

# Multi-Level Wavelet Transformation Based 3d-Breast Mammographic Image Classification

M. Thirumalesh<sup>1</sup>, B. L. Prakash<sup>2</sup>.

<sup>1</sup>Department of ECE, Vignan's Institute of Information Technology, Vizag, A.P, India.

<sup>2</sup>Professor in Department of Electronics and Communication Engineering, Vignan's Institute of Information Technology, Vizag, A.P, India.

## ABSTRACT

*This paper proposes an automatic support system for stage classification of cancer region using probabilistic neural network based on the threshold detection method for medical applications. The detection of the breast cancer is a challenging problem, due to the structure of the cancer cells. This paper presents a segmentation method and wavelet based threshold method for segmenting mammographic images to detect the breast cancer in its early stages. The artificial neural network will be used to classify the status of image that is abnormal or normal. The manual analysis of these samples is time consuming, inaccurate and requires intensively trained person. It will be used as a base for a Computer Aided Diagnosis system for early detection of cancer from mammographic images which will improve the chances of survival for the patient. The experimental result shows that the threshold based segmentation results are more accurate and reliable than other methods. Discrete wavelet transform technique is used for extracting texture features and it decomposed the image into four levels for getting the edge details in horizontal and vertical direction. The Co-occurrence matrix will be determined for these two high frequency subbands for finding the texture features. Probabilistic Neural Network with image and data processing techniques is employed to implement an automated breast cancer classification. Decision making is performed in two stages: feature extraction using Wavelet transformation followed by GLCM and the classification using Probabilistic Neural Network (PNN). The performance of the PNN classifier is evaluated in terms of training performance and classification accuracies. Probabilistic Neural Network gives fast and accurate classification than other neural networks and it is a promising tool for classification of the Tumors.*

**KEYWORDS:** 3-D breast ultrasound, computer-aided diagnosis (CAD), Threshold Estimation Discrete Wavelet Decomposition, PNN, GLCM (Grey level co-occurrence matrix).

## 1. INTRODUCTION

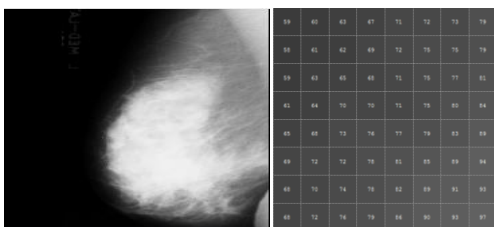
Automated detection and classification of cancers in different medical images is motivated by the necessity of high accuracy when dealing with a human life. Also, the computer assistance is needed in medical institutions due to the fact that it could improve the results of humans in such a domain where the false cases must be at a very low rate. It has been proven that double reading of medical images could lead to better cancer detection. But the cost implied in double reading is very high, that's why good software to assist humans in medical institutions is of great interest now-a-days [1]. Conventional methods of monitoring and diagnosing the diseases rely on detecting the presence of particular features by a human observer. Due to large number of patients in intensive care units and the need for continuous observation of such conditions several techniques for automated diagnostic systems have been developed in recent years to attempt to solve this problem. Such techniques will work by transforming the mostly qualitative diagnostic criteria into a more objective quantitative feature detection problem. The automated detection of breast magnetic resonance images by using some prior knowledge like pixel intensity and some anatomical features is proposed. Currently there are no widely accepted methods therefore automatic and reliable methods for cancer detection are of great need and interest. The applications of PNN in the classification of data for mammogram images are prototyped here for solving the problem of inaccurate detection and classification of breast cancer from the mammographic images.

## 2. BREAST CANCER

Breast cancer is a type of cancer originating from breast tissue, most commonly from the inner lining of milk ducts or the lobules that supply the ducts with milk. Cancers originating from ducts are known as ductal carcinomas, while those originating from lobules are known as lobular carcinomas. Breast cancer occurs in humans and other mammals [2]. While the overwhelming majority of human cases occur in women, male breast cancer can also occur. Worldwide, breast cancer accounts for 22.9% of all cancers (excluding non-melanoma skin cancers) in women. In 2008, breast cancer caused 458,503 deaths worldwide (13.7% of cancer deaths in women). Breast cancer is more than 100 times more common in women than in men, although men tend to have poorer outcomes due to delays in diagnosis. Prognosis and survival rates for breast cancer vary greatly depending on the cancer type, stage, treatment, and geographical location of the patient. Survival rates in the Western world are high; for example, more than 8 out of 10 women (84%) in England diagnosed with breast cancer survive for at least 5 years. In developing countries, however, survival rates are very poor[3].

