

Texture Analysis of Mammographic images

D.A.Kulkarni
Gogte Institute of
Technology
Belgaum-590008
India

Bhagyashree.S.M
Gogte Institute of
Technology
Belgaum-590008
India

G.R.Udupi
Vishvanthrao
Deshapande
Rural Institute of
Technology Haliyal
India

ABSTRACT

Breast cancer is the most common type of cancer among women in the world. Mammography is regarded as an effective tool for early detection and diagnosis of breast cancer. Microcalcification is one of the primary signs of breast cancer. There are various image texture analysis techniques for the detection of the microcalcifications. Screen-film mammography is still the standard method used to detect early breast cancer, thus leading to early treatment. Digital mammography has recently been designated as the imaging technology with the greatest potential for improving the diagnosis of breast cancer. In this work a feature-based approach is used for analysis and classification of malignancy. Gray-level texture and Wavelet co-efficient texture methods are used for feature extraction. Probabilistic Neural Network (PNN) is used for classification of images based on extracted features. The performance of classification by PNN based on features by texture method, wavelet method and combined methods are compared. The Receiver Operating Characteristics (ROC) Analysis is used for performance evaluation.

General Terms

Soft Computing, Image Processing, Computer Aided Diagnosis.

Keywords

Breast Cancer, Microcalcification, Gray-level texture analysis, Wavelet.

1. INTRODUCTION

Breast cancer is leading cause of death among all cancer diseases for middle-aged and older women. Its incidence is increasing as well. Hence, prevention and early diagnosis are very important. Unfortunately, primary prevention is not possible since the cause of this disease is not yet understood. However current methods of treatment are very effective in its early phase. Therefore, removal of the cancer while it is in its early stages is the most promising way to achieve a significant increase in survival rates. At present mammography is shown to be one of the most reliable methods for early detection of breast cancer.

Mammography is the process of using low-dose X-rays to examine the human breast and is used as a diagnostic as well as a screening tool. The goal of mammography is the early detection of breast cancer, typically through detection of characteristic masses and microcalcifications. Microcalcification is considered to be important sign of breast cancer [1]. It has been reported that 30–50% of breast cancer detected and 60–80% of breast carcinomas reveal microcalcification upon histological examinations.

The task of detection of microcalcifications for the diagnosis of breast cancer is difficult one. Dense breasts, improper technical factors, or simple oversight by radiologists may contribute to the failure of detecting microcalcifications. In the following paragraphs, some related works based on mammography are presented.

Liang Shen et al. [2] proposed “Application of Shape Analysis to Mammographic Calcifications”,

which is a set of shape features of individual microcalcifications. They have developed a set of shape factors to measure the roughness of contours of calcifications in mammograms and for use in their classification as malignant or benign. Jay et al. [3] proposed, "Artificial Neural Network Improving the Quality of Breast Biopsy Recommendations". Purpose of this is to determine an artificial neural network (ANN) to categorize benign and malignant breast lesions.

Lee et al. [4] proposed "Classification of clustered microcalcifications using a shape cognitron neural network". A new shape recognition-based neural network built with universal feature planes, called Shape Cognitron (S-Cognitron) is introduced to classify clustered microcalcifications. Chan et al. [5] proposed "Computerized analysis of mammographic microcalcifications in morphological and texture feature spaces", used co-occurrence matrices-based features extracted from region of interest (ROI), containing the microcalcifications.

The aim of this paper is to investigate multiscale texture properties of the tissue surrounding microcalcification for breast cancer diagnosis. Four texture features are extracted from the region of interest of the image. Next a different set of eleven feature are extracted using wavelet method. The extracted feature sets are compared by means of their ability in discriminating malignant from benign tissue using a probabilistic neural network (PNN) classifier. Performance of individual feature sets and the combined scheme is assessed by means of a Receiver Operating Characteristics (ROC) Analysis.

1. PROPOSED DESIGN

The proposed design contains six steps as shown in figure 1.

