A Quantitative Approach for Breast Cancer Segmentation Using Mammogram Images

Shruthi G.K
4th sem, M Tech
Dept. of CS & E
Adichunchanagiri Institute of Technology
Chikkamagalur, Karnataka, India

Arpitha C.N
Assistant Professor
Dept. of CS & E
Adichunchanagiri Institute of Technology
Chikkamagalur, Karnataka, India

Sunitha M.R
Professor
Dept. of CS & E
Adichunchanagiri Institute of Technology
Chikkamagalur, Karnataka, India

***Abstract*—Breast cancer is considered to be a deadliest disease in females nowadays. From past few years CAD (computer aided diagnosis) has gained its popularity in detection of breast cancer. Application of image processing technique in medical field has been increasing day by day. This results in new inventions in diagnosing disease accurately. Accuracy obtained by radiologist in segmentation of breast cancer may tend to decrease when dealing with large volume of images. This paper proposes a method for segmentation of breast cancer based on three categories, twelve classes and six shapes. This study shows the outcome of applying image processing operation like, preprocessing, segmentation and feature extraction. Proposed work uses ROI localization method for segmentation. Comparison results shows that proposed method have proven to achieve high PSNR values when compare to watershed segmentation.**

Index Terms-Image pre-processing, thresholding, morphological-operation, ROI localization.

I. INTRODUCTION

Cancer is a rapid growth of cells in a given region of body. The two important reasons for this rapid growth are mutation and excessive reproduction of cells. Due to this multiplicative growth of cells, there will be uncontrollable growth in organs, these results in formation of tumor. This can happen in any part of the body. If the multiplicative growth of cells takes place in breast, it is said to be breast cancer. When compared to all other different types of cancers, breast cancer is the only one which leads to death in women. The prediction of breast cancer cannot be done in early stages. At the same time, cause is still not understood. So, detection of the breast cancer in its early stage can be helpful to avoid death rates.

The accuracy obtained by image processing researchers in segmentation of breast images for identification of cancer is seemed to less when using different type of segmentation algorithm. If this false segmented region is given to feature extraction phase, results may vary. Due to the false perceptions of radiologists, patient will undergo unnecessary biopsies. Many image processing techniques are proposed for segmentation of breast cancer images. According to the survey made by American Cancer Society, cancer death rate has been fallen to 26%, which was very high in 1991 [1]. This drop in cancer mortality is due the advance in techniques for accurate identification of breast cancer.

In the proposed work the segmentation of mammogram images [2] are done based upon three categories, twelve classes and six shapes. Three categories are, benign, normal and malignant. Six shapes are well-defined circumscribed masses, architectural distortion, asymmetry, normal, other ill-defined masses and speculated masses [3].The proposed method aims at segmenting the breast cancer into different regions, so there should not be any noise in images. So, prior to segmentation, an image should be noise free and its contrast should be increased. The proposed work uses ROI localization technique for segmentation [4]. Tin Kam Ho, created the first ROI localization method for segmentation [5]. In performance analysis, proposed method has proven to be best segmentation in partitioning a mammogram image into its different regions when compare to that of watershed method.

This paper is organized as follows: Section II presents a review on existing techniques for breast cancer pre-processing and segmentation. Section III describes about methodology. Section IV gives details regarding results and discussions. Section V describes about conclusion.

II. LITERATURE SURVEY

A brief survey carried in order to work with the proposed method is discussed in this section. Segmentation of mammograms using watershed is presented in [6]. In this technique segmentation of breast cancer is made based on only its three categories namely, normal, benign and malignant. The pre-processing techniques for breast cancer detection in Mammography images are presented in [7]. In this paper adaptive median filter is utilized to expel spot clamor show in the mammograms. A novel approach for breast disease recognition and division in mammograms is introduced in [8]. In this paper division of tumor locale is finished by utilizing basic picture handling methods, for example, averaging and thresholding. Features are extracted from the segmented region obtained by proposed method. Feature extraction values for breast cancer mammography images are presented in [9]. For the complete identification of breast cancer stages, 23 features are extracted from breast cancer images, taken from MIAS (Mammographic image analysis society). Experimental results show that, the feature extraction values obtained in this method can be used for developing new CAD Computer Aided Diagnosis system.

III. METHODOLOGY

The whole stream of the proposed work is partitioned into three modules as appeared in the Figure.1, the three modules are: Image acquisition, Image pre-processing, Image segmentation and comparison. From the square chart it is apparent that segmentation is the center of this framework. The mammogram image is given as input image which then undergoes image noise reduction and enhancement, segmentation is performed. In the proposed method PSNR values of ROI obtained in segmentation phase taken for performance measure. ROI localization method and watershed segmentation method both are compared and performance graph is obtained finally.

