

# Classification of Mammogram Using Wave Atom Transform and Support Vector Machine Classifier

A.Rajesh<sup>1</sup> and Dr. Mohan Ellappan<sup>2</sup>

<sup>1</sup>Research Scholar, Department of ECE, Shri JIT University, Jhunjhunu, Rajasthan, India

<sup>2</sup> Professor & Head, Department of ECE, Pallavan college of Engineering, Kanchipuram, Tamilnadu, India.

**Abstract-** The most common type of cancer seen women is breast cancer. To improve from this severe disease, monitoring and early detection must be provided and related safety measures must be taken as the first step. During diagnosis, some of the cases may be overlooked depending on fatigue and eyestrain, because determination of abnormalities is a repetitive procedure. In this study, a Computer Aided Diagnosis (CAD) system, using wave atom transform (WAT) algorithm and Support Vector Machine (SVM), is proposed to evaluate mammography images. During the process, the region of interest is defined before applying the method. The system includes a feature extraction approach based on WAT algorithm. In terms of classification, the process has two main stages as classification of normal-abnormal regions and malignant-benign ones.

**Keywords:** Mammography, CAD system, Wave atom transform, SVM

## 1. INTRODUCTION

Today breast cancer is the most frequent form of cancer in women above 40. The World Health Organization's international agency for research on cancer estimates that more than 150,000 women worldwide die of breast cancer each year. Wavelet decomposition and SVM is investigated for classifying the masses as benign or malignant is explained in [1]. Feature extraction and classification are the two stages which are performed by the decision making. Two novel texture features which provides smoothness and the architectural distortion is used to classify the abnormalities in the mammogram is explained in [2]. Three layer back propagation network is used to classify the masses.

A computer aided diagnostic system based on new texture shape feature coding is used to classify the masses is explained in [3]. The mass classification on mammogram presents a great challenge for design of computer aided diagnostic systems due to the complexity of mass background and its mammographic characteristics. SVM is also used for the classification.

Mainly transformation and Principal Component Analysis is also used to reduce the high dimensional data in [4]. The classifiers like SVM and Bayesian Networks are trained and tested using a k-fold cross-validation test method. Ranklet is used as a classification features for classifying the masses in [5]. The dimensionality of the classification problem is reduced by using a recursive feature elimination technique without affecting the classification performances.

Artificial neural network (ANN) is used to classify the masses, which performs benign-malignant classification on region of interest that contains mass is explained in [6]. Texture is the major characteristics of mass classification. The textural features used for characterizing the masses are mean, standard deviation, entropy, skewness, kurtosis and uniformity. This method is used to increase the effectiveness and the efficiency. Texture analysis based on curvelet transform for the classification of mammogram tissues is presented in [7]. The most discriminative texture features of regions of interest are extracted. Then, a nearest neighbor classifier based on Euclidian distance is constructed. The obtained results calculated using 5-fold cross validation.

The approach consists of two steps, detecting the abnormalities and then classifies the abnormalities into benign and malignant tumors.

Computerized schemes to classify masses by optimizing certain criteria to classify cases into one of mutually exclusive classes are developed in [8] and the elevated performance for linear separable problems but poor performance for nonlinear separable data were shown. This classification is based on linear discriminant analysis. The features required to distinguish the benign from malignant masses are abnormality dependent.

## 2. MATERIALS AND METHODS

### 2.1. Wave atom Transform

Wave atom transform is presented by Demanet in 2007. The transformation, obeying the parabolic scaling law, can be considered a variant of 2D wavelet packets. Wave atom transform have two very significant properties. First one is the ability to adapt to arbitrary local directions of a pattern. The second one is the ability to sparsely represent anisotropic patterns aligned with the axes. Wave atoms offer sharp frequency localization than other wave packets. It also has significant sparse expansion for oscillatory functions when compared with wavelets, curvelets and Gabor atoms.

The forms of wave packets, known as wavelets, Gabor, ridgelets, curvelets and wave atoms, are created using two parameters, which are  $\alpha$  and  $\beta$ . These variables symbolize decomposition and directional ability for all wave forms.  $\alpha$  and  $\beta$  values are 1/2 for wave atoms and Figure 1 shows wave packet's support in space and in frequency plane. Here,  $\alpha$  corresponds to the multiscale

structure of the transform and  $\beta$  corresponds to directional selectivity.