2.1 Description of Proposed Design

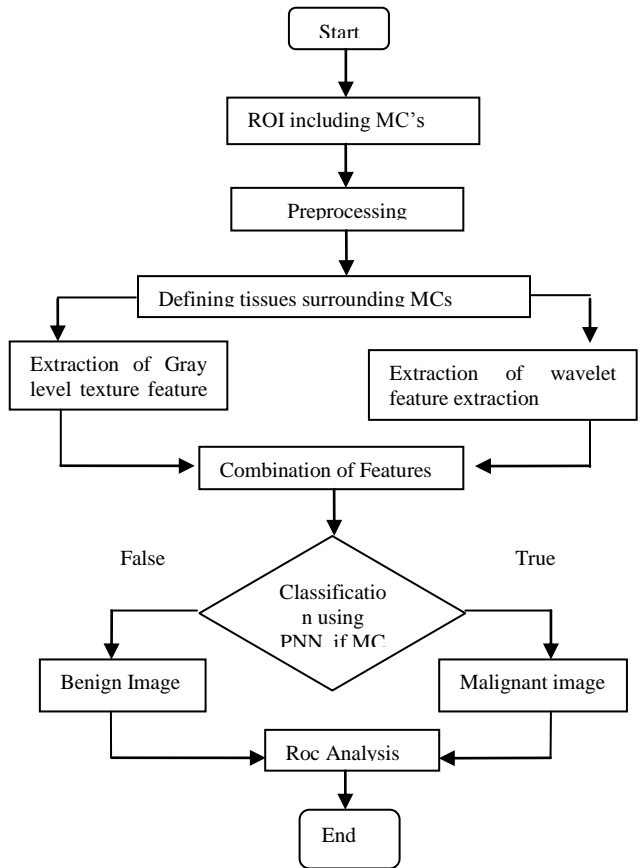


Figure 1 Flowchart of the Classification Scheme

The description for each step is as follows.

Step 1: Region of Interest (ROI)

First consider the region of interest in the pathology part so as to extract the features.

Step 2: Preprocessing

Before taking ROI Preprocessing step is performed. A preprocessing step is performed in order to facilitate the subsequent MC segmentation task. Convert the original image into Negative image. Manually 40x40 matrix is assigned to the negative image to get segmented image. In the segmented image, the pixel having high intensity value in the digital image is selected, then that pixel is compared to the neighboring pixels. The comparison goes on till there is change in the pixel value. This is how region of interest is found.

Step 3: Methods for features extraction

Gray Level Texture Feature Extraction and Wavelet Coefficient Texture Feature are used.

Step 4: Combination of features

In this step, Gray Level Texture Features and Wavelet Coefficient Texture Features are combined.

Step 5: Probabilistic Neural Network

A PNN is used for the classification of tissue surrounding microcalcification

Step 6: ROC Analysis

The ROC is used for the evaluation of the probabilistic neural network.

2. IMPLEMENTATION

In implementation, the steps carried out for selecting the region of interest from the original image are as explained above. The mammogram images may contain suspicious regions that have special texture characteristics. So, to study the texture of these mammogram regions, the first order and second order features are extracted.

2.1 Selection of Region of Interest

The Region of interest in mammogram needs to be selected before the extraction of texture features and classification process. The ROI is selected from the original image. The sample image is as shown in figure 2 with MC's present. The original mammogram with the inverted values called "Negative Image" is as shown in figure 3. It allows to mark the microcalcifications in the original image. Selection of ROI is as follows: First, the pixel having highest intensity value in the digital image is selected, then that pixel is compared to the neighboring pixels. The comparison goes on till there is change in the pixel value. All the pixels having similar intensity form the region of interest. This is how region of interest is found.

The Preprocessing step is carried out as shown in figure 4. The figure 5 shows the output image surrounding tissue region of interest from original image. After finding the region of interest of the images, feature extraction is carried out, as shown in Results.