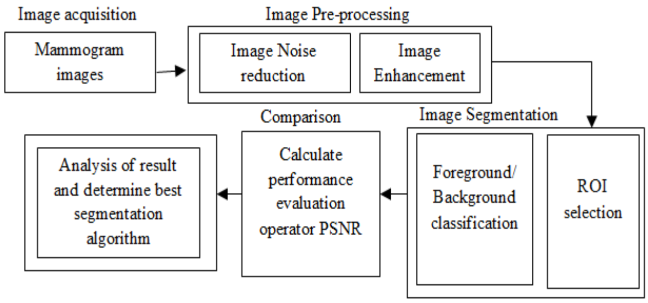

Figure 1: Architecture diagram for the proposed system

A. *Image Acquisition*

With a specific end goal to play out the test examination, test pictures are gathered from mammographic image analysis society (MIAS). The pictures are in dim scale document arrange (PGM - Portable Gray Map). The proposed strategy utilizes smaller than usual MIAS database as it contains finish data about anomalies of each mammographic picture [10].

B. *Image Pre-processing*

The very common noise that can be seen in mammography image is, speckle noise [11]. If image noise is not removed properly, it may lead to wrong classification of disease. This will have strong impact on patient. Pre-processing step is divided into two stages namely image noise reduction and image enhancement.

High intensity labels and additive noise occurred during image acquisition technique will be removed. In our method initially grey scale image is converted to binary image. Resulted image is threshold image. For the purpose of boundary extraction, disk is used as structuring element.

C. *Image Segmentation*

In segmentation step, a noise free and enhanced image is divided into different partitions. Pixels in each partition will be having similar grey scale of multivariate values. Background and foreground region will be separated and ROI is selected in segmentation stage. Adaptive diffusion active contour model is used for selection of ROI [12]. Tumor region is the region of interest for analysis purpose. This is done by selecting a polygonal region of interest. Smoothing of image is done by using function fspecial [13]. By using fspecial function two-dimensional spatial filters are created. fspecial creates Gaussian filters using the equation represented below

$$h_g(n_1, n_2) = e^{-(n_1^2 + n_2^2)/(2\sigma^2)}$$

$$h(n_1, n_2) = \frac{h_g(n_1, n_2)}{\sum_{n_1}\sum_{n_2} h_g} \qquad (1)$$

In the above algorithm, n1 and n2 represents the first and second parameter for the filter. The size of returned filter is determined by parameter h. The default value for sigma is 0.5.

D. *Different techniques for breast image segmentation*

Image segmentation is a process of partitioning the image into different region. In medical image processing segmentation plays very important role. ROI obtained in segmentation phase have adverse effect on feature extraction and classification phase. Different image segmentation techniques are available. Among them watershed segmentation and ROI localization methods use commonly in identification of tumors.

i) *Watershed method for segmentation*

Watershed is a segmentation method based on transformation that can be defined on gray scale image. There are mainly three types of watershed methods namely, flooding based watershed algorithm, and rainfall based watershed algorithm and watershed algorithm based on connected components [14]. Among all three methods watershed based on connected component used commonly in medical image segmentation.

Watershed algorithm

Let A_1, A_2, A_3….A_n be the coordinates sets of points in the minimal region of the image g(x,y).

C (A_i) be the points of coordinates in catchment basin associated with regional minima A_i.

T[n] = { (s,t) | g(s,t) < n }

1.T[n] = Set of points in g(x,y) which are lying below the plane g(x,y) = n.

2. n = flooding stage, varies from min+1 to max+1.
3. min = small value of gray level.
4. max = large value of gray level.

Let $C_n(A_1)$ be the set of points in the catchment basin associated with M_1 that are flooded at stage n.

1. $C_n(A_1) = C(A_1) \cap T[n]$.
2. $C_n(A_i)$ = 1 at location (x,y) if (x,y) Є C(A_i).
3. (x,y) Є T[n], otherwise it is 0.

C[n] is the union of flooded catchment basin portion at the stage n.

1. C [n] $= \cup_{i=1}^{R} Cn(A_1)$.
2. C [Max+1] $= \cup_{i=1}^{R} C(A_i)$.

It keeps on increasing the level of flooding & during this process $C_n(A_i)$& T[n] either increase or remains constant.

