

# Designing of Automated CAD System for Neoplastic Breast Mass Detection

Prarthana Priyadarsini<sup>1</sup>, Prof. Pranabjyoti Haloi<sup>2</sup>

P.G. Student, Department of Electrical Engineering, Jorhat Engineering College, Jorhat, Assam, India <sup>1</sup>

Assistant Professor, Department of Electrical Engineering, Jorhat Engineering College, Jorhat, Assam, India <sup>2</sup>

**ABSTRACT:** This project deals with the development of CAD system for neoplastic breast mass detection from mammograms. As breast cancer remains a second leading cause of cancer deaths among women in many parts of the world. Early detection of breast cancer through periodic screening has noticeably improved the outcome of the disease. Since India is a developing country CAD system has to be developed in a cost effective manner in order to reduce the mortality rate. Commercially available CAD systems use computerized algorithms for identifying suspicious regions of interest (RoI). As presence of pectoral muscle has always been a hindrance in neoplasm detection because of its intensity range which is almost same as masses, thus removal of pectoral muscle is a vital or necessary step to attain proper accuracy in the extraction of RoI . Here MLO view is preferred over other views as it provides a clear view of pectoral muscle. This project gives a novel algorithm for automatic detection and removal of pectoral muscle and auto cropping of neoplastic area (RoI) of masses which has achieved an accuracy of 95% on a set of 282 mammogram images. Further feature extraction of neoplastic area has been done and classified into benign and malignant using SVM classifier with a sensitivity rate of 100%.This composite method has been implemented and applied to mini-MIAS which is one of the most challenging digital database consisting MLO view mammograms.

**KEYWORDS:** Breast cancer, segmentation, feature extraction, classifications, masses, mammograms and support vector machine.

## I. INTRODUCTION

Digital mammography is one of the effective and viable imaging techniques to detect breast cancer. Mammograms can be used to check for breast cancer in women who have no signs or symptoms of the disease. Mammography will detect about 80–90% of the breast cancers in women without symptoms. Computer Aided (CA) detection systems have been developed to aid radiologists in detecting mammographic abnormalities. CAD technology has a positive impact on early breast cancer detection. The detection sensitivity without CAD is about 80% and with CAD it may rise up to 90%[12]. Since the cause of the disease is unknown , early detection and diagnosis is the only key to improve breast cancer control and it can increase the success of treatment, save lives and reduce cost effectively. About 25% of all cancers diagnosed in women are breast cancers and about 20% of all lethal cancers are breast cancers[2]. Being left with the only key of early detection to overcome the life threat due to breast cancer which is found in all types of masses of female population renders the need of development of efficient and cost effective detection system to carry out early diagnosis. The principal difficulty in this task is the lack of a general algorithm that produces good results for all images. Thus, in quest of search for cue to early detection the respective work aimed at developing an automated computer aided breast cancer detection system which segments and classifies masses from a given mammogram which is further classified into benign and malignant.

## II. LITERATURE REVIEW

1. In 2011 [10] Li Liu, Jian Wang, Tianhui Wang presented a paper on Computer-aided method to detect breast boundaries and pectoral muscles on mammograms. A statistical approach based on joint normal distribution is employed to measure the possibility of a pixel in high frequency region. The values provides the knowledge of contour of the breast and pectoral muscle, thus region such as background and pectoral muscles are removed.

# International Journal of Innovative Research in Science, Engineering and Technology

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 6, June 2015

- The proposed method successfully detect skin to air boundaries on all tested mammograms. Proposed method is more effective than wavelet or isocontour method on pectoral muscle detection .
2. In 2011 [11] Dheebea.J and Tamil Selvi.S proposed a paper on system designed to perform screening of digital mammograms for the presence of microcalcifications using support vector machine. Dilation operation is used to enhance the grayscale mammogram image. The Laws texture features are extracted from the image for automatic classification of Micorcalcification clusters, which includes 25 texture features for each pixel value. The algorithm developed here classifies mammograms into normal & abnormal. The proposed scheme can also improve the detection performance significantly with lower false positives rates. A classification accuracy true positive 86.1% is achieved for the database under study (i.e) MIAS Database.
  3. In 2012- PRADEEP N proposed CAD system consists of five principal stages. In the first step, in the input image the ROI is selected by manual cropping process. For the cropped image, image enhancement techniques have been employed to improve the image quality. The segmentation is carried out for the enhanced cropped image by using Local thresholding and Otsu segmentation techniques. Before feature extraction of this segmented image, noise is removed by filling their intensity values by zero. These noises will appear as tiny white spots in the image. Features are calculated for the noise suppressed segmented output by using co-occurrence matrix.

### III.THEORITICAL BACKGROUND

Cancer is an abnormal, continual multiplying of cells. In a normal cell, when DNA gets damaged the cell either repairs the damage or the cell dies. In cancer cells, the damaged DNA is not repaired, but the cell doesn't die like it should. Instead, this cell goes on making new cells that the body does not need. The build up of extra cells often forms a mass of tissue called a lump, growth ,or tumor. **Tumor** is uncontrolled growth of cells which can be either Benign or Malignant. Benign Tumors are not cancerous. Benign tumors may grow larger but do not spread to other parts of the body but still may need to be removed because the local tissue may be damaged. Malignant Tumors are cancerous. Breast cancer is a malignant tumor that starts in the cells of the breast. A malignant tumor is a group of cancer cells that can invade and destroy surrounding tissues and spread (metastazise) to distant areas of the body. Tumor can be easily detected in mammogram as it appears brighter than the rest parenchyma .

