A Novel Approach for Detection of Breast Cancer at an early stage using Digital Image Processing techniques

Sangeetha R
Assistant Professor, Dept. of CSE
PES Institute of Technology, South Campus
Bangalore, India

Dr. Srikanta Murthy K
Professor, Dept. of CSE
Nagarjuna College of Engg.,
Bangalore, India

Abstract— **Breast Cancer is highly predominant in women in today's world. It can start in the breast and can spread to other areas of the body in the course of time. Breast cancer is the second largest disease leading to the death of women. The disease is curable if detected early enough .A lot of research is being done to detect the cancer at the earliest. Early detection at the microcalcification stage can be useful for providing proper treatment at the early stage and saving the patients. Research on breast cancer using digital image processing is not new but lack of proper methods for early detection at microcalcification stage is still a challenge to medical domain and still it is considered as the deadliest. Most of the research work done till now detects the breast cancer at tumor stage and are not accurate to 100% and leads to false positive or false negative results which are highly dangerous. And also they do not provide end to end solution. A novel methodology is proposed in this paper using a combination of different highly efficient techniques of digital image processing which are not yet being implemented in this research area. Using the methodology proposed in this paper it is possible to detect the breast cancer at a very early microcalcification stage itself and the result of this proposed methodology will be of very high accuracy leading to true positive and true negative results. The methodology proposed in this paper provides end to end solution.**

Keywords—Dilation, Otsu's thresholding algorithm, K-Nearest Neighbor Classifier, Bayes Classifier,Gray-level Co-occurrence Matrix, Sobel Edge Detection Mask

I. INTRODUCTION

The previous research work carried out in this area:

In [1] a method was proposed to identify the abnormal masses in mammography image by dividing the preprocessed mammogram into blocks and segmentation was performed through color quantization technique.

In [2] the image was enhanced and segmented to obtain the suspicious regions and then features were extracted by using gray level co-occurrence matrix. These features were used to classify them into benign or malignant masses.

In [3] a method was proposed for segmentation approach that combines watershed algorithm with graph theory based algorithm Grab Cut.

In [4] a method was proposed to enhance the image by applying multilevel wavelet transformation and then the statistical features were extracted. Then the extracted features were used to detect cancer.

Asymmetry is one of the important parameter to identify breast cancer. In [5] a method was proposed to align the images of left and right breast and to identify the asymmetry.

In [6] a method was proposed for increasing the contrast of microcalcification using a method of equalization of histogram of the image.

In [7] a method was proposed for segmentation using watershed algorithm. Then Fourier transformation was used to obtain the relevant features. These features were used to identify the tumors.

In [8] morphological operators were used to segment the image and then Artificial Neural networks was used to classify the cancer cells.

Abnormality in shape is one of important parameter to identify breast cancer. In [9] a method was proposed to identify shape abnormality using Gabor filters.

In [10] a system was proposed in which the mammogram was preprocessed through contrast improvement and grayscale thresholding, segmented using predefined threshold. Then edge detection was performed to display edges over masses and then morphological filtering was performed to obtain the clear and distinct masses. Then the processed mammogram was super imposed on the original mammogram to detect the lesion/tumor out the masses obtained.

In [11] a method was proposed for classifying malignant and benign microcalcification clusters based on morphological topology analysis.

In [12] the mammographic image was segmented using Linde Buzo and Gray(LBG) algorithm to generate a codebook of size 128. Then using the same LBG algorithm the generated code vectors were clustered to form 8 clusters.

II. PROPOSED METHODOLOGY

The proposed methodology consists of following steps:

