MULTI LAYER ARCHITECTURE FOR BREAST CANCER DIAGNOSIS

Suman Mishra
Research Scholar, Department of Electronics and Communication Engineering,
Sathyabama University, Chennai, India
emailssuman@gmail.com

Hariharan Ranganathan
Principal, Rajiv Gandhi College of Engineering,
Chennai, India
ranlal@yahoo.com

Abstract

Breast cancer is one of the dangerous cancers among women. Due to this, the rate of death increases every year. In order to ease the radiologist task and early detection of breast cancer, multilayer architecture based on dyadic wavelet transform and Gaussian Mixture Model (GMM) is proposed in this paper. The chain of processes includes; preprocessing, feature extraction and classification. The need for preprocessing is to remove the noise such as background and patient information in the digital mammograms that affects the classification accuracy of the proposed system. In the feature extraction stage, the textural properties of mammograms are extracted by dyadic wavelet transform in various scale of decomposition. In order to reduce the redundancy of dyadic wavelet coefficients, an efficient averaged sub-band concept is developed. Then the features energy and entropy are extracted from the averaged sub-band and fed into the classifier. The classification consists of series of components; the first layer classifies the given mammogram into either normal or abnormal, the second layer decides the type of abnormalities either mass or microcalcification and the final layer classifies the severity of the abnormality into benign or malignant using GMM classifier. The results show that the average classification accuracy obtained at each layer is more than 95% when using the Digital Database for Screening Mammography (DDSM) database.

***Keywords*: Wavelet Transform; Gaussian Mixer Model; Breast Cancer.**

1. Introduction

Among women, breast cancer is a key syndrome of morbidity and second leading cause of cancer related to death. Anatomical breast coordinate transform to facilitate computerized analysis of mammograms is explained [Brandt *et al.* (2011)]. Gaussian derivative features at four different scales up to the order of three are extracted. The features are selected by greedy sequential search algorithm, extracted inside the breast region of about 1000 random positions. The approximate K Nearest Neighbor (KNN) classifier is used for classification. Classification of clustered microcalcifications (MCs) in digital mammograms is explained [Liyang *et al.* (2005)] using the classifiers Support Vector Machine (SVM), kernel fisher discriminant and relevance vector machine. The features such as the number of cluster, mean effective volume, area, circularity, standard deviation of the effective volume and the mean area of MCs are extracted for each clusters.

Texture properties of the tissue surrounding microcalcification clusters on mammograms for breast cancer diagnosis are presented [Karahaliou *et al.* (2008)]. Initially, images are preprocessed using a wavelet based spatially adaptive method for contrast enhancement. Wavelet and texture features such as Gray level first order statistics features, Gray level co-occurrence matrices features are extracted from the preprocessed images. All extracted textural features were normalized to zero mean and unit standard deviation and subsequently used for classification. Probabilistic neural network is used as classifier.

Feature based detection and classification of MCs in digital mammograms is presented [Yang *et al.* (2007)]. It is based on immune algorithm and SVM classifier. The training set is compressed according to their intra-class and inter-class euclidean distances without losing any support vectors. Immune algorithm based microcalcifications features selector is provided to select an optimal feature subset that can construct the input vectors for the latter SVM training. Then, the compressed and optimized training samples are fed to SVM to make the optimal classification. New Particle Swarm Optimization for feature selection and classification of microcalcifications in mammograms is explained [Geetha and Thanushkodi (2008)]. The Spatial Gray Level Dependence Method (SGLDM) is used for feature extraction. The selected features are fed to a three layer Back Propagation Neural Network (BPNN) with new particle swarm optimization for classification.

A new approach for the detection of microcalcification clusters based on neural networks is presented [Cheran *et al.* (2004)]. The algorithm is composed of three modules; image pre-processing, feature extraction component and BPNN module. The first module uses several algorithms such as H-dome transformation, masking, binarization of grayscale images, connected components and labeling for classification. Feature reduction is carried out and the dimension reduction features are used as input to the BPNN. Ant colony optimization based classification and feature selection of microcalcifications in digital mammograms is explained [Xu *et al.* (2007)]. SGLDM is used for feature extraction and are fed to BPNN for classification.

