IJCSI International Journal of Computer Science Issues, Vol. 8, Issue 1, January 2011
ISSN (Online): 1694-0814
www.IJCSI.org

MRI Mammogram Image Segmentation using NCut method and Genetic Algorithm with partial filters

Pitchumani Angayarkanni[1] M.C.A,M.Phil,(Ph.d)

Assistant Professor ,Department of Computer Science,Lady Doak College, Madurai

ABSTRACT:

Cancer is one of the most common leading deadly diseases which affect men and women around the world. Among the cancer diseases, breast cancer is especially a concern in women. It has become a major health problem in developed and developing countries over the past 50 years and the incidence has increased in recent years. Recent trends in digital image processing are CAD systems, which are computerized tools designed to assist radiologists. Most of these systems are used for automatic detection of abnormalities. However, recent studies have shown that their sensitivity is significantly decreased as the density of breast increases. In this paper , the proposed algorithm uses partial filters to enhance the images and the Ncut method is applied to segment the malignant and benign regions , further genetic algorithm is applied to identify the nipple position followed by bilateral subtraction of the left and the right breast image to cluster the cancerous and non cancerous regions. The system is trained using Back Propagation Neural Network algorithm. Computational efficiency and accuracy of the proposed system are evaluated based on the Frequency Receiver Operating Characteristic curve(FROC). The algorithm are tested on 161 pairs of digitized mammograms from MIAS database. The Receiver Operating Characteristic curve leads to 99.987% accuracy in detection of cancerous masses.

Keywords: *Filters, Normalized Cut, Segmentation, BPN, Genetic Algorithm and FROC.*

1. Introduction:

Breast cancer is one of the major causes for the increased mortality among women especially in developed countries. It is second most common cancer in women. The World Health Organization's International estimated that more than 1,50,000 women worldwide die of breast cancer in year. In India, breast cancer accounts for 23% of all the female cancer death followed by cervical cancer which accounts to 17.5% in India. Early detection of cancer leads to significant improvements in conservation treatment. However, recent studies have shown that the sensitivity of these systems is significantly decreased as the density of the breast increased while the specificity of the systems remained relatively constant. In this work **w**e have developed automatic neuron genetic algorithmic approach to automatically detect the suspicious regions on digital mammograms based on asymmetries between left and right breast image.

In this paper we have introduced the detection of microcalcifications. As one of the early signs of breast cancer , mirocalcifications are tiny granule like deposits of calcium, which appear as small bright spots of mmaograms. Their size varies from 0.1 mm to 1mm. "Cluster: of MCs is defines as a group of three to five MCs within regions. Generally microcalcification clusters are important indication of possible cancer. This algorithm effectively and automatically detect MCs

2. Algorithm Design:

There are four steps involved in the algorithm for the detection MCCs which is shown in the figure.

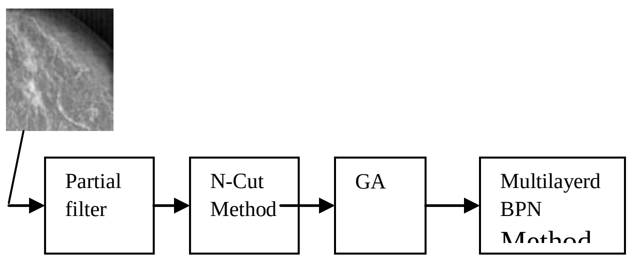


Fig 1: Flow Chart of Algorithm

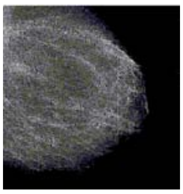
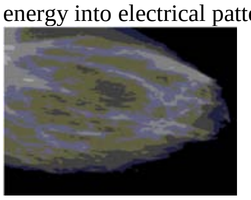
2.1 Partial Filter for Image enhancement:

A filter is a mathematical transformation (called a *convolution product*) which allows the value of a pixel to be modified according to the values of neighbouring pixels, with coefficients, for each pixel of the region to which it is applied. The filter is represented by a table (matrix), which is characterized by its dimensions and its coefficients, whose centre corresponds to the pixel concerned. The table coefficients determine the properties of the filter[1]. The following is an example of a 3 X 3 filter:

1	1	1
1	**4**	1
1	1	1

One of the most important problems in image processing is denoising. Usually the procedure used for denoising, is dependent on the features of the image, aim of processing and also post-processing algorithms [5].

