An Automated Classification of Microcalcification Clusters in Mammograms using Dual Tree M-Band Wavelet Transform and Support Vector Machine

C. Suba
Research Scholar
Vels University, Velan Nagar
Pallavaram, Chennai, Tamilnadu, India

K. Nirmala, Ph.D.
Associate Professor in Computer Science
Quaid-e-Millath Government College
Anna Salai, Chennai, Tamilnadu, India

ABSTRACT
Breast cancer is the second leading cause of cancer deaths after lung cancer. In order to avoid mortality due to breast cancer, an efficient computer aided diagnosis system for early prediction of breast cancer is needed. In this paper, an efficient computerized system is designed for the classification of Microcalcification Clusters (MC) in digitized mammograms. The proposed system uses Dual Tree M-Band Wavelet Transform (DTMBWT) to represent the digital mammogram in a multiresolution manner and Support Vector Machine (SVM) for classification. The extracted sub band energies from DTMBWT decomposed mammograms are used as distinguishable features for the classification of MCs into either malignant or benign by SVM classifier. The results show that the proposed DTMBWT based classification system achieves 91.83%accuracy on Mammographic Image Analysis Society (MIAS) database images.

General Terms
Feature extraction, Classification, Energy computation.

Keywords
Digital mammography, microcalcification, benign, malignant, wavelet transform, support vector machine.

1. INTRODUCTION
Cancer statistics in 2012 show that 4,77,000 men and 5,37,000 women are diagnosed with cancer in India and breast cancer is the most diagnosed carcinoma type in the world. The incidence of breast cancer is increased by 1.1% from 2008 to 2012. Among the cancer patients, 27% were diagnosed with breast cancer [1]. Ant Colony Optimization (ACO) based classification of MCs using mammogram images are described in [2]. The diagonal 'S'matrix obtained from the singular value decomposition of approximation band of wavelet transform is used as feature set along with the selected Jacobi moments using ACO. Without ACO optimization, the combined feature set is also analysed by Manoharanet. al.[3].SVM classifier is employed for MCs classification. The classification of microcalcifications using Jacobi moments is discussed by Lakshmi et al [4]. Jacobi moments are calculated using overlapping window of size 4x4 on abnormal region.

Dyadic wavelet transform and fuzzy shell clustering for the detection MC in mammogram images is presented by Balakumaran et al [5]. Original mammogram images are decomposed by applying dyadic wavelet transform. Then, fuzzy shell algorithm is adopted for the detection of MCs in mammogram. MCs classification in digitized mammograms using soft computing techniques is described by Subash Chandra Bose et al [6]. It consists of four stages; preprocessing, segmentation, feature extraction and classification. Adaptive median filter is used for noise removal in the preprocessing stage. Then, pectoral muscle extraction and MCs detection is done at the second stage by employing fuzzy c means clustering. Consequently, wavelet transform is adopted for statistical feature extraction and neural network is used as classifier for benign and malignant MC classification.

Banumathi et al explained the detection of MCs in mammogram images [7]. Initially, region of interest is extracted from the whole mammogram images by eliminating background information. Before extracting features, noise removal and contrast enhancement technique is applied. Finally, Local binary pattern is adopted for feature extraction and SVM classifier is used for classification into benign and malignant. Daubechies wavelet is utilized by Kohei Arai et al [8] for feature extraction and SVM classifier is adopted for effective binary classification. The combination of multi-scale analysis contourlet transform and neural network is employed for the classification of MCs by Jasmine et al [9]. Features are extracted from the contourlet coefficients and tested by neural networks.

Sara Dehghani et al explained the Statistical and signal processing based classification of MCs [10]. The statistical features such as contrast, correlation, homogeneity and entropy are extracted from the co-occurrence matrix. Radon, Ridgelet and curvelet transform based features are used for SVM based classification. A set of biggest curvelet coefficients is used as feature vector for breast cancer diagnosis by Faye [11]. Nearest neighbour classifier with Euclidean distance measure is used for classification.

