International Journal of Applied Engineering Research
ISSN 0973-4562 Volume 10, Number 11 (2015) pp. 28225-28232

http://www.ripublication.com

A Computer Aided Design For Mass Classification In Digital Mammogram Using Dual Tree M-Band Wavelet Transform and Support Vector Machine

C. Suba[#1], Dr K. Nirmala[*2]
[#1]Research Scholar, Vels University, Chennai, India
[#1]suba_c1978@gmail.com
*[*2]Associate Professor, Quaid-e-Millath Government College, Chennai, India*
*[*2]nimimca @yahoo.com*

Abstract

Breast Cancer is the most common cancer in female after lung and colon cancer. Digital mammography system is the most excellent imaging modality to find breast cancer at the earliest. In this study an efficient computer aided techniques for mass classification in digital mammogram images is proposed using Dual Tree M-Band Wavelet Transform (DTMBWT) and Support Vector Machine (SVM) classifier. The feature extraction process extracts the energy features from the sub-bands of decomposed mammograms and SVM classifier is trained by using the extracted energy features at various level of decomposition. Mammographic Image Analysis Society (MIAS) database images are used for evaluation. Five different types of kernels are used for performance evaluation in SVM classifier. The results show that the proposed system achieves 93.06% accuracy on mass mammograms in MIAS database images.

Keywords: Wavelet Analysis, M-band wavelets, digital mammograms, mass classification, Support Vector Machine

Introduction

Breast Cancer is caused by the irregular and uncontrolled growth of cells. It may cause to death, if the growth of cancer cells is not treated properly. According to American Cancer Society, breast cancer is the first leading cause of cancer death after lung and bowl cancer in the last 10 or more years. The total numbers of female breast cancer cases are steadily increasing from last 25 years and it affects the patients among 20 to 40 years age group [1]. It has become the most hazardous types of cancer among women in the world. The world health organization; International Agency for Research on cancer estimates that more than 400,000 women expired each year and according to American college of Radiology (ACR) statistics, one out of nine women will develop breast cancer during their life cycle [2].

Masses appears in the digital mammogram as bright regions of different sizes, margins (circumscribed, micro lobular, obscured, indistinct, and speculated) and shapes (round, oval, lobular, irregular). They can be either cancerous (malignant) or non-cancerous (benign) [3]. Full field digital mammogram (FFDM) based computer aided mass classification approach is presented in [4]. Morphological and spatial gray level dependence based texture features extracted. Liner Discriminant Analysis (LDA) is adopted for mass classification into benign and malignant tissues.

Cross variogram spectral descriptor based mass classification approach is described in [5]. At first, bilateral registration is performed in pre-processing stage. Then, the features are computed from cross-variogram of digital mammogram. Finally, Support Vector Machine (SVM) classifier is adopted for mass classification. Automated mass classification system is implemented in [6] using digital mammogram images. Texture features are extracted from the Gray Level Co-occurrence Matrix (GLCM) and neural network classifier is employed for the classification of mass into benign or malignant.

Automated detection/diagnosis tools of malignant masses in screening mammography are presented using k-means and SVM classifier [7]. In order to detect the mass lesion, *k*-means segmentation is adopted initially. Then texture features are obtained by employing GLCM technique and classification of mass and non mass region is done by SVM classifier. Wavelet features based mass classification system is implemented using digital mammogram images in [8]. SVM classifier is exploited as classifier for benign and malignant discrimination.

Hybrid approach is developed to segment the malignancy region in mammograms [9]. Noise and artefact are removed in the pre-processing stage. Alarm region generation process with region growing method is used to segment the suspicious region. The segmented region will be examined with Gabor filter in different angles and frequency levels for the classification. Computer aided detection (CAD) systems are used to aid radiologists in detecting the malignant masses in mammograms during screening [10]. It is composed of two folds: first a set of 30 and a set of 73 region based features are extracted for classifier's performance evaluation. Then performance of SVM classifier is compared with neural network, *k*-nearest neighbour and LDA classifiers.

In this study, Dual Tree M-Band Wavelet Transform (DTMBWT) and SVM based mass classification approach is proposed for early detection of breast cancer. Section 2 describes the experimental work of the proposed system. The experimental outcomes of the proposed system are discussed in section 3 using different kernels in SVM classifier on the extracted features and finally, conclusion is made in section 4.

System Design

The proposed system is built based on DTMBWT for feature extraction and SVM for building the classifiers. The proposed mass classification system using digital mammography images constructed by three sequential modules: ROI selection, feature extraction and classification. Figure 1 shows the proposed system design for mass classification.