**Presence of masses:** A mass is demarcated as a space-occupying lesion seen in more than one projection and is usually characterized by its shape and margin. Benign masses are well circumscribed, compact, and roughly circular or elliptical. Malignant lesions usually have a blurred boundary, an irregular appearance, and sometimes are surrounded by a radiating pattern of linear spicules. However, some benign lesions may have a spiculated appearance or blurred periphery. High radiopaque lesion with irregular or ill-defined boundary should be considered with a high degree of suspicion. Circumscribed masses usually have variable sizes with normal dense tissue causing difficulties for mass detection. The sensitivity of mass detection is generally lower because masses are not usually brighter than the background. Spiculated lesions have a central tumour mass that is surrounded by a radiating pattern of linear spicules. Most spiculated lesions are malignant.

**Mammography :** Mammography is one of the available techniques for the early detection of masses or abnormalities which is related to breast cancer. Mammography is the primary imaging modality used for early detection of clinically occult breast cancer. Despite advances in other breast imaging modalities, including ultrasound and magnetic resonance imaging, mammography is still the method of choice. Mammography has been proven to be the most reliable method and it is the key screening tool for the early detection of breast cancer. Mammography is highly accurate, and is in expensive. Masses appear in a mammogram as fine, granular clusters and also masses will not have sharp boundaries, so often difficult to identify in a raw mammogram which a digital mammogram can easily detect.

**Digital mammography** is a specialized form of mammography that uses digital receptors and computers instead of x-ray film to help examine breast tissue for breast cancer[15]. The electrical signals can be read on computer screens, permitting more manipulation of images to theoretically allow radiologists to more clearly view the results[15].Both

# International Journal of Innovative Research in Science, Engineering and Technology

(An ISO 3297: 2007 Certified Organization)

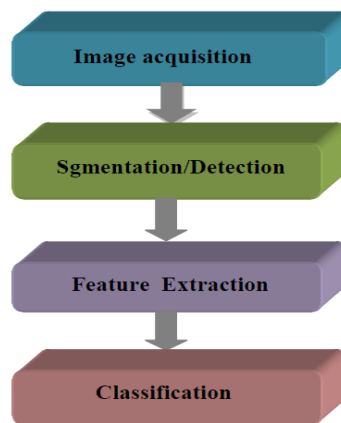
Vol. 4, Issue 6, June 2015

types of mammography involves taking different views of the breast which includes cranial-caudal view(CC), mediolateral-oblique,(MLO), mediolateralview(ML), latero medial(LM).

**Mediolateral Oblique View (MLO):** MLO view is preferred in this project over a lateral 90-degree projection or other views because in medio-lateral oblique (MLO) view mammograms the pectoral muscle is visible as a triangular region of high density at the upper posterior part of the image. Its presence in mammograms poses an additional source of complexity in automated analysis as it may interfere with the results of image processing methods and induce a bias in breast cancer detection. The texture of the pectoral muscle may also be similar to some abnormalities and may cause false positives in the detection of suspicious masses.

## IV. IMAGE PROCESSING METHODOLOGIES

### COMPUTER AIDED DETECTION (CAD SYSTEM)



**Preprocessing methods:** Image pre-processing techniques are necessary, in order to find the orientation of the mammogram, to remove the noise and to enhance the quality of the image. Before any image-processing algorithm can be applied on mammogram, preprocessing steps are very important in order to limit the search for abnormalities without undue influence from background of the mammogram. Breast border extraction and pectoral muscle suppression is also a part of preprocessing. The types of noise observed in mammogram are high intensity rectangular label, low intensity label, tape artifacts etc. The types of noises present in mammogram are represented in Figure

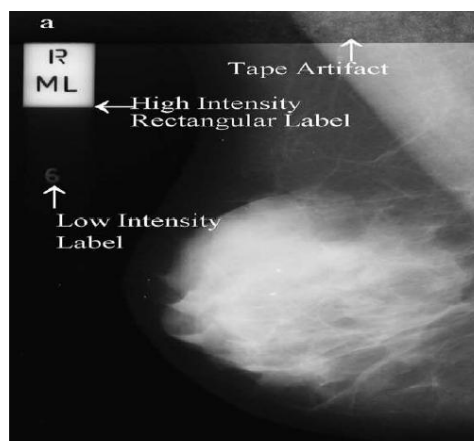


Fig (a) types of noises present in mammogram

## International Journal of Innovative Research in Science, Engineering and Technology

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 6, June 2015

**Gaussian blur:** A Gaussian blur (also known as Gaussian smoothing) is the result of blurring an image by a Gaussian function. The Gaussian blur is a image-blurring filter that uses a Gaussian function for calculating the transformation to apply to each pixel in the image. The equation of a Gaussian function in one dimension is:

$$G(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{x^2}{2\sigma^2}}$$

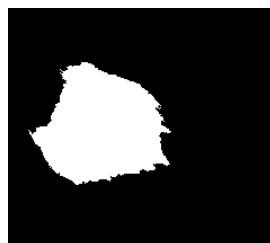
in two dimensions, it is the product of two such Gaussians, one in each dimension:

$$G(x, y) = \frac{1}{2\pi\sigma^2} e^{-\frac{x^2+y^2}{2\sigma^2}}$$

where  $x$  is the distance from the origin in the horizontal axis,  $y$  is the distance from the origin in the vertical axis, and  $\sigma$  is the standard deviation of the Gaussian distribution. When applied in two dimensions, this formula produces a surface whose contours are concentric circles with a Gaussian distribution from the centre point. Values from this distribution are used to build a convolution matrix which is applied to the original image. Each pixel's new value is set to a weighted average of that pixel's neighbourhood. The original pixel's value receives the heaviest weight (having the highest Gaussian value) and neighboring pixels receive smaller weights as their distance to the original pixel increases. This results in a blur that preserves boundaries and edges better than other, more uniform blurring filters.