A novel Computer Aided Diagnosis (CAD) method for detection and classification of MCs based on discrete wavelet transform (DWT) and Adaptive Neuro Fuzzy Inference System (ANFIS) is presented [Karnan et al (2006)]. DWT is used to extract the high frequency signal of the images, and thresholding technique with hysteresis approach is used to locate the suspicious MCs. ANFIS is used to adjust the parameters, making the CAD algorithm more adaptive and the suspicious MCs are classified using multilayer perception. Wavelet decomposition and SVM is investigated for classifying the masses as benign or malignant is explained [Gorgel *et al.* (2009)]. Decision making is performed in the feature extraction by computing the wavelet coefficients and classification using the classifier trained on the extracted features.

A two level hierarchical method is used to classify the masses where Bayesian classifier exists in each level is explained [Viton *et al.* (1996)]. In first level the speculated masses are discriminated from non-speculated masses. In the second level masses with fuzzy edges are separated from well defined edges among the non-speculated edges. Bayesian network uses a probabilistic approach to determine the class conditional probability density functions for background and tumor in breast cancer detection application. Linear discriminant analysis in mixture with stepwise feature selection in [Sahiner at al. (2001)] is trained and tested on morphological features extracted using the machine segmentation and radiologist segmentation. Texture analysis based on curvelet transform for the classification of mammogram tissues is presented [Eltoukhy and Faye (2010)]. 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.

Artificial neural network is used to classify the masses, which performs benign-malignant classification on region of interest that contains mass is explained [Islam et al. (2010)]. 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. A CAD system based on new texture shape feature coding is used to classify the masses is explained [Mencattini *et al.* (2010)]. 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.

In this paper, a novel multilayer architecture for breast cancer diagnosis in a single step is proposed based on dyadic wavelet transform and GMM classifier. The organization of rest of the paper is as follows. In Section 2, the methods and materials involved in the proposed multilayer architecture for breast cancer diagnosis is presented. The proposed feature extraction method is described in Section 3. The experimental results are discussed in section 4. Finally, conclusion is discussed in Section 5.

2. Materials and methods

In this section, a detailed study on dyadic wavelet transform and gaussian mixture model is carried out. The digital mammogram images used in this work is also discussed.

2.1. *Dyadic Wavelet Transform*

The discrete wavelet transform is obtained by iterating an *M* channel multirate filter bank on its low-pass output. The wavelet transform with symmetric basis functions has recently emerged as the next generation standard for various image application areas. Wavelet decomposition arises from iteration of the low-pass filtering and decimation steps of a multirate filter bank. For a true wavelet, one iterates on the low-pass output only, whereas for wavelet packet decomposition, one may iterate on any output. A finite number of iterations will lead to a discrete time multiresolutional analysis with low pass frequency response $\Pi_{k=1}^{n} H_0\left(\frac{\omega}{2^k}\right)$. If the low pass filter h_0 satisfies the orthonormality constraint and "one vanishing moment" (approximation of order 1): $\sum_k h_0[k] = \frac{1}{\sqrt{2}}$, then the infinite product $\lim_{n\to\infty} \Pi_{k=1}^{n} H_0\left(\frac{\omega}{2^k}\right)$ converges to a continuous time function $\hat{\phi}(\omega)$, and inverse Fourier transform $\phi(t)$ is called the scaling function. The scaling function $\phi(t)$ and the wavelet $\psi(t)$ are solution of the dilation equations in Eq. (1).

$$\begin{aligned}\phi(t) &= \sum_k h_0(k)\phi(2t-k),\\ \psi(t) &= \sum_k h_1(k)\phi(2t-k)\end{aligned} \tag{1}$$

And it is orthogonal to its integer translates (under mild additional assumptions). If the filter $h_0(n)$ is FIR, then $\phi(t)$ *and* $\psi(t)$ have compact support. The set of dilates and translates

$$\left\{\psi\left(2^k t-\ell\right)\right\}_{k,\ell\in Z} \tag{2}$$

which forms a tight frame (and in most cases an orthonormal basis) for $L^2(R)$ [Steffenet *et al.* (1993)]. The span of integer translates of the scaling function $\phi(t)$ is the "low-pass" space V_0, the set of scale limited signals [Gopinath *et al.* (1994)]. Any continuous-time function $f(t)$ in V_0 can be expanded as a linear combination

