Computer Aided Segmentation and Classification of Mass in Mammographic Images using ANFIS

K. Yuvaraj[1], **U.S. Ragupathy**[1]

[1] Department of Electronics and Instrumentation Engineering, Kongu Engineering College, Perundurai, Erode, Tamil Nadu, India

Abstract

Background: Breast cancer is one of the leading cancers in woman worldwide both in developed and developing nations as per the records from World Health Organization. Many studies have shown that mammography is very effective tool for the breast cancer diagnosis. Mass segmentation plays an important step for the cancer detection.
Objective: The objective of the proposed method is to segment the mass and to classify the mass with high accuracy.
Methods: The segmentation includes two main steps. First, a rough initial segmentation through iterative thresholding, and second, an active contour based segmentation. The relevant statistical features are extracted and the classification is done by using Adaptive Neuro Fuzzy Inference System (ANFIS).
Results: The proposed mass detection scheme achieves sensitivity of 87.5% and specificity of 100% for a set of twenty two images. The overall segmentation accuracy obtained is 91.30%.
Conclusions: This work appears to be of high clinical significance since the mass detection plays an important role in diagnosis of breast cancer.

Keywords

ANFIS, classification, mammography, mass, segmentation

Correspondence to:

K. Yuvaraj
Department of Electronics and Instrumentation Engineering,
Kongu Engineering College,
Address: Perundurai, Erode, Tamil Nadu, India
E–mail: yuvarajkumarasamy@gmail.com

EJBI 2013; 9(2):37–41
received: January 16, 2013
accepted: April 4, 2013
published: August 30, 2013

1 Introduction

Over the past two decades, cancer has been one of the biggest threats to human life and it is expected to become the leading cause of death over the next few decades. A mammogram is an X-ray of the breast. One can diagnose at very early stage and hence there is a chance for healthy survival. Breast cancer can be traced from mammogram before it can be felt. When mammography is combined with clinical breast exam the chances for finding cancer are even greater. For women with dense breast tissue, digital mammography may be more accurate than standard mammography.

There are many other breast imaging tests that can provide valuable information [1], however, these find difficult to tell the difference between dense breast tissue, benign (non-cancerous) lumps and cancer. And, sometimes they miss tiny calcium deposits that are the earliest sign of a tumour. Thus mammogram plays as a vital tool in breast cancer detection and identification [2]. Mass is defined as space occupying lesion that is described by their shape and marginal properties. A benign is characterized by smooth margination, whereas a malignancy is characterized by an indistinct border that becomes more spiculated with time.

Many researches were done to segment the mass in the mammograms [3, 4, 5]. Of which Active contour based segmentation provides the best way to segment the images whose background and foreground that are statistically different and homogeneous [6]. The contour based segmentation retains the original information of region, edge and shape of the mass [5] and [7].

The region based active contour [8] and [9] model holds good than the edge based contour model [10]. Considering all the above advantages of the region based active contour, the segmentation technique combines the classical segmentation technique with the region based active contour.

The rest of the paper is organized as follows. Section 2 details about the proposed method of mass segmentation. Section 3 discusses about the statistical features extraction and classification from the segmented images. Result and discussion are drawn in Section 4. Conclusion is produced in Section 5.

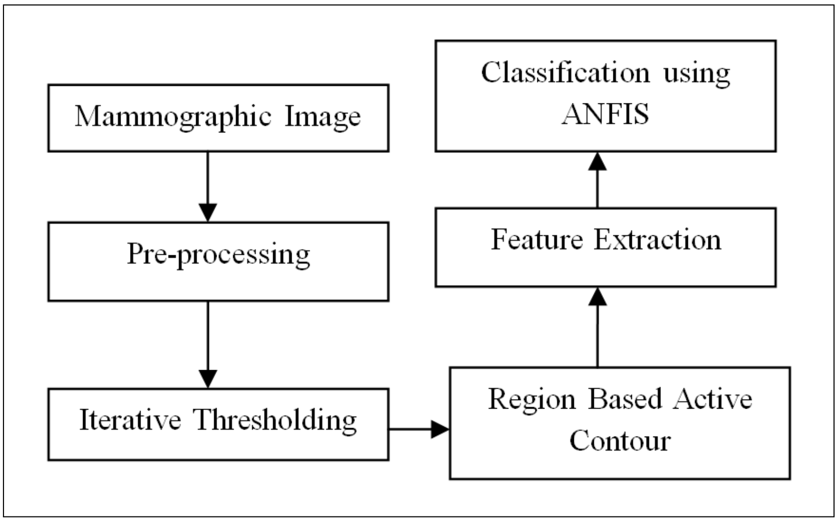

Figure 1: Block Diagram of the Proposed Mass Segmentation and Classification.

