JOURNAL OF CRITICAL REVIEWS
ISSN- 2394-5125 VOL 7, ISSUE 15, 2020

Detection Of Region Of Interest And Stages Using Digital Breast Tomosynthesis Images

Veena M
Assistant professor CSE, PESCE Mandya

M C Padma
Professor & HOD CSE, PESCE Mandya

Dinesh M S
Honarary Professor, PESCE, Mandya

Received: 14 April 2020 Revised and Accepted: 8 August 2020

ABSTRACT
Breast cancer is one of the most prominent cancers in medical diagnosis which is increasing day by day in every year. It also leads to the women's death all over the world. In this paper we are extracting the ROI and identifying stages of Breast Cancer using the Digital Breast Tomosynthesis [DBT] images. Knowing breast cancer in early stages which helps to recover and take a better treatment. This modality is very convenient for the doctors and radiologist to detect or to conduct further procedure.

Keywords- Breast Cancer, Region of Interest [ROI], DBT Images, SVM Classifier, and Gray level Co-occurrence Matrix [GLCM].

1. INTRODUCTION

Breast Cancer is one of the dangerous diseases which are growing rapidly day by day. Breast Cancer can be found in different parts in the Breast. Usually Breast Cancer begins in the ducts which carry milk to the nipple. Sometimes it occurs in the glands that makes breast milk.

Breast Cancer is most common in United State. Usually Breast Cancer is found in older women, and rare in the younger below 40 ages. But research team in US has found 11% of new cases in younger women. Breast Cancer stages are expressed in terms of number on scale 0 – 4. Where stage 0 states that cancer is non-invasive that remain within their original location. Whereas stage 4 express as an invasive which spreads to other part of the body.

The doctor reports the stages of the Breast Cancer in the form of TNM system [9].

Tumor [T]:- This state that how large tumor and where it is located.

Node [N]:- This state that whether the tumor is spread to lymph nodes, if it is spread, than where and how many.

Metastasis [M]:- This state that if cancer is spread to other parts of the body, if it is than where and how much.

Tumor [T]: The stages is split into smaller groups which helps to know even better, they are

T0:- There is no proof of cancer.

T1:- Tumor ≤20mm in sizeable.

T2:- Tumor >20mm but ≤50mm in sizeable.

T3:- Tumor >50mm but ≤80mm in sizeable.

T4:- Tumor >80mm in greatest dimension and also it spread to other parts.

Node [N]: The Regional lymph node [N] includes some following stages, they are

N0:- There is no cancer located in lymph node.

N1:- The lymph node >0.2mm but ≤2mm which spread to 1 to 3 axillary lymph node and it is also called as "Micro metastatic".

N2:- This cancer spread to 4 to 9 axillary lymph nodes.

N3:- This cancer spread to 10 or more axillary lymph node.

Metastasis [M]: The 'M' includes some stages which describe more detail, they are

M0:- This cancer is not metastasized.

M0 (i+):- No clinical or radiographic proof of metastases. The size of this tumor is larger than 0.2mm.

M1:- There is a proof of metastasis.

Digital Breast Tomosynthesis overcome the limitations of the conventional mammogram. Using DBT images ROI and different stages of cancer is detected by using GLCM feature extraction method and classified using multi SVM classifier.

2. LITERATURE SURVEY

As specified by the paper [1] in this approach they have four modules for knowing the ROI and stages of breast cancer. They are Preprocess, Segmentation, Extraction and Classification. The image is done with micro calcification cluster in this method we use sobel edge detection mask, after that using Baye's classifier to say the malignant cancer belongs to which stage.

According to paper [2] the image is preprocessed to improve the quality. The Feature extraction of LBP texture feature, intensity value, texture, and shape and Eigen faces approach is done to the preprocessed image. The final step of that image is classification this is done using support vector machine, decision tree and KNN are used.

According to paper [3] in this survey they are using machine learning techniques to find the breast cancer. After training the data they do preprocessed for that image. The next step is feature extraction is done. In this step they perform one way ANOVA [Analysis Of Variance]. The final stage of this is WBCD [Wisconsin Breast Cancer (Diagnostic) Dataset] using SVM, KNN and Decision tree.