$$f(t) = \sum_n v_n^{(0)} \phi(t-n). \tag{3}$$

The superscript (0) denotes an expansion "at scale level 0." $f(t)$ is completely described by the sequence $\left\{v_n^{(0)}\right\}$. Given such a sequence, its coarse approximation (component in V_1) is computed with low-pass filter of the wavelet filter bank:

$$v_n^{(1)} = \left(\left(v^{(0)} * h_0\right)\downarrow 2\right)[n]. \tag{4}$$

Analogously, the details $\omega_n^{(1)}$ in V_0 V_1 are computed with the high-pass filter $h_1(n)$. Hence, a discrete sequence v_n to be the coefficients of a signal $f(t)$ at some fixed scale, the discrete wavelet transform of v_n will decompose the underlying signal f into a coarse-scale component and detail at several intermediate scales. This decomposition matches multiresolutional models of human and computer vision [Mallat and Stephane (1989)]. In the proposed approach, the dyadic wavelet transform is used to represent the mammograms in multi-scale.

2.2. *Gaussian Mixture Model*

Gaussian Mixture Model (GMM) is commonly used in pattern recognition and machine learning algorithms. This classifier is able to approximate the distribution of the patterns representing the characteristics of a texture in an image. During training, the induction algorithm estimates the mixture of gaussian models that best approximates the distribution of the given values. Formally, a texture pattern is described by a mixture of M gaussian models $\left|\Gamma = \{\gamma_1,....,\gamma_M\}\right.$: the mixture density is a weighted sum of the M component densities. Given an input feature vector $\vec{\chi}$, the conditional probability is computed from the mixture as follows:

$$p\left(\vec{\chi}\,\middle|\,\Gamma\right) = \sum_{m=1}^{M} c_m \cdot \gamma_m(\vec{\chi}) \tag{5}$$

where c_m are the mixture weights and $\gamma_m(\vec{\chi})$ an N-variate Gaussian function. The dimensionality of the gaussian function (N) coincides with the dimensionality of the feature vector $\vec{\chi}$ while the M models, and relative weights, are estimated from the training data using a special case of the expectation-maximization (EM) algorithm [Reynolds and Rose (1995)].

A set of S textures $\{s_1,...,s_S\}$ is represented by S gaussian mixture models $\{\Gamma_1,...,\Gamma_S\}$ A given observation sequence $X = \{x_1,..., x_T\}$ is tested by finding the pattern in a texture which has maximum a posteriori probability. By applying Bayes rule and using the logarithm [Reynolds and Rose (1995)], the probability can be computed as:

$$\hat{S} = \arg\max_{1\le s\le S} \sum \log p\left(\vec{x}_t\,\middle|\,\Gamma_s\right) \tag{6}$$

The GMM can assume different forms depending on the choice of covariance matrix used in the estimation of the N-variate gaussian functions. In this proposed approach GMM is used as a classifier for breast cancer diagnosis.

2.3. *DDSM database*

In this study, a set of mammograms are selected from the DDSM database. It is maintained at the University of South Florida [Heath and Bowyer (1998)]. In total, 270 mammogram images are selected and used to evaluate the proposed multilayer architecture design for breast cancer diagnosis. To train the classifier, 60% of cases in each category are used. The remaining 40% of cases are tested by the classifier. In order to design a best diagnosis system for breast cancer, mammograms with different degree of subtlety and breast densities are chosen. The number of images selected for the analysis is given in Table 1.

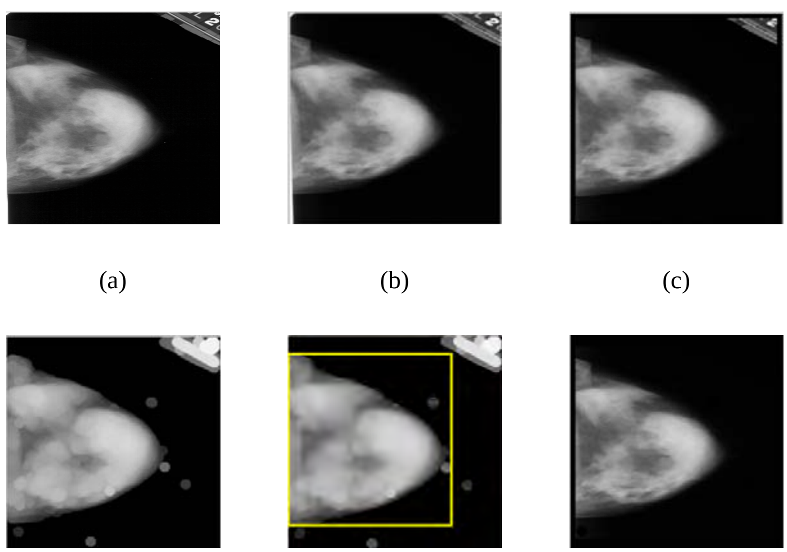
Table 1 Number of images selected for this study