**Image segmentation:** Image segmentation is an important technology for image processing. The goal of image segmentation is to cluster pixels into salient image regions, i.e., regions corresponding to individual surfaces, objects, or natural parts of objects and to simplify and change the representation of an image into something that is more meaningful and easier to analyse. It is the process of partitioning an image into regions and properties.

Local thresholding technique has been proven to provide an easy and convenient way to perform the segmentation on digital mammogram. The segmentation is determined by a single value known as the intensity threshold value. Then, each pixel in the image is compared with the threshold value. Pixel intensity values higher than the threshold will result in a white spot in the output image as shown in the figure below



Fig(b) segmented image after using thresholding technique

### V. OTSU'S THRESHOLDING

**Otsu's method** is used to automatically perform histogram shape-based image thresholding or, the reduction of a gray-level image to a binary image. The algorithm assumes that the image to be thresholded contains two classes of pixels or bi-modal histogram (e.g. foreground and background) then calculates the optimum threshold separating those two classes so that their combined spread (intra-class variance) is minimal.

## International Journal of Innovative Research in Science, Engineering and Technology

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 6, June 2015

Otsu's method forms the thresholded image of the auto cropped ROI by minimizing the intra-class variance and maximizing the interclass variance. The histogram is constructed with  $y_i$  as the number of pixels of gray-level  $i$ ,  $n$  being the maximum gray-level (255 in an 8-bit image). Let the threshold be denoted by  $t$ . The following partial sums are used in this method:

$$A_j = \sum_{i=0}^j y_i \quad \text{for } j=0,1,2,\dots,255$$

$$B_j = \sum_{i=0}^j iy_i$$

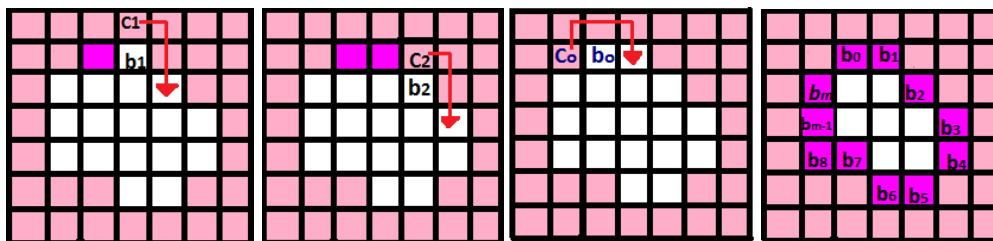
$$\mu_t = \frac{B_t}{A_t}$$

$$v_t = \frac{B_n - B_t}{A_n - A_t}$$

Where,  $\mu$  and  $v_t$  are the centroid of the left and right portion of the histogram separated by threshold  $t$  whose value is obtained from  $j$ .

### VI. MOORE'S ALGORITHM

This algorithm traces the object boundary for binary images whose output is an ordered sequence of boundary points as shown in Fig5.2.2.(2)(a). Given, a binary region  $R$  or its boundary let the starting point,  $b_0$ , be the uppermost, leftmost point in the image that is labelled  $c_0$  i.e. the west neighbour of  $b_0$ . Clearly,  $c_i$  is always a background point. Examine the 8-neighbours of  $b_i$ , starting at  $c_i$  and proceeding in a clockwise direction. This procedure is repeated until  $b=b_0$ .



The steps involved in artefacts removal using Otsu thresholding and Moore's boundary tracing algorithm are as follows:

- a. An orientation check is performed by thresholding the image and calculating the length of white pixel region of the first row from either side upto first white to black transition. Left sided mammograms gives larger length for left side pixel count and vice versa. A flipping action is further carried out if right sided pixel count is higher. This step results a generalized set of left oriented mammograms.
- b. The black border is removed from either side of the mammogram by noting the first black to white transition.
- c. Further, the breast boundary is traced by applying Moore's algorithm on the left- sided Gaussian blurred, Otsu thresholded image.
- d. Finally, all the radiopaque background artefacts like high intensity labels, low intensity labels etc., are eliminated by making the pixels zero outside the boundary traced resulting in breast region localisation, as in Fig.

## International Journal of Innovative Research in Science, Engineering and Technology

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 6, June 2015

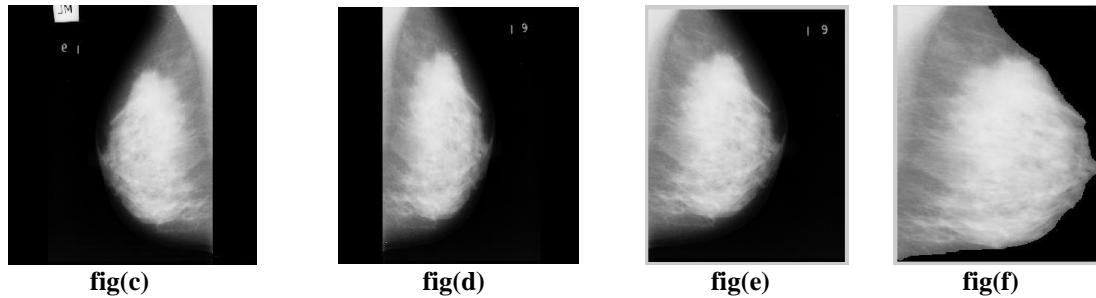


Fig: Step-wise outputs of breast region localisation process with:  
(c) Original digital mammogram(mbd003.pgm)  
(d) Flipped mammogram  
(e) Black border removed mammogram  
(f) Localised breast region with all radiopaque artefacts removed.