2 Proposed Method of Mass Segmentation

Mammographic images are collected from the Imaging centre, Coimbatore and from Mammographic Image Analysis Society (MIAS). The images are subjected to preprocessing which includes cropping (256 X 256) and enhanced by performing histogram equalization. Then the preprocessed image is subjected to iterative active contour based segmentation [11]. From the segmented image Gray Level Co-occurrence Matrix (GLCM) are formulated. Features are extracted from GLCM. These features are fed to the ANFIS for classification.

The block diagram of the proposed iterative active contour based segmentation and classification is shown in Figure 1. The result obtained in various stages of preprocessing are depicted in Figure 2.

The pre-processed image and the iterative thresholded image are shown in Figure 3 (a) and (b) respectively. The average value of the maximum pixel and minimum pixel value of the pre-processed image is set as initial threshold value (T1). The value above T1 is set as foreground and the value below T1 as background. The total number of pixels in the foreground and the sum of all gray values of fore ground is calculated. The total sum divided by the number of pixels gives the average value of fore ground.

The above procedure is repeated for background. Then the average value of the foreground and the background is calculated (T2). If the value of T1 is equal to T2, then T2 is the final threshold value. If T1 is not equal to T2, then T1 is assigned with T2.

The segmentation technique involves initialization of the mask. The square mask of the size 111 x 111 is defined within the mass. This mask is allowed to deform until the minimum energy is encountered. Then, the binary segmented image is obtained.

Figure 4 shows the results obtained in various stages of segmentation. This is mapped with the gray scale pre-

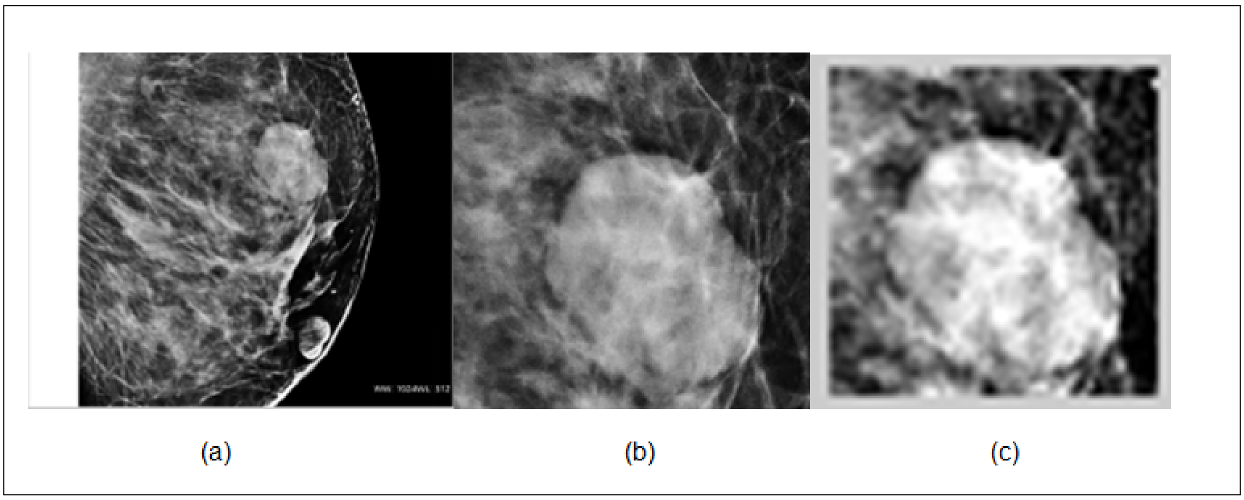

Figure 2: (a) Original Image, (b) Cropped Image and (c) Histogram Equalized Image.