Cases	#of mammogram images
Normal	70
Mass-Benign	50
Mass-Malignant	50
Microcalcification-Benign	50
Microcalcification-Malignant	50

3. Features

Dyadic wavelet transform is used to extract the features where the textural properties can be viewed in multi scale. The transformed image consists of low and high frequency components. For n-level decomposition, dyadic transform construct a low frequency sub-band and 3*n high frequency sub-bands. In order to reduce the redundancy of features space i.e. high dimensional dyadic wavelet coefficients, an efficient averaged sub-band concept is introduced. Before extracting the proposed averaged sub-band features, the image must be free from noises such as background information and labeling.

In preprocessing, bicubic interpolation is used at first to reduce the spatial resolution of the image. Then the border correction is used to remove the non x-ray film region by replacing the gray intensities to zero along the four sides of the images. Finally, the labeling such as patient information is removed by applying morphological dilation. The output image at each step of preprocessing is shown in Figure 1.



(a) (b) (c)

(d) (e) (f)

Figure 1 Preprocessing stages (a) Original image (b) interpolated image (c) border corrected image (d) Morphological dilated image (e) biggest region of the image (f) breast region.

The flow chart of the proposed system is shown in Figure 2.

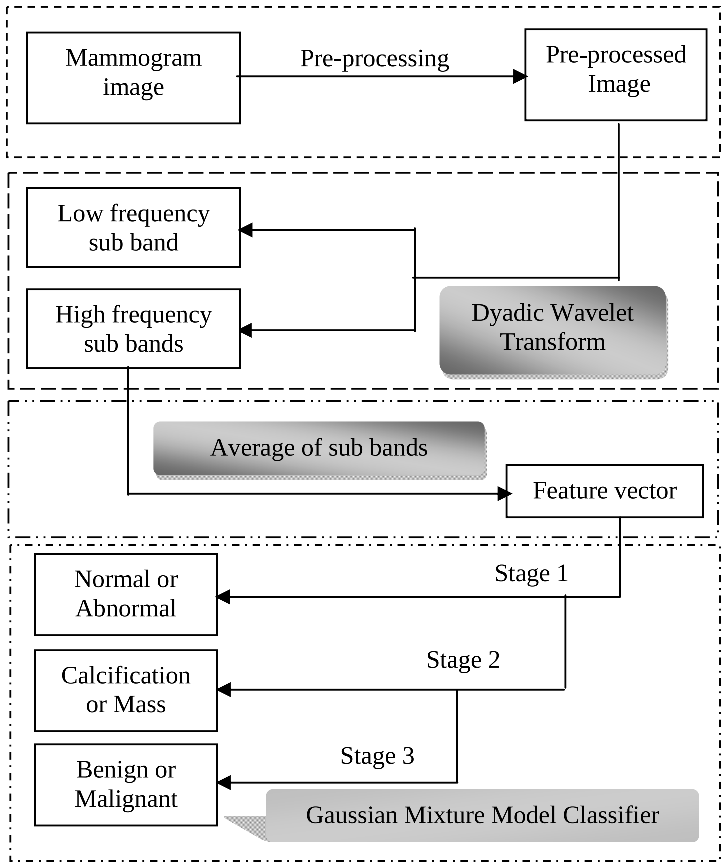
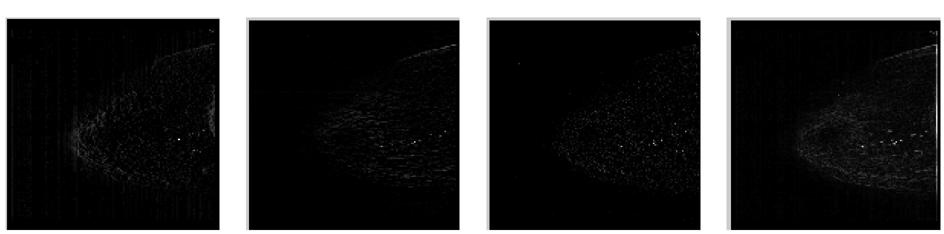


Figure 2 Flowchart of the proposed multilayer architecture for breast cancer diagnosis system.