In this paper, an efficient classification of MCs system is presented using DTMBWT and SVM classifier. The preliminaries of DTMBWT and SVM are discussed in section 2. Section 3 describes the proposed system for MCs classification and their simulation results are discussed in the next section. The conclusion based on the obtained results is made in Section 5.

2. PRELIMINARIES
The basic methodologies behind the proposed classification of MCs system are DTMBWT and SVM. The theoretical backgrounds of both approaches are discussed in this section.

2.1 Dual Tree M-Band Wavelet Transform
The Dual-Tree wavelet transform was initially proposed by Kingsbury [12] and further investigated by Selesnick [13].The M-band dual-tree wavelets prove more selective in the frequency domain than their dyadic wavelet transform. The performance of M-band wavelet Transform is demonstrated

via de-noising comparisons on several image types with various M-band wavelets and thresholding strategies by Caroline Chaux et al [14].

Let us consider a 1D signals belonging to the space $L^2(R)$. The M-band multi-resolution analysis of $L^2(R)$ (with $M \geq 2$) is defined by one scaling function (or father wavelet) $\psi_0 \in L^2(R)$ and (M -1) mother wavelets $\psi_m \in L^2(R)$, $m \in \{1,\ldots M-1\}$.These functions are solutions of the following scaling eqn.

$$\frac{1}{\sqrt{M}}\psi_m\left(\frac{t}{M}\right) = \sum_{k=-\infty}^{\infty} h_m[k]\ \psi_0(t-k) \qquad (1)$$

where the sequence $(h_m[k])\ k \in Z$ are square integrable and are real-valued. The Fourier transform of $(h_m[k])\ k \in Z$ is a 2π-periodic function, denoted by H_m. Hence, the above equation can be expressed in frequency domain as

$$\sqrt{M}\hat{\psi}_m(M\omega) = H_m(\omega)\hat{\psi}_0(\omega) \qquad (2)$$

where $\hat{a}$ is the Fourier transform of a function a. For the set of $\bigcup_{m=1}^{M-1}\left\{M^{-j/2}\,\psi_m(M^{-j}\,t-k), (j,k) \in Z^2\right\}$ function correspond to an orthonormal basis of $L^2(R)$ the following para-unitarity conditions must hold:

$$\sum_{p=0}^{M-1} H_m\left(\omega + p\frac{2\pi}{M}\right)H_m^*\left(\omega + p\frac{2\pi}{M}\right) = M\delta_{m-m'} \qquad (3)$$

where $\delta_m = 1$ if $m = 0$ and 0 otherwise. H_0 is low-pass whereas usually the frequency response H_m, $m \in \{1,\ldots M-2\}$ is a band-pass filter. In this case, cascading the M-band para unitary analysis and synthesis filter banks depicted in the upper branch in Figure 1 allow to decompose to reconstruct perfectly a given signal.

Two-dimensional separable M-band wavelet bases can be derived from the 1D dual-tree decomposition.In a 2D case, there are two bases of $L^2(R^2)$; the first one corresponds to the classical 2D separable wavelet basis while the second one results from tensor products of the dual wavelet basis functions. A discrete implementation of these wavelet decompositions starts from first level to go up to the coarsest resolution level.

2.2 Support Vector Machine

SVM is a set of supervised machine learning approach used for classification, regression and outlier's detection. It is a nonlinear classifier and trained to automatically detect the presence of microcalcifications in a mammogram by El-Naqa et al [15]. It classifies the binary classes by computing a class boundary hyper plane and maximizing the margin in the given training data. It is used for wide range applications such as modern statistical learning theory, digital recognition, object recognition, speaker identification, face recognition and detection, cancer diagnosis, glaucoma diagnosis, gene data analysis and text categorization.