In radiology, computer-aided detection also called computer-aided diagnosis is procedure in medicine that assist doctors in the interpretation of medical images.

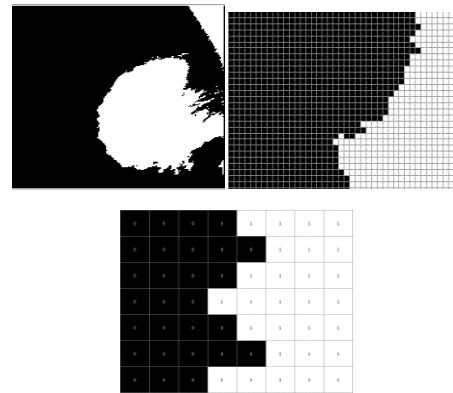
CAD is a relatively young interdisciplinary technology combining elements of artificial intelligence and digital image processing with radiological image processing.



**Figure2.1:** Image of breast cancer & its gray level intensity

A typical application is the detection of a tumor. For instance, some hospitals use CAD to support preventive medical check-ups

in mammography (diagnosis of breast cancer), the detection of polyps in the colon, and lung cancer.

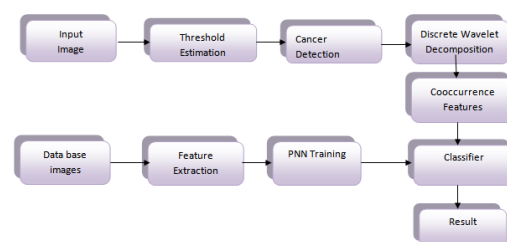


**Figure 2.2:** Segmentation of Breast cancer

Imaging techniques in X-ray, MRI, and Ultrasound diagnostics yield a great deal of information, which the radiologist has to analyze and evaluate comprehensively in a short time. CAD systems help scan digital images, *e.g.* from computed tomography, for typical appearances and to highlight conspicuous sections, such as possible diseases [4].

## 3. ALGORITHM DESIGN

The automated disease identification system is not a single process. This system consists of various modules the success rate of each and every step is highly important to ensure the overall high accurate outputs. the rest of the work is organized as follows.



**Figure3.1:** Block diagram of multi-level image classification

## 4. WAVELET TRANSFORMS ANALYSIS

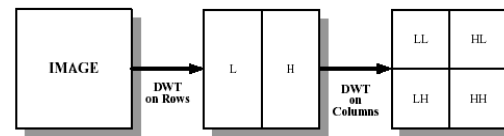
Basically we use Wavelet Transform to analyze non-stationary signals, i.e., signals whose frequency response varies in time, as Fourier Transform is not suitable for such signals. To overcome the limitation of Fourier Transform, Short Time Fourier Transform (STFT) was proposed. There is only a minor difference between STFT and FT. In STFT, the signal is divided into small segments, where these segments (portions) of the signal can be assumed to be stationary. For this purpose, a window function "w" is chosen. The width of this window in time must be equal to the segment of the signal where it is still being considered stationary. By STFT, one can get time-frequency response of a signal simultaneously, which can't be obtained by FT[5]. The short time Fourier transform for a real continuous signal is defined as:

$$X(f, t) = \int_{-\infty}^{\infty} [x(t)w(t - \tau)] e^{-2\pi j f \tau} dt$$

Where the length of the window is (t-τ) in time such that we can shift the window by changing value of t and by varying the value τ we get different frequency response of the signal segments.

### 4.1. 2-D wavelet transforms

The 1-D DWT can be extended to 2-D transform using separable wavelet filters. With separable filters, applying a 1-D transform to all the rows of the input and then repeating on all of the columns can compute the 2-D transform. As depicted in Figure, the four sets are LL, HL, LH, and HH, where the first letter corresponds to applying either a low pass or high pass filter to the rows, and the second letter refers to the filter applied to the columns [6].