Actually, wave atoms are built from tensor products of 1D wave packets. One-dimensional wave packets can be represented as  $\psi_{j,m}^j(X)$ , where  $j, m \geq 0$ , and  $n \in \mathbb{Z}$ . frequency restrictions are  $\pm \omega_{j,m} = \pm \pi 2^j m$  with  $C_1 2^j \leq m \leq C_2 2^j$ . space restrictions is defined as

$$X_{j,n} = 2^j n \tag{1}$$

Two-dimensional wave atoms  $\varphi_\mu(X_1, X_2)$  are constructed with subscript  $\mu = (j, m, n)$ , where  $m = (m_1, m_2), n = (n_1, n_2)$ . 2D orthonormal basis is written as follows:

$$\varphi_\mu^+(X_1, X_2) = \psi_{m_1}^j(X_1 - 2^{-j} n_1) \psi_{m_2}^j(X_2 - 2^{-j} n_2) \tag{2}$$

$$\varphi_\mu^-(X_1, X_2) = \psi_{m_1}^j(X_1 - 2^{-j} n_1) \psi_{m_2}^j(X_2 - 2^{-j} n_2) \tag{3}$$

Where, H is Hilbert transform. The wave atom tight frame is formed by combination of (2) and (3).

$$\varphi_\mu^{(1)} = \frac{\varphi_\mu^+ + \varphi_\mu^-}{2}, \varphi_\mu^{(2)} = \frac{\varphi_\mu^+ - \varphi_\mu^-}{2} \tag{4}$$

**2.2 Support Vector Machine**

Support vector machines (SVMs) are a set of related supervised learning methods that analyze data and recognize patterns, used for classification and regression analysis (Rejani and Selvi, 2009). The standard SVM is a non-probabilistic binary linear classifier, i.e. it predicts, for each given input, which of two possible classes the input is a member of. 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) (Gorgel et al., 2009). SVM has an extra advantage of automatic model selection in the sense that both the optimal number and locations of the basic functions are automatically obtained during training. The performance of SVM largely depends on the kernel. SVM is essentially a linear learning machine. For the input training sample set

$$(x_i, y_i), i = 1..n, x \in R^n, y \in \{-1,+1\}$$

Let the classification hyper plane equation is to be  $(\omega, x) + b = 0$

$$\tag{5}$$

Thus the classification margin is  $2/|\omega|$ . To maximize the margin, that is to minimize  $|\omega|$ , the optimal hyper plane problem is transformed to quadratic programming problem as follows,

$$\begin{cases} \min \phi(\omega) = 1/2(\omega, \omega) \\ s, t, y_i((\omega.x) + b) \geq 1, i = 1,2,..l \end{cases} \tag{6}$$

After introduction of Lagrange multiplier, the dual problem is given by,

$$\begin{cases} \max Q(\alpha) = \sum_{i=1}^n \alpha_i - \frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n y_i y_j \alpha_i \alpha_j K(x_i, x_j) \\ s.t \sum_{i=1}^n y_i \alpha_i = 0, \alpha_i \geq 0, i = 1,2,..n \end{cases} \tag{7}$$

According to Kuhn-Tucker rules, the optimal solution must satisfy

$$\alpha_i (y_i((\omega.x_i) + b) - 1) = 0, i = 1,2,..n$$

That is to say if the option solution is

$$\alpha^* = (\alpha_1^*, \alpha_2^*, \dots, \alpha_n^*)^T, i = 1,2,..n$$

Then

$$w^* = \sum_{i=1}^n \alpha_i^* y_i x_i, b^* = y_i - \sum_{i=1}^n \alpha_i^* \langle x_i, x_j \rangle, j \in \{j | \alpha_j^* > 0\} \tag{8}$$

For every training sample point  $x_i$ , there is a corresponding multiplier. And the sample points that are corresponding to  $\alpha_i = 0$  don't contribute to solve the classification hyper plane while the other points that are corresponding to  $\alpha_i > 0$  do, so it is called support vectors. Hence the optimal hyper plane equation is given by,

$$\sum_{x_i \in SV} \alpha_i y_i (x_i . x_j) + b = 0 \tag{9}$$

The hard classifier is then,

$$y = \text{sgn} \left[ \sum_{x_i \in SV} \alpha_i y_i (x_i . x_j) + b \right] \tag{10}$$

For nonlinear situation, SVM constructs an optimal separating hyper plane in the high dimensional space by introducing kernel function  $K(x, y) = \phi(x) . \phi(y)$ , hence the nonlinear SVM is given by,

$$\begin{cases} \min \phi(\omega) = 1/2(\omega, \omega) \\ s, t, y_i((\omega.\phi(x_i)) + b) \geq 1, i = 1,2,..l \end{cases} \tag{11}$$

And its dual problem is given by,

$$\max L(\alpha) = \sum_{i=1}^l \alpha_i - \frac{1}{2} \sum_{i=1}^l \sum_{j=1}^l y_i y_j \alpha_i \alpha_j K(x_i, x_j) \tag{12}$$

$$s.t. \sum_{i=1}^n y_i \alpha_i = 0, 0 \leq \alpha_i \leq C, i = 1,2,..l$$

Thus the optimal hyper plane equation is determined by the solution to the optimal problem.