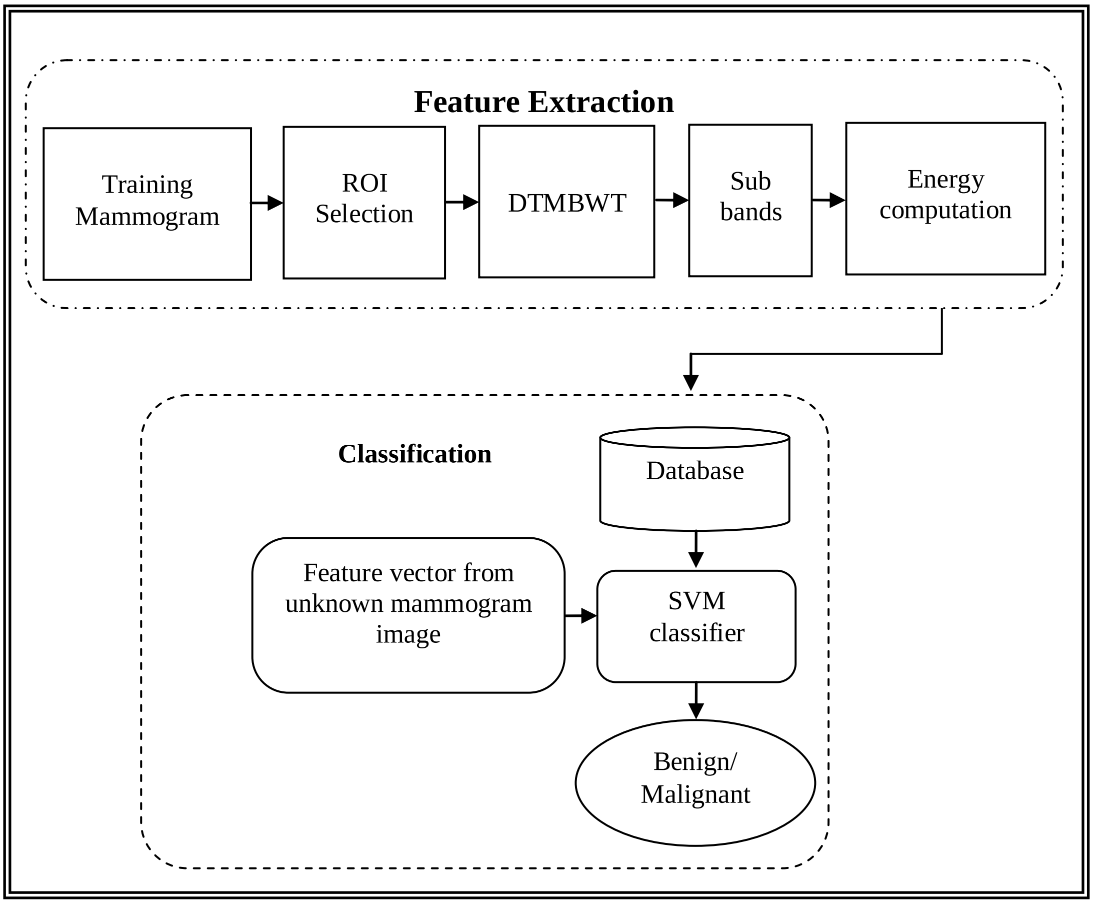
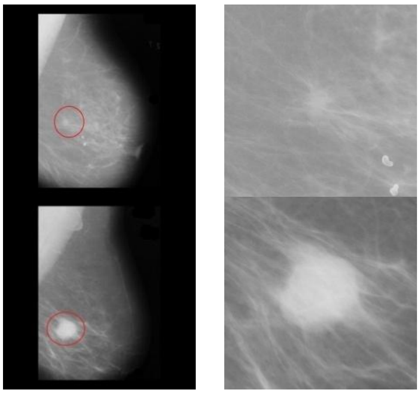


Figure 1: Overall view of the proposed mass classification system using DTMBWT and SVM

Region of Interest (ROI) Selection

The first stage of the proposed mass classification is the selection of region of interest. In order to reduce the computation time and increase the classification accuracy of the system, instead of analysing the acquired digital mammogram, only the region that contains abnormalities of size 256x256 pixels is selected from the original mammogram of size 1024x1024 pixels. Mass lesions appear in digitized mammograms as small to large regions, with intensity values higher than the surrounding regions.

Figure 2 shows the sample mammogram images and its abnormal region. The red circle in the digital mammogram shows the region of abnormalities given by the MIAS database. Top row shows the benign image and bottom row shows the malignant mammogram.



(a) (b)

Figure 2: MIAS Database Images (A) Digital Mammogram (B) Region of Abnormality

Feature Extraction

Feature extraction is considered as significant pre-processing approach for any type of machine learning approaches. Generally, low dimensional representation of an image is called as feature vector. The best discriminative nature of the feature vector provides better characterization of normal and abnormal mammogram patterns. In this study, DTMBWT is used for extracting feature vector from the mammograms.

