Original Research Article Allied sciences

International Journal of Pharma and Bio Sciences

ISSN
0975-6299

EXTRACTING REGION OF INTEREST IN FULL-FIELD DIGITAL MAMMOGRAM WITH GLCM AND GROW CUT REGION BASED SEGMENTATION METHOD.

G.R. JOTHI LAKSHMI*[1] AND DR.ARUN RAAZA[2]

[1]Research Scholar, Department of ECE, Vels University, Pallavaram, Chennai, Tamilnadu, India
[2]Deputy Director, CARD, Vels University, Pallavaram, Chennai, Tamilnadu, India

ABSTRACT

Early breast cancer detection can be done by analyzing mammograms and due to that breast cancer deaths can be decreased. Many methodologies can be followed to detect breast cancers like mammography, MRI, Ultrasound etc. Mammography is a low-cost and simple way of finding breast cancers. Because of early detection, the survival rate of the patient increases and it can be performed easily using digital mammograms because they are easy to capture and manipulate the images, so the abnormalities can be seen more easily. Sometimes, manual reading with expert radiologist will result in misdiagnosis. To avoid this, many computer-aided detection methods have been developed to identify the masses and micro-calcification. Before identifying, the segmentation of mammogram is a vital part and in this paper, the segmentation of mammogram can be done using region growing technique, by calculating five features (mean, dissimilarity, sum average, sum variance and correlation) using Gray Level Co-occurrence Matrix (GLCM). Then by fixing up a threshold value and grow cut method is followed to track the closed region after automated seed point detection. The features extracted are compared with ground truth values to confirm the effectiveness of the proposed method. This results of proposed method show improved performance than existing methods.

KEYWORDS: ***Mammograms, Seeded Region growing (SRG) segmentation, seed selection, GLCM, parameters calculation.***

G.R. JOTHI LAKSHMI*

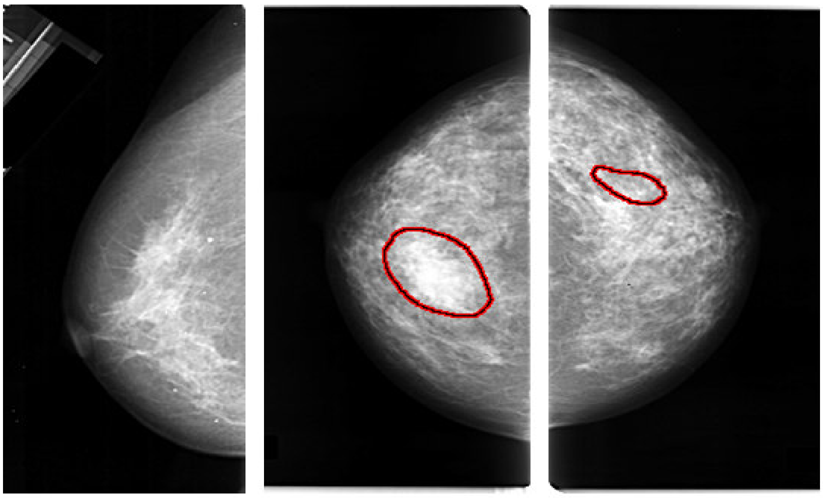
Research Scholar, Department of ECE, Vels University, Pallavaram, Chennai, Tamilnadu, India

***Corresponding author**

Received on: 21-11-2016
Revised and Accepted on: 17-03-2017
DOI: http://dx.doi.org/10.22376/ijpbs.8.2.b430-435

