Combining Support Vector Machine with ReliefF algorithm to classify ultrasound breast tumour images

Telagarapu Prabhakar

Department Electronics & Communication Engineering,
GMR Institute of Technology, Rajam, Andhra Pradesh, India.

***Abstract*—Distinguishing between malignant and benign tumours is difficult and time-consuming task for radiologists. This leads to a high rate of unnecessary biopsies that can be avoided. So, we are proposing a method to effectively differentiate benign and malignant tumours with optimal feature set. Statistical texture features are extracted from tumour region by recent shearlet transform and then to reduce the dimensionality of the feature in order to reduce the computational complexity feature selection methods are applied. Feature selection methods are such as Principal component analysis (PCA), Genetic algorithm and ReliefF based feature selection methods are applied. Using the new reduced feature set abnormalities are classified using KNN and SVM classifier. By comparing the Performance of the feature selection methods ReliefFbased feature selection with SVM classification method gives with the Accuracy of 78.33%, Sensitivity is 84.16% and 80.16% as specificity.**

***IndexTerms*—PCA, Genetic algorithm, Feature selection, ShearletTransform,k-NN,SVM and ReliefF.**

I. INTRODUCTION

Generally mammography is used to detect the breast cancer. However, there are limitations of mammography when the dense breast patients are subjected to the test. Ionizing radiation used as a source which is harmful to the human body. Sometimes physician feels the difficulty in the identification of yearly stages of malignancy with mammography images. So there is a need for another screening test to conform the cancer. Ultrasound imaging modality effectively fulfills that place because of its non-ionizing radiation and low cost [1]. But because of its low contrast images, identification and classification of tumor is difficult with normal eye. Computer Aided Diagnosis (CAD) system is proposed to overcome this limitations by assisting physicians. CAD consists of five process including preprocessing, ROI segmentation, feature extraction, feature selection and classification. In general ultrasound images contains speckle noise that affects the quality of image. Hence speckle reduction has to be carried out in the image preprocessing. Studies have described that various denoising techniques are implemented for the reduction of speckle. Initially median filter is used to reduce the speckle. But it is effective only if the noise is additive [2]. Next to that, speckle is reduced by wiener filter which will effectively reduce the multiplicative noise by converting it into additive by applying a logarithmic function [3-6]. Speckle reduction anisotropic diffusion filters (SRAD) is another technique that is used to reduce the speckle noise. The efficiency of SRAD depends on the accuracy of edge detection and selection of threshold value. In recent day's wavelet and curvelet based denoising techniques are employed to reduce the speckle noise in the ultrasound images [7-10]. Different segmenting methods such as Grey thresholding, Region growing and active contour methods are proposed for Segmenting the region of interest in the ultrasound images and each has its own disadvantage. In this paper commonly used region growing technique is applied to segment the ROI. Feature extraction and selection is the important steps in classifying the benign masses from the malignant. Features are divided into two different category such as texture and morphological. Morphological features are highly dependent on segmentation accuracy. Statistical texture descriptors define the variations in the benign and malignant tumors effectively [11]. Now a day's feature extraction by the transformation is proposed to improve the classification accuracy. Wavelet, Ridgelet and Curvelet based feature extraction methods are compared and the performance is analyzed in the previous work [12-15].

In this paper shearlet transform is used for the feature extraction. Mostly classification accuracy affected when the feature set is large and redundant. To improve the classification accuracy and to reduce the amount of data, feature selection algorithms are used. In the past decades various feature selection methods are proposed to reduce the dimensionality. Genetic PCA and ReliefF based feature selection algorithms are taken into account and its performance is compared by evaluating the reduced feature subset to the two different neural networks such as SVM and KNN.

II. MATERIALS AND METHODOLOGY

B Mode ultrasound images are accessed from BHARATH SCAN CENTRE, Chennai. 40 benign and 30 malignant images are used for further processing. Process flow is explained in the flow chart given in Figure .1

Figure 1. Block diagram for breast cancer detection and classification

Preprocessing

The preprocessing of ultrasound images involves the removal of speckle noise. Ultrasound images are affected by multiplicative noise generated by the number of scatters with random phase called speckle.If f(x,y) is the original image and n(x,y) is the speckle noise component then the noise affected image g(x,y) is represented by

$$g(x,y)=f(x,y)*n(x,y) \quad (1)$$

An image with speckle does not exactly reflect the properties of tissue So there is a necessity to remove the speckle noise without affecting the useful information in the images. In this paper wiener, wavelet and curvelet based denoising techniques are applied to remove the speckle.