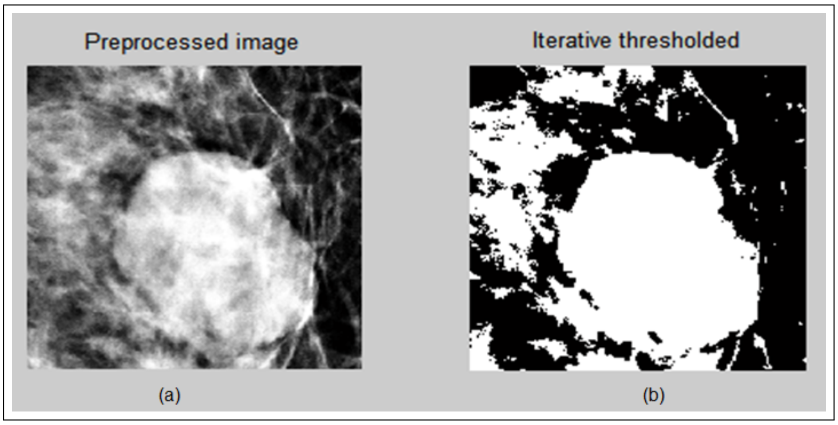

Figure 3: (a) Preprocessed Image and (b) Iterative Thresholded Image.

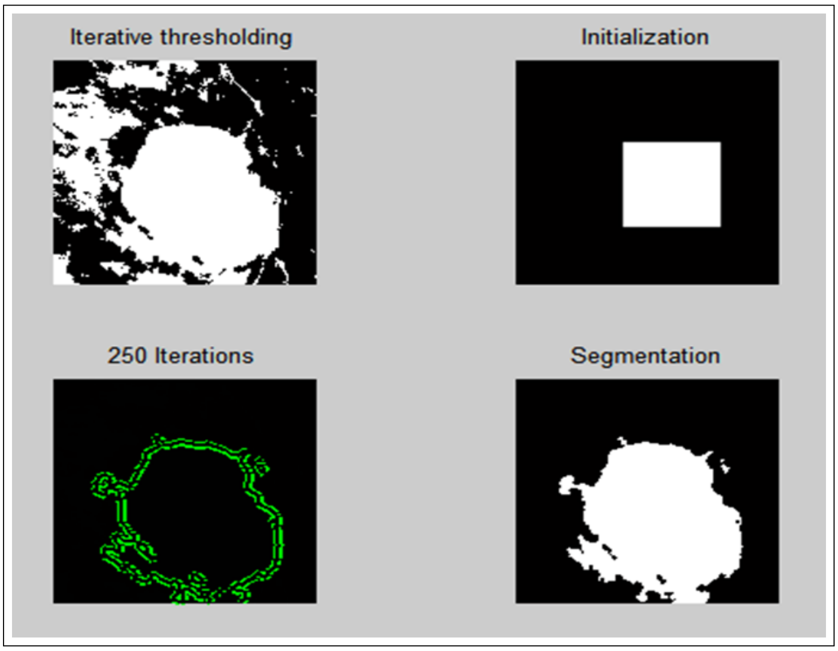

Figure 4: Image Results Obtained in Various Stages of Mass Segmentation.

processed image, and thus the gray level segmented image is obtained as shown in Figure 5.

This method of segmentation is preferred as it holds good when compared to the existing techniques like Fuzzy C Means based segmentation and Level set based segmentation [11].

3 Feature Extraction

From the gray level segmented image as shown in Figure 5 (b), GLCM are formulated. Initially four GLCM matrices are formulated for $0^\circ, 45^\circ, 90^\circ$ and 135° orientations. The average of the above four matrix is formulated as resultant mean GLCM. From the formulated resultant mean GLCM sixteen statistical features are extracted and are listed in Table 1. From these sixteen features the optimal features are obtained by calculating the variance and the features with higher value of variance are selected and are shown in Table 2. These include Contrast, Dissimilarity, Sum average, Sum variance and Auto correlation [12]. For the remaining eleven features since the variance is much lower (approximately equal to zero) and these features are neglected.