According to paper [4] they are used Magnetic Resonance Image [MRI] for breast image, where the breast image consists of chest wall, skin, fatty tissues and glandular tissue. The one of the problem during detection is the big clutters from the tissues in breast skin. To remove this clutter we make use of tumor-free model. To find the position of breast tumor we make use of Confocal Microwave Imaging (CMI) algorithm.

According to paper [5] mammogram image is the breast is taken for the action. The image is preprocessed and done with the ROI extraction with the size 128*128 pixels after that Dual-CT rotting is carry out on the extracted ROI image to get Dual-CT coefficient. For this image feature extraction and analysis is done to know the difference between benign and malignant ones and abnormal and normal ones. The final step is improved KNN classification is used for the extraction image.

As specified by the paper [6] the image is preprocessed the noises and label the image and remove the background in mammogram image. Next ROI selection is done for the tumor region. This tumor is extracted using Gray Level Co-occurrence Matrix [GLCM] based texture analysis.SVM classifier is used.

According to paper [7] in this proposed method is done by using Mammographic Image Analysis Society [MIAS] dataset. This image is of gray scale file format. The first step is to image acquisition. The real MIAS database has been decreases to 200 micron pixel edge and padded so every image is 1024 pixel * 1024 pixel. After that image preprocessed is done and then find the ROI of that image. GLCM features are extracted.

According to paper [8] in mammograms image is preprocessed is done by detecting ROI extracted with normalization by finding normal or abnormal image by applying DCT transform. The features are extracted and improved KNN algorithm is used for classification.

It has been observed from the review that modalities used for revealing breast cancer has its own limitations. Still there is a more scope for improving the accuracy of the results for dense breast with overlapping of breast tissue. With this context the modality DBT is proposed for detecting ROI and different stages of breast cancer for the dense breast and the system architecture is shown in fig 1.

3. PROPOSED METHOD

Our work has been divided into 3 models.

