International Journal of Electronics and Communication Engineering.
ISSN 0974-2166 Volume 10, Number 1 (2017), pp. 45-51

http://www.irphouse.com

Diagnosis of Breast Cancer using Wavelet Entropy Features

M.Chithra Devi[1] and S.Audithan[2]

[1]Research Scholar, Research and Development Centre, Bharathiar University, Coimbatore, India
[2]Professor, Department of Computer Science, PRIST University, Kumbakonam, India

Abstract

Breast cancer affects women than men. The improper or false negative state leads to mortality. The digital mammogram is preferred in the initial stage or general checkup. The radiologist used the mammographic images to find the microcalcification. The discrete wavelet transform algorithm is used to obtain different entropy variations. The supervised K-Nearest Neighbor learning algorithm is applied to classify the abnormalities, which are either benign or malignant. The experimental results show that the wavelet entropy features supports to classify the expected result rapidly and accurately.

Keywords: Breast cancer, Digital Mammogram, Discrete Wavelet Transform, Shanon, Log, Sure entropy, KNN.

1. INTRODUCTION

Breast cancer has been a death-defying disease since 200 BC. It affects both female and male genders [1], [2]. Due to the physiological structure, women are widely affected by this disease. There are various types of diseases under cancer. It cause due to uncontrollable growth of cells. Lack of correct diagnosis and untreated cancers disease can results dangerous illness and death. One such mortal disease seen among women is breast cancer. The early and correct diagnosis using mammographic images are widely preferred. The mammographic images are undergone preprocessing, feature extraction and classification to find the abnormalities [3]. The clustered Daubechies-8 Symlet-8 and Biorthogonal 3.7 filters are applied to enable the clear features of mammographic image. The Discrete wavelet algorithm extracts the features used to pass as input for classification. The Shanon, log and sure entropy values are used to recognize the microcalcifications clearly. The K-nearest neighbor algorithm classifies the breast tumor is either benign or malignant.

2. LITERATURE STUDY

An automatic detection of breast cancer is most important to avoid the unwanted biopsy and to decrease the mortality rate. The simplest method of using mammography images, to identify the disease existence and stage is widely preferable all over the world. Digital mammograms are widely utilized by specialists because it has certain advantages over analog mammograms [3]. The low dose of radiation used for its acquisition and the possibility to store huge quantities of images in digital media are some advantages. Nevertheless, in both analog and digital cases their visual examination depends on the experience and skills of the specialists who observes the images. Researchers did their research using various approaches. Some of the researches are concentrated to provide quality input image at preprocessing stage [4]. Most of the researches are focused at feature extraction stage to avoid computational complexity and to ensure the dimensionality reduction [5]-[9]. Classification level researches are also experimented to classify the feature vectors accurately [10]. Simple and accurate techniques are most required at all the stages in identification system.

3. METHODOLOGY

The proposed method has been implemented in three stages as shown in the figure 1.

