the input image to gray scale format, and then uses threshold to convert this gray scale image to binary. After this the complement of the binary image is taken to obtain the distant transform. Then the watershed function is applied on the complement of distance transform image. But watershed of negative distant transform leads to the serious over segmentation. Thus, this method is not preferred for the segmentation of breast mammogram image. So, after filtering and contrast enhancement, the image is first converted into double format to make it suitable for computing the gradient magnitude. The double format image is used as an input for the segmentation process.

MARKER-CONTROLLED WATERSHED SEGMENTATION

Matlab programming is used to develop algorithm for marker controlled Watershed segmentation. The gradient magnitude is used often to preprocess a gray-scale image prior to using the watershed transform for segmentation which is computed using the linear filtering method (Fu & Mui, 1981). For any gray scale image (x,y), at co-ordinates (x,y), the gradient vector magnitude and angle at which maximum rate of change of intensity level occurs at the specified co-ordinates (x,y) can be computed using the equation (1) and equation (2).

$$g(x,y) = \sqrt{\left(g_1^{\,2}(x,y) + g_2^{\,2}(x,y)\right)} \quad \textbf{(1)}$$

$$\alpha(x,y) = \tan^{-1}\left(g_1(x,y)/g_2(x,y)\right) \quad \textbf{(2)}$$

Where g1(x, y) and g2(x, y) are the gradients in the x and y directions. Magnitude of these gradients is computed using the sobel mask H1 and H2, which are defined by equation (3) given below:

$$H_1 = \begin{bmatrix} -1 & 0 & 1 \\ -2 & 0 & 2 \\ -1 & 0 & 1 \end{bmatrix}, \quad H_2 = \begin{bmatrix} -1 & -2 & 1 \\ 0 & 0 & 0 \\ 1 & 2 & 1 \end{bmatrix} \quad (3)$$

Watershed of simple gradient image does not produce good segmentation result (Vincent & Soille, 1991). There are too many watershed ridge lines that do not correspond to the object in which we are interested. To overcome such a difficulty, a new approach based on the concept of markers is introduced in this project and this approach is called Marker–Controlled Watershed Segmentation. In the present method, the internal markers are produced from the gray scale image and then external markers are found by finding pixels that are exactly midway between the internal markers.

This is done by computing the watershed transform of the distance transformed image of the internal markers. The gradient image is then modified by imposing regional minima at the location of both the internal and external markers.

The next step involves the computation of the watershed transformation of the Marker modified gradient image to produce watershed ridge lines. Finally resulting watershed ridge lines are superimposed on the original image and produce the final segmentation and then the output image produced as a result of Region of Interest (ROI) selection is converted into label matrix. All the parameters of an object are then extracted using the image processing toolbox functions.

EXPERIMENTAL RESULT EVALUATION

Results obtained by applying the proposed algorithms mammogram breast image, that's Take it from Iraqi national program for early detection of breast cancer This image represents the control image for our experiment indicating that mass cancer is detected. Gaussian filter is used to remove the noise and detail form the original mammogram image, the Gaussian filter is very effective in the reduction of impulse and Gaussian noise as shown in Figure 5. Figure 6 shows the segmentation process that performed on the edge map differentiates various regions on the breast depending on their intensity values. Each region has a different intensity value. The fatty tissues, glands, lobules and the ducts display different intensity values and thus can be segregated into different regions. An abnormality such as a mass, tumours or calcifications may be present within the breast has distinctly higher intensity values than the normal tissues of the breast. So in this is figure we need to categorize all the obtained closed structures to their intensity values. The distribution of pixels intensities also vary within each segmented region but the majority the pixels have similar intensity values. So for each region we calculate the arithmetic Mode value for the intensities from the original mammogram and replace those pixels in the region with the computed mode values.

The final segmentation result obtained by proposed Marker based Watershed algorithm is shown in Figure 7. Results obtained by applying the proposed algorithms on digital mammogram image comprised predominantly fatty tissues where abnormalities are present.

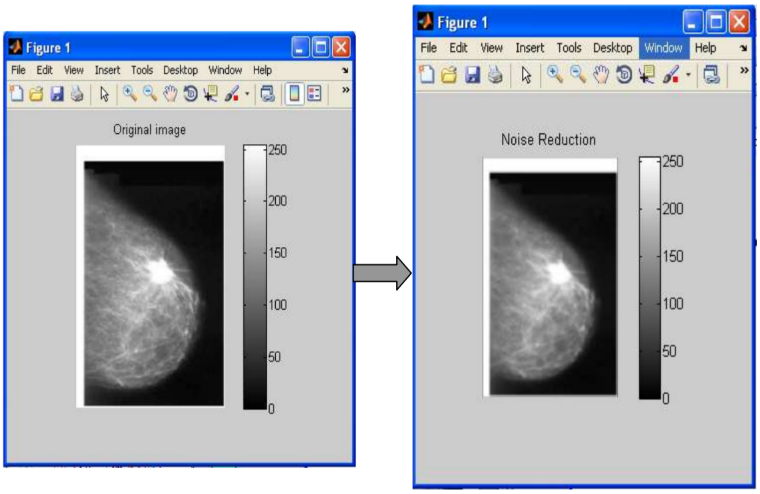

Figure 5: Original mammogram along with the image on which the noise is removed by the Gaussian filter

The accuracy of this technique "Marker–Controlled Watershed Segmentation." was evaluated through quantitative measures derived through the comparison of each segmented mammogram representing an abnormal mass, as "mask" with its corresponding "gold standard". The gold standard is generated by manually segmenting the breast region representing an abnormal mass from each mammogram. The boundary of the regions is then manually traced to extract the abnormal breast region to generate a ground truth (GT) image.