A classification task usually involves with training and testing data which consists of some data instances. Each instance in the training set contains one "target value" (class labels) and several "attributes" (features).Let us consider the training samples $(x_1, y_1), (x_2, y_2), \ldots, (x_n, y_n)$ where x_i in R^d, d-dimensional feature space, and y in {-1,+1}, the class label for n samples. Mathematically, SVM finds the optimal values for the hyper plane parameters w (e.g. w_0) and b (e.g. b_0) that separates the samples. After finding the optimal separating hyper plane, such as $w_0.X + b_0 = 0$, an unseen pattern x_t, can be classified by the decision rule:

$$f(x) = \text{sign}(w_0.X + b_0) \qquad (4)$$

where x is a vector of the dataset mapped to a high dimensional space. Each x_i, belonging as it does to one of two classes, has a corresponding value y_i, classes, while w and b are parameters of the hyper plane that the SVM will estimate. The nearest data points to the maximum margin hyper plane lie on the planes:

$$\begin{aligned}(w.x) + b &= +1 \text{ for } y = +1\\ (w.x) + b &= -1 \text{ for } y = -1\end{aligned} \qquad (5)$$

By rescaling w and b, with no loss of generality, and grouping the above constraints in a single formula:

$$\forall_i, y_i\, f(x_i) \geq 1 \qquad (6)$$

where y = +1for class w_1 and $y = -1$ for class w_2.The optimal separating hyper plane is enforced to separate the two classes of examples with the largest margin because, intuitively, a classifier with a larger margin is more noise-resistant. SVMs identify the data points near the optimal separating hyper plane which are called support vectors. The distance between the separating hyper plane and the nearest of the positive and negative data points is called the margin of the SVM classifier.

3. PROPOSED METHOD

The proposed system for the classification of microcalcification in digitized mammographic image sis composed of two modules; feature extraction and classification. DTMBWT, a variant of wavelet transform is analysed for feature extraction and SVM is employed for classification. Figure 2 shows the proposed system using DTMBWT and SVM classifier.