### VII. PECTORAL MUSCLE REMOVAL AND AUTO-CROPPING OF ROI(REGION OF INTEREST)

The Pectoral muscle suppression step is carried out with the help of Otsu thresholding and slope detection. The steps involved further applied on the localised breast region obtained in previous step are as follows:

(i) The regions with intensity levels close to zero or zero are filled with the mean value of the image.

$$mean\_int = \frac{\sum_{i=1}^c \sum_{j=1}^r img_{i,j}}{N}$$

Where,  $img$ =intensity level of the image for  $i$ th row  $j$ th column pixel.

$N$ =total number of pixels in the image.  $img_{i,k}$ = $mean\_int$

Where,  $\xi < 10$  (very low intensity level).

Then, a thresholding action is applied resulting a binary image  $bw$  which mainly comprise of two thresholded regions, neoplasm and the pectoral region.

ii) End-points of pectoral line are evaluated by noting the first white to black transition for first row and column respectively.

$col=k$ , such that  $img_{1,k}=0$ , Where,  $col$ =column index of the first row for white to black transition.

$row=i$ , such that  $img_{i,1}=0$ , Where,  $row$ =row index of the first column for white to black transition.

iii) Further, the pectoral line co-ordinates are calculated with the help of slope and intercept evaluated from the end-points.

Slope=  $col/row$

where ;  $line\_cl=slope*line\_rw+col$

Where,  $line\_rw$ = row coordinates of the pectoral line, *i.e.*

$line\_rw=1,2,3,\dots, row$ .

$line\_cl$ = column coordinates of the pectoral line.

(iv) Once the co-ordinates are known, the region on the left of the co-ordinates is made black resulting in pectoral region suppression. Further, thresholding of pectoral removed gray level mammogram, results in a binary image comprising only the breast region *i.e.*  $bw\_PecR$ .

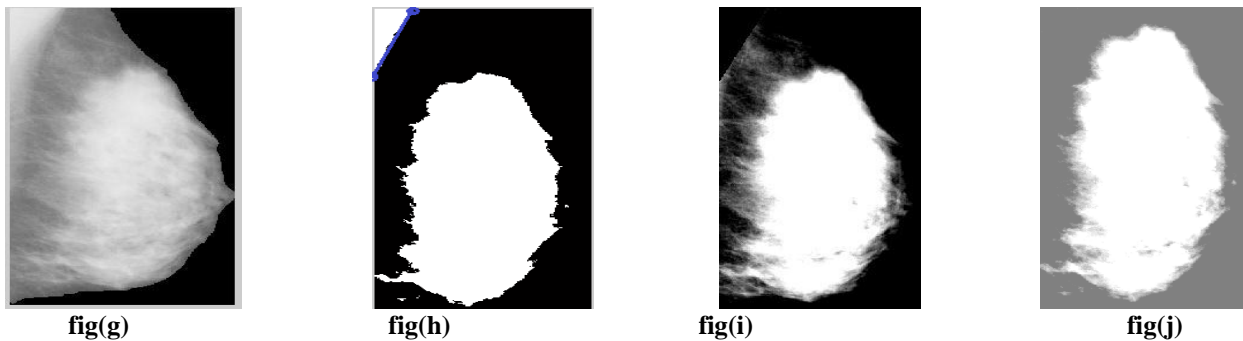
(v) The pectoral muscle removed binary picture  $bw\_PecR$  is then superimposed on the original mammogram image to preserve the essential edges lost in Gaussian blurring. This gives the global segmented ROI.

$globl\_ROI = bw\_PecR * ori\_img$

(vi) Further the global RoI is Otsu thresholded. Morphological filtering is applied to the thresholded image to remove the small area objects leaving behind the largest area object *i.e.* the neoplasm area.



(vii) Finally a rectangular area containing the lesion is cropped out resulting in a précised RoI.



**figures:** Step-wise outputs of breast region localisation process with (g)flipped, cropped and artifact removed image(mdb003.pgm) (h) Otsu thresholded image with slope of pectoral muscle detected (i)Pectoral muscle removed image (j) RoI extracted image.

### VIII. FEATURE EXTRACTION

#### Shape based features:

Several techniques have been developed for characterizing the shape of masses in an image. The important shape based features includes the geometric parameters such as speculation, relative contrast, circulatory, compactness and radial angle are used for classifications.

**Speculation:** Firstly, the centroid is calculated. Euclidean distances are found from the centroid to every 20th pixel on the boundary. Now these radial distances are scanned sequentially and if the difference between the consecutive distances is greater than 5 then the speculation count is incremented by 1.

**Circulatory:** Owing to the fact that the roundness is one of the criteria of benign masses, the probability of masses as being benign is higher when circulatory is higher. The circulatory indicates the circular degree of masses. Here a circle CEq having the same area as the mass is first located at the centre of the mass. The circulatory is found from  $S = \sum 1(x,y) = R$ , Where S is the area of the mass and the centre of the mass, (x,y) is given by:  
 $(\text{perimeter})^2 / 4\pi * \text{area}$

The value of circulatory feature lies between 0 and 1. If the ratio is 1, this means that the mass matches exactly with the Eq C, implying the mass is circular. On the other hand, if the circulatory is much smaller than 1, this implies the mass is far from being circular.