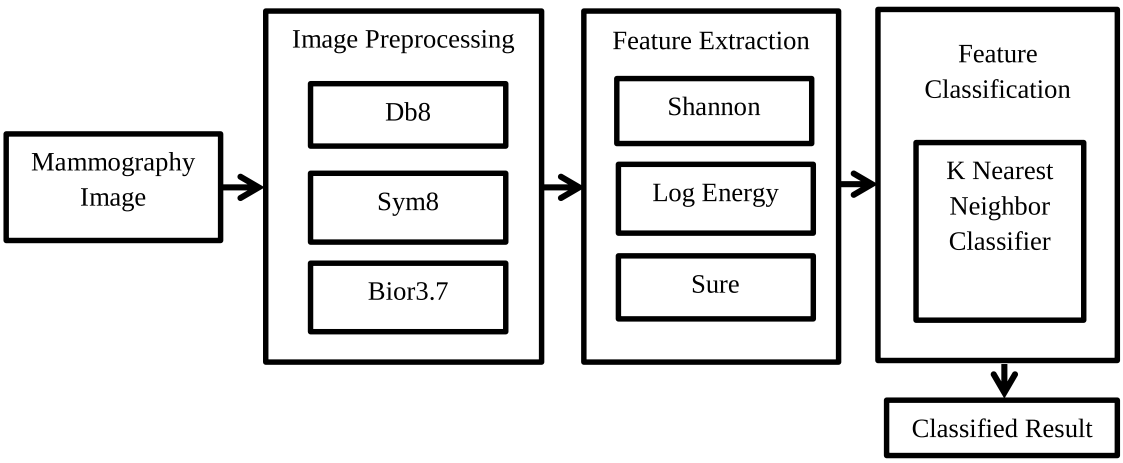

Fig.1 Propose Work Architecture

In the first preprocessing stage, the region of interest (ROI) selection of 32×32 pixels size which identifies clusters of microcalcifications. In the second stage, the feature extraction stage is based on the wavelet decomposition of preprocessed image to compute the important features of each cluster. The last stage is the classification stage, which classify between normal and microcalcifications' patterns. Further it classify between benign and malignant microcalcifications. In preprocessing stage filters such as Daubechies-8 Symlet-8 and Biorthogonal 3.7 are experimented at various levels of decomposition. The preprocessed image is further decomposed by using discrete wavelet transforms. The decomposition of the given signal by DWT is described in terms of their basis functions as given in equation 1.

$$f(x) = \sum_{m=0}^{\infty} \sum_{n=0}^{\infty} C_n^m u_{m,n}(x) \tag{1}$$

where the $u_{m,n}(x) = 2^{-m/2} u(2^{-m} x - n)$ is the translations and dilation of basis function $u(x)$. The wavelet coefficients C_n^m can be computed using a pair of low and high pass filter. A 2D DWT can be implemented by applying 1D DWTs in two dimensions separately. The output for 1 level DWT decomposition is shown in Figure 3 (a). The analysis filter bank splits a discrete image $x(m,n)$ into four sub-bands: one coarse scale (LL) and three fine scales (HL, LH, HH). If this analysis filter bank is iterated on the coarse sub-band, then the spectrum of the original image is divided by the wavelet as shown in Figure 3 (b). The subscript in Figure 2 shows the scale when the sub-bands are created.

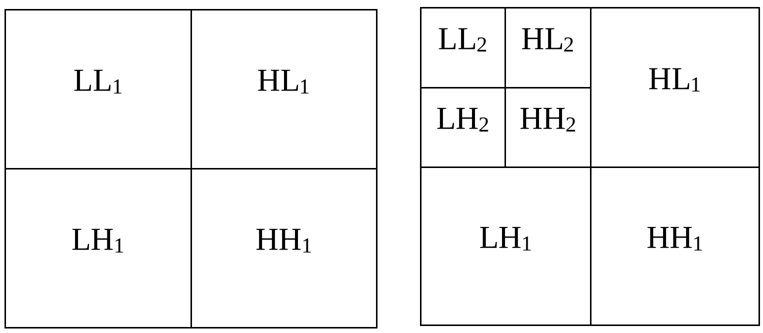

Fig.2 DWT decomposition (a) One level (b) two level

After the representation, the information extraction stage transforms it into a low dimensional space using statistical calculations on them. For mammogram image analysis, different types of entropy features which provide statistical information about the characteristic of benign and malignant are computed. These parameters are extracted from all the sub-images produced by the wavelet decomposition procedure.

The Shannon, Log and Sure entropy features are obtained by using the equation 2, 3, and 4 respectively. Shannon entropy was introduced by Shannon as a basic concept in information theory, measuring the average missing information on a random source. Let X is a random intensity value of an image with associated probability distribution P(X). Then, the Shannon entropy H of X is defined as

$$H(X) = \sum_{i=1}^{n} P(x_i) I(x_i) \tag{2}$$

where, H is a measure of uncertainty in Entropy, X is an Entire Image, x is an image intensity values and P(x) is a probability of image with possible values $\{x_1 \ldots x_n\}$ and I(x) is the function content of X. Alternatively, Log entropy is obtained by using the equation 2. Let x be a random intensity value of an image. Define the entropy H ∞ of X as follows.