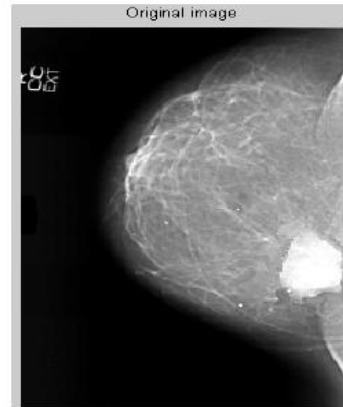


Figure.2 Mammogram Original Image

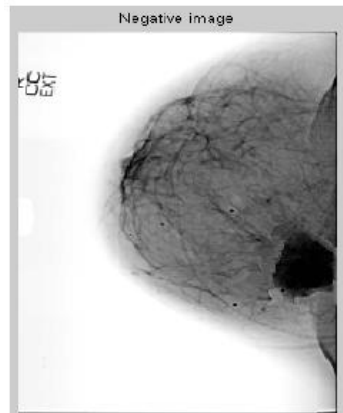


Figure.3 Mammogram negative Image

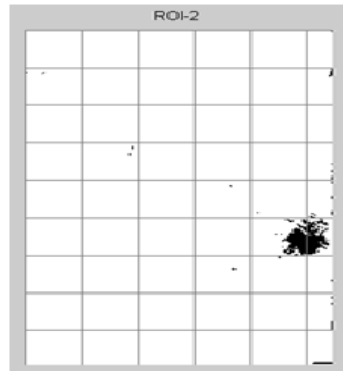


Figure.4 Segmented ROI with MC's

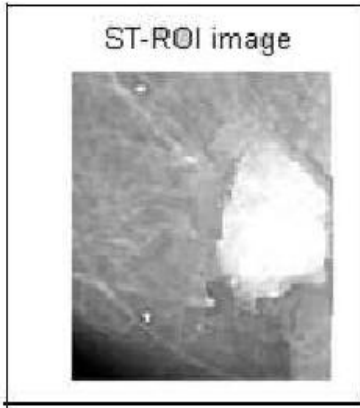


Figure.5 Surrounding Tissue ROI

3. RESULTS

After finding the region of interest of the images, feature extraction is carried out. The feature extraction is carried out at two levels, first order and second order feature extraction. The features are extracted for both Gray-level and Wavelet methods. In the first order feature extraction following features are extracted: Mean, Standard deviation, Variance, and Energy. Second order features extraction has been done as similar to first order features extraction. Second order features are [6] as shown below:

Entry in a normalized gray-tone spatial-dependence matrix,

$p_x(i)$ = i th entry in the marginal-probability matrix obtained by summing the rows of

$$= \sum_{j=1}^{N_g} P(i, j)$$

N_g Number of distinct gray levels in the quantized image.

\sum_i and \sum_j $\sum_{i=1}^{N_g}$ and $\sum_{j=1}^{N_g}$, respectively

$$P_y(y) = \sum_{i=1}^{N_g} p(i, j)$$

$$P_{x+y} = \sum_{i=1}^{N_g} \sum_{\substack{j=1 \\ |i-j|=k}}^{N_g} p(i, j), k = 0, 1, \dots, N_g - 1$$

$$P_{x+y} = \sum_{i=1}^{N_g} \sum_{\substack{j=1 \\ i+j=k}}^{N_g} p(i, j), k = 2, 3, \dots, 2N_g$$

Texture Features are as follows

1. Angular second Movement

$$f_1 = \sum_i \sum_j \{p(i, j)\}^2$$

2. Correlation

$$f_2 = \sum_{n=0}^{N_g-1} n^2 \left\{ \sum_{i=1}^{N_g} \sum_{\substack{j=1 \\ |i-j|=n}}^{N_g} p(i, j) \right\}$$

