

# Detection of malignant tissue in mammography image using morphology based segmentation technique

Sharma N<sup>1</sup>, Manjhi J<sup>2</sup>

<sup>1</sup> Neha Sharma, M Tech Biomedical, Shobhit institute of Engg. and Technology Meerut, India, <sup>2</sup>Jayanand Manjhi, Professor and co-ordinator Biomedical Department, Shobhit university, Meerut, UP, India

**Address for correspondence:** Neha Sharma, Email: nehasharma29bm@gmail.com

## Abstract

Breast cancer is the leading cause of the death among the women. Mammography is the best diagnostic technique for the breast cancer. But not all breast cancer can be seen by mammogram. Although breast cancer can be mortal, people have the highest chances to survive if cancer could be detected at the early stages. But there are certain limitations of the segmentation technique it is difficult to find the effected region perfectly. The proposed work deals with an approach for extracting the malignant masses in the mammography image for the detection of earlier breast cancer. The steps involve in this work are removal of noise from the background information, thresholding and retrieving the largest region of interest, performing morphological operations and extracting the ROI and identifying the malignant masses from the image. This method is compared with Enhancement, edge detection, Region Growing, Watershed Transformation techniques and found more accurate, sensitive, and precise in comparison to the others.

**Keywords-** Mammogram, Malignancy, Morphological, Median filter, ROI, Segmentation, Thresholding

## Introduction

Breast cancer is a malignant tumor that starts in the cells of breast. A malignant tumor is a group of cancer cells that can grow into surrounding tissues or spread to distant areas of the body. The disease occurs almost entirely in women. Signs of breast cancer may include a lump in the breast, a change in breast shape, dimpling of the skin, fluid coming from the nipple, or a red scaly patch of skin. In those with distant spread of the disease, there may be bone pain, swollen lymph nodes, and shortness of breath or yellow skin. Breast cancer can be separated into different types based on the way the cancer cells look under the microscope [1]. Most breast cancer are carcinomas, a type of cancer that start in the cells(epithelial cells) that line organs and tissues like the breast. In fact breast cancers are often a type of carcinoma called adenocarcinoma, which is carcinoma that starts in glandular tissue. Other types of cancers can occur in breast too, such as sarcoma which starts in the cells of muscle, fat or connective tissue. In some cases a single breast tumour can be a combination of different types or be a mixture of invasive and in situ cancer.

And in some rare types of breast cancer, the cancer cells may not form a tumour at all. Breast cancer is the most common cancer among women. About 1 in 8(12%) women in the world will develop invasive breast cancer during lifetime [2]. Mammography is one of the best and commonly used technique in the diagnosis of breast cancer. It is a specific type of imaging that uses a low-dose x-ray system to examine breasts but sometimes it is difficult for the radiologist to detect tumor or cancerous mass so CAD techniques are used for breast masses detection. To detect the malignant region the image is segmented or partitioned into different region [3]. The techniques available for segmentation of mammography images include: Image enhancement technique, histogram equalisation technique, watershed marker based technique, region growing technique, K-means clustering technique, Edge detection, transformation techniques etc. The drawbacks of these techniques are that this generally lacks in providing robustness and imperceptibility requirement. It usually enhances the whole image in a uniform manner, which in many cases produces undesirable results. It is not possible to selectively enhance edges or other required information effectively. Gaussian filters are mostly used

Manuscript received 12<sup>th</sup> February 2016  
Reviewed: 22<sup>nd</sup> February 2016  
Author Corrected: 3<sup>rd</sup> March 2016  
Accepted for Publication 19<sup>th</sup> March 2016

in the removal of white noise. So it is not possible to remove noise completely or properly [4]. Some techniques reduce the details of the image or image information. watershed method did not give that much accurate result as the region extracted from the image is the final extracted result that is not correct. All the drawbacks of existing techniques are completely overcome with the proposed method for segmentation of mammography image i.e. morphological technique and thresholding of image. This technique not only overcomes all the limitations but also increase image quality which is measured with the help of image quality parameters.