The Dual-Tree wavelet transform was initially proposed by Kingsbury [11] and further investigation is done by Selesnick [12]. The M-band dual-tree wavelets prove more selective in the frequency domain than their dyadic wavelet transform. As it is a multiresolution analysis, the feature vector is computed by decomposing the given mammogram using DTMBWT at pre defined levels of decomposition with *M* band filter banks. It produces a number of high frequency sub-bands which depends on the number of level used for decomposition and also the number of filter banks. It is defined by $J \times M^2 - J + 1$, where *J* is number of level of decomposition and *M* is number of band filters. The size of the sub-bands obtained from the decomposition is same as the original image size. Hence, it is very difficult to use the decomposed coefficients as features and also the number of sub-bands increases while increasing the decomposition level. To ease the feature vector computation, energy is computed from each sub band using following eqn.

$$\text{Subband Energy} = \frac{1}{S_h \times S_w} \sum_{i=1}^{S_h} \sum_{j=1}^{S_w} |S(i, j)| \qquad (1)$$

Where S is the sub-band, S_h and S_w is the height and width of the sub-band. This formula is used for extracting all sub-band energies of a given mammogram image.

Classification

The classification stage of the proposed system predicts the abnormal severity of mass into benign/malignant that aids radiologists for early diagnosis and progression of breast cancer in digital mammogram images. The selection of significant feature set not only improves the system performance but also the selection of optimal classifier may increase the classification rate. In order to achieve better classification, supervised learning approach based SVM classifier is adopted.

In order to classify the severity of mass into benign/malignant, energy features are computed using aforementioned feature extraction process. Then, the unknown feature vector is fed into the trained SVM classifier whereas the unknown image is classified into benign/malignant abnormality. The performance of the proposed system is analyzed using five kernel functions in SVM classifier; linear, quadratic, polynomial, Radial Basis Function (RBF) and Multilayer Perceptron (MLP).

Results and Discussions

In order to assess the performance of the proposed mass classification system, totally 56 mass (37 benign and 19 malignant) mammograms are taken from the MIAS database. The proposed mass classification approach is evaluated in terms of classification accuracy which is defined as the ratio of correctly classified mass mammograms into total number of images tested. Also the performance is measured using *k*-fold cross validation approach (*k*=10). The performances of the proposed system on mammograms containing benign, malignant masses are shown in table 1 and 2 respectively. Table 3 shows the average classification accuracy of the proposed system on whole mass mammograms in the MIAS database images.

Table 1: Performance of The Proposed System on Mammograms Containing Benign Masses

Level of Decomposition	Classification Accuracy of Benign Masses (%)				
	SVM Kernel function				
	Linear	Quadratic	Polynomial	RBF	MLP
1	54.52	52.96	55.35	64.60	49.95
2	56.10	65.77	56.14	76.57	45.68
3	56.95	69.68	64.08	78.40	51.88
4	56.33	60.30	77.26	98.74	54.39
5	64.67	**87.38**	75.09	92.36	57.28

Table 2: Performance of the proposed system on mammograms containing malignant masses

Level of Decomposition	Classification Accuracy of malignant Masses (%)				
	SVM Kernel function				
	Linear	Quadratic	Polynomial	RBF	MLP
1	42.92	65.75	72.42	62.50	67.25
2	31.58	59.08	52.58	41.25	**39.75**
3	38.58	58.75	42.17	29.25	42.83
4	44.58	61.08	45.00	15.67	52.50
5	47.08	**98.75**	57.33	13.67	47.58

Table 3: Average Classification Accuracy of Proposed System on Mass Mammograms

Level of Decomposition	Average Classification Accuracy Masses (%)				
	SVM Kernel function				
	Linear	Quadratic	Polynomial	RBF	MLP
1	48.72	59.36	63.88	63.55	54.09
2	43.84	62.43	54.36	58.91	42.72
3	47.62	64.22	53.12	53.82	52.36
4	50.27	60.69	61.13	53.70	53.45
5	51.71	**93.06**	66.21	53.01	52.43

It is observed from the tables 1 to 3 that the proposed mass classification system using DTMBWT energy features achieves a maximum of 93.06% using Quadratic kernel in SVM classifier. The 5th level features produce better classification accuracy in comparison with other features extracted at different level of decomposition. Also it is noted that the Quadratic kernel produces better result than other kernels in SVM classifier. Figure 3 shows the average maximum classification accuracy of the proposed system using five various kernels in SVM classifier.