Denoising by low-pass filtering not only reduces

High Pass Filter	***Low Pass filter***
MRI Mammogram Image Segmentation using NCut method and Genetic Algorithm with partial filters  **Figure 2: Mammogram Image enhanced using high pass filter**	Low pass filtering, otherwise known as "smoothing", is employed to remove high noise from a digital image. Noise is often introduced during the analog-to-digital conversion process as a side-effect of the physical conversion of patterns of light energy into electrical patterns .  **Figure 3: Mamamogram Image Enhancement using Low Pass filter**

the noise but also blurs the edges.

Spatial and frequency domain filters are widely used as tools for image enhancement. Low pass filters smooth the image by blocking detail information. Mass detection aims to extract the edge of the tumor from surrounding normal tissues and background, high pass filters (sharpening filters) could be used to enhance the details of images

3 Image Segmentation:

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. In this we apply Normalized Cut method of segmentation to cluster microcalcification regions[2].

Finally we outline the normalized cut approach of Shi and Malik [13].Here we seek a partition F and G = ~~V~~ F of the affinity weighted,undirected graph (without source and sink nodes). In order to avoid partitions where one of F or G is a tiny region, Shi and Malik propose the normalized cut criterion, namely that F and G should minimize.

$$\mathbf{N(F,G)\equiv L(F,G)(1/L(F,V)+1/L(G,V))}$$

$$\mathbf{L(F,G)=\sum_{\vec{x}_i \in F, \vec{x}_j \in G} a(\overline{x_i,x_j}).}$$

Unfortunately , the resulting graph partitioning problem.

$$F=\arg\min_{F \subset V} N(F,V-F)$$

Note any segmentation technique can be used for generating proposals for suitable regions F, for which N(F, V − F) could be evaluated. Indeed, the SMC approach above can be viewed as using S and T to provide lower bounds on the terms L(F, V) and L(G, V) (namely L(S, V) and L(T, V), respectively), and then using the S-T min cut to globally minimize L(F,G) subject to S C F and T C G. Using this method the microcalcifications are clustered

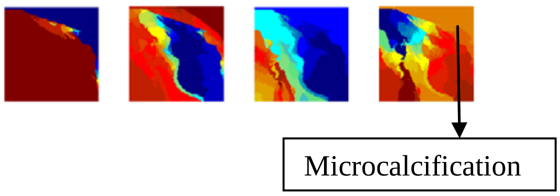


Figure 4: After Normalized Cut Segmentation

The computational efficiency 12.563 seconds on the 160x160 image.

4 Genetic Algorithm:

A partial filtering based normalized cut method is used to generate a image to separate the breast and the non breast region . The GA enhances the breast border . Border detector detects the edges in the binary images , where each pixel takes on either the intensity value of zero for a non border pixel or one for border pixel[3]. Each pixel in the binary map corresponds to an underlying pixel in the original image . In this proposed system , kernel is extracted from border points as a neighborhood array of pixels of the size 3*3 window of binary image. The binary kernels are considered population strings for GA. The corresponding kernels are extracted from gray level mammogram image using spatial coordinate points and the sum of the intensity values are considered as the fitness value . After identifying initial population and the fitness value , the genetic operator can be applied to generate a new population. Reproduction operator produces new string for crossover. Reproduction is implemented as linear search through roulette wheel with slots weighted in proportion to kernel fitness values. In this function, a random number multiplies the sum of population fitness called as stopping point.

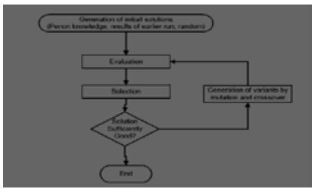


Figure 5: GA

DSS Size	Training hit Rate	Evaluation Hit Rate
10	80.6	69.7
22	81.2	74.3
58	86.2	85.7
100	86.6	82.6

Table 1: DSS Parameter and Performance

Parameter	**Value**
Population Size	400
Maximum number of tournaments	40000
Mutation Frequency	8
Crossover Frequnecy	94
Maximum Program size	516
Instruction Set	{+,-,*,/}