**Compactness:** It is calculated from the perimeter and area of the segmented portion as:  
 $(\text{perimeter})^2 / \text{Area}$

### IX. CLASSIFICATION

Classification into Benign/Malignant is done based on shape features using SVM classifiers:

In SVM first, the input vectors are mapped into a feature space (possible with a higher dimension), either linearly or non-linearly, which is relevant with the selection of the kernel function. Then, within the feature space, a hyperplane is constructed which separates two classes. The choice of SVM is made over other classifiers because SVM training always seeks a global optimized solution and avoids over-fitting, so it has the ability to deal with a large number of features. In the present work Linear type of SVM classifier has been chosen for Benign and Malignant classification where, for example  $\{(x_1, c_1), (x_2, c_2), \dots, (x_n, c_n)\}$  are the training instance-label pairs with  $c_i$  either 1 or -1, a constant denoting the class to which the point  $x_i$  belongs. Each  $x_i$  is a  $p$ -dimensional real vector, with  $p$  being the number of features selected. The separating hyperplane can be represented by the equation

$$x \cdot w - b \leq -1 \quad x \cdot w - b = 0$$

Where,  $w$ =vector normal to the hyperplane

$b$ =bias, or the distance of the hyperplane from the origin

# International Journal of Innovative Research in Science, Engineering and Technology

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 6, June 2015

**Validation measures** like sensitivity and specificity are the standard validation or accuracy check of the robustness of the proposed techniques or algorithms. In the proposed CAD system, for detection of masses it has been grouped into two classes: one class is malignant or the abnormal class and the other is the normal class. The following definitions are used to describe correct classification or misclassification for each class:

- (i) True Positive (TP) = An abnormal classified as abnormal
- (ii) True Negative (TN) = A normal classified as normal
- (iii) False Positive (FP) = A normal classified as abnormal
- (iv) False Negative (FN) = An abnormal classified as normal

The performance criteria of a radiologist or a diagnostic system are generally appraised by two indices: **sensitivity** and **specificity**.

**Sensitivity or True Positive Fraction (TPF):** The sensitivity of a recognized test is the fraction of number of abnormal classified as abnormal over total number of abnormal cases.

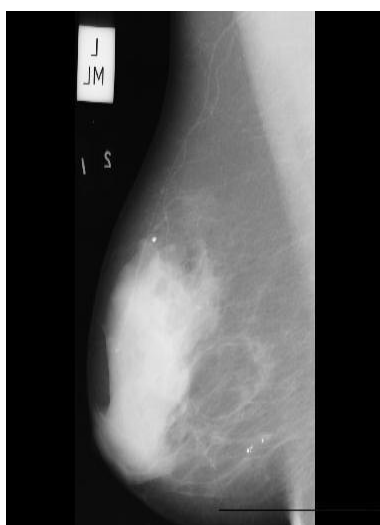
**Specificity or True Negative Fraction (TNF):** The specificity of a test is the fraction of healthy cases over the total of normal case.

The goal of normal mammogram identification is to maximize TNF with very low FNF, which is the same as minimizing FPF with very high TPF.

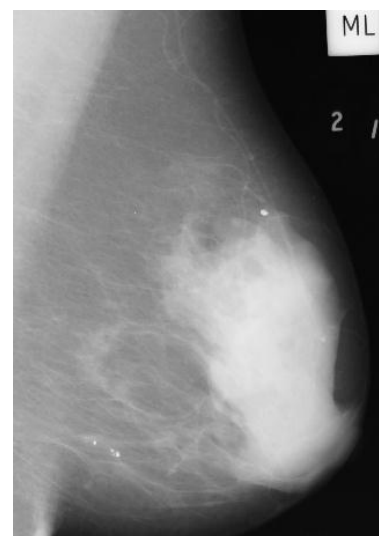
## X. RESULTS AND DISCUSSIONS

### PRE-PROCESSING:

Pre-processing step includes black border removal and flipping of the image (from right to left) to generalize the pre-processed dataset. After black border removal, the image comprise of various radiopaque background artefacts like high intensity labels, low intensity labels etc., are eliminated by making the pixels zero outside the boundary traced resulting in breast region localization



Fig(i) original mammogram



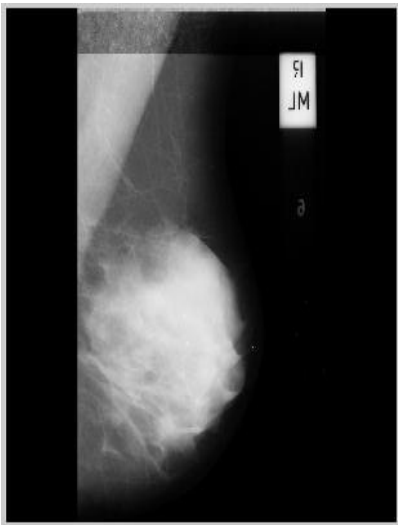
Fig(ii) flipped mammogram



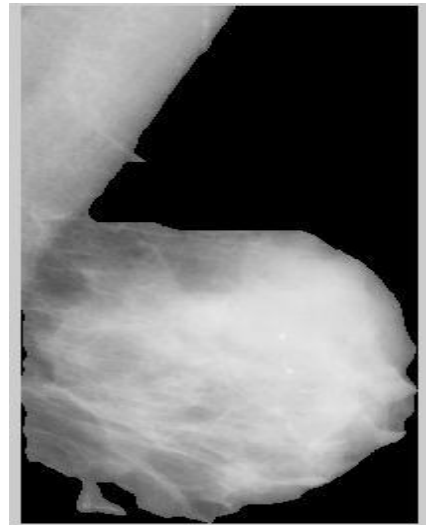
## International Journal of Innovative Research in Science, Engineering and Technology