Algorithm initializes at C [min + 1]=T[min +1], and then proceeds recursively assuming that at step n C[n-1] has been constructed. Q is set of connected components in T[n].

ii) ROI localization method for segmentation

The specific procedure of breast tumor ROI localization is as follows:

1. A fixed window (denoted by W0) same as breast image window in size is used to crop a finger–vein candidate region in CCD imaging plane.
2. A predefined window (denoted by W1) is used to locate a sub region in W0. This can decrease the effect of useless background.
3. The pixel values at each row image are accumulated in the sub region W1.
4. The extreme row-sum is pinpointed to roughly denote the position of the distal inter-phalangeal joint.
5. Three points P1, P2, and P0 are positioned along the identified baseline. The points P1 and P2 denote the crossing of the joint baseline and the breast borders, respectively. In the meantime, the point P0 positions for the median of the segment flanked by P1 and P
6. Based on point P0, a window, denoted by W2 is used to crop a ROI image from the breast image region.

IV. RESULTS AND DISCUSSIONS

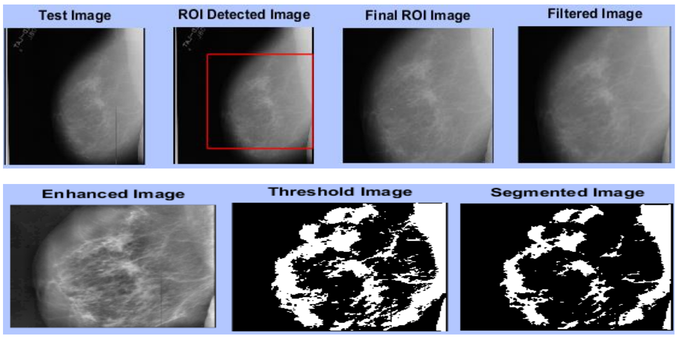

Figure 2: Output results for pre-processing and segmentation phase for breast image with benign feature

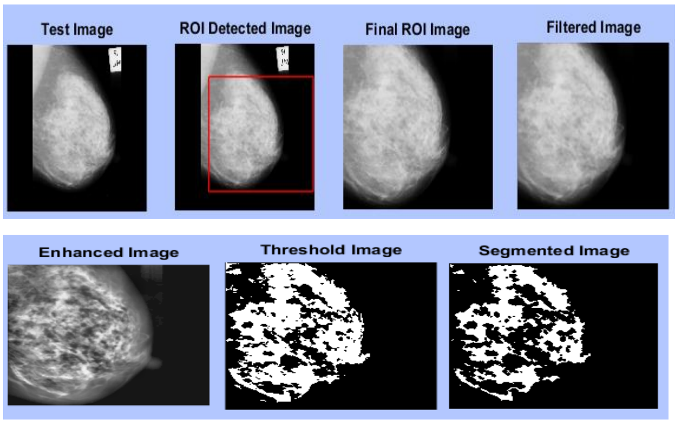

Figure 3: Output results for pre-processing and segmentation phase for breast image with malignant feature

The proposed method takes totally 250 images from (MIAS) database for analysis purpose [15]. Among them, ten images with different shapes and categories are taken for report generation. Each image will undergo preprocessing and segmentation. Figure 2 and 3 shows results of preprocessing and segmentation of breast cancer image with benign and malignant features respectively.

A.PSNR

Peak signal-to-noise ratio is also termed as PSNR, is an engineering term corresponds to ratio between the full possible power related to signal and the power of degrading noise that affects the fidelity of its depiction. Because many signals have a very wide dynamic range, PSNR is usually expressed in terms of logarithmic decibel scale. The formula for the representation of PSNR is depicted in equation 2.

$$\mathrm{PSNR}=10\log_{10}\frac{(L-1)^2}{\frac{1}{N^2}\sum_{x=0}^{N-1}\sum_{y=0}^{N-1}[f(x,y)-f*(x,y)]^2} \qquad (2)$$

where L is the number of gray levels (e.g., for 8 bits L=256). f (x, y). The original image, f* (x, y) is the decompressed image, x, y.