Table 2:Parameter Setting for Genetic Programming

5 Generating the Asymmetric Image:

After the images were aligned, bilateral subtraction was performed [47,48] by subtraction was performed by subtracting the digital matrix of the left breast image from the digital matrix of the right breast image. Microcalcification in the right breast image have positive pixel values in the image obtained after subtraction, while microcalcification in the left breast image have negative pixel values in the subtracted image. As a result, two new images were generated: one with positive values and the other with negative values. The most common gray level was zero, which indicated no difference between the left and right images. Simple linear stretching of the two generated images to cover the entire available range of 1024 gray levels was then calculated. The difference between corresponding pixels contains important information that can be used to discriminate between normal and abnormal tissue. The asymmetry image can be thresholded to extract suspicious regions. To generate FROC curve, the asymmetry image is thresholded using ten different intensity values ranges from 50-150. Figure 6 shows a asymmetry image and connected regions extracted based on thresholding to obtain a progressively larger number of high difference pixels.

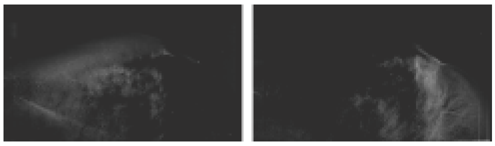


Figure 6 Asymmetric images

Two different techniques are used in the interpretation of mammogram. The first technique consists of systematic search of each mammogram for visual pattern symptomatic tumors. Such as, a bright, approximately circular blob with hazy boundary might indicate the presence of a circumscribed mass. The second technique, the asymmetric approach , consists of systematic comparison of corresponding regions in the left and the right breast.

6 BPN training:

In addition, a backpropagation artificial neural network (BP-ANN) was also developed and evaluated on the same data. The parameters for ANN training were published before. Figure 5 compare the ROC curves for the LGP and the BP-ANN algorithms respectively. The BP-ANN yielded an ROC area index of Az=0.88±0.01. Our GP approach achieved a statistically significantly better performance with Az=0.91±0.01.

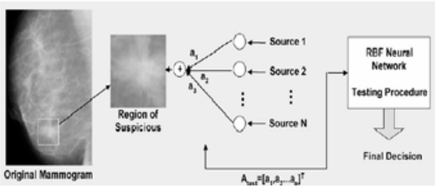
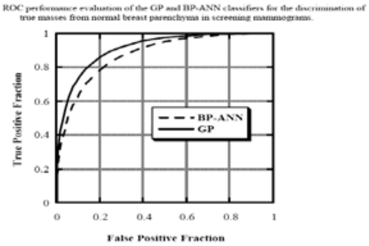


Figure 7a) Steps involved in automated Classification using Ant Colony Optimization



7. ROC curve:

Finally the technique was evaluated on the mammograms randomly selected from the non-suspicious section of the data base. The method outlined small regions in 5 out of the 15 non suspicious mammograms. The areas identified were generally very small compared to those in abnormal mammograms

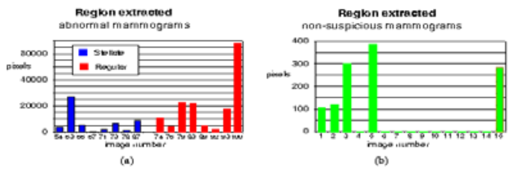


Figure 8 Lesion Areas detected for Abnormal and Non-Suspicious cases (large image extracts). [Figures (a) and (b) are presented at different ordinate scales]

Fig 8(a) shows the extracted areas for the abnormal lesions. (Image sequence 54 - 87 are stellate lesions and 74 to 100 are regular masses). We first establish whether these represent two different populations, by applying a Mann-Whitney (Wilcoxon rank sum) non-parametric test, since it is unrealistic to presume any specific underlying distribution. Median values are 450 and 1450 pixels respectively which produce a confidence level of 85% that the two data sequences emanate from distinct populations. Since this is not significant at normally acceptable levels we can compare the abnormals as a single distribution against the non-suspicious set, Fig 8(b). Using the same test, median values of 5500 and 10 pixels for the two distributions are established, giving a confidence level of greater than 97.5% that the two distributions are different, suggesting that our protocols are an effective method of area detection.

8. Results & Discussion:

In our proposed algorithm the mammogram is segmented Partial filter based enhancement and Ncut method based segmentation with clustering using Genetic Algorithmic system that optimizes the Maximizing a Posterior Probability(MAP)[4] . The Neuron Genetic Algorithm based image segmentation method is a process seeking the optimal labeling of the image pixels. Labeling process consists of assigning same label to the kernels having similar patterns. Kernel is a 3*3 window of neighborhood pixels. The Optimum label is the one which minimizes the MAP estimate.. The system is trained using Multilayered feed forward Network was found to be 99.99% accurate.