**3. EXPERIMENTAL SETUP**

The proposed method consists of two stages namely feature extraction stage and classification stage. The two stages are described below:

**3.1.Feature Extraction stage:**

Feature extraction is an essential pre-processing step to pattern recognition and machine learning problems. The main objective of this study is to distinguish between the types of mass as benign or malignant. MIAS database, a benchmark database used by many researchers is taken to evaluate this study. The original mammograms in MIAS are very big size (1024×1024 pixels). The whole

image consists of 50% of noise. In order to remove this unwanted noise, a cropping operation is done manually before the feature extraction by choosing the given center of abnormality as the center of the ROI. The size of ROI image is 256\*256. In the proposed system, the wave atom coefficients are used as features to classify the images. The original image is decomposed by using the WAT. All coefficients are considered as features for the classification process. This method is applied to all the training images and the feature vectors are stored in the database called as Feature DataBase. The energy of each sub-band of the image I is calculated by using the formula in

$$Energy_e = \frac{1}{RC} \sum_{i=1}^R \sum_{j=1}^C |I_e(i, j)|$$

where  $I_e(i, j)$ , the pixel value of the eth sub-band and R, C is width and height of the sub-band respectively. Figure 1 shows the original mammogram image (mdb134) and Fig. 1b shows the cropped image. The block diagram of the proposed method is shown in figure.2.

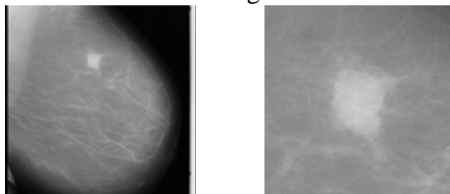


Fig.1 (a) Original Image (b) Extracted ROI

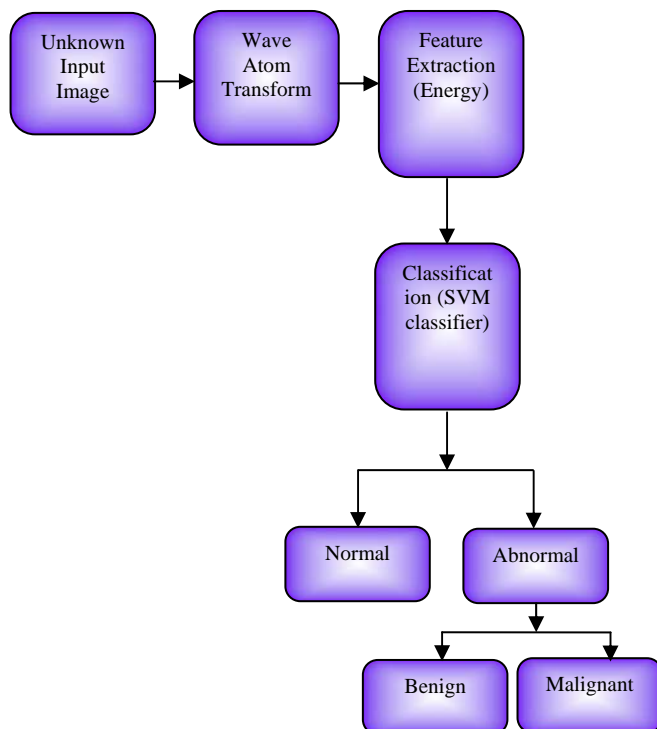


Fig.2. Block diagram of the proposed method

**3.2 Classification stage:**

Classification phase executes two phases. In the first one, the classifier is applied to classify mammograms into normal and abnormal cases. Then the mammogram is considered as abnormal if it contains tumor (mass). Finally, the abnormal mammogram is classified into malignant or

benign in the final stage. In this classification stage, SVM classifier in every phase is trained at specific number of training set in each category. Table 1 shows the number of training and testing images.

**Table 1.** Number of images used to train SVM Classifier

Images	Normal	Abnormal	Benign	Malignant
Training	47	37	25	13
Testing	70	56	37	19

**4. EXPERIMENTAL RESULTS**

The assessment is carried out in relation with two stages of mass classification system namely (i) normal/abnormal classification (ii) benign/malignant classification.

