Neural Systems Approach for Ammography Finding by Utilizing Wavelet features

A. Mallareddy[1], A. Priyanka[2]

[1]Research Scholar (JNTUH), Department of Computer Science & Engineering,
Professor & HOD(CSE) Sri Indu Institute of Engineering & Technology,
Sheriguda(M), Ibrahimpatnam(M), RR Dt. Hyderabad – 501510, India

[2]M.Tech Scholar (CSE), Department of Computer Science & Engineering,
Sri Indu Institute of Engineering & Technology,
Sheriguda (M), Ibrahimpatnam (M), RR Dt. Hyderabad – 501510, India

Abstract: ***We present an application of artificial neural networks to mammographic images, aimed at improving early detection of sensitivity to breast cancer. The proposed application consists of two main steps: a pre-treatment step whose role is to extract the characteristics of the available mammographic images using the wavelet and co-occurrence (GLCM) matrices approach and a classification step based on an artificial neural network that uses these characteristics as input vectors for its training algorithm. The output of the training phase of this model is a categorization of the pre-treated images into three main groups: normal, benign and malignant. After the training phase, the network can be used in order to label new and unseen images as one of these three types.***

Keywords: Breast Cancer, Neural Networks, Mammographic Images, Wavelet process, GLCM Classification

1. Introduction

Breast cancer is currently one of the leading causes of mortality among women throughout the world. Statistically speaking, the breast cancer represents about 25% of all diagnosed cancers in women, and approximately 20% of all fatal cancers[1]. In Morocco, for example, some studies have shown that approximately one out of every eight women might be affected with breast cancer during her lifetime and half of all affected women might die of this disease.

To reduce these high percentages, medical specialists insist on the necessity for early detection of this disease; and according to most of them mammography remains the only reliable and practical method to achieve this goal. Mammography[4] is a medical examination that uses low doses of X-rays for producing practical images of the internal structure of the breast[2][3]. This is why mammogram-based diagnosis is highly recommended in order to increase the chances of cure for early affected patients. However, one of the main characteristics of mammographic images is their high spatial resolution, which makes the task of detection very hard for radiologists. Hence the necessity for developing new tools that could help specialists in deciding based only on mammographic images, whether a woman has an early stage of breast cancer or not.

The main goal of this paper is the proposition of such a tool in the form of an intelligent system capable of filtering and classifying digital mammographic images. The wavelet process of this system transforms each mammographic image into a space frequency domain of that image. These sub bands are given as inputs to gray level co-occurrence matrix whose components represent the image features useful for discriminating between the three categories. The classification process submits these features as inputs to a multilayer perceptron that learns, using the back propagation training algorithm, to separate them into normal, benign and malignant categories. After this training step, the system is tested on unseen images in order to assess its ability to recognize the category of each new image it receives as input.

2. Pre-processing and Feature Extraction of Raw Images

Image pre-processing is one of the subjects that have been widely studied in the area of medical image classification because the classification results may depend on the quality of the used images. This is why the first step of our system is dedicated to the problems of pre-processing and feature extraction. As an example, fig. 1 shows a filtered image, which consists in a digitized gray scale image.

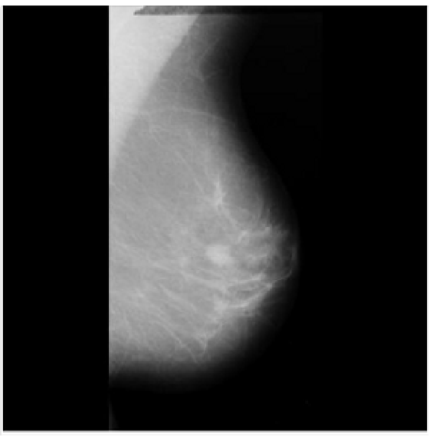

Figure 1: A sample example of filtered mammographic image

The pre-processing problem consists in extracting from each filtered image the best vector of characteristics that can be used as an input for the classification process. For this, we apply a segmentation method that allows the partitioning of each mammographic image into two regions. This method consists in using an intensity threshold, C, as a reference to which each pixel value in the image should be compared[6]. According to this comparison all pixels that have a gray level value greater than are coded by 1 and all others by 0.