B. Comparison graph

Table 1: PSNR values for segmentation using Watershed and ROI localization

Images from MIAS database	Watershed PSNR	ROI PSNR
Image 1	24.722	30.334
Image 2	23.456	29.025
Image 3	23.728	27.354
Image 4	23.645	23.668
Image 5	21.78	27.841
Image 6	21.842	20.454
Image 7	19.887	28.809
Image 8	25.624	29.227
Image 9	24.896	26.563
Image 10	22.357	24.556

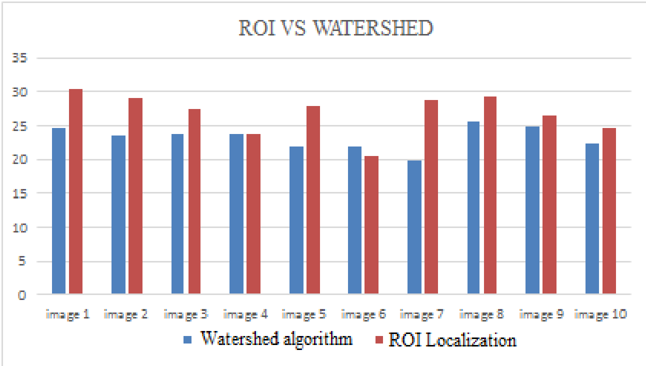

Figure 4: PSNR graph for segmentation phase

Table 1 represents the PSNR values for segmentation using watershed and ROI localization. Analysis is done on 250 images with different abnormalities. For purpose of generating comparison graph ten images are taken with different features. Comparison graph in the figure 4 shows that ROI localization method has achieved high PSNR value. High PSNR value indicates better the quality of image.

V. CONCLUSION

Breast cancer is a deadliest disease in women whose symptoms cannot be found in its starting stages. Early and exact determination of breast disease assumes a vital part in keeping away from death rates. The proposed strategy utilizes computer aided diagnosis method as finding technique for investigation of breast growth utilizing mammogram pictures. In this paper ROI localization methods partition the image into different regions without any loss. The method has been tested over 250 different kinds of images, and proved that ROI localization method has high PSNR value than watershed segmentation.

REFERENCES

1. American Cancer Society. Breast Cancer Facts & Figures 2016-2017. Atlanta: American Cancer Society, Inc., 2015.
2. Alhaddi, B., Zu'bi M.H. and Suleiman H.N. "Mammogram Breast Cancer Image Detection Using Image Processing Function," Information Technology Journal 6(2), 217–221, 2007.
3. www.wiau.man.ac.uk/services/MIAS/MIASweb.html
4. Leo Breiman. "Random Forests-Random Features,"Technical Report 567, Department of statistics, University ofCalifornia, Berkeley, September 1999.
5. Verikas, A., Gelzinis, A. and Bacauskiene, M., "Mining data with random forests: A survey and results of new tests," Pattern Recognition, 44, 330-349. 2011.
6. Breiman, L. "Random forests," Machine Learning Journal Paper, 45, pp. 5-32, 2001.
7. R. Ramani, N.Suthanthira Vanitha, S. Valarmathy, "The Pre-Processing Techniques for Breast Cancer Detection in Mammography Images", I.J. Image, Graphics and Signal Processing, vol. 5, 47-54, 2013.
8. Michael A. Wirth and Alexei Stapinski. "Segmentation of the breast regionin mammograms using snakes." IEEE . 2004.
9. Anupa Maria Sabu, D. Narain Ponraj and Poongodi, "Textural Features Based Breast Cancer Detection: A Survey," International Journal of Engineering and Future Technology" vol. 3, no. 9, Sep2012.
10. Vijaya Kumar Gunturu and Ambalika Sharma "Contrast Enhancement of Mammographic Images Using Wavelet Transform,"IEEE, 2010.
11. Huai Li, K. J. Ray Liu and Shih-Chung B. Lo, "FractalModeling and Segmentation for the Enhancement of Microcalcifications in Digital Mammograms," IEEE transcations on medical imaging, vol. 16, no. 6, December 1997.
12. Sunitha M R, H S Jayanna, Ramegowda, "Human tracking using edge and camera pattern feature for denser crowded videos," International Journal of Engineering and Future Technology, vol. 8, no. 8, pp. 9 – 17, 2016.
13. S. Deepa and V. SubbiahBharathi, "Textural feature extraction and classification of mammogram images using CCCM and PNN," IOSR Journal of Computer Engineering, vol. 10, Issue 6, 2013, pp. 07-13.
14. D.Lavanya and K.Usha Rani, "Performance Evaluation of Decision Tree Classifiers on Medical Datasets". International Journal of Computer Applications, vol. 26, no. 4:1-4, July 2011.
15. G. Ravi Kumar, Dr G. A. Ramachandra and K. Nagamani. "An Efficient Prediction of Breast Cancer Data using Data Mining Techniques." International Journal of Innovations in Engineering and Technology (IJIET) 2.4 (2013): 139.