**Figure4.1:** Block Diagram of DWT(a) Original Image (b) Output image after the 1-D applied on Row input (c) Output image after the second 1-D applied on row input

The Two-Dimensional DWT (2D-DWT) converts images from spatial domain to frequency domain. At each level of the wavelet decomposition, each column of an image is first transformed using a 1D vertical analysis filter-bank. The same filter-bank is then applied horizontally to each row of the filtered and sub sampled data.

One-level of wavelet decomposition produces four filtered and sub sampled images, referred to as sub bands. The upper and lower areas of Fig., respectively, represent the low pass and high pass coefficients after vertical 1D-DWT and sub sampling. The result of the horizontal 1D-DWT and sub sampling to form a 2D-DWT output image is shown in Fig.

We can use multiple levels of wavelet transforms to concentrate data energy in the lowest sampled bands[4]. Specifically, the LL sub band in fig 2.1(c) can be transformed again to form LL2, HL2, LH2, and HH2 sub bands, producing a two-level wavelet transform. An (R-1) level wavelet decomposition is associated with R resolution levels numbered from 0 to (R-1), with 0 and (R-1) corresponding to the coarsest and finest resolutions.

## 5. GRAY-LEVEL CO-OCCURRENCE MATRIX

At first the co-occurrence matrix is constructed, based on the orientation and distance between image pixels. Then meaningful statistics are extracted from the matrix as the texture representation. Haralick proposed the following texture features: Energy, Contrast, Correlation, Homogeneity,

For example; with an 8 grey-level image representation and a vector  $t$  that considers only one neighbor, we would find

	0	1	2	3	4	5	6	7
0	0	0	0	0	0	0	0	0
1	0	1	2	0	0	0	0	0
2	0	1	0	2	0	0	0	0
3	0	0	1	1	0	0	0	0
4	0	1	0	0	1	0	0	0
5	0	0	0	0	0	0	0	0
6	0	0	0	0	0	0	0	0
7	0	0	0	0	0	0	0	0

Figure5.1: Classical Co-occurrence matrix

**5.1 Energy:** It is a gray-scale image texture measure of homogeneity changing, reflecting the distribution of image gray-scale uniformity of weight and texture [3].

$$E = \sum \sum p(x,y)^2$$

Where  $p(x,y)$  is the GLCM

**5.2 Contrast:** Contrast is the main diagonal near the moment of inertia, which measure the value of the matrix is distributed and images of local changes in number, reflecting the image clarity and texture of shadow depth.

$$\text{Contrast } I = \sum \sum (x-y)^2 p(x, y)$$

**5.3 Entropy:** It measures image texture randomness, when the space co-occurrence matrix for all values is equal, it achieved the minimum value.

$$S = \sum \sum p(x, y) \log p(x, y)$$

**5.4 Correlation Coefficient:** Measures the joint probability occurrence of the specified pixel pairs.

$$C = \sum \sum ((x - \mu_x)(y - \mu_y)p(x, y) / \sigma_x \sigma_y)$$

**5.5 Homogeneity:** Measures the closeness of the distribution of elements in the GLCM to the GLCM diagonal.

$$H = \sum \sum (p(x, y) / (1 + |x-y|))$$

## 6. Probabilistic Neural Networks

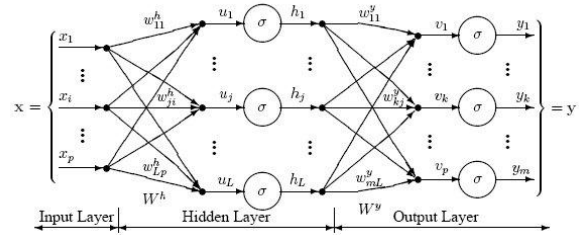


Figure6.1: Algorithm of Probabilistic neural network

This network has an input layer (on the left) with three neurons, one hidden layer (in the middle) with three neurons and an output layer (on the right) with three neurons. There is one neuron in the input layer for each predictor variable. In the case of categorical variables,  $N-1$  neurons are used to represent the  $N$  categories of the variable [4].

**6.1 Input Layer:** - A vector of predictor variable values  $(x_1...x_p)$  is presented to the input layer. The input layer (or processing before the input layer) standardizes these values so that the range of each variable is -1 to 1. The input layer distributes the values to each of the neurons in the hidden layer. In addition to the predictor variables, there is a constant input of 1.0, called the *bias* that is fed to each of the hidden layers; the bias is multiplied by a weight and added to the sum going into the neuron.