1. Image preprocessing
2. Image segmentation
3. Feature extraction and classification

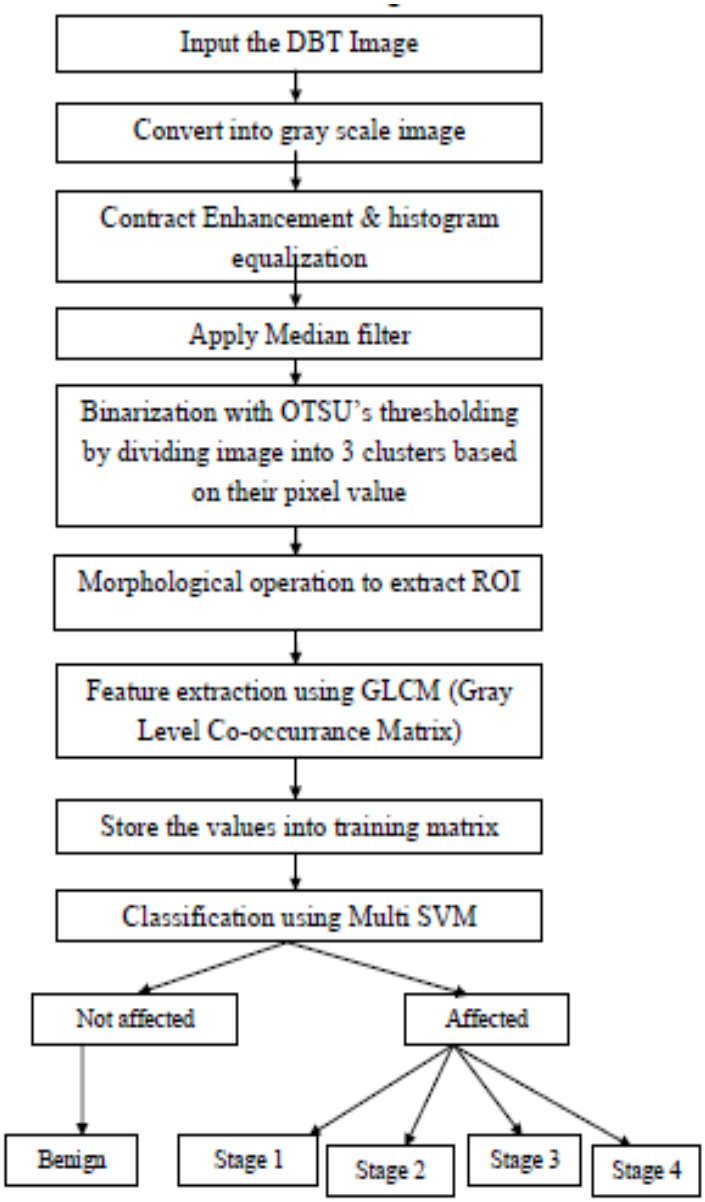


Fig 1: OVER ALL SYSTEM ARCHITECTURE

3.1 Image Preprocessing

The DBT images are preprocessed to remove the noisy element to get better quality of the image. To de-noise the image we are using median filter and then resize it and convert that image to gray scale image as shown in fig 2.

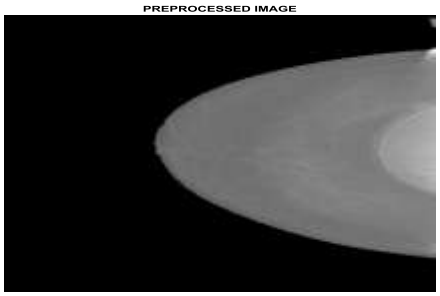


Fig 2: PREPROCESSED IMAGE

3.2 Image segmentation

OTSU method is used for segmentation OTSU Thresholding performs well for elongated / rectangular faults. To discard background image we are using morphological operation. After the segmentation process the image is collected in database for the further process of ROI extraction as shown in fig 3.

There are two types of Thresholding Methods

Global Thresholding: Global Threshold values rely on gray values.

Local Thresholding: It divides original image into several sub regions

OTSU is a type global threshold depends only on gray values of image. The pre-processes image is divided into three clusters namely black, gray & white. Based on the pixel the cluster is divided into three parts.

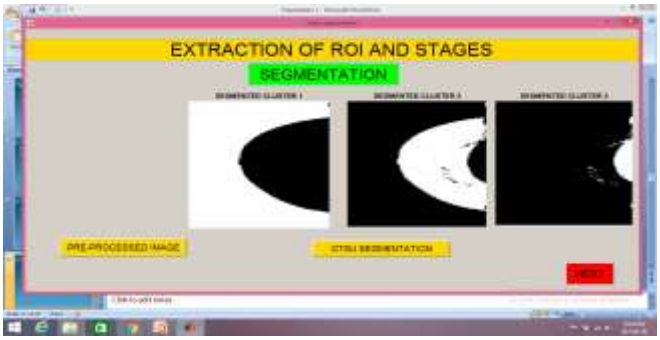


Fig 3: OTSU SEGMENTATION IMAGE

Morphological image processing:

This operation is used for improving the thickness of border of the image. We have four types of Morphological operation; they are Dilation Erosion Opening Closing. For our work we are using Dilation and Closing operation as shown in Fig 4 and the extracted ROI is shown in Fig 5.

Dilation: Dilation builds on object huge by adding pixels around its edge. This procedure depends on the structuring element. Dilation of an Image 'A' by a structured element 'B' is written by **A⊕B.**

$$g(x, y) = \begin{cases} 1 \text{ if 'B' hits 'A'} \\ 0 \text{ Otherwise} \end{cases}$$

Closing: This operation generally removes small holes and fills gaps in the contour and it is denoted by A•B as shown in Fig 4.

$$\mathbf{A \bullet B = (A \oplus B) \ominus B}$$

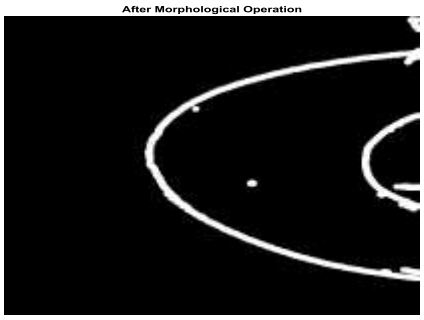


Fig 4: MORPHOLOGICAL OPERATION OF THE BREAST.