4 Result and Discussion

The statistical features are extracted from 61 images. Table 3 illustrates the number of images considered for training and testing ANFIS classifier. Number of training samples considered are 39. For testing 22 samples are considered. The membership functions are used to convert the input values into fuzzy values. Various types of membership function is used which includes triangular, trapezoidal, generalised bell, gaussian and gaussian2. Among which the testing error for the triangular membership function is 0.4603 which is lesser when compared with

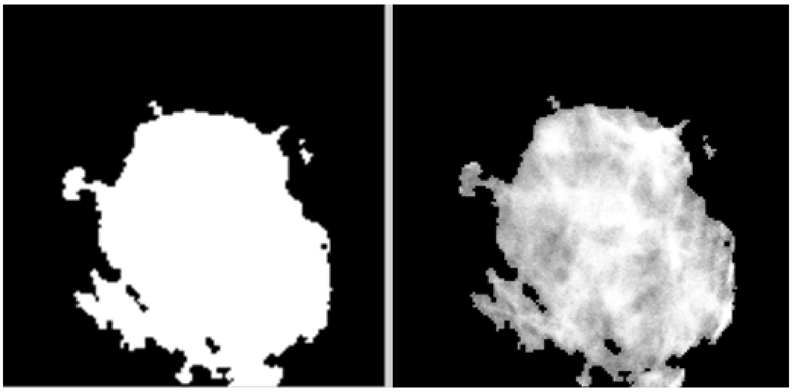

Figure 5: (a) Binary Segmented Image and (b) Gray Level Segmented Image.

other membership functions (0.5701-trapeziodal, 1.1941-gbell, 0.8864-gaussian, and 1.483-gaussian2).

Table 2: Sample Features Extracted.

Features / Image Category	Normal mdb047	Benign mdb025	Malignant mdb264
Contrast	0.3658	0.3125	0.7847
Dissimilarity	0.0550	0.0446	0.1202
Sum average	7.4156	7.0528	7.1306
Sum variance	97.8043	93.1580	92.7942
Auto correlation	34.1819	32.5812	32.6761

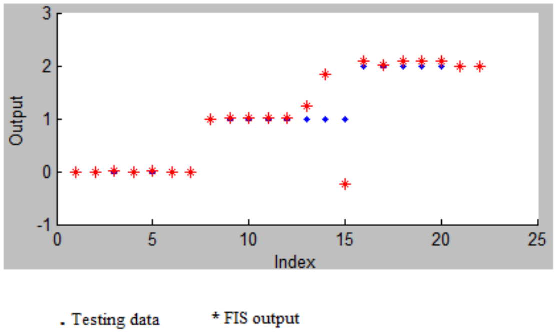

Figure 6: ANFIS-Classification Plot.

Table 1: Feature Selection from GLCM.

Feature Reduction / **Feature Extracted from GLCM**	Normal mdb123 (x)	Benign mdb69 (y)	Malignant mdb267 (z)	Mean $M = (x+y+z)/3$	Variance $\lvert((M-x)+(M-y)+ +(M-z))/3\rvert$
Contrast	0.3806	0.4526	0.0732	0.2988	**0.18**
Correlation	-1.3026	-1.2566	-1.2613	-1.2735	0.0194
Dissimilarity	0.0568	0.0709	1.0459	0.3904	**0.4378**
Energy	0.5347	0.4869	0.4844	0.502	0.01228
Entropy	-0.4232	-0.3958	-0.3939	-0.4043	0.0117
Homogeneity	0.9922	0.9894	0.9869	0.9895	0.0054
Maximum Probability	0.6438	0.4977	0.5036	0.5483	0.0636
Sum average	9.0698	6.9194	7.1435	7.7109	**1.0859**
Sum variance	120.6634	90.7037	93.4084	101.5918	**12.7143**
Sum entropy	0.3298	0.4143	0.4213	0.3884	0.039
Difference variance	0.3806	0.4526	0.6032	0.4758	0.0789
Difference entropy	0.0618	0.0866	0.0990	0.0824	0.0138
Information measure of correlation	0.9844	0.9868	0.9867	0.9859	0.0010
Inverse difference normalized	0.9961	0.9950	0.9936	0.9949	0.0008
Inverse difference moment normalized	0.9966	0.9959	0.9945	0.9956	0.0005
Auto correlation	41.6196	31.8959	32.8250	35.4478	**4.1145**

Hence triangular membership function is chosen. For each input three triangular membership functions are used. Total number of rules framed is 729. The classification result obtained through ANFIS is shown in the Figure 6.

Table 3: Mass classification Result by ANFIS.