$$H\infty(X) = -\log \left(\sum\{P_x(x)^2\}\right) \tag{3}$$

where, X is an Entire Image, x is an image intensity value, Log minimize the values. Px is the probability of intensity values. Similarly, the Surface entropy is also based on statistics about the distribution of local surface normals. We are interested in regions of maximal diversely oriented normals, since they show promise to be stably located at transitions of multiple surfaces or capture entire (sub-) structures that stick out of the surroundings. To identify such regions, we measure the entropy

$$H(XE) = |C_i| \leq \varepsilon \rightarrow e(s) = \sum_i \min\left(C_i^2, \varepsilon^2\right) \tag{4}$$

where, XE is a random variable characterizing the distribution of surface normal orientations occurring within a region of interest E ⊆ R 3. It extracts interest points where this entropy measure achieves local maxima, i.e. where XE is most balanced. The classification algorithm KNN supports to classify the suspicious region fall under any one of the two states. First state is normal and the second state is finding either benign or malignant. It is a supervised learning algorithm [11], [12]. The purpose of this algorithm is to classify a new object based on attributes and training samples [13],[14]. The KNN algorithm used neighborhood classification as the prediction value of the new query instance. The classification result helps to avoid unwanted biopsy.

3.1 Proposed Algorithm

Step 1: Collect breast mammogram images.

Step 2: Preprocess the image using db8, sym8 and bior3.7 filters to remove noise.

Step 3: Apply discrete wavelet transform to decompose the preprocessed image.

Step 4: Apply Shannon Entropy, Log Energy Entropy and SURE entropy to segment the abnormal region.

Step 5: Calculate the processed image accuracy for all the entropy measures.

Step 6: Classify the Compare the accuracy of entropy which gives highest accuracy.

4. EXPERIMENTAL STUDY

The digital mammograms analyzed in this work are by using the mini-MIAS database (Mammographic Image Analysis Society) constructed by the UK National Breast Screening Programme [15]. This database contains 297 images acquired to healthy women and 25 images of patients with microcalcification clusters. The accuracy obtained at various levels of decomposition is specified in the figure 2.

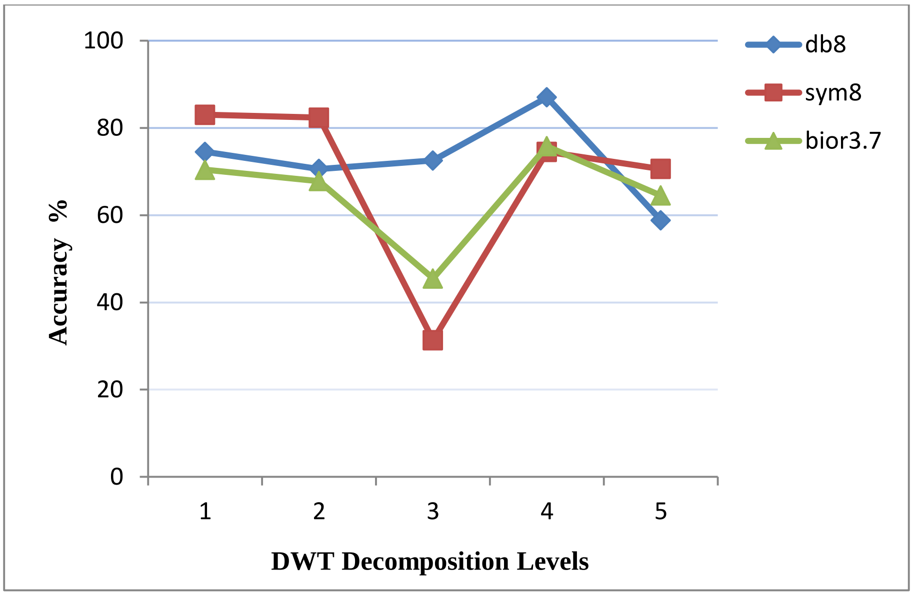

Fig. 2. Accuracy obtained at different level of decomposition

Experimental results shows that the discrete level of decomposition at fourth level provides better performance than other levels of decomposition. The classification accuracy using KNN is having maximum true positive and least false negative.