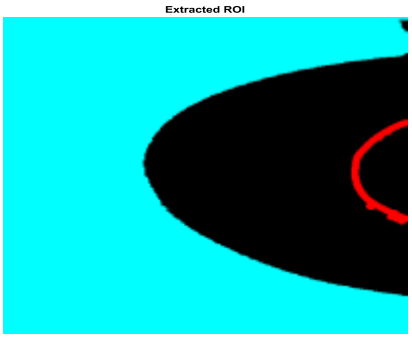


Fig 5: EXTRACTION OF ROI

3.3 Feature extraction and classification

For feature extraction we are using GLCM. This algorithm has many parameters like area, shape, texture and density features. After locating the ROI, the classification of the cancer is recognized. GLCM features with the formula are shown in Table 1.

Table.1. FEATURE EXTRACTION FORMULAE

Sl.No	GLCM Feature	Formula
1.	*Contrast*	$\sum_{i,j=0}^{N-1} P_{i,j}(i-j)^2$
2.	*Correlation*	$\sum_{i,j=0}^{N-1} P_{i,j}\left[\frac{(i-\mu_i)(j-\mu_j)}{\sqrt{(\sigma_i^2)(\sigma_j^2)}}\right]$
3.	*Dissimilarity*	$\sum_{i,j=0}^{N-1} P_{i,j}\lvert i-j\rvert$
4.	*Energy*	$\sum_{i,j=0}^{N-1} P_{i,j}^2$
5.	*Entropy*	$\sum_{i,j=0}^{N-1} P_{i,j}(-\ln P_{i,j})$
6.	*Homogeneity*	$\sum_{i,j=0}^{N-1} \frac{P_{i,j}}{1+(i-j)^2}$
7.	*Mean*	$\mu_i = \sum_{i,j=0}^{N-1} i(P_{i,j})\ ,\ \mu_j = \sum_{i,j=0}^{N-1} j(P_{i,j})$
8.	*Variance*	$\sigma_i^2 = \sum_{i,j=0}^{N-1} P_{i,j}(i-\mu_i)^2\ ,\ \sigma_j^2 = \sum_{i,j=0}^{N-1} P_{i,j}(j-\mu_j)^2$
9.	*Standard Deviation*	$\sigma_i = \sqrt{\sigma_i^2}\ ,\ \sigma_j = \sqrt{\sigma_j^2}$

4. METHODOLOGY

Algorithm for Extracting ROI and finding the Stages of Breast Cancer:

Purpose: Finding ROI and different stages.

Input: DBT Image.

Output: Stages of Breast Cancer with ROI.

1. Select the DBT Image
2. Apply Histogram Equalization and convert RGB image to gray scale for more convenient.
3. Apply Median Filter to remove noise elements to get high frequency pixels.
4. Apply Morphological operation to the preprocessed image to get ROI.
5. Compute Histogram and Probabilities of each intensity level.
6. Set up initial $w_i(0)$ and $\mu_i(0)$
7. Set all possible thresholds t=1, Maximum intensity.

1. Update w_i and μ_i
2. Compute Area

8. Desired threshold corresponds to the maximum Area and minimum Area.
9. Desired threshold = multiply by constant 2.
10. The given input has divided into 70:30 for training and testing image.
11. Based on the features the image is classified as T1, T2, T3 stages and benign using multi SVM.
12. Training set represented by {(**x**1,y1), . . . , (**x**l ,yl)} of cardinality l, where **x***i* ∈Rd and

yi∈ {1, . . . ,k}, the formulation proposed is given as follows

$$\min_{\mathbf{w}_m \in \mathcal{H}, b \in R^k, \xi_i \in R^{l \times k}} \quad \frac{1}{2}\sum_{m=1}^{k} \mathbf{w}_m^T \mathbf{w}_m + C \sum_{i=1}^{l} \sum_{t \neq y_i} \xi_{i,t}$$

$$\text{subject to} \quad \mathbf{w}_{y_i}^T \varphi(\mathbf{x}_i) + b_{y_i} \geq \mathbf{w}_t^T \varphi(\mathbf{x}_i) + b_t + 2 - \xi_{i,t},$$

