© 2019 IJSRSET | Volume 6 | Issue 3 | Print ISSN: 2395-1990 | Online ISSN : 2394-4099
Themed Section : Engineering and Technology

Investigation on Micro-Calcifications for Breast Cancer Via DWT and BPNN

Dr. J Madhavan *[1], Dr. Bhaludra R Nadh Singh* [2], Dr. Bremiga Gopalan*[3]
*[1]Professor, Department of ECE, Bhoj Reddy Engineering College for Women, Vinay Nagar, Hyderabad,Telangana, India
*[2]Professor, Department of CSE, Sphoorthy Engineering College, Hyderabad,Telangana, India
*[3]Assistant Professor, Department of ECE, Bhoj Reddy Engineering College for Women, Vinay Nagar, Hyderabad, Telangana, India

Abstract - A high-sensitivity computer-aided diagnosis algorithm which can detect and quantify micro-calcifications for early-stage breast cancer. The algorithm can be divided into two phases: image reconstruction and recognition on micro-calcification regions. For Phase I, the suspicious micro-calcification regions are separated from the normal tissues by wavelet layers and Renyi's information theory. The Morphology-Dilation and Majority Voting Rule are employed to reconstruct the scattered regions of suspicious micro-calcification. For Phase II, total 31 descriptors which mainly includes shape inertia, compactness, eccentricity and grey-level co-occurrence matrix are introduced to define the characteristics of the suspicious micro-calcification clusters. In order to reduce the computation load, principal component analysis is used to transform these descriptors to a compact but efficient expression by linear combination method. The efficacy of back-propagation neural network classifier exhibits its superiority in terms of high true positive rate (TP rate) and low false positive (FP rate) rate, in comparison to other classifier.

Keywords: Breast Cancer, Wavelet Transform, Principal Components Analysis, Neural Network

Introduction

Breast cancer is currently one of the leading causes of death among women worldwide. Breast cancer is a cancer that starts in the cells of the breast in women and men. Worldwide, breast cancer is the second most common type of cancer after lung cancer and the fifth most common cause of cancer death. Thus, early detection is the key for improving breast cancer prognosis. However, current methods of treatment are very effective against breast cancer in its early phase. Therefore, removal of the cancer while it is still in its early stages is the most promising way to achieve a significant change in the current breast cancer situation. Mammography is the most common procedure for detecting non-palpable cancers. Micro-calcification is an important feature to diagnose early-breast cancer. However, most medical doctors apply different standardsand personal experience to examine X-ray images of breast cancer so that a certain degree of overlook or incorrect judgment is always emerged. Computer-aided diagnosis system provides an auxiliary but scientific method, which is capable to process large amount of X-ray images, to doctors for valuable comparison and double-check. Though numerous computer-aided diagnosis systems have been reported, yet how to upgrade the sensitivity and reduce computation burden is still considerably required. In order to enhance the pixels of micro-calcification image, I.applied wavelet's coefficients reconstruction method to suppress the background noise [1]. I used descriptors of spatial domain and frequency domain as the input of neural network classifier

IJSRSET2293197 | Received : November 13, 2018 | Revision: 25 February 2019 | Accepted May 10, 2019 | Published May 16, 2019 | May-June-2019 [6 (3) : 470-478]

[2].for select 15 features such as area, compactness, boundary and etc. to extract potential individual micro-calcification objects [3].

In this paper, a high sensitivity computer aided system is developed. It mainly consists of three key techniques including wavelet transform, principal component analysis and back-propagation neural network classifier [4]. Wavelet transform is employed to find the suspicious micro-calcification clusters of each image by eliminating background noise. However, the threshold value for each image to distinguish foreground and background pixels is almost case-by-case so that Renyi's information theory is applied for separation of suspicious microcalcification clusters from the normal tissue pixels. In our work, the micro-calcification regions are defined by 31 descriptors. By principal component analysis, the X-ray images to be inspected can be converted into 9 principal components so that not only the computation burden can be greatly reduced but also the complexity of diagnosis algorithm can be much simplified. Finally, back-propagation neural network classifier is utilized to identify exact microcalcification image.
The proposed diagnosis system exhibits superior performances in terms of high true positive rate and low false positive rate, by applying to different database images.