Image Category	No. of training images	No. of testing images	Error
Normal	16	7	0
Benign	13	8	2
Malignant	10	7	0

With the information available from Table 3, sensitivity, specificity and overall accuracy of the proposed classification system are calculated by using the equations (1), (2) and (3) respectively. These are used as performance metrics for the classification. The values are tabulated in Table 4.

$$\text{Sensitivity} = \frac{\text{Number of true positive decisions}}{\text{Number of actual positive cases}} \times 100 \quad (1)$$

$$\text{Specificity} = \frac{\text{Number of true negative decisions}}{\text{Number of actual negative cases}} \times 100 \quad (2)$$

$$\text{Accuracy} = \frac{\text{Number of correct decisions}}{\text{Total number of cases}} \times 100 \quad (3)$$

Table 4: Performance Metrics of the Classifier.

Metrics	Value (%)
Sensitivity	87.5
Specificity	100
Accuracy (Overall)	91.30

The sensitivity obtained is 87.5%, the specificity obtained is 100% and the overall classification accuracy obtained through ANFIS is 91.30%.

5 Conclusion

In this paper a new computer aided mass segmentation and classification scheme is proposed. Five statistical features were selected for the classification of mass. The classification is done by using ANFIS. The implementation of the proposed method was carried out using MATLAB. Result shows that the suggested features can give acceptable accuracy for the classification of mass. The result shows that the overall accuracy is 91.30%. With increase in the number of test samples the accuracy may still be improved to a great extent.

References

[1] Kuo H C, Giger M, Reiser I, Boone J, Lindfors K, Yang K. Handbook of Breast Imaging. 2012; 697-704.

[2] Samir Kumar Bandyopadhyay. Survey on Segmentation Methods for Locating Masses in a Mammogram Image. An International Journal of Computer Applications, Article 6. 2011; 9(11):25-28.

[3] Sahiner B, Petrick N, Chan H P, Hadjiiski LM, Paramagul C, Helvie MA, Gurcan MN. Computer-Aided Characterization of Mammographic Masses: Accuracy of Mass Segmentation and its Effects on Characterization. IEEE Transaction on Medical Imaging. 2001; 20(12): 1275-1284.

[4] Sahiner B, Chan H P, Petrick N, Helvie M A, Hadjiiski LM. Improvement of Mammographic Mass Characterization using Spiculation Meausures and Morphological Features. Medical Physics. 2010; 18(7): 1455-1465.

[5] Jun Liu, Jianxun Chen, Xiaoming Liu, Lei Chun2, Jinshan Tang1, Youping Deng. Mass Segmentation using a Combined Method for Cancer Detection. An International Conference on Bioinformatics and Computational Biology. 2010; 5(3):S6.

[6] Peyman Rahmati A, Andy Adler A, Ghassan Hamarneh B. Mammography Segmentation with Maximum Likelihood Active Contours. An International Journal on Medical Image Analysis. 2012; 16(6): 1164-1186.

[7] Thiruvankadam, Acharyya, Neeba, Jhunjhunwala, Ranjan S. A Region Based Active Contour Method for Extraction of Breast Skin line in Mammograms. An International Conference on Medical Imaging: from Nano to Macro. 2010;189-192.

[8] Bhagwati Charan Patel, Sinha GR. Energy and Region based Detection and Segmentation of Breast Cancer Mammographic Images. An International Journal of Image, Graphics and Signal Processing. 2012; 6(2): 44-51.

[9] Alierza Vard, Kamal Jamshidi, Naser Movahhedinia. Small Object Detection in Clustered Image using Correlation based Active Contour Model. Journal on Pattern Recognition Letters. 2012; 33(5): 543-553.

[10] Yu Zhang, Noriko Tomuro, Jacob Furst, Daniela Stan Raicu. Image Enhancement and Edge-based Mass Segmentation in Mammogram. Proceedings of Medical Imaging 2010: Image Processing. 2010; 7623:76234P-76234P-8

[11] Yuvaraj K, Ragupathy U S. Mammographic Mass Segmentation based on Iterative Active Contours. Proceedings of Biomedical Systems, Signals and Images (BSSI 2012) 2012.

[12] Andrew F. Laine, Sergio Schuler, Jian Fan and Walter Huda. Mammographic Feature Enhancement by Multiscale Analysis. An International Conference on Medical Imaging. 1994; 13(4):725-740.