$$\xi_{i,t} \geq 0,$$

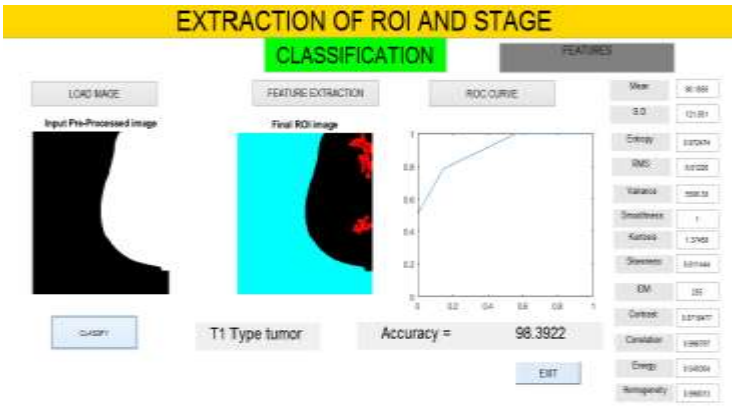
$$i = 1, \ldots, l, \; t \in \{1, \ldots, k\} \setminus y_i.$$

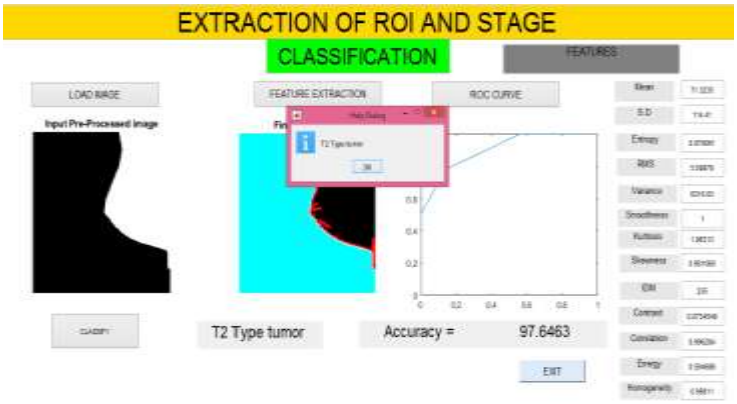
The resulting decision function is $\text{Argmax}_m f_m(\mathbf{x}) = \text{argmax}_m (\mathbf{w}T_m\phi(\mathbf{x}) + b_m)$

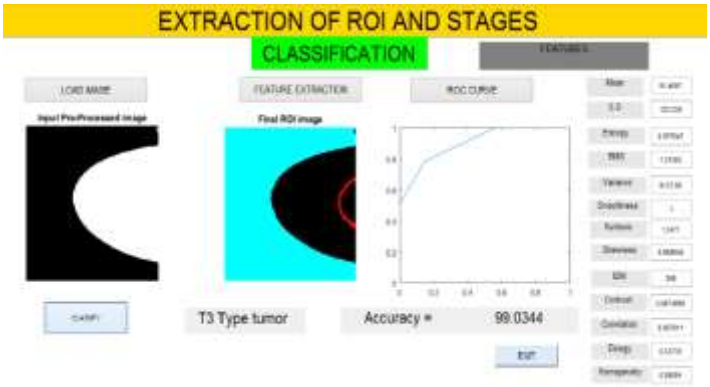
13. Based on the maximum and minimum pixel value ROC curve is plotted.
14. Based on the ROC curve accuracy is calculated.

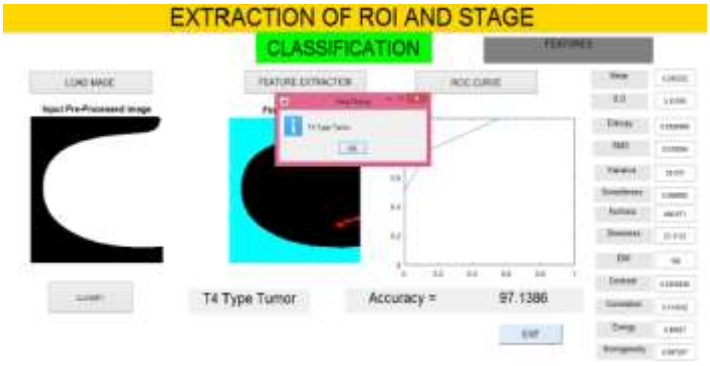
5. EXPERIMENTL RESULTS AND DISCUSSION:

The outcome is based on the DBT image. The stages of Breast Cancer are obtained on the basis of Region of Interest (ROI). The data set which we are used to find out ROI and Stages is publically available in UPMC Breast Tomography and FFDM collection database of tomosynthesis images. The result is shown in Fig 6.