3. Contrast

$$f_3 = \frac{\sum_i \sum_j (ij) p(i, j) - \mu_x \mu_y}{\sigma_x \sigma_y}$$

4. Sum of Squares

$$f_4 = \sum_i \sum_j (i - \mu)^2 p(i, j)$$

5. Inverse Difference Moment

$$f_5 = \sum_i \sum_j \frac{1}{1 + (i - j)^2} p(i, j)$$

6. Sum Average

$$f_6 = \sum_{i=2}^{2N_g} i P_{x+y}(i)$$

7. Sum Entropy

$$f_7 = \sum_{i=2}^{2N_g} (i - f_6) p_{x+y}(i)^2$$

8. Sum Variance

$$f_8 = - \sum_{i=2}^{2N_g} p_{x-y}(i) \log\{P_x(i)\}$$

9. Entropy

$$f_9 = - \sum_i \sum_j p(i,j) \log(p(i,j))$$

10. Difference Entropy

$$f_{10} = \text{variance of } p_{x-y}.$$

11. Difference Variance

$$f_{11} = - \sum_{i=0}^{N_g-1} P_{x-y}(i) \log\{P_{x-y}(i)\}$$

3.1 First order feature extraction

The following tables show the features for first order. Tables 1, 2 and 3 below show the values of Statistical features for Cancerous, Normal and ROI images. The table 3 shows the features for ROI images have higher values as compared to normal images. Because the ROI images have high intensity value pixels in the image.

Table 1. First order Statistical features for Cancerous Images

Images	Mean	StdDev	Variance	Energy
Cmam1	72.09	44.436	1974.54	99.862
Cmam2	18.337	26.745	715.318	99.85
Cmam3	49.115	33.925	1150.91	99.925
Cmam4	34.401	24.446	597.60	99.002
Cmam5	37.671	23.886	570.55	91.913

Table 2. First order Statistical features for Normal Images

Images	Mean	StdDev	Variance	WEnergy
Nom1	93.222	15.37	236.25	99.295
Nom2	39.73	19.936	397.456	93.884
Nom3	44.78	29.612	876.849	92.56

Nom4	37.669	13.075	170.951	86.329
Nom5	61.788	14.374	206.625	96.323

Table 3. First order Statistical features for ROI's Images

Images	Mean	StdDev	Variance	Energy
ROI1	180.549	19.865	394.635	99.927
ROI2	96.427	83.741	7012.61	99.86
ROI3	140.25	103.44	10700.2	99.981
ROI4	161.403	27.815	773.657	99.864
ROI5	152.868	52.904	2798.827	92.209

Figure 6 shows the graph for analysis of Normal, Cancerous and ROI's images for the first order Statistical features. Figure 6, it shows the features of ROI images have higher value as compared to normal images.

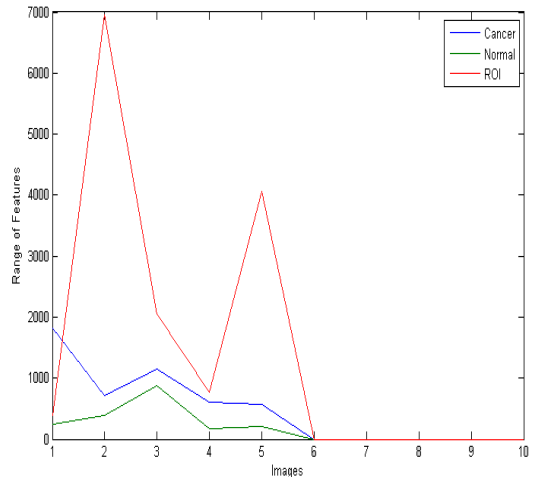


Figure.6 Graph for Comparison of Statistical features of Cancer, Normal & ROI images

3.2 Classification

Probabilistic Neural Network (PNN) is used here as a classifier. The PNN is trained using seventy images in round robin fashion. The seventy images consist of thirty five cancerous and the remaining normal images.