Segmentation

Ultrasound image segmentation is difficult because of its small change in contrast at boundaries and discontinuities. Selection of segmentation method depends on the types of the image and its application. In this work region growing method is used to segment the ROI because of its simplicity.

Feature Extraction

In this work recent wavelet domain transform called shearlet transform is used to extract the feature set. Statistical features such as Entropy, energy, sum average, sum variance, cluster tendency, correlation and contrast, mean, angular second moment, are extracted by shearlet transformed coefficients.Applications of shearlet transform involve the multi resolution analysis of signals in higher dimensions. Shearlet transform is a affine systems which are parameterized by scaling, translation with shearing parameter. Discrete shearlet transform is represented by

$$\Psi_{a,s,t}(x)=2^{-3j/4}\,\psi\left(\left(B_{a,s}^{-1}(x-t)\right)\right) \quad (2)$$

Where a- scaling parameter t - Translation parameter ; s- Shear parameter. B matrix is defined by

$$\begin{pmatrix} a & \sqrt{as} \\ 0 & \sqrt{a} \end{pmatrix}$$

ψ is a mother wavelet function which is choose based on the condition

$$\psi(\xi_1, \xi_2) = \psi_1(\xi_1)\,\psi_2(\xi_2/\xi_1). \quad (3)$$

Where, ψ_1-wavelet function, ψ_2- bumb function.

Since shearlet decomposition is implemented by multidirectional decomposition followed by multi scale decomposition. Shear filters designed by the Mayer window is applied on the images to achieve multi directional decomposition. Multi level scaling is performed in each directional information which produces the shearlet coefficients. Because of the directional sensitivity in various scales, the shearlet transform is used to extract the features in this work.

Feature Selection

If all the features derived from each sub band is considered for classification it leads to computational complexity. It also reduces the accuracy of classification and increase the time consuming. Feature selection methods are proposed to avoid such a problems with improved accuracy. In this work Genetic algorithm and PCA based feature selection methods are used to reduce the dimensionality.

(a). Principal component analysis

PCA is a non parametric method of extracting relevant information from higher dimensional data sets. It converts the set of variable into a set of linearly uncorrelated variable called principal components. Let's consider the training feature set $X=[x_1,x_2,......x_i]$, then the derived feature set Y is represented by

$$Y=X^T P \quad (4)$$

Where P is a principal component matrix.

Principal component has its maximum variance with respect to the data set. Second principal component is selected in the direction which maximizes the variance among all directions orthogonal to the first. In similar way another principal components are also selected. This will reduce the unwanted features in effective manner.

(b) Genetic algorithm

Features derived from the shearlet transforms are taken as a input to the genetic algorithm. Optimal feature set is derived by genetic operations of subset selection, subset evaluation, stopping criteria and result evaluation.Features are taken as constant size population of individuals and each individual is evaluated based on the fitness function. High performance individuals are selected as population for next generation which keeps many of the features of its parents. New feature set for the next generation is also evaluated by two important operation of genetic algorithm called cross over and mutation.Cross over is performed by randomly selecting a point from two feature set called parents and interchanging the segment next to the selected point to create new individuals.

Mutation is done by changing one or more components in the selected individuals. This entire process is continued until the stopping criterion is satisfied. Then the selected subsets are evaluated with the prior knowledge about the data set. Small derived subset from GA is used for the classification process.

(c) Relieff algorithm

Kononenko proposed filter type feature selection method called Relief algorithm. It has been widely used because of its simplicity and efficiency. Basic concept of relief algorithm is to make the good feature should make the same category closed, and keep the simple in different categories off.Working of relief algorithm based feature selection. Initially Select sample X randomly from training data set. Then the sample with k- nearest neighbor is selected from the same class with distance H called nearest Hit. Similarly sample with k nearest neighbor is selected from other class with distance M called nearest Miss.If the distance between X and H is minimum compared with X and M then that feature is well differentiated compared with others.

$$Diff(x,X,H) < Diff(x,X,M) \quad (5)$$

If the feature satisfies the condition mentioned above which results in improved weight values. The distance between X and M is minimum compared with X and H which results in reduction of the feature weight. Weight of that feature is updated and process is continued n times.Features with maximum weight are selected for the classification.