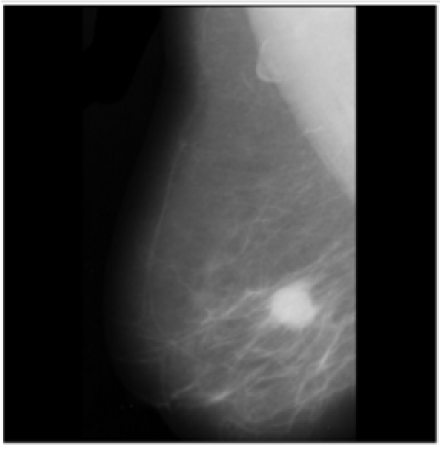

Figure 2 :(a) before segmentation

Figure 2 (b): After segmentation

2.1 Wavelet Process

In the present study, DWT is used to extract the mammogram feature vectors by applying multilevel of decompositions which reduce the number of values used as a classifier input and at the same time keeps the main features of the image details. Different mother functions from Daubechies family are used in the multiresolution analysis. The basic idea of the discrete wavelet transform (DWT) [8] [9] is approximating a signal through a set of basic mathematical functions. The continuous wavelets transform (CWT) of a function f using a wavelet function basis is defined as:

$$f(a,b)=\int f(x)\varphi_{a,b}(x)dx$$

While $\psi(x)$ is the mother wavelet functions. The basis of wavelet function is obtained by scaling and shifting a signal mother wavelet function.

$$\varphi_{a,b}(x)=1/\sqrt{a}\varphi((x-b)/a)$$

Where a is the scale factor and b is the shift value. The mother wavelet should only satisfy the zero average condition (i.e. $\int\psi(x)\,dx = 0$). The DWT is obtained by taking $a = 2$ and$b\in Z$. In the case of 2D signal (i.e. images), the 2D analysis can be performed as a product of two 1D basis functions as shown in the following equation:

$$\varphi_{a1,b1,a2,b2}(x1, x2) = \varphi_{a1,b1}(x1).\varphi_{a2,b2}(x2)$$

This yield a multiresolution decomposition of the signal into four sub bands called the approximation (low frequency component) and details (high frequency component). The approximation **A** is a low resolution of the original image. The details are the coefficients that are neglected during approximation for the horizontal **H**, vertical **V** and diagonal direction **D**. The decomposition process can be iterated with successive approximation being decomposed in turn (multilevel decomposition). While using wavelets decomposition, important information that represents the structure of original data can be captured. Also, wavelets can capture both texture and information efficiently.

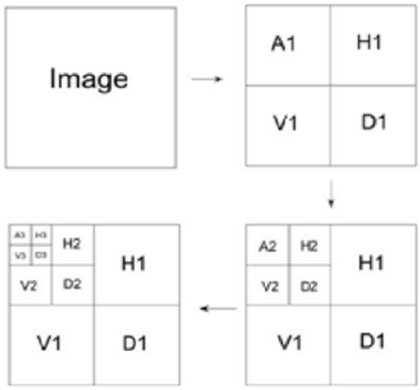

Figure 3: DWT multilevel decomposition of an image

Here we are applying Wavelet process to the given mammography image by third level decomposition. The resultant bands

Figure 4: Multi level wavelet decomposition

2.2 GLCM Process

In statistical texture analysis, texture features are computed from the statistical distribution of observed combinations of intensities at specified positions relative to each other in the image. According to the number of intensity points (pixels) in each combination, statistics are classified into first-order, second-order and higher-order statistics. The Gray Level Coocurrence Matrix (GLCM) method is a way of extracting second order statistical texture features.

The approach has been used in a number of applications, Third and higher order textures consider the relationships among three or more pixels[11]. These are theoretically possible but not commonly implemented due to calculation time and interpretation difficulty.
A GLCM is a matrix where the number of rows and columns is equal to the number of gray levels, G, in the image. The matrix element P (i, j | Δx, Δy) is the relative frequency with which two pixels, separated by a pixel distance (Δx, Δy), occur within a given neighbourhood, one with intensity 'i' and the other with intensity 'j'. The matrix element P (i, j | d, θ) contains the second order statistical probability values for changes between gray levels 'i' and 'j' at a particular displacement distance d and at a particular angle (θ). Using a large number of intensity levels G implies storing a lot of temporary data, i.e. a G × G matrix for each combination of (Δx, Δy) or (d, θ). Due to their large dimensionality, the GLCM's are very sensitive to the size of the texture samples on which they are estimated. Thus, the number of gray levels is often reduced.

2.3 Design Flow

Proposed diagnosis system which consists of three levels of classification is based on two phases; learning phase and testing phase. In the learning phase three ANN's are trained using feature vectors extracted from wavelets coefficients [5][7]. The target is based on expert radiologist diagnosis data. In test phase, the previously trained ANN's are used to diagnosis a query mammogram.