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 6, June 2015

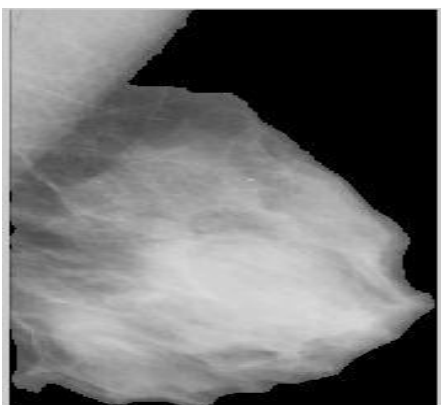


fig(iii): Mammogram with black border and tape artefacts

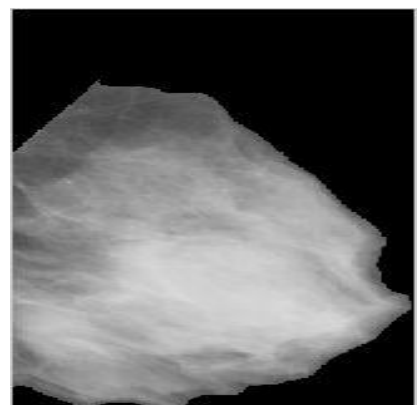


fig(iv): Label removed and boundary extracted image

After label removal step the high intensity portion remaining except the lesion is the pectoral muscle. Pectoral muscle removal is the crucial pre-processing step for proper lesion segmentation which is successfully obtained as shown in fig



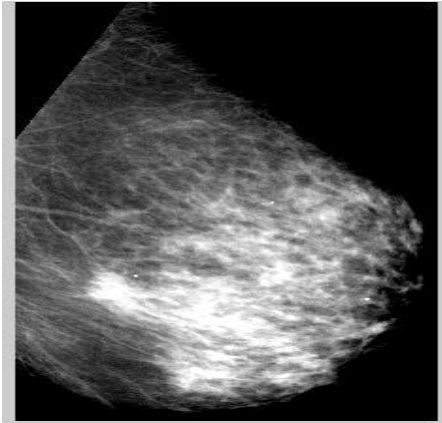
fig(v): Label removed mammographic image(mdb001.pgm)



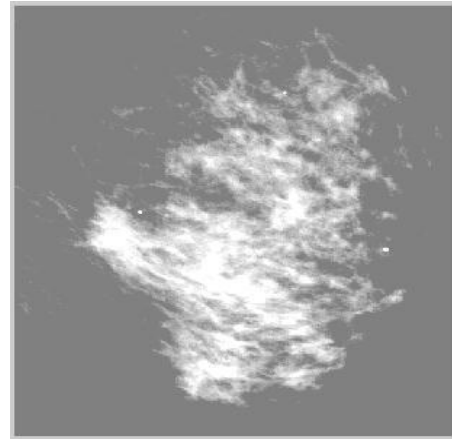
fig(vi): pectoral muscle removed image

### ROI Extraction

After successful removal of pectoral muscle and other artefacts, contrast of the image is enhanced so that the relevant region or the region of interest (ROI) can be easily extracted using the automatic cropping algorithm. The Figures below shows a set of pectoral muscle removed images, with their corresponding ROI extracted image fill with the mean value of the image.

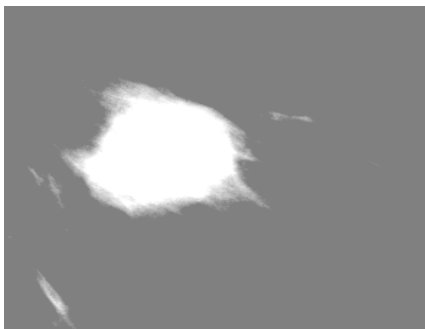


Fig(vii) : Pectoral muscle removed mammogram  
mdb058.pgm



fig(viii): Region of interest extracted mammograms

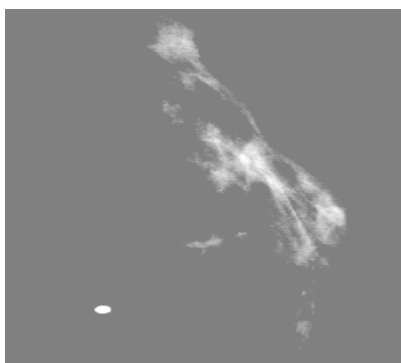
**SEGMENTATION USING OTSU'S METHOD**



**fig(ix)mdb028.pgm**



**fig(x)**



**fig(xi)mdb075.pgm**



**fig(xii)**

fig(ix, xi) :ROI extracted image(malignant)  
fig(x, xii):Corresponding otsu threshold image

## International Journal of Innovative Research in Science, Engineering and Technology

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 6, June 2015

### Extraction of features:

1. Shape based features like Speculation, Compactness and Circulatory ratio are evaluated for 86 masses(abnormal) and 20 non-masses(normal) from otsu segmented images.
2. Shape features (Speculation, Compactness and Circulatory ratio)for set containing mass set i.e. malignant and benign candidates combined. Shape based features of masses evaluated for 49 benign and 37 malignant from otsu' segmented image.