## Materials and Methods

For image segmentation the MATLAB software of 7.1 version is used. And one sample of mammography image is taken from MIAS database. The image is a true color image with size 1024x1024. The proposed research intends to develop an image processing algorithm to extract the malignant mass from the mammography image for the detection of breast cancer. The work proposed is based on the following procedure: (a) removing the noise with median filter and the background information. (b)Applying thresholding and retrieving the largest region of interest (ROI). (c) Performing the morphological operations and extracting the ROI and identifying the malignant mass from the screened images of the breast. This method was tested on the mammography image taken from the MIAS database and implemented using MATLAB code. The technique is compared with other existing technique i.e. enhancement technique, watershed, region-growing. This technique consist of two processes:

**A. Pre-processing-** Pre-processing is an important issue in low-level image processing. The underlying principle of pre-processing is to enlarge the intensity difference between objects and background and to produce reliable representations of breast tissue structure [5]. The main objective of this process is to improve the quality of the image to make it ready to further processing by removing the unrelated and surplus parts in the background of the mammogram. Breast border extraction and pectoral muscle suppression is also a part of pre-processing. Pre-processing include following steps:

i. *Initial cropping:* cropping is one of the most important part of pre processing step. `imcrop` command is used for cropping the image. Cropping is done on the

raw mammogram image taken from the MIAS data base. In MIAS database each image is clipped or padded so that it is 1024x 1024 pixels which is the size of all raw mammographic image. Therefore the dark padding and the region outside the breast are excluded from the image. After cropping only the region of interest is kept for further processing of image.

ii. *Intensity adjustment:* Image enhancement techniques are used to improve an image, where "improve" is sometimes defined objectively (e.g., increase the signal-to-noise ratio), and sometimes subjectively (e.g., make certain features easier to see by modifying the colors or image's intensity values).[5]

iii. *CLAHE (Contrast limited adaptive histogram equalization):* CLAHE is a modification of the original histogram equalization operating on small regions in the image, called *tiles*, rather than the entire image. Each tile's contrast is enhanced, then the neighbouring tiles are combined using bilinear interpolation (directly implemented by MATLAB) to eliminate artificially induced boundaries.

iv. *Noise reduction:* After contrast enhancement noise is reduced from the image using median filter. A median filter is a nonlinear filter efficient in removing salt and pepper noise median tends to keep the sharpness of image edges while removing noise. Median filtering corresponds to non-linear smoothening which is used to reduce random noise. This smoothening filter allows edge preservation when compared to other smoothening filter. This filtering ranks the image pixels in a certain neighbourhood of certain pixel and replaces the value of this by median of neighbourhood.

**B. Image segmentation-** The most important stage of mass detection CAD schemes is to separate the suspicious regions that may contain masses from the background parenchyma i.e., to partition the mammogram into several non-overlapping regions, then extract regions of interests (ROIs), and locate the suspicious mass candidates from ROIs. The suspicious area is an area that is brighter than its surroundings, has almost uniform density, has a regular shape with varying size, and has fuzzy boundaries. Segmentation of masses from the rest of the image can be a complex process due to the diversity of mass characteristics from one image to the other.[5,6]

**C. Remove background information-** In order to remove the background information such as wedges and

labels in the mammogram images, first convert the grayscale or color image to binary image by using threshold technique and morphological operations. Thresholding is the simplest method of image segmentation. From a grayscale image, thresholding can be used to create binary images. After the gray scale mammogram images are converted into binary, remove the noise or background information from the binary image by removing all the smaller objects except the largest mammography part.

**D. Thresholding-** Here Histogram thresholding and slicing techniques are used to segment the image. They may be applied directly to an image, but can also be combined with pre- and post-processing techniques. Thresholding is also known as binarization of image. Thresholding approaches segment scalar images by creating a binary partitioning of the image intensities.

**E. Removing connected components corresponding to the pectoral muscles-** This is a supplemental stage to remove the high intensity areas that are known to be non-mass regions like the pectoral muscle and any artifact labels that are attached to the borders of the image. White neighborhoods all over the image that are connected to the image's borders are removed during this stage, resulting in a darker image.