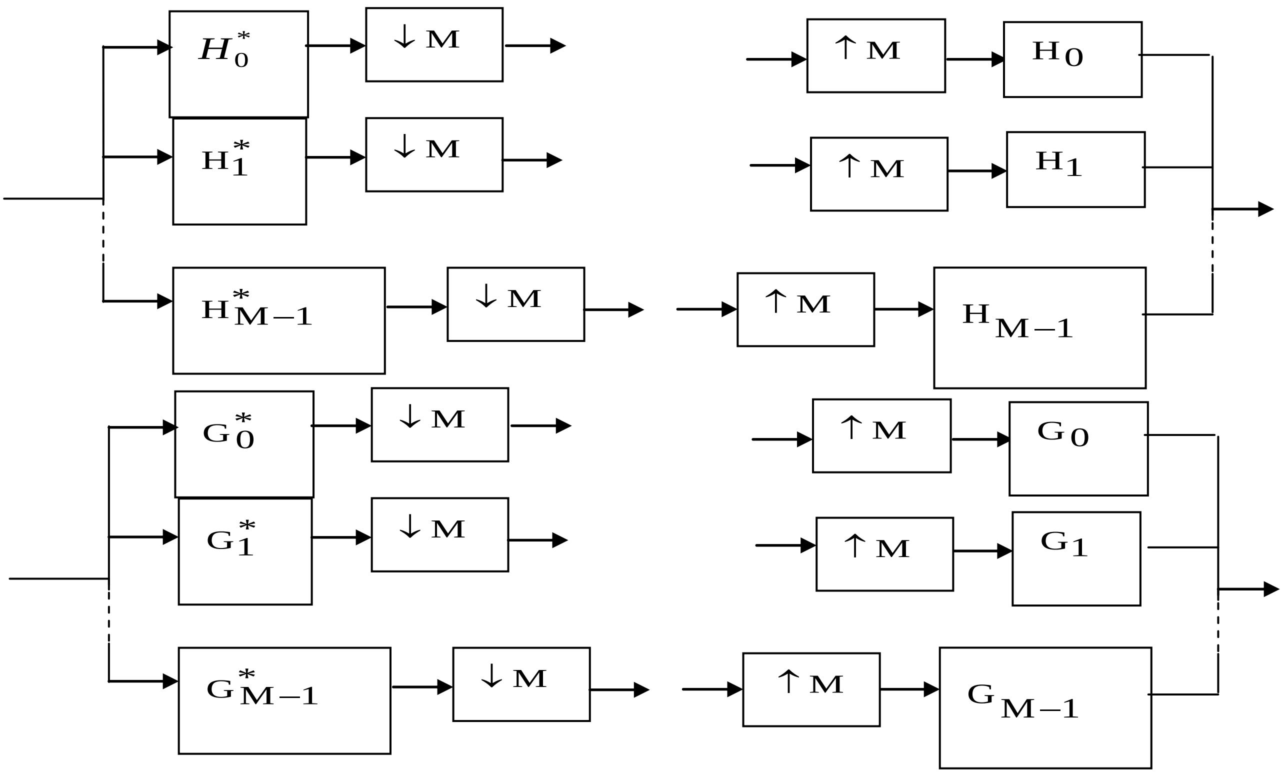


Fig 1: Analysis/synthesis M-band para unitary filter bank

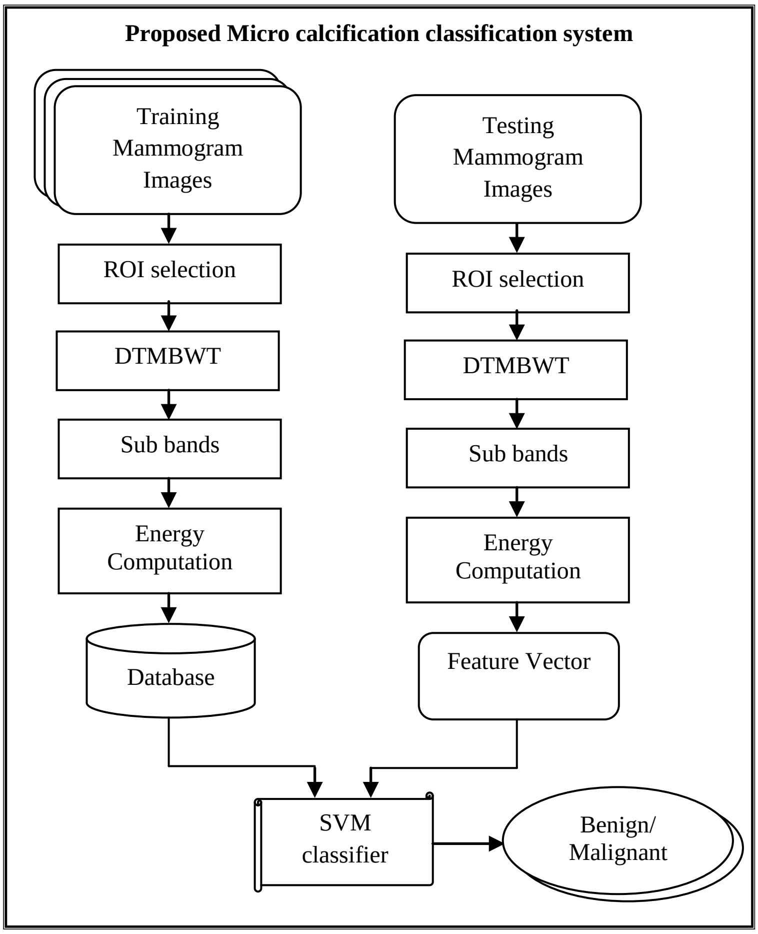


Fig 2: Proposed classification system using DTMBWT and SVM

3.1 Feature Extraction

Feature extraction is an important pre-processing and machine learning approach for all pattern recognition system. In this stage, the best discriminating features for the classification task is extracted. The proposed feature extraction process starts with decomposing the given mammogram by DTMBWT. As it is a multi-resolution analysis, pre defined levels of decomposition (*J*) up to 5th level are used for feature extraction. The decomposition by DTMBWT yields sub bands by filtering the given mammogram using *M* band filter banks. The number of sub-bands created by DTMBWT depends on the number of level used for decomposition and also the number of filter banks used. It is defined by $J \times M^2 - J + 1$, where *M* is number of band filters and *J* is number of level of decomposition.

