==========
Engineering & Technology in India www.engineeringandtechnologyinindia.com
Vol. 1:3 April 2016
==========

MAMMOGRAPHIC MICROCALCIFICATION SEGMENTATION USING FUZZY C MEANS CLUSTERING

Mekala.S., M.E.
Nishanthi.C., M.E. Student

==========

Abstract

Breast cancer is one of the major causes of death among women. Mammography is the main test used for screening and early diagnosis. Early detection performed on X-ray mammography is the key to improve breast cancer prognosis. This paper presents a research on mammography images using Morphological operators and Fuzzy c – means clustering for cancer tumor mass segmentation. The first step of the cancer signs detection should be a segmentation procedure able to distinguish masses and micro calcifications from background tissue using Morphological operators and finally fuzzy c- means clustering (FCM) algorithm has been implemented for intensity – based segmentation. This method does not require any manual processing technique for classification, thus it can be assimilated for identifying benign and malignant areas in intelligent way. Moreover it gives good classification responses for compressed mammogram image. The goal of the proposed method is twofold: one is to preserve the details in Region of Interest (ROI) at low bit rate without affecting the diagnostic related information and second is to classify and segment the micro-calcification area in reconstructed mammogram image with high accuracy. The experimental result shows that the proposed model performance is good at achieving high sensitivity of 97.27%, specificity of 94.38% .

Key words: Mammography, Micro calcification, Segmentation, Fuzzy c- means clustering.

I. INTRODUCTION

Breast cancer is the most common malignancy that affects women worldwide and is the

leading cause among non-preventable cancer death. The American Cancer Society (ACS) estimates that on an average, in every 15 minutes five women are diagnosed with breast cancer. It is also estimated that one in eight women will be diagnosed with this disease in her lifetime, and 1 in 30 will die from it. Breast cancer is the second most prevalent cancer among Indian women, the first being cervical cancer. In the age group of 30-70 years, one in fifty eight women are affected by this disease and the occurrence is mainly seen in the urban areas. Mammography is the best technique for reliable detection of early, non- palpable, potentially curable breast cancer. As a result of the increasing utilization of mammographic screening, the mortality rate due to this disease was observed to decrease for the first time in 1995. Since the interpretation of mammograms is a repetitive task that requires much attention to minute details, the opinion of radiologists may vary. To overcome this difficulty, during the past decade, the use of image processing technique for Computer Aided Diagnosis (CAD) in digital mammograms has been initiated. This has increased diagnostic accuracy as well as the reproducibility of mammographic interpretation.

II. MICROCALCIFICATION IDENTIFICATION

To investigate the potential correlation between the topology of micro-calcification clusters and their pathological type a series of micro- calcification graphs are constructed to describe the topological structure of micro-calcification clusters at different scales. A set of graph theoretical features are extracted from these graphs for modeling and classifying micro-calcification clusters.

The identification methodology consists of four main phases: estimating the connectivity between micro-calcifications within a cluster using morphological dilation at multiple scales; generating a micro-calcification graph at each scale based on the spatial connectivity relationship between micro- calcifications; extracting multi-scale topological features from these micro-calcification graphs; and using the extracted features to build classifier models of malignant and benign micro-calcification clusters. All image analysis development work was done within MATLAB 2014a.

Morphological dilation is performed on each individual micro-calcification using a disk- shaped structuring element at multiple scales. Here the scale corresponds to the radius of the structuring element measured in pixels. The effect of multi-scale morphological dilation on a micro-calcification cluster is shown. It can be seen that the multi-scale morphological dilation continuously absorbs neighboring pixels into individual micro- calcifications resulting in a change in the connectivity between micro-calcifications within the cluster.

A. MICROCALCIFICATION GRAPH GENERATION

The topology of micro-calcification clusters is represented in graph form. A micro-calcification graph is generated based on the spatial connectivity relationship between micro-calcifications within a cluster. In a micro-calcification graph, each node represents an individual micro-calcification, and an edge between two nodes is created if the two corresponding micro-calcifications are connected or overlap in the 2-D image plane. The node locations in the graphs are in accordance with the original spatial distribution of micro-calcifications within the two clusters, and the node sequences are consistent with those in Fig.3.1, which are sorted in a left-to- right and bottom-to-top direction.

B. MULTI SCALE FEATURE EXTRACTION

After generating micro-calcification graphs over a range of scales, a set of graph theoretical features can be extracted to capture the topological properties of micro-calcification clusters. These features will constitute the feature space for the classification of malignant and benign clusters. Before extracting the topological features of micro- calcification clusters, the following definitions for general graphs are provided. Micro-calcification cluster is more connected. A set of values of the maximum vertex degree against scale which also have an increasing trend from small to large scales are tend toward stability when reaching the maximum value. Similarly, as indicated by the average vertex degree, the maximum vertex degree values for the malignant cluster are also larger than those of the benign cluster. The resulting values of the average vertex eccentricity against scale are plotted. At the first few scales, most micro- calcifications are isolated from others in the cluster, which results in small average eccentricity values (the eccentricity of isolated vertices is set to 0).