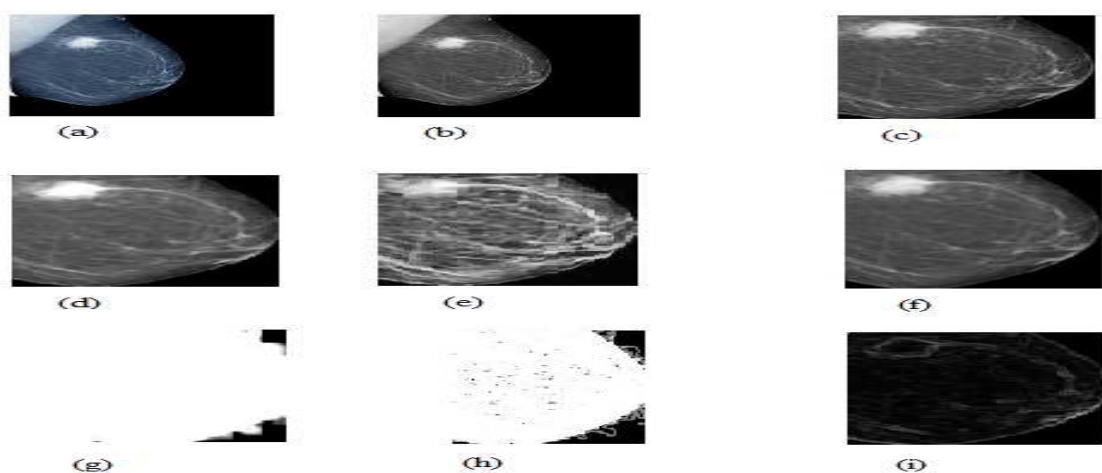
**F. Elimination of noise by locating connected segment larger area-** Once the mammogram image has obtained, result after the series of steps of the proposed method, images. In obtained images noise is observed. To eliminate this noise will seek the connected component with the greatest area, and the rest is

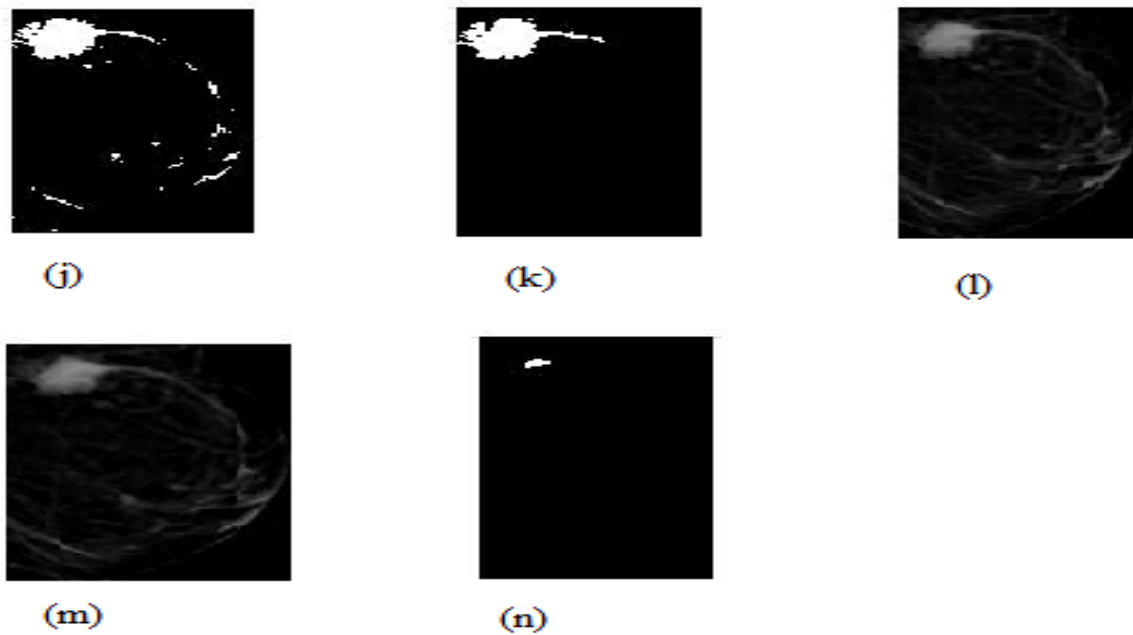
eliminated, so that with this process, the resulting image is segmentation of abnormality present in mammography. With this process the smaller binary objects are removed from the mammography image leaving behind the largest part.