EXTRACTION OF ROI AND STAGE
CLASSIFICATION
FEATURES
LOAD IMAGE
FEATURE EXTRACTION
ROC CURVE
Input Pre-Processed image
Final ROI image
T1 Type tumor
Accuracy =
98.3922
EXIT


EXTRACTION OF ROI AND STAGE
CLASSIFICATION
FEATURES
LOAD IMAGE
FEATURE EXTRACTION
ROC CURVE
Input Pre-Processed image
T2 Type tumor
Accuracy =
97.6463
EXIT


EXTRACTION OF ROI AND STAGES
CLASSIFICATION
FEATURES
LOAD IMAGE
FEATURE EXTRACTION
ROC CURVE
Input Pre-Processed image
Final ROI image
T3 Type tumor
Accuracy =
99.0344
EXIT


EXTRACTION OF ROI AND STAGE
CLASSIFICATION
FEATURES
LOAD IMAGE
FEATURE EXTRACTION
ROC CURVE
Input Pre-Processed image
T4 Type Tumor
Accuracy =
97.1386
EXIT

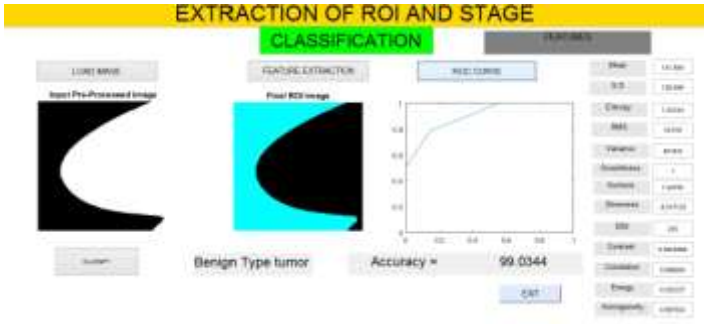


Fig 6: ROI AND DIFFERENT STAGES OF DBT IMAGES.

6. CONCLUSION AND FUTURE ENHANCEMENT

According to the Literature Survey, when we compare with the other modalities DBT works well with more accuracy to spot the size, location and shape of the Breast Tumor. Especially DBT is more concurred for women who has dense breast. The experimentation is conducted on the UPMC Breast Tomography and FFDM collection database of tomosynthesis images. Each modalities have their own limitation, thus fusion of information retrieved from different modalities may be helpful to improve the accuracy. Further we are implementing fusion technique to fuse the information from two modalities to obtain more accurate results.

REFERENCES

[1] A Novel Approach for Detection of Breast Cancer at an early stage using Digital Image processing techniques, Author : Sangeetha R, Dr. Srikanta Murthy K. (ICISC-2017) 2017 IEEE.

[2] Study of Breast Cancer Detection Methods using Image Processing with Data Mining Techniques. Author: Dr D Devakummari, V.Punithavathi. IJPAM, ISSN:1311-8088 Volume 118 No 18 2018, 2867-2873.

[3] A Comparative Study of Breast Cancer Diagnosis Using Supervised Machine Learning Techniques. Author: Madhuri Gupta & Bharat Gupta, IEEE, ISBN;978-1-5386-3452-3.

[4] Early Breast Cancer Detection via UWB Based on MRI-derived Model. Author: Xia Xiao, Qinwei Li, Liang Wang. IEEE.

[5] Early Detection and Classification of Breast Cancer. Author V.Vishrutha and M. Ravishankar, Vol,1 AISC 327.

[6] Breast Cancer Diagnosis Using Image Processing and Machine Learning for Elastography Images. IEEE 2019.

[7] A Research on Detection and Classification of Breast Cancer using k-means GMM and CNN Algorithms. Author: S. Shamy and J. Dheeba (IJEAT) August 2019.

[8] Classification of Region of Interest in Mammograms using Dual Contourlet Transform and Improved KNN. Author: Min Dong, Zhe Wang, Cheghui Dong Hindawi 2017.

[9] https://www.cancer.net/cancer-types/breast-cancer/stages.