C. HIERARCHICAL CLUSTER ANALYSIS

The stepwise procedure attempts to identify relatively homogeneous groups of cases based on selected characteristics using an algorithm either algomerative or divisive, resulting to a construction of a hierarchy or treelike structure depicting the formation of clusters. This is one of the most straightforward method.HCA are preferred when the sample size is moderate (under 300 – 400, not exceeding 1000).

D. CLASSIFICATION

The main difficulty here is the accurate extraction of a band of pixels around the segmented mass. Claridge and Richte used the Polar co- ordinate transform (PCT) to map lesions into polar co-ordinates. A spiculation measure was then computed from the PCT images to discriminate between circumscribed and spiculated masses. Hadjiski classified masses as benign or malignant using texture features computed from the RBST image. They tested the performance of a hybrid classifier consisting of an adaptive resonance theory network cascaded with LDA. They used a set of manually segmented ROIs and reported a higher accuracy with the hybrid classifier than with a back propagation neural network or LDA. Pohlman used six morphological features to classify masses as benign or malignant. To segment the lesions, they used an adaptive region growing technique, which required the selection of manual seed points.

III. PROPOSED WORK

The relative fibroglandular tissue content in the breast, commonly referred to as breast density, has been shown to be the most significant risk factor for breast cancer after age. This work presents a novel multi-class fuzzy c-means (FCM) algorithm for fully-automated identification and quantification of breast density, optimized for the imaging characteristics of digital mammography. The proposed algorithm involves adaptive FCM clustering based on an optimal number of clusters derived by the tissue properties of the specific mammogram, followed by generation of a final segmentation through cluster agglomeration using linear discriminant analysis. When evaluated on 80 bilateral screening digital mammograms, a strong correlation was observed between algorithm- estimated PD% and radiological ground-truth of $r=0.83$ ($p<0.001$) and an average Jaccard spatial similarity coefficient of 0.62. These

results show promise for the clinical application of the algorithm in quantifying breast density in a repeatable manner. Besides of analysing performance of AIS, we also compared the performance of AIS as a data reduction method on two real world classification problems. They are Diabetes Disease and Breast Cancer classification problems. The related datasets were taken from the UCI data mining repository. In these experimentations however, train & test data partitioning was conducted in a different way. Firstly, the training and testing data were determined and then the training data were reduced with AIS and FCM methods. In this step, data were reduced so that approximate compression ratios of 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, 95%and 98% were obtained.

A. METHODOLOGY

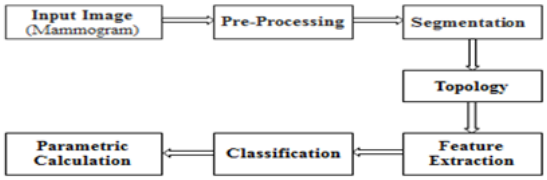

Fig.1Basic Flow Diagram

B. PREPROCESSING

The preprocessing phase of the proposed system is focused at removal of channel noise, enhancing the contrast and for removal of the background of mammogram images. The ROI containing abnormalities are separated from the background and further features are computed from the ROI. Channel noise is considered as salt and pepper noise which is removed using median filter (Jae, 1990) whereas histogram equalization technique (Nunes*et al*., 1999) is used to enhance the contrast and Otsu Global threshold (Otsu, 1979) is used for extracting the background from ROI

C. SEGMENTATION

This section details the segmentation of mammograms for identifying the cancer in breasts. The proposed approach utilizes mathematical morphology operations for the segmentation. The morphological operations are applied on the grayscale mammography images to segment the abnormal regions. Erosion and dilation are the two elementary operations in

Mathematical Morphology. An aggregation of these two represents the rest of the operations . The symbols _, Θ, o, and • , respectively denote the four fundamental binary morphological operations: dilation, erosion, opening, and closing. A function $f(x, y)$ denotes the image, where (x, y) _ $R2$ or Z2 ,or simply f , and the function $h(x, y)$, or h will act as the structuring element.

D. FUZZY C-MEANS ALGORITHM

Fuzzy C-means algorithm is also called as ISODATA. It was most frequently used in pattern recognition. Fuzzy C-mean is the method using in clustering. It is using one piece of data to belong to two or more clusters. It always based on minimization of objective functions to achieve a good classification.

E. SEGMENTATION ALGORITHM USING FUZZY-C MEANS ALGORITHM

Step 1: Read the input mammogram image and decide the number of clusters C.

In this C=3

Step 2: Assign the value of ε(threshold) and number of iteration as T.

Step 3: Assign the cluster centres,

$V^{(i)}=[V^{(i)},V_2^{(i)}\ldots.V_c^{(i)}]$

Step 4: Evaluate the degree of membership function

Step 5: Evaluate the centres of clusters $^{(q+1)}$

Step 6: If $\|V^{(q+1)} - VV^{(q)}\| <$ i or the number of iteration q>1 then write the Output as clustering output, or else q=q+1 go to step4

Step 7: Extract the cancerous area from clustered output; perform morphological operations to calculate the area of the cancerous region.