The next step after preprocessing is to decompose the preprocessed image by dyadic wavelet transform. As the detailed information's such as edges and singularities are provided by the high frequency sub-bands, the averaged sub-band concept is applied only to those sub-bands. The averaged sub-band coefficient is computed by using the following Eq. (7)

$$Avg_{i,j} = \frac{1}{3}\sum_{i=1}^{R}\sum_{j=1}^{C} H^{1}_{i,j} + H^{2}_{i,j} + H^{3}_{i,j} \quad (7)$$

Where $H^{k}_{i,j}$ represents the k^{th} high frequency sub-band of a particular decomposition level and (*i*,*j*) is the coefficient location. *R* and *C* are the width and height of the sub-band. Figure 3 shows the averaged high frequency coarsest sub-bands obtained at two-level decomposition. From the averaged sub-band, the features such as energy and entropy are computed and their performance is evaluated by GMM classifier.

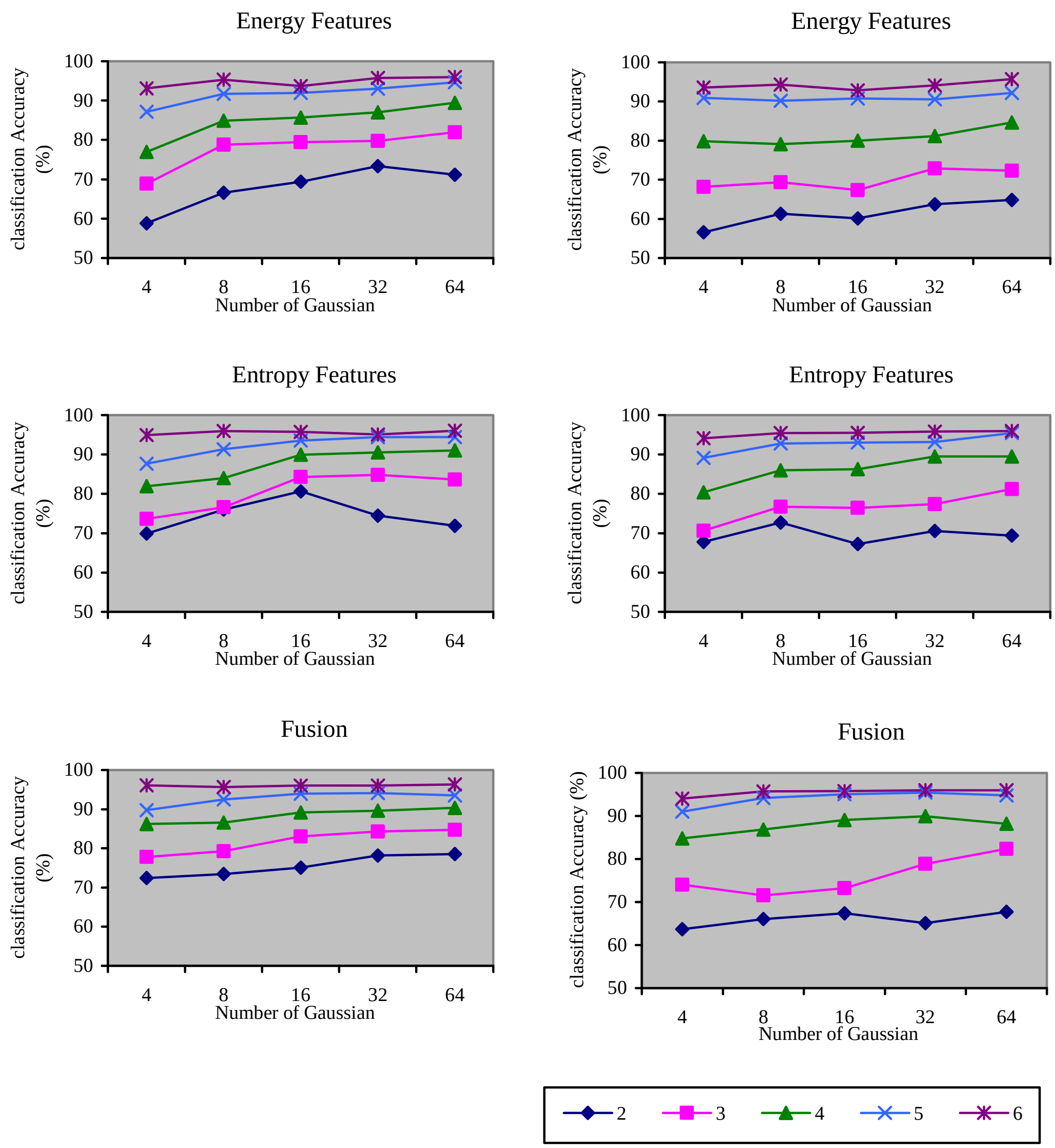


(a) (b) (c) (d)