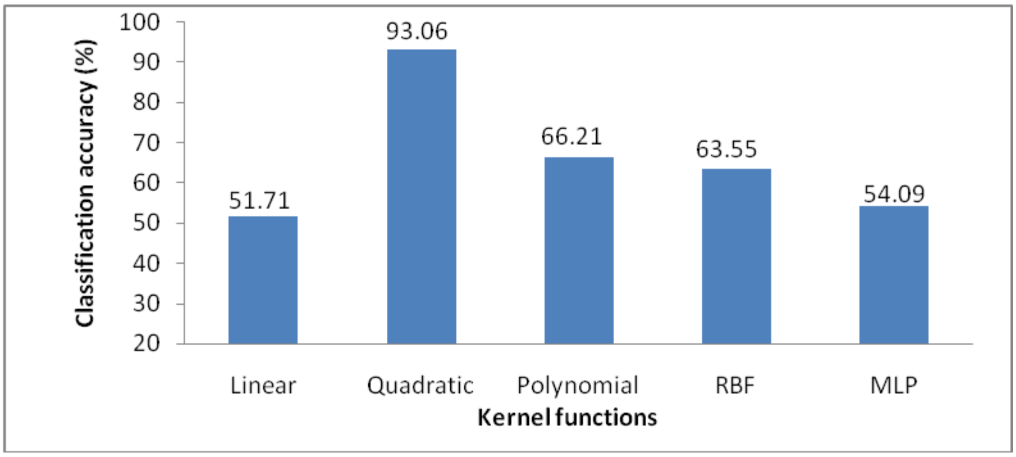


Figure 3: Kernel Functions Vs. Classification Accuracy

Conclusion

In this study, an automated mass classification system using DTMBWT and SVM is presented using digital mammograms. The proposed system makes use of DTMBWT and SVM for feature extraction and classification respectively. A low dimensional energy features from the DTMBWT sub-bands are used for training the SVM classifier. In order to classify an unknown mass mammogram into benign/malignant, a hyperplane is created by the SVM classifier using the training benign/malignant features. Also, different kernels are used for obtaining better classification results. Experimental results show that the proposed system achieves maximum classification accuracy of 93.06% using quadratic kernel function.

References

[1] American Cancer Society Learn about Breast Cancer available at http://www.cancer.org.

[2] Jasmine, J. L., Baskaran, S., and Govardhan, A., 2011, “An automated mass classification system in digital mammograms using contourlet transform and support vector machine”, International Journal of Computer Applications, 31, pp.54-54.

[3] Gouda, W., Selim, M. M., and Elshishtawy, T., “An Approach for Breast Cancer Mass Detection in Mammograms”, Al Azhar University Engineering Journal, pp.1-11.

[4] Wei, J., Sahiner, B., Hadjiiski, L. M., Chan, H. P., Petrick, N., Helvie, M. A., and Zhou, C., 2005, “Computer-aided detection of breast masses on full field digital mammograms”, Medical physics, 32(9), pp. 2827-2838.

[5] Ericeira, D. R., Silva, A. C., De Paiva, A. C., and Gattass, M, 2013, “ Detection of masses based on asymmetric regions of digital bilateral mammograms using spatial description with variogram and cross-

variogram functions", Computers in biology and medicine, 43(8), pp.987-999.

[6] Rekha, C., and Gayathri, G., 2014, "Approaches For Automated Detection And Classification Of Masses In Mammograms", International Journal of Engineering and Computer Science, 3(11), pp. 9097-9011.

[7] de Oliveira Martins, L., Junior, G. B., Silva, A. C., de Paiva, A. C., and Gattass, M., 2009, " Detection of masses in digital mammograms using K-means and support vector machine", ELCVIA: electronic letters on computer vision and image analysis, 8(2), pp.39-50.

[8] AbuBaker, A., 2012, "Mass lesion detection using wavelet decomposition transform and support vector machine", International Journal of Computer Science & Information Technology, 4(2), pp. 33-46.

[9] Meenalosini, S., Janet, D. J., and Kannan, D. E., 2012, "Segmentation Of Cancer Cells In Mammogram Using Region Growing Method And Gabor Features", International Journal of Engineering Research and Applications, 2(2), pp.1055-1062.

[10] Lesniak, J., Hupse, R., Kallenberg, M., Samulski, M., Blanc, R., Karssemeijer, N., and Székely, G., 2011, "Computer aided detection of breast masses in mammography using support vector machine classification", International Society for Optics and Photonics , pp.79631K-79631K.

[11] Kingsbury, N., 2001, "Complex wavelets for shift invariant analysis and filtering of signals, Applied and computational harmonic analysis", 10(3), pp.234.253.

[12] Selesnick, I.W., 2004, "The double-density dual-tree DWT", IEEE Transaction on Signal Processing, 52(5), pp.1304.1314.