CHARACTERIZATION OF BREAST TISSUES IN COMBINED TRANSFORMS DOMAIN USING SUPPORT VECTOR MACHINES

B.N. Prathibha[1] and V. Sadasivam[2]

Department of Computer Science and Engineering, Manonmaniam Sundaranar University, Tamil Nadu, India
E-mail: [1]prathibhashri@yahoo.in, [2]vs_msu@yahoo.com

Abstract
Mammography is a well established imaging technique for showing tissue abnormalities of breast and has been proven to reduce death rate due to breast cancer in screened populations of women. The proposed method classifies the breast tissues according to severity of abnormality (benign or malign) using combined transforms domain features. In this paper two such domains are explored, Discrete Cosine Transform - Discrete Wavelet Transform (DCT-DWT) and Discrete Cosine Transform - Stationary Wavelet Transform (DCT-SWT). The method is tested on 221 mammogram images from the MIAS database. The combined transform domain features proves to be a promising tool for precise classification with SVM classifier. The DCT-DWT domain yields 96.26% accuracy for discrimination between normal-malign samples comparing to DCT-SWT which gives 93.14%. The novelty of the proposed method is demonstrated by comparing with nearest neighbor classification technique.

Keywords:
Combined Transforms, Mammograms, SVM, Nearest Neighbor Classifier

1. INTRODUCTION

Although many years of effort have been spent in improving surgical and radio therapeutic techniques, the mortality rate from breast cancer remains appalling. It is commonly conceded that only early detection is the best means of reducing this mortality [1] and mammography has finally evolved as a means of achieving this purpose.

In addition to providing better images, the digital mammograms have resulted in significant reduction in radiation dose. With the importance of an emphasis on mammography in the role of early detection of breast cancer, it is of utmost importance that meticulous techniques be used. Factors that affect image quality include equipment, image-recording system, processing compression of the breast, and the technologist's skill in positioning the patient. The interpretation skills of the radiologist are limited by a suboptimal image. A poor quality image or poor positioning can account for many of the cancers missed by mammography. The Computer-Aided Detection (CAD) systems help radiologists in improving the quality of the image for visual representation. CAD's can also be used in place of a second reviewer, thus economically and technically aids radiologists in giving accurate diagnostics.

2. RELATED WORK

Many studies have been made on problems of breast cancer diagnosis, based on digital mammograms. Most of these algorithms are concentrated on discrimination between, benign and malignant microcalcifications (MCC) or between benign and malignant masses and some on detection of MCC and masses or on shape analysis of MCC and masses. A few studies have been made on the classification of normal and abnormal mammograms, detection of architectural distortion and asymmetric density. B. Zheng et al. [2] have proposed a scheme that uses a neural network with spatial domain features and Discrete Cosine Transform (DCT) based features as inputs to obtain a spectral entropy based decision criteria. The DCT features are used by Farag A. and Mashali S. [3] in order to discriminate normal mammograms from mammograms with microcalcifications where classification is done with a three-layer back propagation neural network. Essam A Rashed et al. [4] used a region based multi resolution analysis for distinguishing malign and benign tumors. The biggest wavelet coefficients were used as the discriminating feature, achieving 99.5% of successful classification. The Euclidean distance measure is used to design the classifier. The high frequency signals of mammograms are extracted by using DWT by Weidong Xu et al. [5] in order to find the presence of micro calcifications. The CAD algorithm used the Adaptive network based Fuzzy Inference system (ANFIS) for more preciseness. April Khademi and Shridhar Krishnan [6] used the statistical features from wavelet domain using the Shift Invariant DWT or the Stationary Wavelet Transform (SWT) with scale invariant representation, giving a generalized framework for medical image analysis. They classified small bowel images, retina images as well as mammogram images obtaining 85%, 82.2% and 69% accuracy respectively.