While decomposing the given mammogram by DTMBWT, the size of yielded decomposed image is equal to input mammogram size. Hence, it is very difficult to select the features from the high dimensional coefficients of the decomposed mammogram. To avoid this, the energy value of each sub bands is computed. The sub-band energy is nothing

$$\text{Subband Energy} = \frac{1}{S_h \times S_w} \sum_{i=1}^{S_h} \sum_{j=1}^{S_w} \left| S(i,j) \right| \qquad (7)$$

where S is the sub-band, S_h and S_w is the height and width of the sub-band. Using the aforementioned formula, all sub-band energies are computed and collectively called as feature vector.

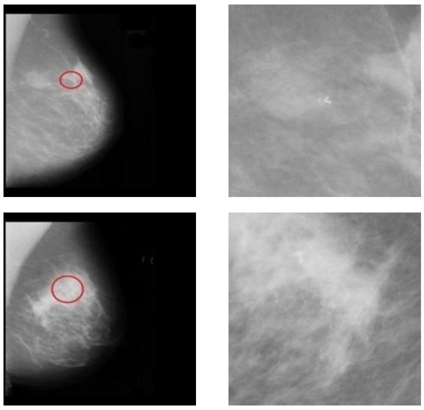
3.2 Classification Stage

In addition to efficient feature extraction algorithm, the selection of appropriate classifier improves the performance of the classification system. In this study, SVM classifier is used as classifier. It is a statistical learning theory based pattern recognition technique. A hyperplane is formed using the extracted features of benign and malignant samples by the SVM classifier which classifies the features optimally.

In this stage, two class SVM classifier is designed to classify the given microcalcification clusters into either benign or malignant. The extracted sub band energy features from DTMBWT for all training samples are fed to the classifier for training. The performance of SVM classifier is analysed using five kernel functions; linear, quadratic, polynomial, Radial Basis Function (RBF) and Multilayer Perceptron (MLP). To classify the unknown mammogram, the proposed DTMBWT sub-band energy features are extracted and given to the trained SVM classifier.

4. RESULTS AND DISCUSSIONS

MIAS digital mammogram database images are used in this study. It contains mammograms of different types of abnormalities; microcalcification, circumscribed masses, speculated masses and normal images. The proposed system considers only the microcalcification images as it is very difficult for the radiologists to classify the microcalcification clusters as benign or malignant. There are 25 microcalcification images (12 benign and 13 malignant) of size 1024x1024 pixels available in MIAS database. The locations of abnormalities are also available. In order to reduce the computation time and increase the classification accuracy of the system, instead of analysing the entire mammogram only the region of size 256x256 pixels that contains abnormalities is selected. Figure 3 shows the sample mammogram images and region of abnormality. The red circle in the entire image shows the region of abnormalities given by the MIAS database. Top row shows the benign image and bottom row shows the malignant image.



(a) (b)

Fig 3: MIAS database images (a) entire mammogram (b) region of abnormality

All the microcalcification images in the MIAS database are considered in this study. As the number of images for classification is only 25, k-fold cross validation is used to test the SVM classifier. The performance of the overall classification accuracy achieved by the proposed classification system corresponding with five SVM kernel functions are shown in Table 1 to 5 respectively. The accuracies shown in the tables are the average accuracy of 10 folds.