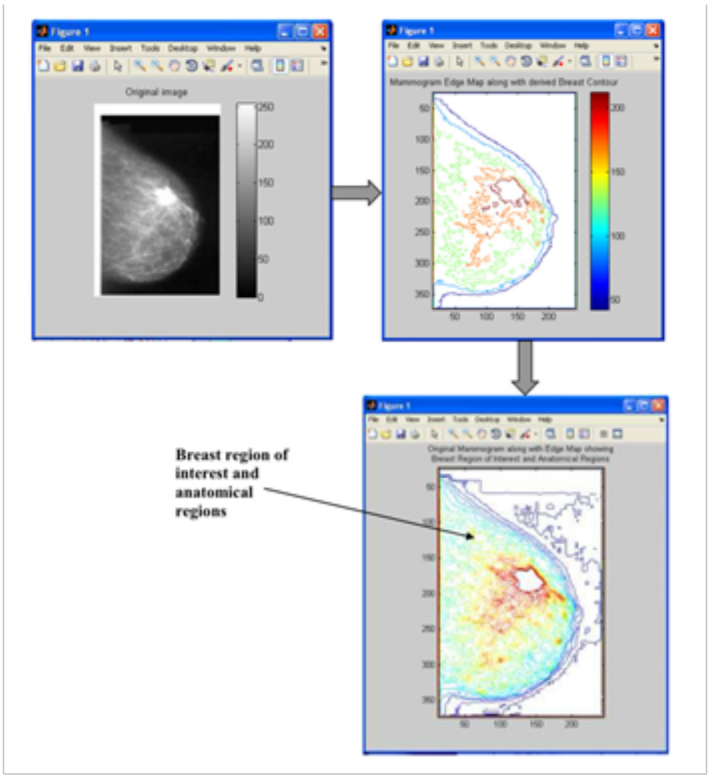

Figure 6: Original mammogram along with edge map derived breast contour and showing breast region of interest and anatomical regions

ANALYSIS

Comparing the median filter, adaptive min max and wiener filter we got high PSNR for all the images tested as shown in the figure.

Advantage

The proposed method of pre processing prevented with removal of back ground artefacts, pectoral muscle suppression and image quality enrichment helps much in early detection.

Disadvantages

Of 77 breast cancer diagnosis, 10 would be cases of over diagnosis. Over diagnosis is the discovery of a cancer that would never have been detected without screening.

Constraints

There are 2 main challenges with diagnosing breast cancer with a mammogram.

1- False negative scan appears normal even though cancer is present.

2- False positive is scan looks abnormal even though no cancer is present.

How to reduce the false negative and false positive?

Computer learn by example and in order to each to computer how to interpret mammogram, we had to show many example, it will eventually learn to get the thing right.

RESULTS

1- The first photo is getting the image.

2- The image then convert to BW (black and white) image using rgb2 gray function. it's converts gray-scale image to a binary image.

3- The region growing method used to reduce the pectoral muscle pay. The proposed method helps to select the seed point automatically.

4- Next,we will delete the background, so that we can select the muscle.

5- The image quality was enhanced by using wiener filter, and CLAHE filter. The evaluation of image quality is crucial for medical imaging systems such as compression, transmission and enhancement.

CODING OF MATLAB THAT USING IN DETECTION

1

```
clc;
clear;
close all;
warning off;
i=imred('mdb107.pgm');
subplot(3,3,1);
imshow(i);
title('orignal photo');
try
i=rgb2gray(i);
end
```

2

```
z=im2bw(i,0.1);
subplot(3,3,2);
imshow(z);
title('orignal B&w);
```

3

```
info=region props(z);
a=cat(1, info.Area);
[m,1]=max(a);
x=info (1). centroid;
bw2=bwselect(z,x(1),x(2),8);
i=immultiply(i,bw2);
subplot(3,3,3);
imshow(i)
title('Getting the breast and muscle ');
```

4

```
[x,y]=size(z);
tst1=zeros(x,y);
r1=[ ];
m=1;
for j=1:x
if z(j,:)==tst1(j,:)
r1(m)=j;
m=m+1;
end
end
r2=[ ];
m=1;
for j=1:y
if z(i,j)=tst1(i,j)
r2(m)=j;
m=m+1;
end
end
i(:,r2)=[ ];
i(r1,i)=[ ];
subplot (3,3,4);
imshow (i);
title ('After deleting back ground ');
if i(1,1)=0
c=3;
r=3;
else
r=3;
c=size (1,2)-3;
end
r2=im2bw(1,0.5);
bw3=bwselect(r2,c,r,8);
bw3=~bw3;
ratio=min(sum(bw3)/sum(z2));
```