Ibrahim Faye et. al. [7] proposed a method for classification of mammogram using wavelet features, achieving 98% accuracy. The Discrete Wavelet Transform (DWT) is used by Jianhua Yao et. al. [8] for breast tumor analysis on Dynamic Contrast Enhanced Magnetic Resonance images. They obtain 0.989 and 0.984 area under the curve values for different training and testing datasets. The support vector machine is used as the classifier for breast tissue classification. The curvelet transform is used as a feature extraction technique by Mohamed Meselhy Eltoukhy et. al. [9] for breast cancer diagnosis in digital mammograms. The Euclidean distance is then used to construct a supervised classifier. The experimental results gave 98.59% classification accuracy rate, which indicate that curvelet transformation is a promising tool for analysis and classification of digital mammograms. An evaluation on the performance of CAD systems is done by Sheila Timp et. al. [10] with double reading on radiologist's diagnostics in classifying benign and malign masses. They conclude that CAD systems with temporal analysis have the potential to help radiologists in discriminating benign and malign masses.

Studies show that the extensive use of transforms due to their sparse representation and multi resolution capacity. Each transform is good at extracting any one specific type of feature from an image and an image may be comprised of different

features. Hence, the proposed system combines the transforms so that cascaded transforms compensates the drawbacks of each other if any, giving more preciseness in analysis of mammograms. Rest of the paper is divided into 3 sections, section 3 discusses the methodology used, section 4 discusses the experiments and section 5 gives the conclusion.

3. PROPOSED METHODOLOGY

The proposed methodology studies the performance of multi transform features for discriminating breast tissues according to abnormality severity. The four cornerstones of the proposed method are as shown in Fig.1. During the pre processing step, the mammograms are contrast enhanced by using Contrast Limiting Adaptive Histogram Equalization (CLAHE) algorithm and transferred to spectral domain using Discrete Cosine Transform (DCT). The wavelet features are then extracted using DWT and SWT which gives rise to two sets of features. The most distinguishing features from each domain are selected using Principal component Analysis (PCA) and fed to the SVM classifier individually for further classification.
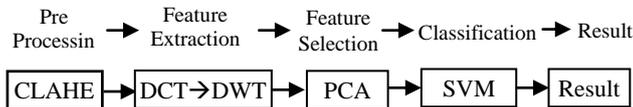

Fig.1. Flowchart of Proposed method

3.1 MAMMOGRAM ACQUISITION AND PREPROCESSING

The proposed method is implemented using MATLAB 7.5 using the mammography data taken from the Mammographic Image Analysis society (MIAS) [11]. It contains mammogram images of size 1024×1024 pixels with ground truth information about the abnormalities, i.e., type of cancer, severity of the diagnosis (Benign or Malign), center coordinates of location of the abnormality and radius of the circle enclosing the abnormality.

Before feature extraction the mammograms are contrast enhanced to improve the quality. Enhancement has the ability to make more visible unseen or barely seen features of a mammogram. A generalization of Adaptive Histogram Equalization (AHE), contrast limiting AHE (CLAHE) is used. It has more flexibility in choosing the local histogram mapping function. CLAHE [12] operates on small regions in the image, called tiles, rather than the entire image. The contrast, especially in homogeneous areas, can be limited to avoid any noise amplification, which might be present in the image by selecting the suitable clipping level of the histogram. The clip level specifies the contrast enhancement limit, the higher the value of clip level, more the contrast and vice versa.

3.2 FEATURE EXTRACTION

The mammogram images are discrete cosine transformed first, and then wavelet coefficients are extracted. The Discrete Cosine Transform (DCT) has strong capability to compress all energy of an image; hence one can often reconstruct a sequence very accurately from a few coefficients. The discrete cosine transform [13] of an N×N image, f(x, y) is defined in Eq.(1).

$$F(u,v) = C(u)C(v)\sum_{x=0}^{(N-1)} f(x,y)\cos\frac{(2x+1)u\pi}{2N}\cos\frac{(2y+1)v\pi}{2N} \quad (1)$$

where, $C(u) = C(v) = \dfrac{1}{\sqrt{N}}$, for $u, v = 0$,

$C(u) = C(v) = \sqrt{\dfrac{2}{N}}$, for $u, v = 1, 2, ... N - 1$

and $f(x, y)$ is the intensity of the pixel in row i and column j of original image f.