II. RECONSTRUCTION OF SUSPICOUS MICRO-CALCIFICATION REGIONS

The flowchart of Phase I, named as reconstruction of the suspicious micro calcification's regions, is shown as Fig. 1.

2.1 Wavelet platform

In order to enhance the pixel of microcalcification image, wavelet transform is employed to construct wavelet hyper-plane by taking the advantage of time-frequency analysis and multi resolution analysis. At first, the digital images are transformed into wavelet platform Wk by wavelet layers [5]:

$$Wk=\alpha Hk+\beta Vk+\gamma Dk \quad (1)$$

where Hk , Vk and Dk are the sub-band wavelet coefficients of horizontal details, vertical details and diagonal details respectively, corresponding to the k wavelet platform. α, β and γ are the associated weights respectively. After preliminary trials by signal processing and comparisons, the wavelet function DAUB 4 from Daubechies family, has the best sensitivity to outstand micro-calcification [6]. As a matter of fact, the micro calcification is a high frequency component in image. The diagonal details are much more sensitive to high frequency content of image. Therefore, the weight of diagonal details is set as the largest in our work, compared with the ones for horizontal details and vertical details.

2.2 Threshold value

By image processing, whether the individual pixel belongs to background or foreground can be determined by threshold value. Therefore, it is necessary to construct the optimal threshold value of wavelet coefficient for each sheet of X-ray image. In this research, Renyi's information theory is applied to define the optimal threshold value. Suppose wavelet analysis has transformed the interested digital image into wavelet coefficientsby sequence from 0 to K . By Renyi's information theory, the probability density function of normalized histogram p(k) is nothing but the combination of the probability function of foreground $\varphi(k)$ and the probability function of background ϕ (k)

$$p(k)=\varphi(k)\phi(k) \ . \quad (2)$$

In general, φ (k) and ϕ (k) are mostly corresponding to normal tissues and micro-calcification tissues respectively. By Renyi's information [7], the optimal threshold value of wavelet coefficient is determined by maximizing the distance function between φ (k)and ϕ (k). The details can be found in [7].

2.3 Reconstruction of Suspicious Micro-calcification Regions

First of all, the suspicious microcalcification pixels of wavelet platform in the specific analysis layers are dilated to the entire domain by the approach of Morphology-dilation. However, the probability for partial normal tissue mistaken as the micro-calcification is also raised due to the augmentation of suspicious micro- calcification pixels. Hence Majority Voting Rule is employed to define the boundary of suspicious micro- calcification to separate from the normal tissue pixels. In general, the third layer or fourth layer of the Wavelet transform is adopted for further analysis later on because a certain level of true positive rate has to be preserved.That is, the first and second layers of Wavelet transform are disregarded.

In order to ensure high sensitivity, a few of key image processing parameters have to be selected so that the accuracy of the computer-aided diagnosis can be improved. For Phase I, they are Wavelet Coefficient Resolution (WCR), disk shaped structuring element of pixels radius for Morphology dilation, RSE and the weight of diagonal details, γ. Without loss of generality, suppose the weight of horizontal details, α, and the weight of vertical details, β, are both equal to 1. Next step is to determine the wavelet coefficient resolution. From Fig. 2, it is evident to find that the highest true positive rate can be obtained if the wavelet coefficient resolution is set as 2048. In fact, at the same time the lowest false positive rate can be achieved, in comparison to WCR equal to either 256,1048 or 4096

As for the disk shaped structuring element of pixels radius, RSE, it is selected as 12 via the computer analysis of the effect on TP rate and FP rate, shown in Fig. 3. On the other hand, though higher true positive rate can be achieved by setting higher value for weight of diagonal details, γ, unfortunately the false positive rate is increased as well. By trade-off, the value of γ is selected as 4.2 in our work. By verification from the clinical diagnosis, the computer-aided image-processing diagnosis system for Phase I exhibits satisfactory performance in terms of true positive rate 97.19% but poorly for false positive rate 18.32%. This has to be further compensated by Phase II.