Figure 3 (a) (b) and (c) high frequency sub-bands of two-level decomposition (d) averaged sub-band.

4. Experiments and Results

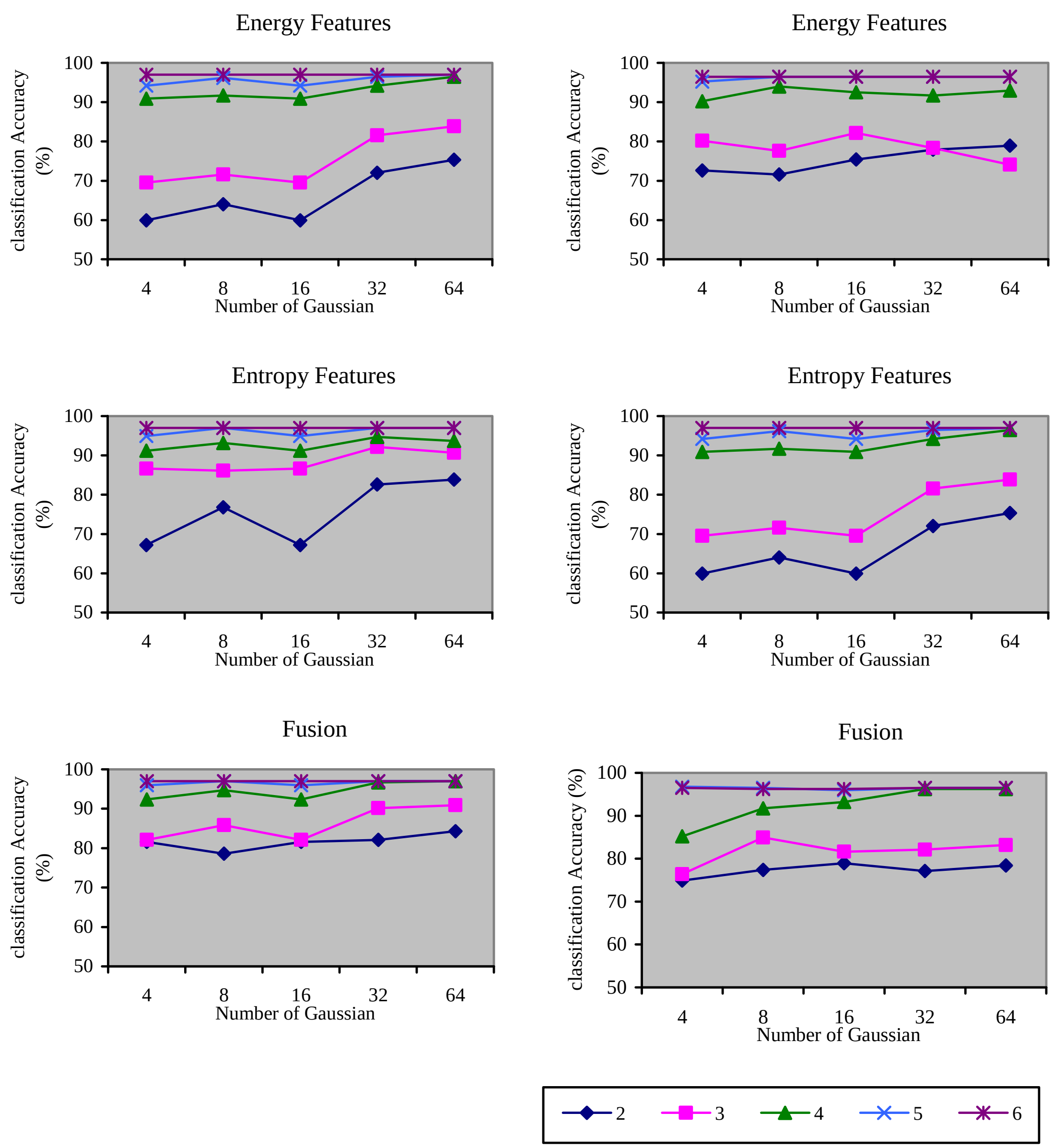
In this study, breast cancer is diagnosed by designing three-layer architecture using Dyadic wavelet transform. In each layer, the classification steps are carried out by GMM classifier. The classifier in the first layer helps to determine whether the given mammogram is affected by breast cancer or not. The two most important mammogram abnormalities are microcalcification and mass. If the output of the first layer classifier is abnormal then the second layer classifier is triggered and tested for abnormalities of the input mammogram image. The most significant classifier is in the third layer that determines the abnormal severity into either benign or malignant. To evaluate the complete system performance, the classification accuracy at each scale is computed. Figure 4 shows the classification accuracy obtained by the first two layers of the proposed approach.