Classification

Classifying the benign and malignant tumors of breast cancer in ultrasound image is a challenging task. The selected feature set by various feature selection methods is given to two different classifier such as KNN and SVM classifier to evaluate the classifier performance.k-NN operates based on the distance measure as an evaluating function. If feature set is represented by X and the label is represented by Y then the training set is represented by

$\{(X^1,Y^1),(X^2,Y^2)....,(X^n,Y^n)\}$. Where each X is a *d* dimensional feature vector $(x_i^1,x_i^2,\ldots\ldots x_i^d)$. When an unknown feature set (X^{new},Y^{new}) is applied, k closest training points to $x^{new}$is calculated by Euclidean distance(ED).

$$ED= [(x_j^1- x_{new}^1)^2+ \ldots + (x_j^d- x_{new}^d)^2]^{1/2} \quad (6)$$

Data is classified to the cluster which forms the maximum points with respect to other.

SVM classifier is machine learning system works based on the principle of statistical learning theories. Lets the training set is $S\{(X_i,Y_i)_{1\leq i\leq N}\}$ where $X_i \in R^n$ and $Y_i \in \{-1,1\}$ as label then SVM defines the hyper plane which maximize the distance between two class. If f(x) is the optimal hyper plane to separate the two class, it is represented by

$$f(x)= sgn (\sum \alpha_i y_i K(x_i,x)+b) \quad (7)$$

Where α_i is a Lagrange multiplicand and $K(x_i,x)$ is the kernel function. This forms the optimal hyper plane to separate the test data with maximum accuracy.

III. Results and Discussion

a) Pre Processing

Ultrasound images are collected from Bharath scan center, Chennai and the images are resized with 512x512. Images are converted into gray scale and denoising methods are applied to reduce the speckle noise. In this work Weiner, Wavelet, and Curvelet based denoising techniques are used to remove the speckle noise in ultrasound images.Wavelet and curvelet based denoising techniques are performed by applying threshold in the wavelet domain coefficients. Reconstruction was carried out in the thresholded coefficient matrix to obtain the denoised image. Fig 4.1 shows the ultrasound input image and Fig 4.2 shows denoised images by various denoising methods.

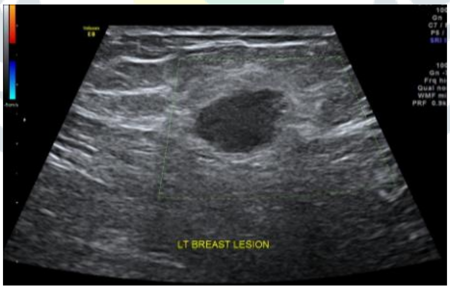

Figure 2. Ultrasound Input image

(a) (b) (c)

Figure3 Results of denoisedImage (a) Weiner filtered Image (b) Wavelet denoisedImage. (c) CurveletdenoisedImage.

From the results it is observed that wiener filter removes the speckle by smoothing. But as a result of smoothing the edge information in the images also get affected. It is also observed that the result of wiener filtering depends on the window size of wiener filter.Wavelet and Curvelet based denoising techniques produce comparatively good results than wiener filter. In these methods low frequency components and high frequency components are separated by multi scaling with translation. Vishu threshold based global thresholding is applied on the coefficient matrices to reduce the speckle. Comparing with wavelet, curvelet has higher directional sensitivity. This implies the higher edge preserving capability compared with wavelet.

Table 1. Performance comparison of Denoising techniques

Filtering methods	Benign images (PSNR)dB		Malignant images (PSNR)dB	
	Min	Max	Min	Max
Wiener filter	34.6595	39.1078	34.2981	39.3076
Wavelet thresholding	38.5931	45.2463	38.2131	45.3904
Curvelet thresholding	75.7839	79.979	75.7441	80.1755

Performance of this denoising procedure is evaluated by the measure of PSNR. From the result it is observed that curvelet based denoising produce better PSNR compared to other because of its higher directional geometry and also it is observed that the edges of images are effectively preserved.

b) Segmentation

The denoised ultrasound images are segmented by region growing algorithm. Ultrasound images with tumours are taken as input and the seed point is selected manually. Threshold value is fixed by the trial and error method and the region growing method is implemented.The input image with tumour is segmented by using region growing method and the segmented output image is shown in Figure 4.