III. RECOGNITION OF MICRO-CALCIFIATION REGIONS

During Phase I, the suspicious micro calcification regions are reconstructed. But a number of normal tissue pixels still remain in the ROI (region of interest). As mentioned at the tail of last section, the false positive rate, 18.32%, is still too high to be acceptable. In Phase II, Neural Network classifier is employed to further eliminatethe normal tissue pixels. The flowchart of Phase II is shown in Fig. 4.

3.1 Detection of Micro-Calcification Objects

In digital image analysis, a few descriptors are definitely necessary to define the feature of micro-calcification regions ([8], [9]). A set of 31 descriptors, including shape inertia I, III and IV, compactness, eccentricity and grey-level co-occurrence matrix, is introduced to define micro-calcification in our work.

3.2 Principal Component Analysis

Principal Component Analysis (PCA) is aimed to describe micro-calcification regions by just a few descriptors or variables that still preserve the major features of micro-calcification regions so that computation load can be greatly reduced and the efficiency of image processing can be improved. The flowchart of PCA is shown in Fig. 5. First of all, the parameters and weights of a training set have to be constructed by the X-ray images of the ill patients with micro-calcification already-verified by the hospitals. PCA transforms the original descriptors into a few independent principal components. In fact, each principal component is a linear combination of N original descriptors as follows:

Suppose we have a random vector population x, where

$$\mathbf{x} = (x_1, \ldots, x_n)^T \qquad (3)$$

and the mean of that population is denoted by

$$\mu_{\mathbf{x}} = E\{\mathbf{x}\}$$

and the covariance matrix of the same data set is

$$\mathbf{C}_x = E\{(\mathbf{x} - \mu_{\mathbf{x}})(\mathbf{x} - \mu_{\mathbf{x}})^T\}$$

In order to reduce this computation load and unnecessary numerical errors, merely the first M principal components are employed for PCA. In our work, the contribution index[10] of first 9 principal components, out of 31 descriptors, has been achieved up to 99.98% so that M is equal to 9 in this paper.

The transpose matrix A is formed from eigenvectors with respect to the first M principal components as follows:

$$A = [e_1 \; e_2 \cdots e_M]^T$$

The principal components would be used as the inputs of the Neural Network classifier in next section.

3.3 Back-propagation Neural Network Classifier

The back-propagation neural network is used as the mathematical model by multi-layers perception mechanism which connects the artificial neurons in layer-by-layer sense, shown in Fig. 6. I1,I2 , I N and Y1, Y2 Yl are the inputs and outputs respectively. N and l are the numbers of node of the input layer and the output layer respectively. By minimizing the mean square error between input and expected value through back propagation learning, the network parameters for the network such as weight and bias can be obtained. In this work, two hidden layers, which have 33 nodes and 6 nodes respectively, are employed. The numbers of input layer nodes and output layer nodes are 9 and 1 respectively. If the output Y > 0, the pattern is classified as the micro-calcification cluster. The recognition flowchart of the back-propagation neural network is shown in Fig. 7.

3.4 Results and Discussion

Totally 26 ROI images are clustered into 716 regions, which includes 121 micro-calcification regions, by Phase i. Two ROI images for two patients are shown in Stage "A" in Table 1. If the original descriptor, instead of applying PCA, is directly used as the input of the classifiers, then back propagation neural network can always preserve the micro-calcification clusters. This can be observed in Stage "B" in Table 1. This is also verified by the true positive rate and false positive rate, shown in Table 2. As a result, if PCA and BPNN are both applied, the diagnosis performance can be much enhanced. The associated true positive rate and false positive rate are 97.12% and 7.89% respectively.