INTRODUCTION

In recent years, the mortality rate among women increases accounting for more than 1.6% of deaths due to the occurrence of breast cancer.[21] Nowadays it becomes the most common cancer detected in women worldwide. The incidence of breast cancer in India is keeping on increasing.[22] soon it will become number one cancer in females pushing cervical cancer to the second spot.[22] The situation could be understood by going through recent data from Indian Council of Medical Research (ICMR). It gives an alarming report that one in 22 women in India is likely to suffer from breast cancer during her lifetime. [22]In America, the situation is bad than India that one in eight being a victim of this deadly cancer. [23]But the consoling fact related to breast cancer is early identification of breast cancer that increases the survival rate. So it is mandatory to identify the incidence of breast cancer at an early stage through computer aided techniques. This could be done by analyzing one of the breast image modalities mamographies. A Mammogram is a picture of breast captured by using low dense X-rays. For detecting breast pathology, mammography is used as radiographic examination. It is important to classify the mammograms into masses and micro-calcification which are again classified into benign and malignant. Before doing the classification, it is needed to segment of mammograms[14,15]. Normal mammogram with benign cyst and mammogram with cancer images are shown in figure 1.



Normal mammogram   Benign cyst (not cancer)   Malignant cyst (cancer)

Figure 1
Normal and affected mammograms [24]

Rafael Gonzalez et al, mentioned three techniques for segmenting an image are follows thresholding based, region based and morphological watershed-based method[1,2]. Out of three regions based segmentation gives a better output, because the region based method starts with the selection of appropriate seed point. Liu et al proposed a novel mammogram segmentation method based on normalized cut method. After preprocessing, the texture features of the mammogram are extracted, and setting up the weights function with these features; then, the normalized cuts method is used to find the partitions of the mammogram.[3]Aghdam HHet al proposed a probabilistic adaptive thresholding method that uses texture information and its probability to find the most probable threshold values for specific parts of the mammogram.[4] In digital mammogram, to segment the suspicious region growing method with alarm region generation process is used.[5] In the work of kumar et al in order to realize automatic and adaptive segmentation threshold selecting, they used a new uncertainty theory-Cloud Model.[7] In that model the uncertainty of image was considered and concepts from characteristics of the region to be segmented were extracted. Nagi et al proposed an integrated algorithm using morphological processing and seeded region growing technique, to extract RoI from digital mammogram by suppressing pectoral mussels.[8] Dantulwar et al discussed a single seeded region growing technique to segment the images & the results were optimized using fuzzy logic by calculating four performance parameters.[9]

SEGMENTATION

The main motive of image segmentation is to get various features of the images. These extracted image features can be merged or split to build objects of interest. The built objects can be analyzed and interpreted further. Image segmentation divides an image into groups of pixels .These group of pixels must be homogenous based on some criterion. The outcome of segmentation is the splitting up of the image into connected areas. Thus segment is concerned with partitioning an image into meaningful regions. Segmentation can be done based on two basic properties: Discontinuities and similarities. Discontinuities use a technique for dividing an image based on sudden changes in gray-scale levels. Using discontinuities technique, it is possible to find isolated points, lines and edges in an image and boundary of an image can be estimated using edge detection.[10,11] Similarities are based on finding out connected pixels based on their properties using thresholding, region growing and region splitting and merging techniques.

Seeded region growing algorithm (srga)

Since digital mammogram image consisting of different intensity value, the segmentation process performed on the edge map differentiates various regions on the breast. Each region has a dissimilar intensity value. The different parts of breast image like lobules, glands, ducts and fatty tissues have different intensity values and these various parts can be separated into different regions. Any one of the breast parts may consist a lesion like mass, tumors or calcifications. This lesion region has definitely higher intensity values than the normal tissues of the breast. So it is necessary to identify the closed structures based on their intensity values and to classify the obtained closed structures.[12] The pixels intensities' distribution also varies within each segmented region but the majority of the pixels have similar intensity values. So for each region in the original mammogram, it is necessary to find the arithmetic Mode value for the available intensities and replace those pixels in the region with the computed mode values.