1. Image Preprocessing
2. Image Segmentation
3. Feature Extraction
4. Feature Classification

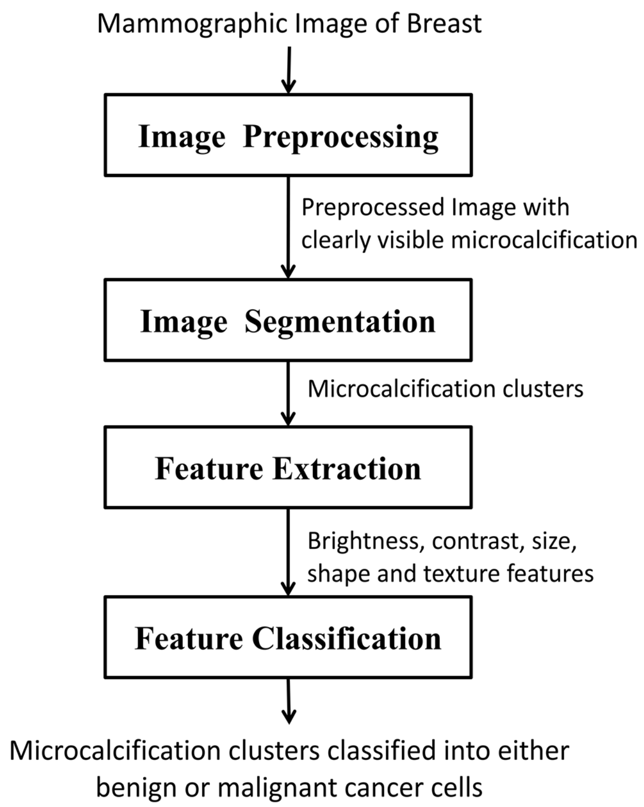

Figure 1.Block diagram of proposed methodology

Image Preprocessing

This is mainly performed to remove the noise in the image.
This can be achieved by following steps:

1. By using median filters.
2. Then the contrast in the resultant image should be increased by setting a certain threshold value. The threshold can be determined using histogram. This can be achieved by histogram modified contrast limited adaptive histogram equalization(HM-CLAHE) method.
3. Then the contrast of microcalcification cells can be further enhanced using dilation which is a morphological operator.

Dilation of image A by structuring element B is given as,

$$A \oplus B = \{Z \,|\, (\hat{B})_Z \cap A \neq \emptyset\}$$

Where B is reflected about its origin and this reflection is shifted by z.

At the end of pre-processing stage, the noise in the image will be eliminated and the contrast of microcalcification cells will be enhanced to a greater extent, due to which they will be clearly visible in the image.

Image Segmentation

The segmentation of the image is performed to segment the microcalcification cells. This achieved by using morphological operator and applying Otsu's thresholding algorithm. This can be done as follows:

1. The image can be filtered with morphological white top hat for detecting microcalcification cells which are small bright particles in a slow varying background. The top hat will remove the background without reducing the microcalcification cells
   Top hat transformation = image - Open(image)
2. Fixing an optimal threshold to segment the filtered image separating microcalcification cells from the background using Otsu's thresholding algorithm.

Otsu's thresholding algorithm:

a. Compute the histogram of the image. Let each gray level have probability p_i.
b. Calculate the aggregate sum $P_1(k)$ for k =0,…,L-1

$$P_1(k) = \sum_{i=0}^{k} p_i$$

c. Calculate the aggregate mean $m_1(k)$ for k =0,…,L-1

$$m_1(k) = \frac{1}{P_1(k)} \sum_{i=0}^{k} ip_i$$

d. Calculate the overall intensity mean m_G

$$m_G = \sum_{i=0}^{L-1} ip_i$$

e. Calculate variance between the class σ_B^2 for k =0,…,L

$$\sigma_B^2 = P_1(k)\,(m_1(k) - m_G)^2 + P_2(k)\,(m_2(k) - m_G)^2$$

f. Find the value of k for which $\sigma_B^2(k)$ is maximum. Determine this as the Otsu threshold k*.

g. Obtain separability measure ŋ*

$$ŋ^* = \frac{\sigma_B^2\,(k^*)}{\sigma_G^2}$$

Otsu's thresholding algorithm will segment the microcalcification cells from the image with a very high accuracy.

3. The identified microcalcification cells can then be grouped into clusters using K- Nearest neighbor clustering algorithm.

For the k-NN classifier the nearest k features are considered, and the new feature is assigned the label of the most frequently occurring label in the k nearest neighbors.

Feature Extraction

Many features such as brightness, contrast, size, shape and texture can be extracted from microcalcification clusters by using following methods:

1. Obtaining statistical moments of the gray level histogram of the region. The nth moment of statistics of gray level histogram can be obtained using the equation:

$$\mu_n(v) = \sum_{i=0}^{A-1} (v_i - m)^n\, p(v_i)$$

Where v is a discrete random variable, $p(v_i)$ is an estimate of the probability of value v_i (i=0,1,2,…….A-1) occurring, and m is

$$m = \sum_{i=0}^{A-1} v_i\ p(v_i)$$

2. Constructing gray level co-occurrence matrix. Co-occurrence matrix describes the texture by finding the occurrence of certain gray levels. Various descriptors can be obtained from the co-occurrence matrix including the maximum probability, entropy, uniformity.