IV. CONCLUSION

An innovative diagnosis algorithm to detect and quantify micro-calcification is presented and verified by DDSM database where all the results available. There are two phases for the entire diagnosis by imageprocessing. Phase I is called reconstruction of the suspicious micro-calcification in this research. By applying Wavelet transform and Renyi's information theory, the suspicious micro-calcification regions in the X-ray image can be efficiently reconstructed by elimination of most normal tissue pixels and noise. PhaseII is called recognition of the micro-calcification. By applying Principle Components Analysis (PCA), merely 9 principle components, out of 31 descriptors, are constructed to define the feature of the micro-calcification tissues. In order to quantify and recognize micro-calcification pixels, the Back-Propagation Neural Network (BPNN) classifier is developed. The BPNN classifier consists of an input layer whose inputs are nothing but the selected 9 principle components describing the geometry of each suspicious micro-calcification cluster, two hidden layers, with 33 nodes and 6 nodes orderly and an output layer whose output is the binary data: 1 for micro- calcification and 0 for normal tissue cluster. As for recognition capability, the back-propagation neuralnetwork classifier exhibits superior performance, true positive rate is greatly raised up to 97.12% and TP/FP Rate is greater than 12.31% if PCA and BPNN are both employed. That is, the diagnosis system proposed in our work exhibits high sensitivity.