MATERIALS AND METHODS

In proposed methodology, Seeded region growing with GLCM was used to do perfect segmentation on mammogram images. Acquisition of mammogram images was done through the MIAS database. The corrupted image was enhanced using adaptive volterra filter. Then Grey level co-occurrence matrix was developed to calculate five parameters. The parameters were mean, sum average, sum variance, dissimilarity, correlation. Based on threshold value and five parameters, a seed point was selected. By following the region growing method, to perform iterations to identify the closed region & the proposed method gave perfect region of interest. The flow –chart for the proposed methodology is as shown below in figure 2.

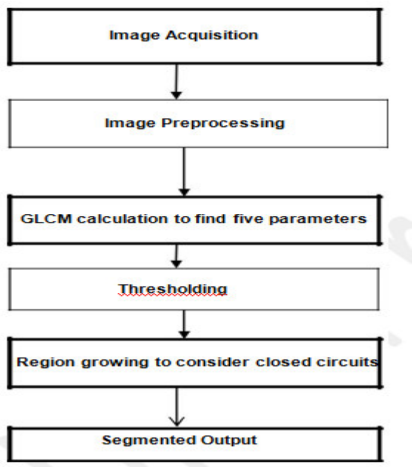


Figure 2
Flow chart for proposed methodology

Image Acquisition

The mammogram images were obtained from MIAS (Mammogram Image Analysis Society) database[25]. In mammogram image, the lesion region has high intensity than the other region.

Image Preprocessing

Due to poor hardware X-ray system, the mammogram images may be corrupted by noise. To eliminate the noise, the mammogram image was passed through adaptive volterra filter that improves the appearance of the corrupted image. The two dimensional quadratic system with input x[n1, n2] and output y[n1, n2] is governed by the equation

$$y[n1,n2] = \sum_{m11=0}^{N1-1}\sum_{m12=0}^{N2-1}\sum_{m21=0}^{N1-1}\sum_{m22=0}^{N2-1} h1[m11,m12,m21,m22]$$
$$x[n1-m11,n2-m12]x[n1-m21,n2-m22] \quad (1)$$

Equation (2) can be represented in the matrix form as

$$Y[n1,n2]=X^T[n1,n2]H_2X[n1,n2] \quad (2)$$

The quadratic kernel H2 has N1N2×N1N2 elements and each element consists of N_2^2submatricesH (i, j) with N1 × N2 elements.

GLCM calculation to find the parameters

From the enhanced mammogram image, the grey level co-occurrence matrix (GLCM) was calculated by considering the neighborhood pixel intensity. From this GLCM matrix, a set of descriptors (parameters) were calculated for characterizing the GLCM. GLCM gives an idea that how co-occurring pairs of pixels are spatially related in various orientations with reference to distance and angular spatial relationships. GLCM is computed by displacement vector'd' and rotational angles 0° , 45°, 90° and 135°. GLCM extracts features based on pixel and its neighbor pixel in an image.

The Haralick textural features calculated were
Mean
Sum of Average
Sum of Variance
Dissimilarity and Correlation

Thresholding

Based on the measured textural features, a thresholding was fixed to select a seed point.

Seeded Region Growing Technique

After seed selection, by comparing the characteristics of seed pixel with nearby pixels based on the intensity of the seed point, the group of similar characteristics pixels were identified to find the Region of Interest (RoI).The comparison of seed intensity with neighboring pixel intensity and adding the neighboring pixel to the region of interest was repeatedly done till no more pixels can be added in the region based on the pixel intensity. Thus the segmented region was separated out.

RESULTS

The collected mammogram images from MIAS data base were processed as per the proposed flow. The input image was shown in figure 3.The segmented images were obtained using proposed method was shown in figure 4 and by using other existing techniques were shown in figure 5, figure 6 and figure 7. From the database a set of 50 images were collected and the output was verified for all the 50 images. The segmented images based on SRG technique was compared with the segmented output obtained through other sobel threshold[26], gradient threshold[27] and watershed sobel gradient methods.[28] Moreover the five parameters were calculated using proposed technique and other existing techniques as shown in table 1. The comparison of the obtained parameters using SRG and other three existing techniques was shown in figure 8.