The discrete wavelet transform (DWT) highlights structural, geometrical and directional features of objects in an image. To compute wavelet transform [14] the original image is blurred by a low-pass convolution filter to get a low resolution image with loss of some information. A high-pass convolution filter is then applied on the original image and a "detail image" is obtained and added to the low resolution image to balance the loss of information. The convolution and decimation steps are computed recursively on the previously acquired low-resolution stream for the required decomposition level.

The SWT, also called un-decimated wavelet transform is very similar to DWT. In DWT, the translated version of a signal is not the same as the original signal, which is due to lack of shift invariance, but the SWT achieves shift-invariance by removing the down samples thus obtaining a redundant decomposition.

3.3 FEATURE SELECTION

The higher dimensional feature space may result in deterioration of the classifier performance hence a set of minimum number of features, with high discriminating power is selected. The feature set is normalized before feature extraction. PCA is the [15] commonly used statistical technique for dimensionality reduction, in the sense that it replaces a large set of observed variables with a smaller set of new variables in such a way that they highlight their similarities and differences. PCA involves the calculation of the Eigen value decomposition of a data covariance matrix or singular value decomposition (SVD) of a data matrix, usually after mean centering the data for each attribute. The results of a PCA are usually discussed in terms of the principal component coefficients, also known as loadings and the component scores.

3.4 CLASSIFICATION

Classification is done with SVM classifiers. SVM is a non-parametric, supervised classifier, [16] using labeled examples to build models for classification of new data. Given a training set, $(x_1, y_1), (x_2, y_2), \ldots (x_s, y_s)$ such that $x_i \in R^P$ are feature vectors of dimensionality P and $y_i \in \{+1, -1\}$ are labels for a binary classifier. SVM require the solution of the following optimization problem,

$$\min\left(\frac{1}{2}w^T w + C\sum_{i=1}^{N}\xi_i\right) \text{ subject to, } y_i(w^T\Phi(x_i) + b) \geq 1 - \xi$$

where, w and b are hyper plane parameters, $C > 0$ is the penalty parameter and $\Phi(x)$ is a function to map vector x into a higher

dimensional space. $K(x_i, x_j) = \Phi(x_i)^T\Phi(x_j)$ is called the kernel function.

4. EXPERIMENTS AND RESULTS

The case sample analyzed consists of 221 ROI extracted from 221 individual mammograms of mini-MIAS database. The dataset is divided into three groups; normal-benign (52-55), normal-malign (52-55), and benign-malign (52-55). The proposed system studies the performance of different filters in hybrid transform domain and combines the result.

Fitness criteria:

The degree of reliability of the classification is found with the confusion matrix and some of the figures of merit associated with it. First, by defining arbitrarily which of the two sets is positive (P) or negative (N), the following four quantities were defined: true positives (TP) is the number of positive ROI that are correctly classified to class positive; true negatives (TN) is the number of negative ROI that are correctly classified to class negative; false positives (FP) is the number of negative ROI that are incorrectly classified to be class positive; and false negative (FN) is the number of ROI that are classified to be negative despite they are of class positive. The Sensitivity, Specificity, Accuracy and Matthews's correlation coefficient (MCC) are considered as fitness criteria and calculated as follows,

$$Sensitivity = \frac{TP}{(TP + FN)}$$

$$Specificity = \frac{TN}{(FP + TN)}$$

$$Accuracy = \frac{(TP + TN)}{(P + N)} \text{ and}$$

$$MCC = \frac{(TP.TN - FP.FN)}{\sqrt{(TP + FP)(TP + FN)(TN + FP)(TN + FN)}}$$

The MCC is considered along with accuracy since it considers failure classification rate along with successful classification rate where as accuracy considers only successful classification rate. Table.1 gives the figures of TP, FN, TN, FP of the data sets used and the classification power of the proposed system in terms of sensitivity, specificity, accuracy and MCC, for leave-1-out cross validation.