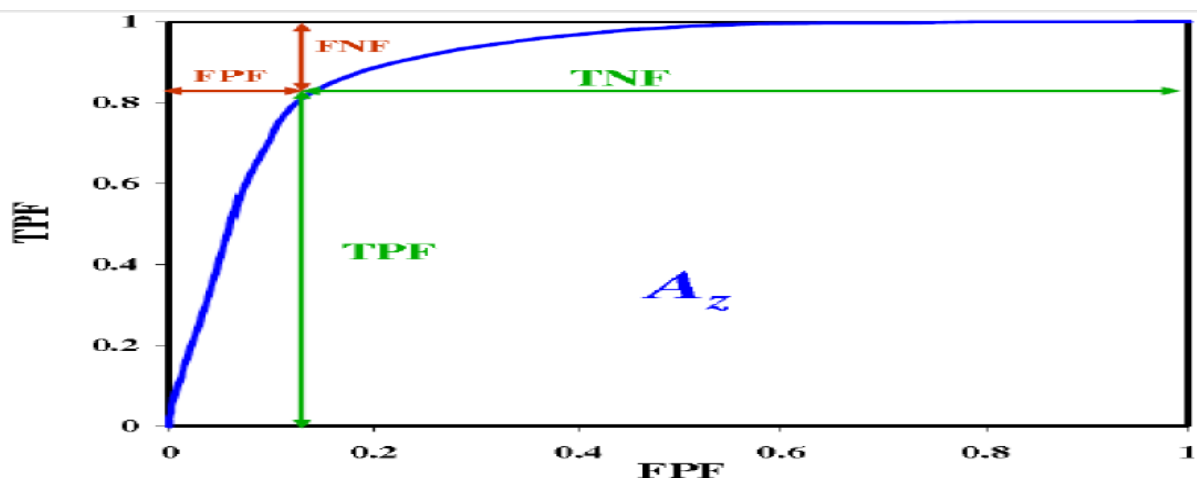


Fig : A General ROC Curve with TPF, FPF, TNF, and FNF Labelled

### Classification and validation measure

Classification is performed to step up the sensitivity and specificity by initially classify and screen out the non- mass candidates from mass candidates using shape feature. Then from remaining set of mass candidates the malignant neoplasm are detected and segregated from benign type of neoplasm using shape features. SVM classifier is used for classification and the classification results are cross-validated using hold-out type cross-validation method. Cross-validation is performed over twenty nine iteration and the corresponding validation measures for the algorithm are evaluated along with their average.

Classification and validation measures using the shape based features evaluated over the Otsu based segmented image.

Table I. Maximum values of SVM Classification Validation Measures Based on Shape Based Features Of mass and non-mass.

Maximum Accuracy	0.7500
Maximum specificity	0.2857
Maximum sensitivity	1.000
Average sensitivity	0.8571

# International Journal of Innovative Research in Science, Engineering and Technology

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 6, June 2015

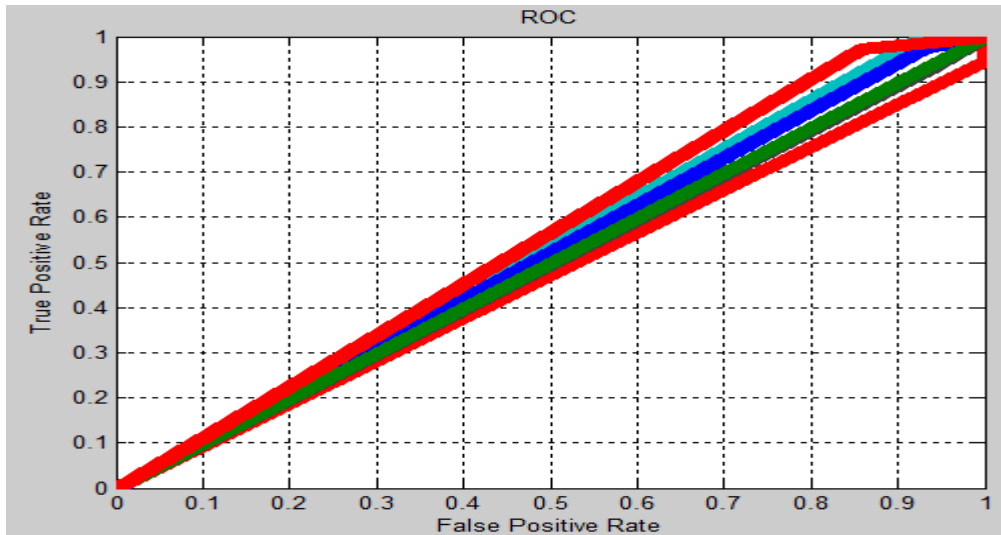


Fig: ROC Curves over twenty nine Iterations for Shape Based SVM Classification of Mass and Non-mass.

Table II. Maximum values of SVM Classification Validation Measures Based on Shape Based Features Of malignant and benign mass.

Maximum Accuracy	0.7143
Maximum specificity	0.2143
Maximum sensitivity	1.000
Average sensitivity	0.8214

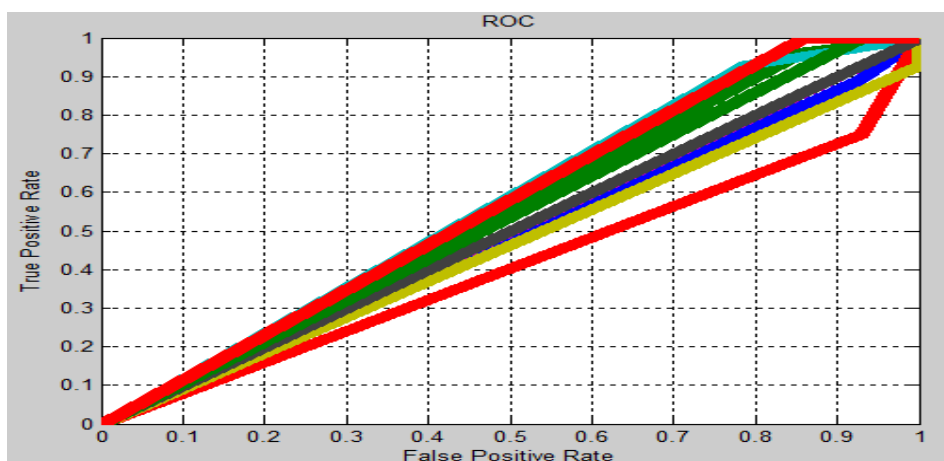


Fig: ROC Curves over twenty nine Iterations for Shape Based SVM Classification of Malignant/ Benign mass.