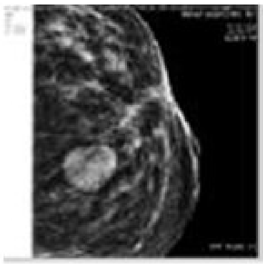


Figure 3
Input image

Figure 4
Segmented output using SRG

Figure 5
Segmented output using Sobel threshold

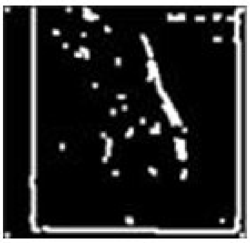


Figure 6
Segmented output using Gradient threshold

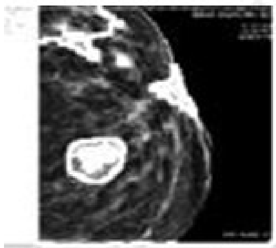


Figure 7
Segmented output using Watershedsobel gradient

Table 1
Parameters Calculation for SRG and Other methods

Textural Features / Segmentation Methods	Mean	Sum of Average	Sum of variance	Dissimilarity	Correlation
Seed Growing (SRG) method	**5.01e+02**	**1.33e+01**	**2.89e+02**	**2.49e-01**	**9.93e-01**
Sobel threshold[26]	3.24e+02	5.36e+00	1.31e+02	5.39e-01	8.24e-01
Gradient Threshold [27]	5.01e+01	2.27e+00	4.29e+00	9.16e-01	7.26e-01
Watershed Sobel gradient[28]	4.94e+01	4.9e+00	1.19e+02	3.69e-01	9.75e-01

The parameters related with SRG has significantly high compared with other techniques except dissimilarity.

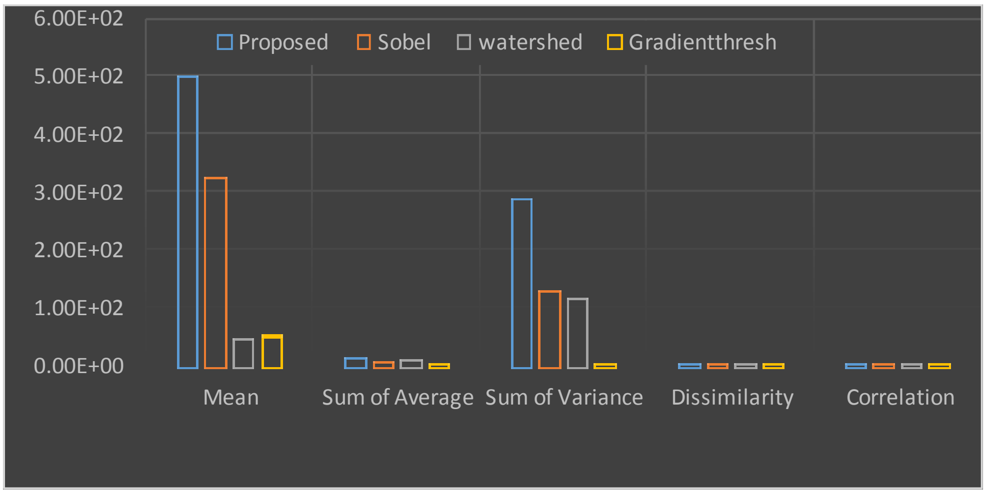


Figure 8
Comparison of Parameters

From Table 1, we can conclude that Haralick texture features mean, sum of average, sum of variance, dissimilarity, and correlation which were obtained from seeded region growing technique were significantly different than that of the other existing techniques sobel threshold, gradient threshold and watershed sobel gradient. The mean, sum of average, sum of variance and correlation were significantly higher than the mean, sum of average, sum of variance and correlation values obtained from other three techniques. Whereas the calculated value of dissimilarity of seeded region growing had significantly less than the dissimilarity values of the other three existing techniques. As a whole, the RoI obtained through seeded region growing was perfect and clear to process further when compared with the output with the other three techniques.