The performance of the proposed system is carried on 70 normal images and 37 mass images. Among the 37 abnormal images, there are 25 benign and 13 malignant images available. All the images are considered for the classification test. Table 2 shows the success rates of SVM method for the classification of images as normal and abnormal. Table 3 shows the success rates of SVM method for the classification of images as normal and abnormal.

At first, the dataset is classified as normal and abnormal. Then the accuracy is calculated for both normal and abnormal images. After that the same process is done for both benign and malignant images.

**Table2.** Success rates of SVM method for the classification of images as normal and abnormal.

Wave atom transform +SVM classifier(WAT+SVM)			
Scale	Accuracy	Sensitivity	Specificity
1	100	1	1
2	100	1	1
3	96	0.92	1
4	92	0.88	1

**Table3.** Success rates of SVM method for the classification of images as benign and malignant.

Wave atom transform +SVM classifier(WAT+SVM)			
Scale	Accuracy	Sensitivity	Specificity
1	100	1	1
2	90	0.90	0.90
3	78	0.76	0.70
4	66	0.63	0.63

**5. CONCLUSION**

This paper describes a CAD system for recognizing breast cancer in ROIs of digital mammograms. The study also investigates the presentation of the system with wave atom transform, and SVM method. The best success rates in this work were obtained with the coefficients at scale of 1, 2 and 3 by using SVM. These results demonstrate that wave atom transform and SVM are useful and prevailing methods to distinguish the mammographic images as normal, benign and malignant.

### REFERENCES:

- [1] Pelin Gorgel, Ahmet Sertbas, Niyazi Kilic, Osman N. Ucan and Onur Osman, "Mammographic Mass Classification Using Wavelet Based Support Vector Machine", *Journal of Electrical & Electronics Engineering*, Vol. No 9(1), 2009, pp 867-875
- [2] S. Baeg and N. Kehtarnavaz, "Texture Based Classification of Mass Abnormalities in Mammograms", *Proceedings of 13th IEEE Symposium on Computer-Based Medical Systems*, 2000
- [3] Yuan Chen and Chein I Chang, "New Texture Shape Feature Coding-Based Computer Aide Diagnostic Methods for Classification of Masses on Mammograms", *Proceedings of the 26th Annual International Conference of the IEEE*, September 2004
- [4] Maurice Samulski, Nico Karssemeijer, Peter Lucas and Perry Groot, "Classification of mammographic masses using support vector machines and Bayesian networks", *Proceedings of SPIE on Medical Imaging*, 2007
- [5] Matteo Masotti, "Exploring Rank lets Performances in Mammographic Mass Classification Using Recursive Feature Elimination", *Proceedings of the 16<sup>th</sup> IEEE Signal Processing Society Workshop on machine learning for signal processing*, 2006 pp 265-269
- [6] Mohammed J. Islam, Majid Ahmad and Maher A. Sid-Ahmed, "An Efficient Automatic Mass Classification Method in Digitized Mammograms Using Artificial Neural Network", *International Journal of Artificial Intelligence & Applications (IJAIA)*, Vol. No 1(3), July 2010
- [7] Mohamed Meselhy Eltoukhy and Ibrahima Faye, "Curvelet Based Feature Extraction Method for Breast Cancer Diagnosis in Digital Mammogram", *IEEE International Conference on Intelligent and Advanced Systems*, June 2010, pp 1-5.
- [8] H.D. Cheng, X.J. Shi, R. Min, L.M. Hu, X.P. Cai, H.N. Du (2006) "Approaches for automated detection and classification of masses in mammograms", *Pattern Recognition*, Vol. No 39, pp 646-668.
- [9] B. Sahiner, N. Petrick, H.P. Chan (2001) "Computer-aided characterization of mammographic masses: accuracy of mass segmentation and its effects on characterization", *IEEE Transactions on Medical Imaging*, Vol. No 20(12), pp 1275-1284
- [10] Jasmine, J.S.L., S. Baskaran and A. Govardhan, 2011. An automated Mass classification system in digital mammograms using contourlet transforms and support vector machine. *Int. J. Comp. Appl.*, 31: 54-54.
- [11] Rejani, Y.A. and S.T. Selvi, 2009. Early detection of breast cancer using SVM classifier technique. *Int. J. Comput. Sci. Eng.*, 1: 127-130.
- [12] Laurent Demanet and Lexing Ying, "Wave atoms and time up scaling of wave equations", *Journal Numerische Mathematik*, Vol.113, no.1, pp. 1-71, 2009