5. CONCLUSION

Researches are still in process at various dimensions by several researchers to find the solution for breast cancer diagnosis at screening level using mammographic images efficiently and rapidly. The proposed approach, using different entropies at various levels of decomposition provides fine features. The KNN classifier results high performance with 87 percentage of accuracy. The research may be further focused towards a clustered classification to improve more accuracy.

REFERENCE

[1] Strickland, R.N and Hahn, H.L., 1996, “Wavelet Transforms for Detecting Microcalcifications in Mammograms”, IEEE Transactions on Medical Imaging, 15(2), pp.218–229.

[2] Bruce, L. M., and Adhami, R.R., 1999, “Classifying mammographic mass shapes using the wavelet transform modulus-maxima method”, Medical Imaging IEEE Transactions, 18(12), pp.1170-1177.

[3] Chen, C.H and Lee, G., 1997, "On Digital Mammogram Segmentation and Microcalcification Detection using Multiresolution Wavelet Analysis”, Graphical Models and Image Processing, 59, pp. 349–364.

[4] Mario Mustra, Jelena Bozek, and Mislav Grgic, 2009, “Breast Border Extraction and Pectoral Muscle Detection using Wavelet Decomposition”, IEEE, pp.1428-1435.

[5] Balakumaran, T., Vennila, I., and Gowri Shankar, C., 2010, “Detection of Microcalcification in Mammograms using Wavelet Transform and Fuzzy Shell Clustering”, International Journal of Computer Science and Information Security, 7(1), pp.121-125.

[6] Nizar Ben Hamad, Khaled Taouil and Med Salim Bouhlel. 2013, “Mammographic Microcalcifications Detection using Discrete Wavelet Transform”, International Journal of Computer Applications, 64(21), pp.17-22.

[7] Yashashri, G., Garud and Neha G.Shahare, 2013, “Detection of Microcalcifications in Digital Mammogram using Wavelet Analysis”, American Journal of Engineering Research, 2(11), pp.80- 85.

[8] Shanmugavadivu, P., Sivakumar, V., and Suhanya, J., 2014,“Wavelet Transformation-Based Detection of Masses in Digital Mammograms”, International Journal of Research in Engineering and Technology, 3(2), pp.131-138.

[9] C.H.Chen, and G.Lee, 1997, “Image Segmentation using Multitiresolution Wavelet Analysis and Expectation Maximum (EM) Algorithm for mammography”, International Journal of Imaging System and Technology, 8(5):491-504.

[10] Pelin Gorgel, Ahmet Sertbas, Niyazi Kilic, Osman N.U., and Onur Osman, 2009, “Mammographic Mass Classification using Wavelet based Support Vector Machine”, Journal of Electrical &Electronics Engineering, 9(1), pp.867-875.

[11] Coomans, D., and Massart, D.L., 1982, “Alternative k-nearest Neighbor Rules in Supervised Pattern Recognition”, Analytica Chimica Acta, 136, pp.15-27.

[12] Kuncheva, L.I., 1995, “Editing for the K-nearest Neighbors Rule by a Genetic Algorithm”, Pattern Recognition Letters, 16, pp.809-814.

[13] Mico, L., Oncina, J., and Carrasco, R.C., 1996, “A Fast Branch and Bound Nearest Neighbour Classifier in Metric Space”, Pattern Recognition Letters, 17, pp.731-739.

[14] Bremner, D., Demaine, E., Erickson, J., Iacono, J., Langerman, S., and Godfried, P.M., 2005, “Output-Sensitive Algorithms for Computing Nearest-Neighbour Decision Boundaries”, Discrete and Computational Geometry, 33(4), pp.1-11.

[15] http://www.mammoimage.org.