3.3 Receiver Operating Characteristics (ROC) Analysis

To evaluate the PNN, the ROC analysis is carried out. A receiver operating characteristics (ROC) graph is a technique for visualizing, organizing and selecting classifiers based on their performance. ROC analysis has been extended for use in visualizing and analyzing the behavior of diagnostics systems.

The following graph shows the ROC Analysis. High performance for ROC analysis is obtained by the proposed combined scheme $Az = 0.940$ as compared with individual methods.

The performance analysis of Wavelet method has an $Az = 0.90$ and The performance analysis of Gray-level method has an $Az = 0.85$.

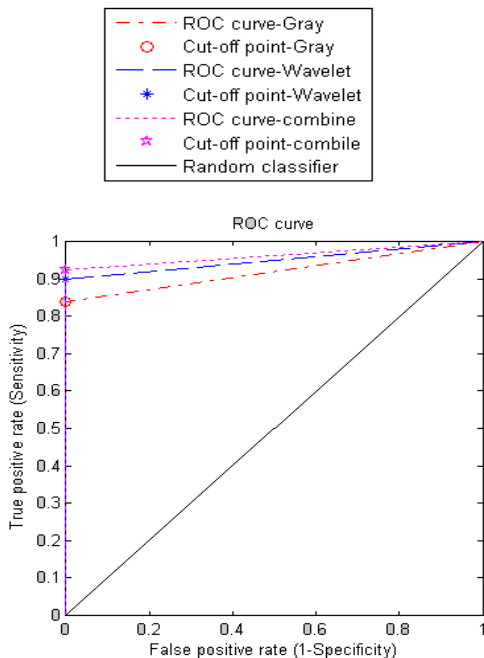


Figure.7 ROC Analysis

4. CONCLUSION

Here a feature-based approach for analysis and Classification of mammographic images is presented. The ROI of the original images have been selected. The texture features are extracted

using Wavelet and Gray-level methods. The features are classified using PNN. The performance evaluation of classification using PNN is done using ROC analysis. The high performance is obtained by proposed combined scheme (Wavelet plus Gray-level) as compared with individual one (Wavelet and Gray-level)

5. REFERENCES

- [1] Cristina Juarez L., Volodymyr Ponomaryov, J. Luis Sanchez R "Detection of Microcalcifications in Digital Mammograms Images Using Wavelet Transform" CERMA, vol. 2, pp.58-61, Electronics, Robotics and Automotive Mechanics Conference (CERMA'06), 2006.
- [2] L. Shen, R. M. Rangayyan, and J. E. L. Desautels, "Application of shape analysis to mammographic calcifications," IEEE Trans. Med. Imag. vol. 13, no. 2, pp. 263–274, Jun. 1994.
- [3] J. A. Baker, P. J. Kornguth, J. Y. Lo, and C. E. Floyd, "Artificial neural network: Improving the quality of breast biopsy recommendations," Radiology, vol.198, no.1, Jan. 1996.
- [4] S. K. Lee, P. Chung, C. I. Chang, C. S. Lo, T. Lee, G. C. Hsu, and C.W Yang, "Classification of clustered microcalcifications using a shape cognitron neural network", Neural Network, vol 16, no 1, pp121-132, Jan 2003.
- [5] H. P. Chan, B. Sahiner, K. L. Lam, N. Petrick, M. A. Helvie, M. M. Goodsitt, and D. D. Adler "Computerized analysis of mammographic microcalcifications in morphological and texture feature spaces", Medical Physics Journal, vol 25, issue 10, 2007
- [6] Haralick, K.Shanmugam, and I.Dinstein," Texture features for Image Classification", IEEE Transaction, Syst., Man., Cybern. , vol SMC-3, no 6, pp 610-621, Nov 1973.
- [7] D. F. Specht, "Probabilistic neural networks," Neural Network, vol. 3, no. 1, pp. 109–118, 1990.