Table.1 shows the performance of the classifier in discriminating normal to malign, normal to benign and from benign to malign samples using DCT-DWT domain features. It is clear that the proposed method works well in classifying normal from malign samples (96.26) than classification of normal from benign samples (93.33) and benign from malign samples (91.18).

Table.2 shows the results of classifier using DCT-SWT domain features for classifying the same dataset. It is clear that the DCT-SWT domain performs well in classifying normal from malign samples as the DCT-DWT domain. The Fig.3 shows the effect of clip limit used in CLAHE, on classification accuracy, as discussed in section 3.1. The value of the clip limit is varied from 0.01 to 0.05 and is observed that from 0.01 to 0.03 (lower contrast level) the accuracy is high and stable (93.33) and from 0.04 (high contrast level) the accuracy drops (90.65).
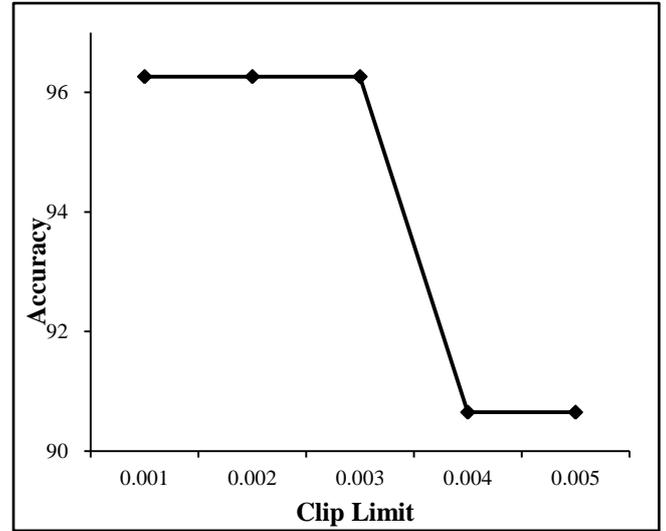

Fig.2. Effect of Clip Limit on Accuracy

The classification rate of the proposed system is compared with the Nearest Neighbor classification technique for the same dataset [14]. It is shown that the SVM classifier gives best results comparing Nearest Neighbor classification techniques for all the three datasets.

Table.1. Performance Analysis on different samples

Samples	TN	TP	FN	FP	Sensitivity	Specificity	Accuracy	MCC
Normal-Malign	49	54	1	3	98.18	94.23	96.26	0.93
Normal-Benign	48	50	3	4	94.34	92.31	93.33	0.87
Benign-Malign	47	46	4	5	92.00	90.38	91.18	0.80

Table.2. Performance Analysis on different datasets

Samples	TN	TP	FN	FP	Sensitivity	Specificity	Accuracy	MCC
Normal-Malign	46	49	3	4	92.00	94.23	93.14	0.86
Normal-Benign	42	47	5	8	84.00	90.38	87.25	0.75
Benign-Malign	44	47	5	6	88.00	90.38	89.22	0.78

Table.3. Performance evaluation with Nearest Neighbor classifier

Samples	Sensitivity	Specificity	Accuracy	MCC
Normal-Malign	98.08	90.91	94.39	0.89
Normal-Benign	90.91	90.38	90.65	0.81
Benign-Malign	87.27	76.92	78.50	0.58

5. CONCLUSION

The proposed method classifies normal and abnormal mammograms in hybrid transforms domain, using SVM classifier. The combined transform space has shown a great potential for interpreting useful diagnostic information from mammograms more accurately. In the current experiments, DCT-DWT domain features emerged as the best discriminator when compared with DCT-SWT domain. The use of preprocessing such as normalization, feature selection further improves the classification accuracy considerably. The quality of the training set and level of contrast enhancement used are perhaps the most important factors to be considered.

In future work, the performance of other state-of-the art classifiers will be studied in discriminating types of malignant abnormalities in mammograms in hybrid transform domain.