**G. Perform subtraction-** Image subtraction or pixel subtraction is a process whereby the digital numeric value of one pixel or whole image is subtracted from another image which makes the background disappear leaving only the target. It simply compares the previous frame image with the current one. Image subtraction is a tool for transient object discovery and characterization,  $Z = \text{imsubtract}(X, Y)$  subtracts each element in array Y from the corresponding element in array X and returns the difference in the corresponding element of the output array Z. X and Y are real, non sparse numeric arrays of the same size and class, or Y is a double scalar. In this paper subtraction is done on two images, the final segmented image and the gray scale image. After subtracting segmented image from the gray scale image the final image of the tumor is obtained.

**H. Erosion-** Erosion is one of two fundamental operations in morphological image processing from which all other morphological operations are based. It was originally defined for binary images, later being extended to grayscale images, and subsequently to complete lattices. The erosion, of f by a flat structuring element b at any location (x, y) is defined as the minimum value of the image in the region coincident with b when the origin of b is at (x, y). Erosion removes small-scale details from a binary image but simultaneously reduces the size of regions of interest, too.

## Result and Discussion





**Figure 3 (a) Original mammography image (b) Gray- scale image (c) Cropped image (d) intensity adjusted image (e) CLAHE image (f) median filtered image (g) entropy image (h) standard deviation image (i) local range of an image (j) binary image ( k) ROI region (l) pectoral muscle removed image (m) background subtracted image (n) final segmented image**

The quality of the image is obtained by calculating image quality parameters shown in table below: The overall accuracy is the ratio between the total number of correctly classified regions and the test set size (total number of region s). Hitting denotes the ratio of correct segmentation while missing denotes the ratio of missing mass. Over hitting denotes the ratio of false mass segmented, Relative hitting denotes the relative correct ratio against segmentation result, relative missing denotes relative missing ratio against segmentation result. Sensitivity is a true positive measure that refers to the proportion of image that contains cancerous mass which have been classified correctly. Specificity is a true negative measure that refers to the proportion of image containing a cancerous mass that have been incorrectly classified [7].

**Table- 3.1 Diff Image Performance Parameters of Improved Morphological Technique.**

Accuracy	Sensitivity	Precision	specificity	Missing	Over hitting	Relative missing	Kapps
77%	89%	82%	29%	10%	19%	9%	85%

When the values of improved morphological technique is compared with the values obtained from other technique i.e. Enhancement, edge detection, Region Growing, Watershed Transformation techniques it is concluded that this technique is better than other technique and removes all the drawbacks and limitation of existing technique. This technique provide robustness and with this technique it is possible to enhance edges and the technique is efficient for noisy images.

**Conclusion and future work:** Identifying breast cancer is a challenging problem in medical image processing and medical fields. Digital mammogram on a particular segmentation algorithm makes it difficult to identify the breast cancer accurately. The acquisition parameters also influence the quality of the image. Mammography image segmentation using this technique is efficient in getting the malignant breast region and extracting the malignant mass. For future work it may planned to develop an algorithm to acquire a smoother breast region for image processing, and it can also be used to study the behavior of the breast, to improve the edge detection, to classify between breast regions and masses, and the algorithm can also detect abnormalities in segmented breast and which can produce more accurate results by increasing image quality parameters than proposed method.