3. Applying Sobel edge detection mask. The Sobel edge detection mask are

-1	-2	-1
0	0	0
1	2	1

-1	0	1
-2	0	2
-1	0	1

The features extracted using these methods will be of very high quality and accuracy and they can be used for classification.

Feature Classification

The extracted features can be used to classify the microcalcification clusters into either benign or malignant cancer cells using Bayes' Classifier (which is a very robust classifier).

Decision function of Bayes' Classifier:

$$d_j(x) = p\left(\frac{x}{\omega_j}\right) P(\omega_j)$$

for j = 1,2,……W classes and where x is a pattern vector which is assigned to the class of largest numerical value, $p\left(\frac{x}{\omega_j}\right)$ is the probability of pattern x coming from ω_j, $P(\omega_j)$ is the probability of occurrence of class ω_j.

III. Expected outcome

The proposed methodology will identify the microcalcification clusters at a very early stage and classify them as benign or malignant cancer cells with a very high accuracy and will lead to true positive and true negative results.

References

[1] Indra Kanta Maitra, Sanjay Nag, Prof. Samir Kumar Bandyopadhyay, "Identification of Abnormal Masses in Digital Mammography Images", *International Journal of Computer Graphics,* **2(1), (2011)**

[2] Khuzi A Mohd, Besar R, Wan Zaki WMD, Ahmad NN, " Identification of masses in digital mammogram using gray level co-occurrence matrices", *Biomed Imaging IntervJ,* **5(3),** 1-13 **(2009**)

[3] Yan Zhang, Xiaoping Cheng, "Medical image segmentation based on watershed and graph theory", *Image and Signal Processing (CISP)*, 3, 1419-1422(2010)

[4] Vidhya M., Sangeetha N., Vimalkumar M.N., Helenprabha K., "Early stage detection of cancer in mammogram using statistical feature extraction", Recent Advancements in Electrical, Electronics and Control Engineering (ICONRAEeCE), 401-404 (2011)

[5] Alina Sultana, MihaiCiuc, RodicaStrungaru and Laura Florea., " A New Approach in Breast Image Registration", *International Conference on Intelligent Computer Communication and Processing*, 149-154 (2010)

[6] M. Sundaram, K. Ramar, N. Arumugam, G. Prabin., "Histogram Based Contrast Enhancement for Mammogram Images", *International Conference on Signal Processing,Communication, Computing and Networking Technologies*,628-503 (2011)

[7] Essafij, R. Doughrij,s. M'hiri k. Ben romdhane and f. Ghorbelm, "Segmentation and classification of breast cancer cells in histological images" ,*International Conference of the IEEE Engineering in Medicine and Biology Society*, (2010)

[8] Heba Al-Hiary, Basim Alhadidi, " An implemented approach for potentially breast cancer detection using extracted features and artificial neural networks", Computing and Informatics, Vol. 31, 2012, 225–244

[9] Banik, S., Rangayyan, R. M., & Desautels, J. E. L. "Detection of Architectural Distortion in Prior Mammograms Using Measures of Angular Dispersion,Medical measurements and application" proceedings(MeMeA),2012 *IEEE international symposium*

[10] Naishil N. Shah, Tushar V. Ratanpara, C.K.Bhensdadia, " Early Breast Cancer Tumor Detection on Mammogram Images" , International Journal of Computer Applications (0975 – 8887) Volume 87 – No.14, February 2014

[11] Z. Chen, A. Oliver, E. Denton, and R. Zwiggelaar. "Classification of Microcalcification Clusters Using Topological Structure Features", Proceedings of the 16th Medical Image Understanding and Analysis[MIUA2012], pp. 37-42, Swansea, UK, 09-11 July, 2012.

[12] Kekre H.B., Sarode Tanuja K. and Gharge Saylee M., " Tumor Detection in Mammography Images using Vector Quantization Technique", *International Journal of Intelligent Information Technology Application*, **2(5),** 237- 242 **(2009)**

[13] "Feature Extraction and Image Processing", Mark S. Nixon and Alberto S. Aguado, , Second Edition, Academic Press(imprint of Elsevier)

[14] "Digital Image Processing for Medical Applications", GEOFF DOUGHERTY, First Publication, Cambridge University Press

[15] "Digital Image Processing", Rafael C. Gonzalez and Richard E. Woods, Third Edition, Prentice Hall Publication