REFERENCES

[1] Ellen Shaw De Paredes, "*Atlas of Mammography*", 3rd Edition, Lippincott Williams & Wilkins, 2007.

[2] B. Zheng, Wei Quan and Laurence P. Clark, "Digital Mammography: Mixed Feature Neural Network with Spectral Entropy Decision for Detection of Micro calcifications", *IEEE Transactions on Medical Imaging*, Vol. 15, No. 5, pp. 589-597,1996.

[3] Farag A and Mashali S, "DCT based features for the detection of microcalcifications in digital mammograms", *IEEE Midwest Symposium on Circuits and Systems*, Vol. 1, pp. 352-355, 2003.

[4] Essam A Rashed, Ismail A Ismail and Sherif I. Zaki. "Multiresolution mammogram analysis in multilevel decomposition", *Journal of Pattern Recognition Letters*, Vol. 28, No. 2, pp. 286–292, 2007.

[5] Weidong Xu, Lihua Li and Shaofang Zou, "Detection and classification of Micro calcifications based on DWT and ANFIS", *Proceedings of 1st International Conference on Bioinformatics and Biomedical Engineering*, pp. 547-550, 2007.

[6] April Khademi and Shridhar Krishnan, "Medical image texture analysis: A case study with small bowel, retinal and mammogram images", *Proceedings of Canadian conference on Electrical and Computer Engineering*, pp. 1949-1954, 2008.

[7] Ibrahim Faye, Brahim Belhaouri and Mohamed Meselhy Eltoukhy, "Digital Mammograms Classification Using a Wavelet Based Feature Extraction Method", *Proceedings of the Second International Conference on Computer and Electrical Engineering*, pp. 318-322, 2009.

[8] Jianhua Yao, Jeremy Chen, and Catherine Chow. "Breast tumor analysis in DCE MRI using texture features and wavelet transforms", *IEEE Transactions of Selected Topics in Signal Processing*, Vol. 3, No.1, pp. 94–100, 2009.

[9] Mohamed Meselhy Eltoukhy, Ibrahima Faye and Brahim Belhaouari Samir, "Breast cancer diagnosis in digital mammogram using multiscale curvelet Transform", *Computerized Medical Imaging and Graphics*, Vol. 34, No. 4, pp. 269-276, 2010.

[10] Sheila Timp, Celia Varela, and Nico Karssemeijer, "Computer-Aided Diagnosis with Temporal Analysis to Improve Radiologists' Interpretation of Mammographic Mass Lesions", *IEEE transactions on information technology in Biomedicine*, Vol. 14, No. 3, pp. 803-808, 2010.

[11] http://peipa.essex.ac.uk/ipa/pix/mias/

[12] S.M. Pizer and E.P. Amburn, "Adaptive histogram equalization and its variations", *Computer Vision, Graphics, And Image Processing*, Vol. 39, No. 3, pp. 355-368, 1987.

[13] Gonzalez, Woods, and Eddins, "*Digital Image Processing Using MATLAB*" , Prentice Hall, 3rd edition, 2004.

[14] S Mallat, "A theory for multiresolution signal decomposition: The wavelet representation", *IEEE Transactions on Pattern Analysis Machine Intelligence*, Vol. 11, No. 7, pp. 674–693, 1989.

[15] Weilong Chen, Meng Joo Er, and Shiqian Wu, "PCA and LDA in DCT domain", *Pattern Recognition Letters*, Vol. 26, No. 15, pp. 2474-2482, 2005.

[16] Ethem Alpaydin and Fikret Gurgen, "Comparison of statistical and neural classifiers and Their Applications to Optical Character Recognition and Speech Classification", In Londes, C.T. (ed.), Image Processing and Pattern Recognition Academic Press, pp. 61-88, 1996.

[17] B.N. Prathibha and V.Sadasivam, "Multi Resolution Texture Analysis of Mammograms Using Nearest Neighbor Classification Techniques", *International journal of Information Acquisition*, Vol. 7, No. 2, pp. 109-118, 2010.