DISCUSSION

The proposed method used to separate the RoI exactly from mammogram images. Nayak et al proposed a method for segmenting the suspicious region in digital mammograms based on un-decimated wavelet transform and adaptive thresholding techniques.[18] Hamissi et al presented a method for detection of abnormal masses in digital mammogram. In pre-processing they followed three steps 1. Noise removal 2. Suppression of artifacts, labels and pectoral muscles3.enhance the breast region. Then Region of Interest was indentified using an adaptive segmentation procedure based on K means Clustering followed by a Merging Regions method. In the segmented region, statistical and textural features were calculated by using gray level co-occurrence matrices (GLCM) and a Decision Tree Classification was performed to isolate normal and abnormal regions in the breast tissue[19]. Yuvraj et al proposed an automated method for mass segmentation using seed selection by extracting textural features and seeded region growing technique.[20]

CONCLUSION

The output images show that the seeded region growing technique is very much useful to find the Region of Interest in an accurate manner when compared to other existing techniques. The textural features extracted using GLCM with SRG has significant values when compared with other methodology and these parameters are very much useful for classification of mammogram images. These parameters are considered as input to further classification algorithms. In this algorithm, the mammogram has been segmented using Region based segmentation. This algorithm is a fully automated which provides a breast contour segmentation. The algorithm is applied to different set of image to provide the segmented output. The advantage of using region growing technique is that in this method based on the similar defined properties the region of interest is exactly separated. The seed point is determined based on the criteria we want to make. The results obtained through region growing technique is very sensitive to initial seed selection. So it is important to find the accurate initial seed for image segmentation.

CONFLICT OF INTEREST

Conflict of interest declared none.

REFERENCES

1. Leavline EJ, Sutha S, Singh DA. Wavelet domain shrinkage methods for noise removal in images: A compendium. IJCA 2011 Nov;33(10):28-32.
2. Gonzalez RC, Woods RE. Digital image processing. Prentice Hall, 2009.
3. Jain AK. Fundamentals of digital image processing. Prentice-Hall, Inc.; 1989 Jan 1.
4. Liu L, Zeng F. Digital mammogram segmentation based on normalized cuts. InGlobal High Tech Congress on Electronics (GHTCE), 2012 IEEE 2012 Nov 18 (pp. 127-130). IEEE.
5. Aghdam HH, Puig D, Solanas A. Adaptive Probabilistic Thresholding Method for Accurate Breast Region Segmentation in Mammograms. InICPR 2014 Aug 24 (pp. 3357-3362).
6. S.Meenalosini, Dr.J.Janet, and Dr.E.Kannan, Segmentation Of Cancer Cells In Mammogram Using Region Growing Method And Gabor