V.FIGURES AND TABLES

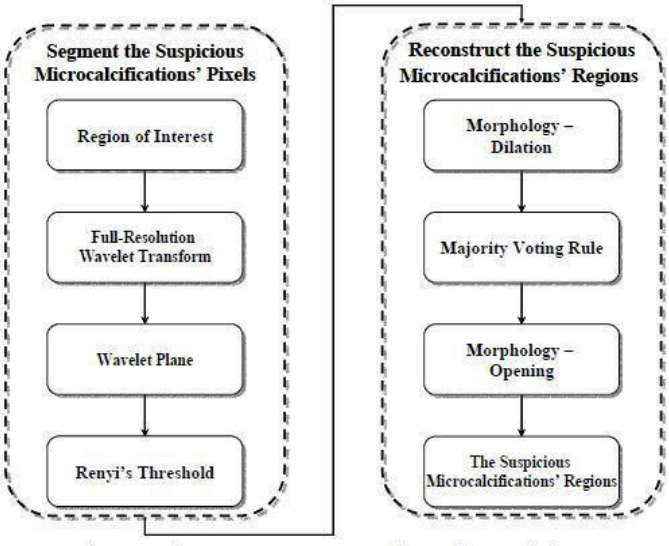


Fig. 1 Phase I: Reconstruction of Suspecious Micro-calcification Regions

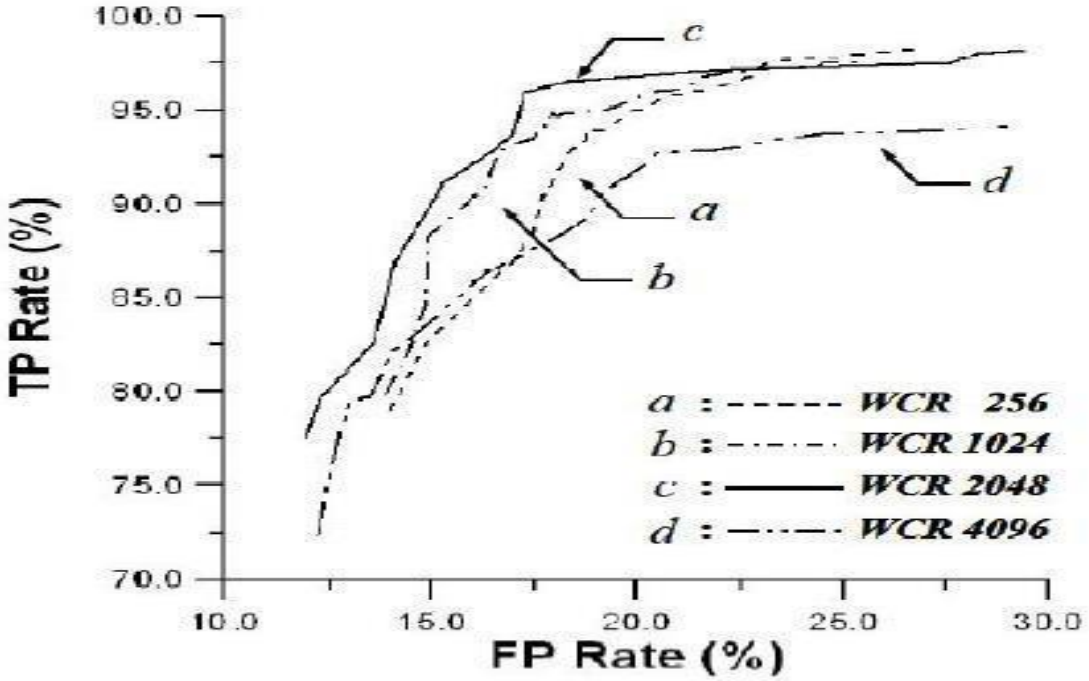


Fig. 2 ROC under Choice of Wavelet Coefficients Radius

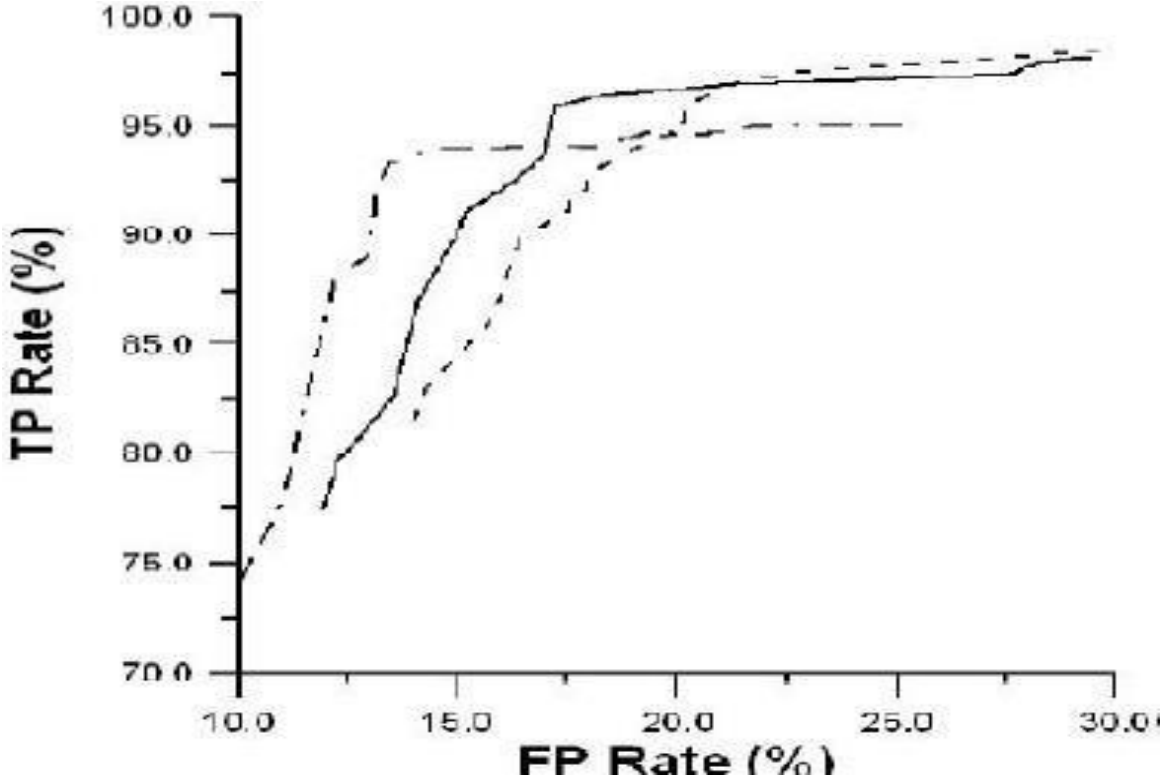


Fig. 3 ROC under Choice of Disk Shaped Structuring Element Pixels Radius

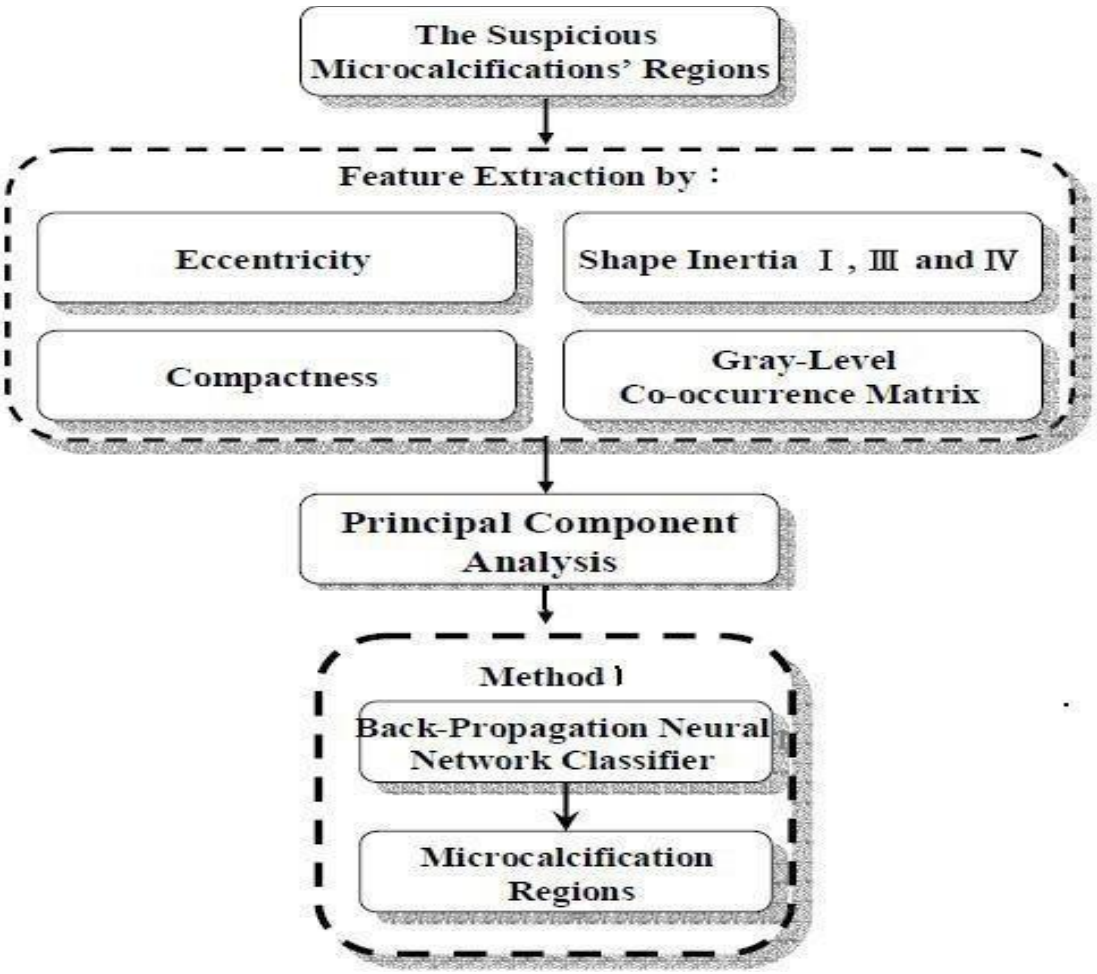


Fig. 4 Phase II: Recognition of Micro-calcification Regions

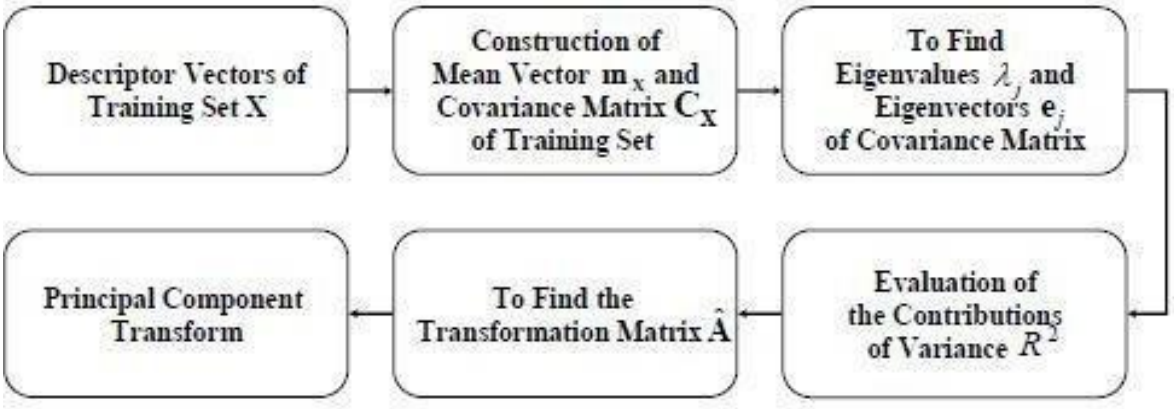


Fig. 5 Flowchart of Principal Component Analysis

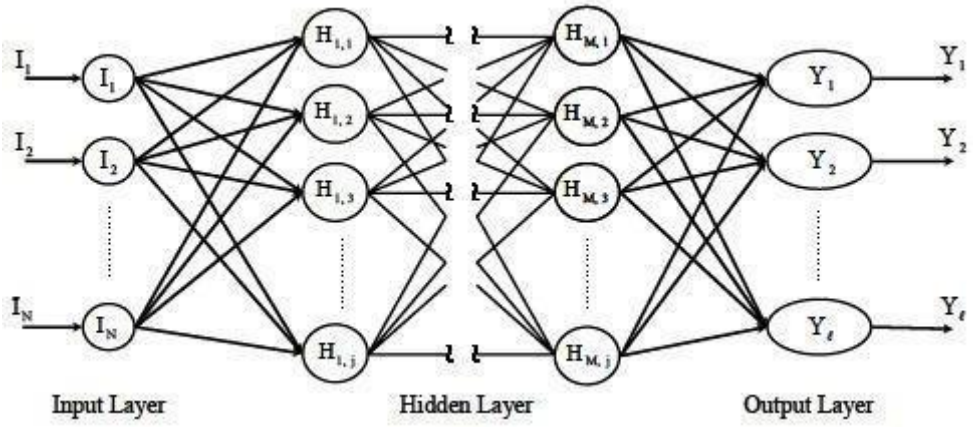


Fig. 6 Back-propagation Neural Network Classifier

Table 1 Stages of Diagnosis on Microcalcification Images

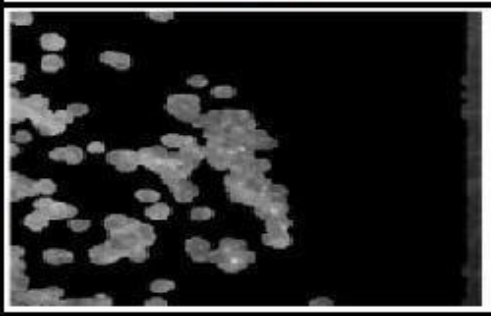
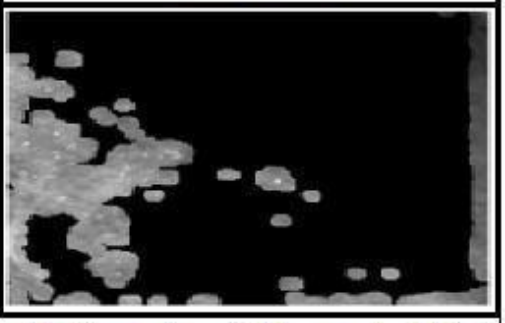