**6.2 Hidden Layer:**-Arriving at a neuron in the hidden layer, the value from each input neuron is multiplied by a weight  $(w_{ji})$ , and the resulting weighted values are added together producing a combined value  $u_j$ . The weighted sum  $(u_j)$  is fed into a transfer function,  $\sigma$ , which outputs a value  $h_j$ . The outputs from the hidden layer are distributed to the output layer.

**6.3 Output Layer:**- Arriving at a neuron in the output layer, the value from each hidden layer neuron is multiplied by a weight  $(w_{kj})$ , and the resulting weighted values are added together producing a combined value  $v_j$ . The weighted sum  $(v_j)$  is fed into a transfer function,  $\sigma$ , which outputs a value  $y_k$ . The  $y$  values are the outputs of the network [6].

If a regression analysis is being performed with a continuous target variable, then there is a single neuron in the output layer, and it generates a single  $y$  value. For classification problems with categorical target variables, there are  $N$  neurons in the output layer producing  $N$  values, one for each of the  $N$  categories of the target variable.

### 7. Implementation of GUI:

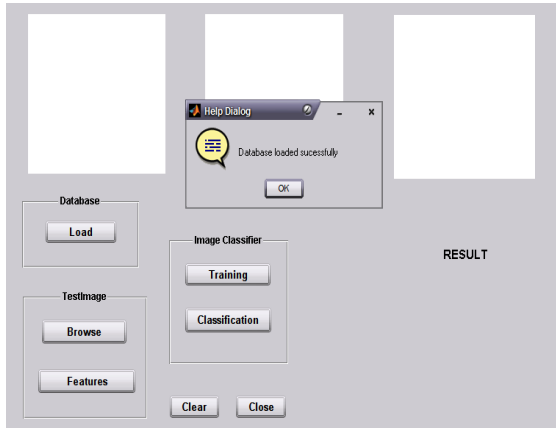


Figure7.1: uploading the data from the database

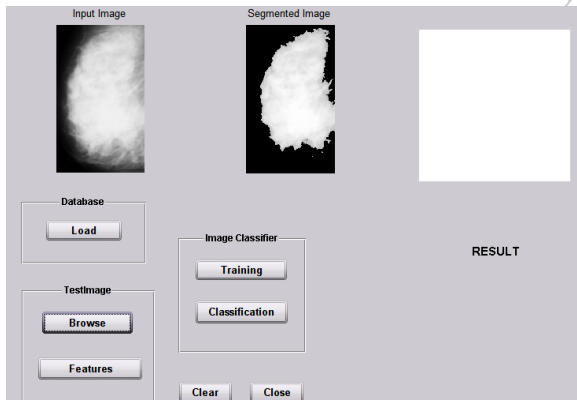


Figure7.3: reading the input image

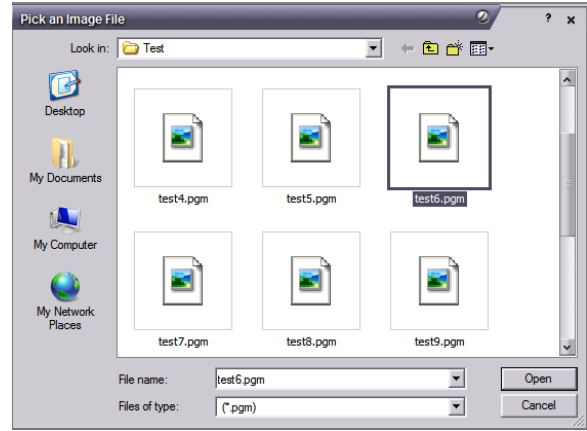


Figure7.2: selecting of input image

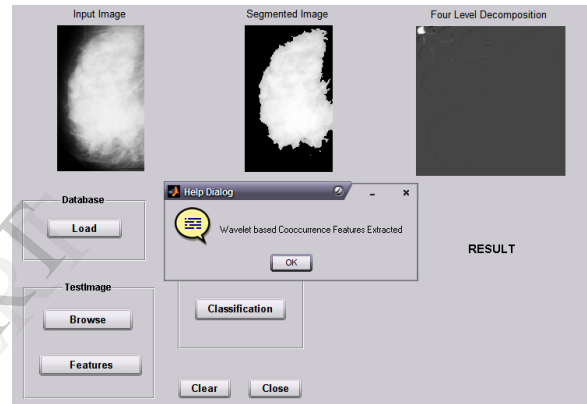


Figure7.4: wavelet based co-occurrence feature extraction

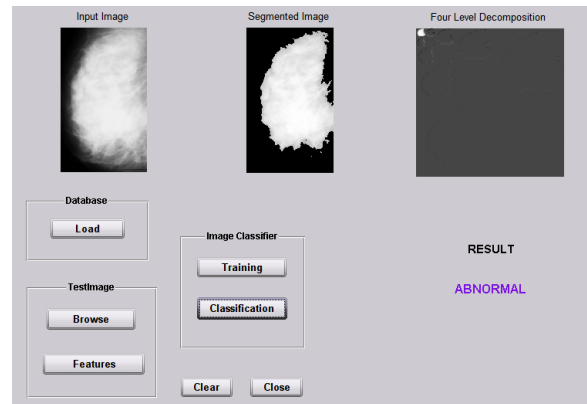


Figure7.5: after extraction of feature result of input image



## 8. CONCLUSION

This project implemented an automatic breast cancer image classification using texture features and it will be classified effectively based on artificial neural network. Here, probabilistic neural network was used for classification based on unsupervised learning using wavelet statistical features and target vectors. The threshold was estimated from smoothing details of images accurately for effective breast cancer segmentation. In addition with, the statistical features are extracted from co-occurrence matrix of detailed coefficients of segmented images. These features are useful to train a neural network for an automatic classification process. Finally this system is very useful to diagnose the diseases from mammographic images for early detection of cancer.