## Appendices

Formulae:

$$\text{Accuracy} = \frac{\text{True Positive Region} + \text{True Negative Regions}}{\text{False Positive} + \text{true positive} + \text{true negative} + \text{False Negative}} \dots\dots\dots \text{i}$$

$$\text{sensitivity} = \frac{\text{True Positive Region}}{\text{True Positive Regions} + \text{False Negative Regions}} \dots\dots\dots \text{ii}$$

$$\text{specificity} = \frac{\text{True negative Region}}{\text{True negative Regions} + \text{False positive Regions}} \dots\dots\dots \text{iii}$$

$$\text{Precision} = \frac{\text{True positive Region}}{\text{True positive Regions} + \text{False positive Regions}} \dots\dots\dots \text{iv}$$

$$\text{Missing} = \frac{\text{false negative Region}}{\text{True positive Regions} + \text{False negative Regions}} \dots\dots\dots \text{v}$$

$$\text{over - hitting} = \frac{\text{false positive Region}}{\text{True positive Regions} + \text{False negative Regions}} \dots\dots\dots \text{vi}$$

$$\text{Relative missing} = \frac{\text{false negative Region}}{\text{True positive Regions} + \text{False positive Regions}} \dots\dots\dots \text{vii}$$

$$\text{Kapps} = \frac{2 * \text{hitting}}{2 * \text{hitting} + \text{missing} + \text{over hitting}} \dots\dots\dots \text{viii}$$

T.P- True positive region is the region of image that contains cancerous mass.

T.N- True negative region is the region that does not contain malignant mass.

F.P- False positive region refers to portion of image incorrectly classified as cancerous mass.

F.N- False negative region refers to the portion of image incorrectly classified as not a portion of cancerous mass.

## Acknowledgement

The work presented in this thesis could not have been done without the help of all of whom I want to thank. First, It gives me immense pleasure to express my deepest gratitude and indebtedness to Dr Jayanand, my thesis supervisor, Associate Professor, and Coordinator Centre for Biomedical Engineering., Shobhit University, Meerut, for his patient guidance and support under whose inspiration, encouragement and guidance I have successfully completed this thesis work. I am truly very fortunate to have the opportunity to work with him.

I found his guidance to be extremely valuable. My heartiest thanks goes to my co-supervisor Mrs. Shiva faculty, Assistant professor During my project, I was immensely helped by her and I am thankful to her for all the help she provided. I submit my gratitude to my parents for without their support nothing could be done.

**Funding:** Nil,

**Conflict of interest:** None.

**Permission of IRB:** Yes

## References

1. Guyton, A.C., & Hall, J.E.(2000), *text book of medical physiology*(10<sup>th</sup> ed).
2. American Cancer Society. Breast Cancer Facts & Figs. American Cancer Society, Inc, 2013. Available online at: <http://www.cancer.org/acs/groups/content/@epidemiologysurveillance/documents/document/acspc-036845.pdf>. cited on 23th March 2016.
3. Bronzino, J. *The biomedical engineering handbook* (2<sup>nd</sup> ed., vol.1). US. 2000.CRC press.

4. H.P. Chan, B. Sahiner, K.L. Lam. Computerized analysis of mammographic microcalcifications in morphological and texture feature spaces. *Med. Phys.* vol -25, no. 10, 2007-2019.
5. N. Petrick, H. P. Chan, B. and D. Wei. An adaptive density weighted contrast enhancement filter for mammographic breast mass detection. *IEEE Trans. Med. Image*, vol. 15, pp. 59-67, 2000.
6. A. Sahakyan, H. Sarukhanyan. Segmentation of the breast region in digital mammograms and detection of masses”, *International journal of advanced computer science and applications*, pp.102, vol- 3, no. 2, 2012.
7. Warfield, S. K. Zou and K. H.Wells. Simultaneous truth and performance level estimation (STAPLE): an algorithm for the validation of image segmentation. *IEEE Transactions on Medical Imaging*.vol.23, no.7, July2004.
8. Yasmin, M. Sharif, and S. Mohsin. Survey Paper on Diagnosis of Breast Cancer Using Image Processing Technique. *Research Journal Of Recent Sciences*, vol. 2, no. 10, pp. 88-98, October 2013.
9. I.K. Maitra, S. Nag and S.K. Bandhopadhyay. Identification of abnormal masses in digital mammography images. *International Journal Of Computer Graphics*, vol. 2, no. 1, 2011.

.....

#### How to cite this article?

Sharma N, Manjhi J, Detection of malignant tissue in mammography image using morphology based segmentation technique: *Int J Med Res Rev* 2016;4(4):635-640.doi: 10.17511/ijmrr.2016.i04.27.

.....