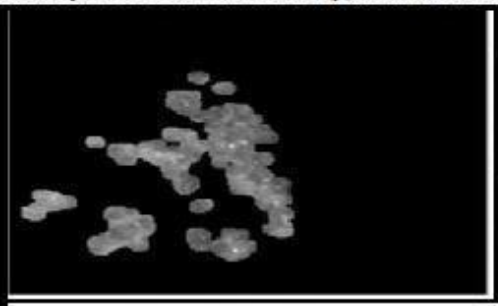
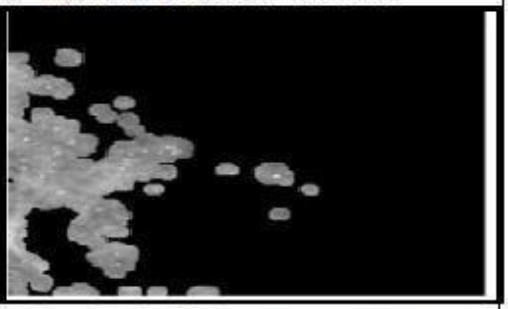
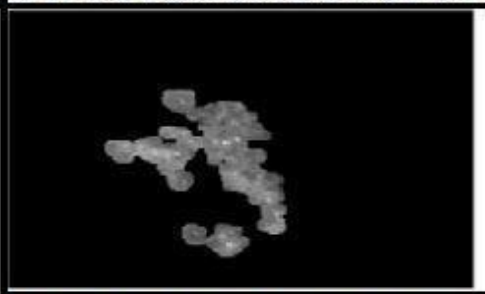
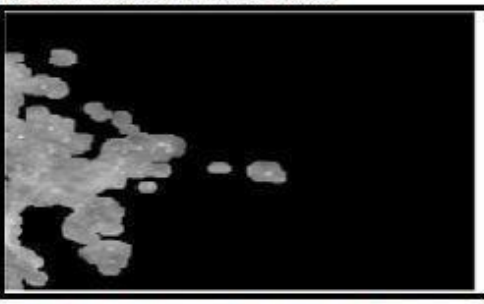
	Case #1	Case # 2
A		
	Wavelet Coefficients Resolution is 2^{11} and Disk Shaped Structuring Element Pixels Radius is 12.	
B		
	The original descriptor vector to be the Input of Back-propagation Neural Network Classifier.	
C		
	Principal Component Spatial Vector to be the Input of Back-propagation Neural Network Classifier.	

Table 2 Performance by Back-propagation Neural Network Classifier

Number of Principal Components	TP Rate (%)	FP Rate (%)	TP Rate/ FP Rate
5	95.46	8.32	11.47
7	96.71	8.40	11.51
9	97.12	7.89	12.31
15	96.60	8.45	11.43

References

1. 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, 2008.