(a) (b) (c) (d)

Figure 4 Process of segmentation (a) Input image (Malignant) (b) Region Growing method for Input image (c) Binary image (d) Segmented image

It is observed that the region growing algorithm effectively segments the tumours in the ultrasound images. It is also observed that the selection of threshold and the maximum region growing distance from the seed point plays a major role in efficiency of segmentation in region growing algorithm.

c) Feature Extraction

In the feature extraction, shearlet transform is applied in the segmented region of interest and the image is decomposed with three level scaling. It forms the six different coefficient matrixes in the frequency domain.

Nine statistical texture features described earlier are extracted from each of the coefficient matrix. Set of 54 features is extracted from 70 images in which 40 is benign and 300 is malignant. Most of these features are redundant and it may reduce the classifier performance. So the feature selection methods are applied to reduce the dimensionality of the feature set.

d) Feature Selection

Feature selection is carried out by Genetic algorithm, Relieff algorithm and PCA to reduce the redundant features. By each feature selection methods, best 15 features from the total features are taken for the classification.

Genetic algorithm

In Genetic algorithm fitness function is fixed such that it reduces the redundant information in the feature set and maximizes the information in each feature. Stopping criteria is fixed as a maximum generation. Selection process is continued until it exceeds the stopping criteria.

Relieff algorithm

Relieff algorithm selects the features with maximum nearest hit and minimum nearest loss. This is achieved by sorting the features with respect to their weights. It is also observed that the Relieff algorithm does not concern about the redundancy nature of the features.

Principal component analysis

In PCA based feature selection, PCA is applied to the feature set to find the principal component. Based on the best score, the features are ranked with their indices which form new selected the features. Table 4.4 shows the selected features by principal component analysis.

e) Classification

A total of 40 image samples in which 50 random set are consider as training and 20 image samples are considered for testing**.** Selected features by different feature selection methods were respectively fed to the classifier for differentiating benign from malignant tumor. KNN and SVM were adopted here to evaluate the consistency of these texture features. In SVM Linear kernel is used to detect the optimal hyper plane. The performance of the classification is evaluated by three common indices called accuracy, sensitivity, specificity which are represented as

Accuracy=TN+TP/TP+FN+TP+FP

Sensitivity=TP/TP+FN

Specificity=TN/TN+FP

Where TP-True Positive; TN- True Negative; FP- False Positive; FN- False Negative.

Sensitivity defines the probability of the affected individual as affected. Specificity define that the probability of classification declares an unaffected individual as unaffected. Quality of the system is measured by the term Accuracy. Table.2 and 3infer that the performance of the feature selection methods when KNN and SVM classifiers are used respectively for the classification.

Table 2 Performance of the KNN classifier

Feature reduction methods	Accuracy (%)	Sensitivity (%)	Specificity (%)
Without feature reduction	71.66	76.66	79.16
Genetic algorithm	73.33	76.17	79.17
PCA	75.00	77.83	80.83
ReliefF	76.67	76	82.50

Table 3 Performance of the SVM classifier

Feature reduction methods	Accuracy (%)	Sensitivity (%)	Specificity (%)
Without feature reduction	73.33	75	78.5
Genetic algorithm	75	80	76
PCA	76.67	80	77.5
ReliefF	78.33	84.16	80.16

Result shows that the classification accuracy improved while using the feature selection methods. Performance of the genetic algorithm based feature selection is poor compared with PCA. This is because of the feature selection mainly depends on the fitness function. Relieff based feature selection method does not concern about redundancy of features. But in PCA based feature selection each selected feature has its maximum variance with other features. In the classifier point of view misclassification in KNN is high compared with SVM classifier. This is because of its poor learning and the classification depends on the distance measure of each feature. In SVM classifier, optimal hyper plane is created based on the training data set and classification only depends on the support vectors not on all the features.