F. PERFORMANCE EVALUATION OF FUZZY C-MEANS

The most important difference is that in FCM, each has a weighted associated with a specify cluster so, a point doesn't have cluster as much as a little or more association to

cluster. It was very determine by increase distance to the center of the cluster. FCM will tend to run slower than K-means, since it is actually doing more work. Every point is calculated with each cluster, and many operations are involved in each evaluation.

In this proposed work we make these differences or weakness our strong point for full detection of breast cancer from this we were able to find the messes and the cancer area.

G. FEATURE EXTRACTION

As the case for masses, the features used for the diagnosis of calcification can be viewed as either capturing morphological or texture information. Researchers have reported that morphology is one of the most important clinical factors in calcifications diagnosis. Features for calcification classification can also be organized in terms of whether the describe properties of the cluster as a whole or of the individual calcifications that make up the cluster.

IV. RESULT

The proposed method is implemented by using MATLAB software in programming level. MATLAB is a data analysis and visualization tool which has been designed with powerful support for matrices and matrix operations. As well as this, MATLAB has excellent graphics capabilities, and its own powerful programming language. When dealing with mammograms, it is known that pixels of tumor regions tend to have maximum allowable digital value. Based on this information, morphological operators are used such as Dilation is used to detect the possible clusters which contain masses. Image features are then extracted to remove those clusters that belong to background or normal tissue as a first cut..The fuzzy c-means clustering algorithm is used as a segmentation strategy to function as better classifier & aims to class data into separate groups according to their characteristics.

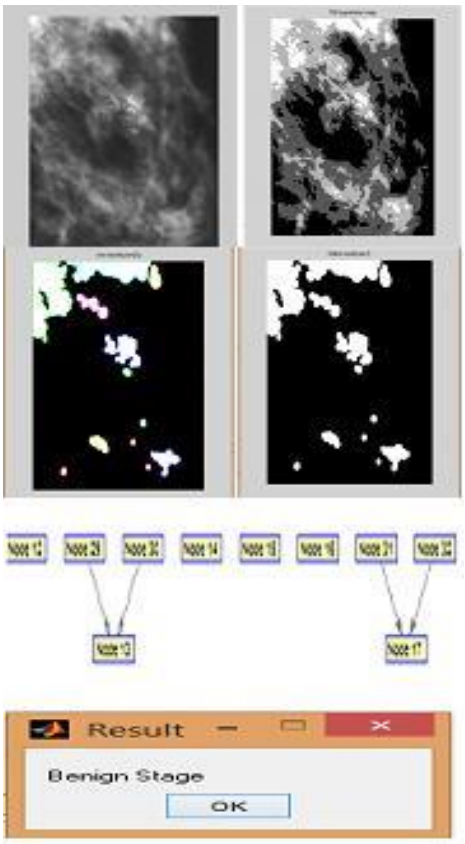

Fig 2. Processing and indication stages

V. CONCLUSION

The early diagnosis through regular screening and timely treatment has been shown to prevent cancer. In this paper we have presented a novel approach to identify the presence of breast cancer mass in mammograms. The proposed work utilizes fuzzy c- means clustering for clear identification of clusters. The FCM is a new approach, using this we have successfully detected the breast cancer masses in mammograms. This result indicates that this system can facilitate the doctor to detect breast cancer in the early stage of diagnosis process.

==========

REFERENCES

[1] American Cancer Society, Global Cancer Facts & Figures, 2nd ed. Atlanta, GA, USA: Amer. Cancer Soc., 2011, pp. 11–12.

[2] Eurostat, Health Statistics Atlas on Mortality in the European Union. Office for Official

Publications of the European Union, 2009, pp. 91–93.

[3] American Cancer Society, Breast Cancer Facts & Figures 2011– 2012. Atlanta, GA, SA: Amer. Cancer Soc., 2011, pp. 11–15.

[4] National Comprehensive Cancer Network, NCCN Clinical Practice Guidelines in Oncology: Breast Cancer Screening and DiagnosisNCCN Clinical Practice Guidelines in Oncology: Breast Cancer Screening and Diagnosis, 2012.

[5] Cancer Research U.K. (2012, Aug.) "Who is screened for breast cancer?"

[6] L. Shenet al., "Application of shape analysis to mammographic calcifications," IEEE Trans. Med. Imag., vol. 13, no. 2, pp. 263–274, Jun. 1994.

[7] P. Dhawanet al., "Analysis of mammographic micro calcifications using gray-level image structure features," IEEE Trans. Med. Imag., vol. 15, no. 3, pp. 246–259, Jun. 1996.

[8] H. D. Cheng et al., "Computer-aided detection and classification of micro calcifications in mammograms: A survey," Pattern Recog., vol. 36, no. 12, pp. 2967–2991, 2003.

[9] Y. Ma, et al., "A novel shape feature to classify micro calcifications," in Proc. 17th IEEE Int. Conf. Image Process., 2010, pp. 2265–2268.

==========

Mekala. S., M.E.
mekala.4138@gmail.com

Nishanthi. C., M.E. Student
nishachinnasamy93@gmail.com

Department of Electronics & Communication Engineering
Sri Subramanya College of Engineering & Technology
NH - 209, Sukkamanaickenpatti
Palani 624615
Tamil Nadu
India