Features, IJERAVol. 2, Issue 2,Mar-Apr 2012, pp.1055-1062
7. Senthilkumar B, Umamaheswari G, Karthik J. A novel region growing segmentation algorithm for the detection of breast cancer. In Computational Intelligence and Computing Research (ICCIC), 2010 IEEE International Conference on 2010 Dec 28 (pp. 1-4). IEEE.
8. Nagi J, Kareem SA, Nagi F, Ahmed SK. Automated breast profile segmentation for ROI detection using digital mammograms. InBiomedical Engineering and Sciences (IECBES), 2010 IEEE EMBS Conference on 2010 Nov 30 (pp. 87-92). IEEE.
9. Dantulwar S, Krishna RK. Performance Analysis Using Single Seeded Region Growing Algorithm. IJIRAE2014 Jul;1(6):2349-163.
10. Maitra IK, Nag S, Bandyopadhyay S. Automated digital mammogram segmentation for detection of abnormal masses using Binary homogeneity enhancement algorithm. J Comput Sci Eng (IJCSE). 2011 Jun;2(3):416-27.
11. Zade AS, Wanjari M, Chowhan AP. Detection of Cancer Cells in Mammogram Using Seeded Region Growing Method and Genetic Algorithm. Journal of Science & Technology. 2014:15-20.
12. Nanayakkara RR, Yapa YP, Hevawithana PB, Wijekoon P. Automatic breast boundary segmentation of mammograms. Int J. Soft Comput. Eng.(IJSCE). 2015;5(1).
13. Udaya GG. Efficient Segmentation and Classification of Micro calcification Using PSO and SVM in Mammographic Images.
14. Bandyopadhyay SK. Survey on segmentation methods for locating masses in a mammogram image. IJCA 2010 Nov;9(11):25-8.
15. Angayarkanni SP,Banu Kamal NA, Thavavel V Mammogram MRI image segmentation and classification. Journal of Theoretical & Applied Information Technology.JATIT 2013 Nov 20;57(2).
16. Sahakyan A, Sarukhanyan H. Segmentation of the breast region in digital mammograms and detection of masses. IJACSA Page. 2012 Feb(102).
17. M. Khan, Image Segmentation Methods: A Comparative Study," IJSCE., vol. 3, no. 4, pp. 84–92, 2013.
18. Nayak A, Ghosh DK, Ari S. Suspicious lesion detection in mammograms using undecimated wavelet transform and adaptive thresholding. In Advanced Computing Technologies (ICACT), 2013 15th International Conference on 2013 Sep 21 (pp. 1-6). IEEE.
19. Hamissi S, Merouani HF. Novel fully automated computer aided-detection of suspicious regions within mammograms. InInnovative Computing Technology (INTECH), 2012 Second International Conference on 2012 Sep 18 (pp. 153-157). IEEE.
20. Yuvaraj K, Ragupathy US. Automatic Mammographic Mass Segmentation based on Region Growing Technique. In3rd International Conference on Electronics, Biomedical Engineering and its Applications (ICEBEA'2013) April 2013 Apr 1 (pp. 29-30)
21. https://www.cancer.org/content/dam/cancer-org/research/cancer-facts-and-statistics/breast-cancer-facts-and-figures/breast-cancer-facts-and-figures-2015-2016.pdf
22. http://www.jagranjosh.com/current-affairs/icmr-study-revealed-doubling-of-incidence-of-breast-cancer-in-metropolitan-cities-1318998184-1
23. www.breastcancer.org/symptoms/understand_bc/statistics
24. http://marathon.csee.usf.edu/Mammography/DDSM/thumbnails/normals/normal_02/case0202/A-0202-1.html
25. http://marathon.csee.usf.edu/Mammography/Database.html
26. Makandar A, Halalli B. Threshold based segmentation technique for mass detection in mammography. J Comput. 2016 Nov 1;11(6):472-8.
27. Nithya R, Santhi B. Classification of normal and abnormal patterns in digital mammograms for diagnosis of breast cancer. International Journal of Computer Applications. 2011 Aug 31;28(6):21-5.
28. Zaheeruddin Z, Jaffery ZA, Singh L. Detection and shape feature extraction of breast tumor in mammograms. InProceedings of the World Congress on Engineering,(2) 2012.

Reviewers of this article

Dr. S. Jerritta

Associate Professor, Department of ECE, Vels University, Chennai, India.

Asst.Prof.Dr. Sujata Bhattacharya

Assistant Professor, School of Biological and Environmental Sciences, Shoolini University, Solan (HP)-173212, India

Prof.Dr.K.Suriaprabha

Asst. Editor , International Journal of Pharma and Bio sciences.

Prof.P.Muthuprasanna

Managing Editor , International Journal of Pharma and Bio sciences.

We sincerely thank the above reviewers for peer reviewing the manuscript