5

```
if ratio>=1
i=immultiply (i,bw3);
else
z2=im2bw (i,0.75);
bw3=bwselect (z2,c,r,8);
ratio 2=min(sum(bw3)/sum(z2));
if round (ratio 2)==0
1v1=graythresh(i);
z2=im2bw(i,1.75*1v1);
bw3=bwselect (z2,c,r,8);
bw3=~bw3;
i=immultiply (i,bw3);
else
bw=~bw3;
i=immultiply (i,bw3);
end
end
subplot (3,3,5);
imshow (i);
title ('Getting only the Breast ');
```

6

```
mask=fspecial('average',[3 3];
SNR=0.2;
i=deconvwnr(i,mask,SNR);
subplot (3,3,6);
imshow (i);
title ('weiner Filter ');
```

7

```
i=adapthisteq(i);
subplot (3,3,7);
imshow (i);
title ('clathe Filter '), [m,1]=max(a);
x=info (1). centroid;
bw2=bwselect(z,x(1),x(2),8);
i=immultiply(i,bw2);
subplot(3,3,3);
imshow(i)
title('Getting the breast and muscle ');
z2=im2bw (i,0.75);
bw3=bwselect (z2,c,r,8);
ratio 2=min(sum(bw3)/sum(z2));
if round (ratio 2)==0
1v1=graythresh(i);
z2=im2bw(i,1.75*1v1);
bw3=bwselect (z2,c,r,8);
bw3=~bw3;
i=immultiply (i,bw3);
else
bw=~bw3;
i=immultiply (i,bw3);
end
end
subplot (3,3,5);
imshow (i);
title ('Getting only the Breast ');
```

Figure 7: Shape of cancer tumour by Matlab

CONCLUSION

The developed algorithm is used to know about the location and size of the of mass in mammograms having abnormal regions. This method uses the user defined input parameters such as threshold for the analytical calculations. The optimal value of the threshold is highly dependent on shape and location of the tumor as well as on the different views of the images. The accurate mass detection was achieved by the proposed algorithm Marker–Controlled Watershed Segmentation. The result obtained shows that the proposed algorithm has a higher computational efficiency along with the good performance for image segmentation problems ascompared to the conventional tool. The algorithm used in this work is marker dependent which in turn depends upon the selected value of threshold. In this work, the optimal value of threshold is selected interactively. But this algorithm can be made more efficient and fast by providing the automatic value of threshold. This makes it faster and robust. In future, this work can be extended to classify the tumors using any of the techniques such as Fuzzy Expert Systems, Neural Network System, ANFIS and Hidden Markov Model. CAD system helps to identify the diminutive changes in the breast. The proposed method of pre processing presented with removal of back ground artefacts, pectoral muscle suppression and image quality enrichment helps much in early detection. Our results show that, significant filter mask for wiener filter and CI is 0.2 for CLAHE are in fluenciny factors for enhancement of mammography. The proposed method tested on images of mini mias database, ROI extracted from all the images accurately and proved to be suitable for CAD system of early detection of breast cancer. The proposed method definitely can be considered for automated detection of abnormal benign, malignant and micro calcification. And CAD system helps to identify the diminutive changes in the breast. The proposed method of pre processing presented with removal of back ground artefacts, pectoral muscle suppression and image quality enrichment helps much in early detection. Our results show that, significant filter mask for wiener filter and CI is 0.2 for CLAHE are in fluenciny factors for enhancement of mammography. The proposed method tested on images of mini mias database, ROI extracted from all the images accurately and proved to be suitable for CAD system of early detection of breast cancer. The proposed method definitely can be considered for automated detection of abnormal benign, malignant and micro calcification.

REFERENCES

Center for Disease Control and Prevention (n.d.).

Dhawan, P. A. (2003). *Medical Imaging Analysis*. Hoboken, NJ: Wiley-Interscience Publication.

Fu, K.S. & Mui, J.K. (1981). A Survey of Image Segmentation. *Pattern Recog., 13*(1), 3-16. http://dx.doi.org/10.1016/0031-3203(81)90028-5.

Highnam, R. & Brady, M. (1999). *Mammographic Image Analysis*. Kluwer Academic Publishers.

IEEE Engineering in Medicine and Biology (IEEE EMBS) (n.d.).

Kekre, H. B., Sarode, T. K., & Gharge, S. M. (2009). Tumor detection in mammography images using vector quantization technique. *International Journal of Intelligent Information Technology Application*, *2*(5), 237-242.

National Cancer Control Programs. (2002). *Policies and managerial guidelines* (2nd ed.). Geneva, World Health Organization.

Parkin, D. M., Bray, F., Ferlay, J., & Pisani, P. (2005). Global cancer statistics, 2002. *CA: A Cancer Journal for Clinicians*, *55*(2), 74-108.

Vincent, L. & Soille, P. (1991). Watersheds in Digital Spaces: An Efficient Algorithm based on Immersion Simulations. *IEEE Trans. Pattern and Machine Intell., 13*, 583-598. http://dx.doi.org/10.1109/34.87344.