## REFERENCES

- [1]. N. Kwak, and C. H. Choi, "Input Feature Selection for Classification Problems", IEEE Transactions on Neural Networks, 13(1), 143-159, 2002.
- [2]. E. D. Ubeyli and I. Guler, "Feature Extraction from Doppler Ultrasound Signals for Automated Diagnostic Systems", Computers in Biology and Medicine, 35(9), 735-764, 2005.
- [3]. D.F. Specht, "Probabilistic Neural Networks for Classification, mapping, or associative memory", Proceedings of IEEE International Conference on Neural Networks, Vol.1, IEEE Press, New York, pp. 525-532, June 1988.
- [4]. D.F. Specht, "Probabilistic Neural Networks" Neural Networks, vol. 3, No.1, pp. 109-118, 1990.
- [5] Kornel, P., Bela, M., Rainer, S., Zalan, D., Zsolt, T. and Janos, F., "Application of neural network in medicine", *Diag. Med. Tech.*, vol. 4, issue 3, pp: 538-54, 1998.
- [6] Orr M.J.L., Hallam J., Murray A., and Leonard .T, "Assessing rbf networks using delve," *International Journal of Neural Systems*, vol. 10, issue 5, pp. 397-415, 2000.
- [7] Yu-Len Huang, Jeon-Hor Chen, Wu-Chung Shen, "Diagnosis of hepatic tumours with texture analysis in non enhanced computed tomography images", *Academic Radiology*, Vol. 13, pp 713-720, 2006.

## AUTHOR'S INFORMATION



Mr. B.L. Prakash obtained his B.Tech. degree from Nagarjuna University, in the year 1988. He obtained his M.E. degree from Andhra University, Visakhapatnam in the year 2003. He became a Life member of I.S.T.E. in the year 1989, Life member of Institute of Engineers in the year 2001 and Life Fellow of I.E.T.E. in the year 2007. Presently working as a Professor in the department of ECE, VIIT, Visakhapatnam. He owns 11 papers in various National and International conferences and journals. He is the Vice-Chairman of I.E.T.E. Visakhapatnam center.



Mr. M. Thirumalesh obtained his B.Tech (Electronics and Communication Engineering) Degree from Bhimavaram Institute of Engineering & Technology, affiliated to JNTU Kakinada in 2011. Presently he is pursuing M.Tech (Digital Electronics & Communications Systems) at Vignans Institute of Information and Technology, Duvvada, affiliated to JNTU Kakinada.