2. B. Zheng, W. Qian and L. P Clarke, "Digital mammography:mixed feature neural network with spectral entropy decision for detection of microcalcifications," IEEE Trans. Med. Imag., vol.15,no. 5, pp. 589-597, 1996.
3. S. Yu and L. Guan, "CAD System for the Automatic Detection of Clustered Mictocalcifications in Digitized Mammogram Films,"IEEE Trans. Med. Imag., vol. 19, no. 2, pp. 115-126, 2013.
4. H. W. Chen, "Image Processing of Micro-calications for Earlystage Breast Cancer via Wavelet Analysis and Neural Network,"M.S. Thesis, Dept. Mechanical Eng., National Cheng-Kung Univ.,Tainan, Taiwan, R.O.C., 2008.
5. A. Hojjatoleslami., L. Sardo and J. Kittler, "An RBF Based Classifier for the Detection of Microcalcification in Mammograms with Outlier Rejection Capability," IEEE Trans. Med. Imag., vol.21, no. 12, pp. 1379- 1384, 1997.
6. T. C. Wang. and N. B Karayianmis, "Detection of Microcalcification in Digital Mammograms Using Wavelets,"IEEE Trans. Med. Image., vol. 17, no. 4, pp. 498-509, 1998.
7. B. Giuseppe, C. Angelo and P. Antonio, "Computer Aided Detection of Microcalcifications in Digital Mammograms," Computers in Biology and Medicine, vol. 30, pp. 267-286, 2015.
8. S. Yu and S. Guan, "A CAD System for the Automatic Detection of Clustered Microcalcifications in Digitized Mammogram Films," IEEE Trans. Med. Image., vol: 19, no: 2, pp: 115-126, 2015.
9. R. M. Haralick, K. Shanmugam and I. Dinstein, "Textural Features for Image Classification," IEEE Trans. on Systems, Man, and Cybernetics, vol. 3, no. 6, pp. 610-621,1973.