Table 1. Performance of the proposed system using linear kernel in SVM

Level of Decomposition	Classification Accuracy (%)		
	Benign	Malignant	Average
1	41.33	72	56.67
2	52.67	72.5	62.58
3	74.17	69.17	**71.67**
4	61.33	54.33	57.83
5	64.67	50.67	57.67

Table 2. Performance of the proposed system using quadratic kernel in SVM

Level of Decomposition	Classification Accuracy (%)		
	Benign	Malignant	Average
1	51	48.33	49.67
2	69.67	61.67	65.67
3	70.33	56	63.17

4	56.67	63.83	60.25
5	62.17	52.67	57.42

Table 3. Performance of the proposed system using polynomial kernel in SVM

Level of Decomposition	Classification Accuracy (%)		
	Benign	Malignant	Average
1	56.5	52.83	54.67
2	67.5	73.17	70.33
3	94.33	89.33	91.83
4	75.67	72.67	74.17
5	61.17	52.5	56.83

Table 4. Performance of the proposed system using RBF kernel in SVM

Level of Decomposition	Classification Accuracy (%)		
	Benign	Malignant	Average
1	61.67	43.67	52.67
2	66.33	55.83	61.08
3	75.5	58.33	66.92
4	60	47.17	53.58
5	52.5	52	52.25

Table 5. Performance of the proposed system using MLP kernel in SVM

Level of Decomposition	Classification Accuracy (%)		
	Benign	Malignant	Average
1	41.83	67.83	54.83
2	42.17	63.83	53
3	44.67	59.67	52.17
4	41.67	58.17	49.92
5	39.83	68.33	54.08

It is observed from Table 1 to 5 that the proposed system achieves 91.83% maximum classification accuracy using polynomial kernel in SVM classifier. Among the five kernels, the performance of MLP kernel is worst than other kernels with only 54.83%. Figure 4 shows the performance of the proposed system at various level of decomposition. 'L' indicates the level of wavelet decomposition.

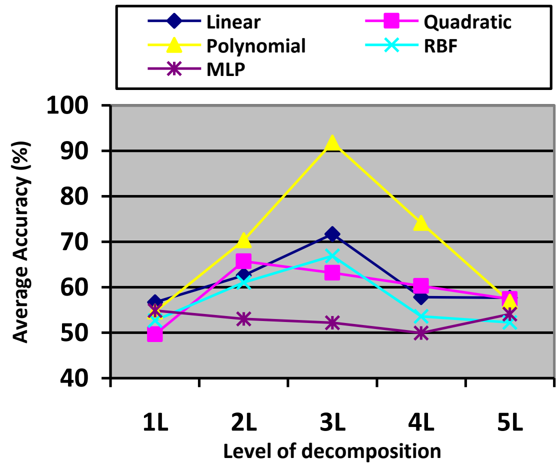


Fig 4: Performance of the system at various level of decomposition

It is observed from the figure 4 that the maximum classification is achieved by extracting DTMBWT sub-band energy features at lower level of decomposition. At higher level of decomposition, due to redundant data creation, the extracted features do not perform well for classification. Also it is noted that the performance of the proposed system is better at 3rd level of decomposition with polynomial kernel. Figure 5 shows graphical representation of the proposed microcalcification classification system performance in association with different types of kernels used in SVM classifier.