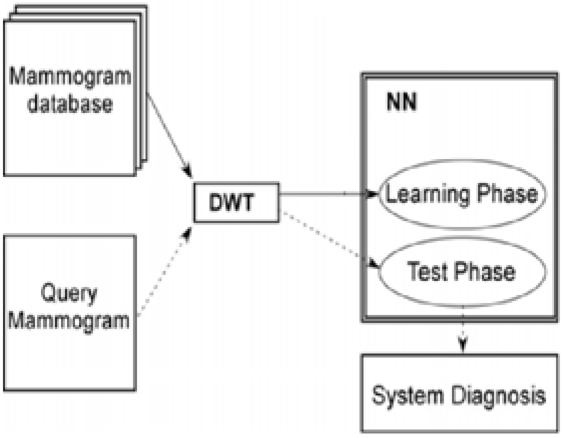

Figure 5: Proposed architecture

3. Conclusion

Using the multilevel decomposition of DWT to extract the features of the digital mammograms is a promising technique to extract features for diagnosis purpose. The proposed diagnosis system achieves good results in classifying the mammograms. Future work could focus on changing the size of features vector and improving the effectiveness the classifier. The presented results in the previous section appear as a positive achievement compared with previous results in the literature. The shortage of this technique is that it needs a supervised classifier with huge database to achieve these high percentage results which is not practically available all the time.

References

[1] P. Pisani, D. M. Parkin and J. Ferlay, "Estimates of the worldwide mortality from eighteen major cancers in 1985: Implications for prevention and projections of future burden," Int. J. Cancer, vol. 55, pp. 891–903,1993.
[2] J. E. Meyer, T. J. Eberlein, P. C. Stomper and M. R. Sonnenfeld, "Biopsy of occult breast lesions: Analysis of 1261 abnormalities," J. Amer. Med. Assoc., vol. 263, pp. 2341–2343, 1990.
[3] E. A. Sickles, "Periodic mammographic follow-up of probably benign lesions: Results in 3,184 consecutive cases," Radiology, vol. 179, pp. 463–468, 1991.
[4] E. A. Sickles, "Breast Imaging: From 1965 to the present," Radiology, vol. 215, pp 1-16, 2000.
[5] Y. Chitre, A. P. Dhawan, and M. Moskowitz, "Artificial neural network based classification of mammographic microcalcifications using image structure features," Int. J. Pattern Recogn. Artif. Intel. vol. 7, no. 12, pp. 1377–1402, 1993.
[6] W. E. Polakowski, D. A. Cournoyer, S. K. Rogers, M. P. DeSimio, D. W. Ruck, J. W. Hoffmeister, and R. A.Raines, "Computer-aided breast cancer detection and diagnosis of masses using difference of Gaussians and derivative-based feature saliency," IEEE Trans. Med. Imag., vol. 16, pp. 811–819, June 1997.
[7] S. K. Kinoshita, P. M. A. Marques, A. F. F. Slates, H. R. C. Marana, R. J. Ferrari, and R. L. Villela, "Detection and characterization of mammographic masses by artificial neural network," in Proc. 4th Int.Workshop Digital Mammography, Nijmegen, the Netherlands, June 1998, pp. 489–490.
[8] E. A. Rashed, I. A. Ismail, and S. I. Zaki, "Multiresolution diagnosis of digital mammograms using artificial neural networks," AJSE, vol. 4, no. 1, pp. 59-65, Aug. 2005.
[9] E. A. Rashed, I. A. Ismail, and S. I. Zaki, "Multiresolution mammogram analysis in multilevel decomposition," Submitted to Pat. Rec. Let., 2005.
[10] C. B. R. Ferreira, and D. L. Borges, "Analysis of mammogram classification using a wavelet transform decomposition," Pat. Rec. Let., vol. 24, pp. 973–982, 2003.
[11] S. Liu, "The Analysis of Digital Mammograms: Spiculated Tumor Detection and Normal Mammogram Characterization," Ph.D. Thesis, School of Electrical and Computer Engineering, Purdue University, May 1999.
[12] H. D. Cheng, X. Cai, X. Chen, L. Hu, and X. Lou, "Computer-aided detection and classification of micro calcifications in mammograms: a survey," Pat. Rec., vol. 36, pp. 2967-2991, 2003.

Authors Profile

Prof. A. MallaReddy is Head of the Department, Computer Science & Engineering, Sri Indu Institute of Engg. & Technology, Sheriguda(Vi),IBP(M),RR Dist-501510. Prof. A.MALLAREDDY, has done his

M. Tech in Computer Science & Engineering from JNTUH and Pursuing Ph.D (CS) in Cloud Security from JNTUH. He has 09 years of teaching experience. Currently working as Professor and Head of the Department CSE at Sri Indu Institute of Engineering &Technology, Sheriguda(Vi),Ibrahimpatnem(M) R.R.Dist, He has guided many projects for UG and PG students.He is a Life member of CSI.

A Priyanka is M. Tech Scholar, Computer Science & Engineering, Sri Indu Institute of Engineering & Tech. Sheriguda(Vi),IBP(M),RR Dist-501510. He has done B.Tech in Computer Science & Engineering from JNTUH and Pursuing M.Tech from JNTUH. She is self-motivated, hard-worker and detail oriented. I am interested in research in new things