# International Journal of Innovative Research in Science, Engineering and Technology

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 6, June 2015

## Discussion

The result of proper extraction of ROIs proves the successful execution of pre-processing steps like black border, label and pectoral muscle removal which has able to reach accuracy level of 95%. Further to screen out the non-mass candidates from the mass ones as well as malignant mass from benign masses, segmentation and classification is carried out. SVM classification is performed based on shape features namely speculation, compactness, circulatory ratio. Maximum sensitivity of 100% has been achieved for twenty nine iteration of hold out cross-validation processes.

## XI. CONCLUSION

Results of proper extraction of ROIs prove the successful execution of pre-processing steps like label removal and pectoral muscle removal. Overall accuracy for extraction of region of interest is 95%. For higher detection rate of malignant neoplasms the non-mass candidates are successfully screened out from mass candidates based on their shape features exhibited, then malignant mass are screened out from benign masses based on shape based features. For better detection of Malignant and Benign, Gaussian blurring is applied along with Otsu's thresholding method. SVM classification is performed on three types of shape based feature sets, which includes, speculation, circulatory ratio and compactness which has able to reach sensitivity rate of 100%.

## XII. ACKNOWLEDGEMENT

This research work is made under the guidance of Prof.. Pranob Haloi (Assistant professor) of JEC, Assam. I, Prarthana Priyadarsini would like to thank my guide and my teachers & my dear friends for helping me in this research work and for their guidance and constant encouragement towards the work

## REFERENCES

- [1] WHO Cancer Fact Sheets. (2009). <http://www.who.int/mediacentre/factsheets/fs297/en/index.html>
- [2] American Cancer Society, "Cancer Facts and Figures 2003", Atlanta, GA: American Cancer Society, 2003.
- [3] BI-RADS Lexicons [http://www.acr.org/departments/stand\\_accred/birads](http://www.acr.org/departments/stand_accred/birads)
- [4] Rafael C. Gonzalez, Richard E. Woods, "Digital Image Processing," Pearson Education (Singapore) Pte Ltd, Indian Branch, 482 F. I. E. Patpargang, Delhi 110092, India 2004.
- [5] M. Arfan Jaffar<sup>1</sup>, Bilal Ahmed<sup>1</sup>, Ayyaz Hussain<sup>1</sup>, Nawazish Naveed<sup>1</sup>, Fauzia Jabeen<sup>2</sup> and Anwar M. Mirza<sup>1</sup> "tumor detection in digital mammogram images" Fourth International Conference on Innovative Computing, Information and Control.(2009)
- [6] Mohammed A. Alolfe , Wael A. Mohamed, Abou-Bakr M. Youssef, Ahmed S. Mohamed and Yasser M Kadah "computer aided diagnosis (CAD) system based on combined support vector (SVM) and Linear discriminant analysis(LDA) classifier for detection and classification of breast cancer in digital mammograms" Biomedical Engineering Department,Cairo university.(2009)
- [7] Y.Ireaneus Anna Rejani+, Dr.S.Thamarai Selvi proposed a method which includes the mammogram image was filtered with Gaussian filter based on standard deviation and matrix dimensions such as rows and columns.(2009)
- [8] N. Saltanat, M. A. Hossain, Member IEEE and M. S. Alam School of Computing, Informatics and Media University of Bradford, BD7 1DP, UK." An Efficient Pixel Value based Mapping Scheme to Delineate Pectoral Muscle from Mammograms"EU Erasmus Mundus Project - eLINK (east-west Link for Innovation, Networking and Knowledgeexchange) under External Cooperation Window - Asia Regional Call (EM ECW - ref. 149674-EM-1-2008-1-UK-ERAMUNDUS). (2010)
- [9] Mohamed Meselhy Eltoukh , Ibrahim Faye, Brahim Belhaouari Samir Electrical and Electronics Engineering Department,2Fundamental and Applied Sciences Department,Universiti Teknologi PETRONAS,Perak, Malaysia "Automatic Detection of Breast Masses in Digital Mammograms Using Pattern Matching" 2010 IEEE EMBS Conference on Biomedical Engineering & Sciences (IECBES 2010), Kuala Lumpur, Malaysia, 30th November - 2nd December 2010.(2010)
- [10] Dheeba.J Research Scholar,Dept. of Computer Science and Engineering Anna University, Tirunelveli, TN, India"Classification of Malignant and BenignMicrocalcification Using SVM Classifier"Proceedings of ICETECT 2011.(2011).
- [11] PRADEEP N ,Department of Computer Science and Engineering, Bapuji Institute of Engineering and Technology,Davangere, Karnataka, INDIA." Segmentation and Feature Extraction of Tumors from DigitalMammograms" Computer Engineering and Intelligent Systems ISSN 2222-1719 (Paper) ISSN 2222-2863 (Online) Vol 3, No.4, 2012(2012)
- [12] R.B Dubey & Hanmandlu "A CAD System for detection and classification of Mammograms", Transactions on Neural Network,IEE Computational Intelligence Society,TNN-2011-P-2845. (2011)
- [13] Mini-MIAS database of mammograms. <http://peipa.essex.ac.uk/info/mias.html>
- [14] US breast cancer statistics(2012). [http://m.breastcancer.org/symptoms/understand\\_bc/statistics](http://m.breastcancer.org/symptoms/understand_bc/statistics)
- [15] Factsheet on mammogram, by National Cancer institute. [www.cancer.gov/cancer-topics/factsheet](http://www.cancer.gov/cancer-topics/factsheet)
- [16] Mammography, Wikipedia. <https://en.wikipedia.org/w>
- [17] Diagnostic mammography, Breast Cancer Screening Imaging Exam San Diego, Imaging Healthcare files.