IV. CONCLUSION

Feature selection is effective technique for differentiating a benign from malignant tumours while huge numbers of features are taken into account. The aim of this work is to evaluate the optimal feature selection methods for classification of breast cancer in ultrasound images. In which texture features are extracted by shearlet transform and the feature selection was done by PCA, Relieff and Genetic algorithm. The classification was done by KNN and SVM classifier for the comparison. Comparision results shows that the ReliefF based feature selection with SVM classification method gives with the Accuracy of 78.33%, Sensitivity is 84.16% and 80.16% as specificity.

REFERENCES

[1] G.Abed Heshmati, JamshidShanbehzadeh, Roya Amjadifard, “ReliefF-based Feature Selection for Automatic Tumor Classification of Mammogram Images ”, IEEE international conference on machine vision and Image processing ,pp.1-5(2011).

[2] Changming Zhu1, Jun Ni1, Yanbo Li1, and GuochangGu, “Speckle Noise Suppression Techniques for Ultrasound Images”, IEEE International Conference on Internet Computing for Science and Engineering, pp.122-125,2009.

[3] Prabhakar, Telagarapu, and S. Poonguzhali. "Feature based active contour method for automatic detection of breast lesions using ultrasound images." Applied Mechanics and Materials. Vol. 573. Trans Tech Publications, 2014..

[4] Feihu Yang, Weiqing Cheng, Renfu Dou and Ningning Zhou, “An Improved Feature Selection Approach Based on ReliefF and Mutual Information” International Conference on Information Science and Technology,pp.246-250,2011.

[5] F.yousefiRizi, H.AhmadiNoubari, and S.K. Setarehdan, “Wavelet Based Ultrasound Image Denoising: Performance analysis and Comparison” IEEE international conference pp.3917-3920, 2011.

[6] K VenkatrayuduDevarapu, SubrahmanyamMurala , and Vinod Kumar, “Denoising of Ultrasound Images using Curvelet Transform”, IEEE proc,vol.3,pp.447-451,2010.

[7] Prabhakar Telagarapu, Poonguzhali S, “Analysis of Contourlet Texture Feature Extraction to Classify the Benign and Malignant Tumors from Breast Ultrasound Images” , International Journal of Engineering and Technology,Vol 6 No 1,pp.293-305,2014.

[8] Rahil Garnavi, Mohammad Aldeen, and James Bailey, “Computer-Aided Diagnosis of Melanoma Using Border- and Wavelet-Based Texture Analysis” , IEEE Transactions on information technology in biomedicine, vol. 16, no. 6, pp.1239-1252, 2011.

[9] Shichong Zhou, Jun Shib, JieZhub, Yin Cai, and Ruiling Wang, “Shearlet-based texture feature extraction for classification of breast tumor in ultrasound image”, journal on Biomedical Signal Processing and Control 8 ,pp,688-696, 2013.

[10] Wang-Q Lim, “The Discrete Shearlet Transform, A New Directional Transform and Compactly Supported Shearlet Frames”, IEEE Transactions on image processing, vol. 19, no. 5, pp.1166-1179, 2010.

[11] Wen-Jie Wu, Shih-Wei Lin, and Woo Kyung Moon, “Combining support vector machine with genetic algorithm to classify ultrasound breast tumour images”, journal on Computerized Medical Imaging and Graphics 36, pp.627– 633, 2012.

[12] Xiangjun Shi, H.D. Cheng, Liming Hu, WenJu, Jiawei Tian, “Detection and classification of masses in breast ultrasound images” Elsevier in digital Signal Processing 20 (2010) 824–836

[13] Xu Liu, ZhiminHuo and Jiwu Zhang, “Automated segmentation of breast lesions in ultrasound images” proc.IEEE, pp.7433-7435.

[14] Yi Chen, Yihua Lan, and Haozheng Ren, “A Feature Selection Method Base on GA for CBIR Mammography CAD”, IEEE International Conference on Intelligent Human-Machine Systems and Cybernetics, pp.175-178, 2012.

[15] Yuanyuan Wang, Jialin Shen, Yi.Guo, and Weiqi Wang, “Computerized Classification of Breast Tumors with Morphologic and Texture Features of Ultrasonic Images”, IEEE proc, pp 23-28, 2008.