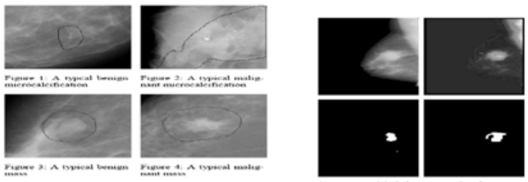


Fig 9 Original image, “voteboard”, “coarse” segmentation, “fine” Segmentation

MIAS Category	No. of Image Pairs	No. of Abnormalities
Normal	**53**	**-**
Circumscribed	**21**	**26**
Masses:Speculated Masses	**19**	**18**
III-Defined Masses	**14**	**15**
Architectural	**18**	**19**
Distortion: Asymmetry	**14**	**14**
Calcification	**22**	**25**
Total	**161**	**117**

Table 3: Tested Pairs Vs. Abnormalities

Author & Reference	Methods	Computational Time	Computational Efficiency	Detection Rate
Ferrari & Rangayyan	Directional Filtering with Gobar Wavelet	12 Sec	O(n)	74.4 %
Lau and Bischor	Asymmetry Measure	15 Sec	O(log n)	80.0 %
Sallam and Bowyer	Unwrapping Technique	10 Sec	O(n log n)	81.6 %
Proposed Approach	NCut Segmentation with GA based Neural network	2 S	O(n log n)	99.9 %

Table 4: Detection Rate and Efficiency proposed

9. Conclusion:

The proposed algorithms are tested on 161 pairs of digitized mammograms from Mammographic Image analysis Society(MIAS) database. A free response receiver operating characteristic (FROC) curve is generated for the mean value of the detection rate for all the 161 pairs of mammograms in the MIAS database, to evaluate the performance of the proposed method. There is no doubt that for the immediate future mammography will continue to play a major role in the detection of breast cancer. The ultimate objective of this paper was to identify tumor or masses in breast tissue[5]. Since hamartomas consists of normal breast tissue with abnormal proportions and the first step was try to identify the different tissue type in mammography with normal breast tissue. The important features have been extracted from the Normalized cut method of the each sub image using various statistical techniques. The Genetic algorithm has been implemented and the breast border was identified from the clustered image. The tests that were carried out using a set of 117 tissues samples, 67 benign and 50 malignant. The result analysis has given a sensitivity of 99.8%, a specificity of 99.9% and an accuracy above 99.9%, which means encouraging results. The preliminary results of this approach are very promising in characterizing breast tissue.

References:

[1] Bosch. A.; Munoz, X.; Oliver.A.; Marti. J., ***Modeling and Classifying Breast Tissue Density in Mammograms,*** Computer Vision and Pattern Recognition, 2006 IEEE Computer Society Conference on Volume 2, Issue , 2006 Page(s): 1552 – 15582.

[2] Dar-Ren Chena, Ruey-Feng Changb, Chii-Jen Chenb, Ming-Feng Hob, Shou-Jen Kuoa, Shou-Tung Chena, Shin-Jer Hungc, Woo Kyung Moond,*Classification of breast ultrasound images using fractal feature, ClinicalImage*, Volume 29, Issue4, Pages 234-245.

[3] Suri, J.S., Rangayyan, R.M.: *Recent Advances in Breast Imaging, Mammography,and Computer-Aided Diagnosis of Breast Cancer*. 1st edn. SPIE (2006)

[4] Hoos, A., Cordon-Cardo, C.: *Tissue microarray pro.ling of cancer specimens and cell lines: Opportunities and limitations*. Mod. Pathol. 81(10), 1331–1338 (2001)

[5] Lekadir, K., Elson, D.S., Requejo-Isidro, J., Dunsby, C., McGinty, J., Galletly, N.,Stamp, G., French, P.M., Yang, G.Z.: *Tissue characterization using dimensionality reduction and .uorescence imaging*. In: Larsen, R., Nielsen, M., Sporring, J. (eds.) MICCAI 2006. LNCS, vol. 4191, pp. 586–593. Springer, Heidelberg (2006).

About the Author:

Pitchumani Angayarkanni Sekaran is Assistant Professor at Lady Doak College, Madurai,Tamil Nadu,India. My research interests include “Image Processing and data mining”.