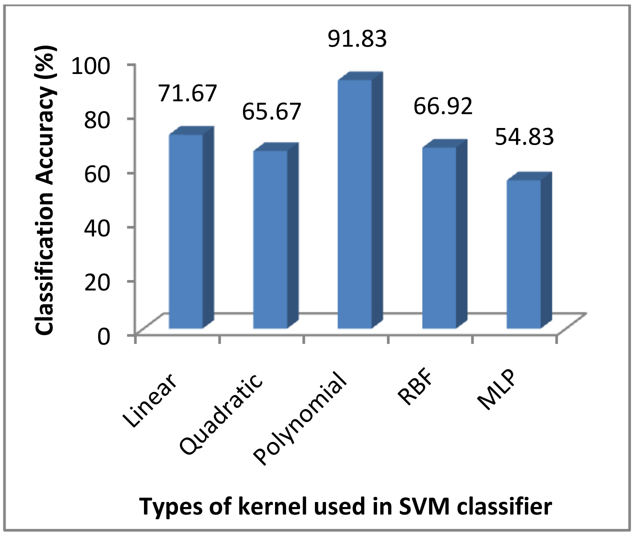


Fig 5: Types of kernel in SVM classifier vs. Classification accuracy

5. CONCLUSION

In this study, automated classification of microcalcification clusters in digital mammogram is presented. The proposed system is considered as a pattern recognition system where DTMBWT and SVM are used as feature extraction and classification techniques. The sub band energies obtained from DTMBWT at a predefined level of decomposition is given to train the SVM classifier. Five different types of kernels are used for performance evaluation. Experimental results show that the proposed system achieves maximum classification accuracy of 91.83% using polynomial kernel function. In future, the proposed system can be extended for the classification of mass abnormalities in digital mammograms.

6. REFERENCES

[1] Priya, D, and Muthulekshmi M, 2015. Review of cancer statistics in India. International Journal of Advances in Signal and Image Sciences, 1(1), 1-4.

[2] Lakshmi, S.RN.V.S, and Manoharan, C, 2011. Wavelet Analysis and Orthogonal Moments based Classification of Microcalcification in Digital Mammograms. Journal of Computer Science, 7(10), 1541-1544.

[3] Manoharan, C., and Lakshmi, S.RN.V.S. 2010. Classification of Micro Calcifications in Mammogram using Combined Feature Set with SVM. International Journal of Computer Applications. 11(10), 30–34.

[4] Lakshmi, S.RN.V.S, and Manoharan, C. 2011.An automated system for classification of micro calcification in mammogram based on Jacobi moments. International Journal of Computer Theory and Engineering, 3(3), 431-434.

[5] Balakumaran, T., and Vennila, ILA, 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), 121-125.

[6] Bose, J.S.C.K.R, Kumar, S and Karnan, M. 2012. Detection of Microcalcification in Mammograms using Computing Techniques. European Journal of Scientific Research, 86(1), 103-122.

[7] Bhanumathi, R., and Suresh, G.R. 2013. Detection of Microcalcification in Mammogram Images using Support Vector Machine based Classifier. ITSI Transactions on Electrical and Electronics Engineering, 1(2), 2320-8945.

[8] El-Naqa, I., Yang, Y., Wernick, M. N., Galatsanos, N. P., and Nishikawa, R. M. 2013. Improvement of Automated Detection Method for Clustered Microcalcification Based on Wavelet Transformation and Support Vector Machine. International Journal of Advanced Research in Artificial Intelligence, 2(4), 23-28.

[9] Jasmine, J.S.L., Govardhan, A., Baskaran.S. 2010. Classification of Microcalcification in Mammograms using Nonsubsampled Contourlet Transform and Neural Network. European Journal of Scientific Research, 46 (4). 531-539.

[10] Dehghani, S., and Dezfooli, M.A. 2011. Breast Cancer Diagnosis System Based on Contourlet Analysis and Support Vector Machine. World Applied Sciences Journal, 13(5), 1067-1076.

[11] Faye, I. 2010. Breast cancer diagnosis in digital mammogram using multiscalecurvelet transforms. Conference on Computerized Medical Imaging and Graphics, 269-276.

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

[13] Selesnick, I.W. 2004. The double-density dual-tree DWT. IEEE Transaction on Signal Processing, 52(5), 1304-1314.

[14] Chaux, C., Duval, L., and Pesquet, J. C. 2006. Image Analysis using a Dual-Tree M-Band Wavelet Transform. IEEE Transactions on Image Processing, 15(8), 2397-2412.

[15] El-Naqa, I., Yang, Y., Wernick, M.N., Galatsanos, N.P., and Nishikawa, R. 2002. Support vector machine learning for detection of microcalcifications in mammograms. IEEE International Symposium on Biomedical Imaging, 201-204.