(a) (b)

Figure 4 (a) Performance of Layer-1 classifier (b) performance of Layer-2 classifier.

The average classification accuracy of the first layer and second layer of the proposed approach is shown in Figure 4 (a) and (b). It is inferred that, the classification accuracy increase as the level of decomposition increases and also number of gaussian increases. This is also true for second and third layer of classification due to the reason that the dyadic wavelet transform produces more detailed coefficients at higher level of decomposition and the discriminating power of GMM classifier is increased by increasing the number of gaussian. Figure 5 (a) and (b) shows the average classification accuracy of abnormal severity for microcalcification and mass respectively.



(a) (b)

Figure 5 (a) Performance of Layer-3 classifier of microcalcification system (b) performance of stage 3 classifier of mass system.

As the dyadic wavelet transform is multi scale algorithm in nature, the proposed system is tested by extracting the features at various decomposition levels. In this study, up to sixth decomposition level is used. The performance analysis shows that the proposed multilayer approach achieves maximum average classification accuracy at sixth level dyadic decomposition and number of gaussian used is 64. Beyond the sixth level of decomposition does not affect the performance of the proposed approach and hence stopped at sixth level. Table 2 shows the number of high frequency sub-bands available at each scale, the number of averaged sub-bands and number of extracted features.

Table 2 Number of extracted features and sub-bands at each scale of decomposition.

Level of decomposition	#of high frequency sub-bands	# of averaged sub-bands	# of energy features	# of entropy features	# of fused features
2	6	2	2	2	4
3	9	3	3	3	6
4	12	4	4	4	8
5	15	5	5	5	10
6	18	6	6	6	12

5. Conclusion

In this paper, a novel multilayer architecture for the classification of digital mammogram based on dyadic wavelet transformation and GMM is proposed. Initially, mammogram images are decomposed by dyadic wavelet transform with various level of decomposition from 2 to 6 after preprocessing. The high dimensional feature space of high frequency sub-bands are reduced by the averaged sub-band concept and the energy and entropy features are extracted. In order to increase the classifier accuracy to detect breast cancer, three layer architecture is designed by using GMM classifier at each layer. The sequences of breast cancer classification at each layer are; normal/abnormal, microcalcification/mass and benign/malignant. The performance of the proposed multilayer architecture shows significant improvement in diagnosing breast cancer of over 95%.

References

[1] Brandt,S.S.; Karemore,G.; Karssemeijer,N.; Nielsen,M. (2011): An anatomically oriented breast coordinate system for mammogram analysis." , IEEE Transactions on Medical Imaging ,**30**(10), pp.1841-1851.

[2] Liyang,W.; Yang,Y.; Nishikawa,R.M.; Jiang,Y. (2005): A study on several machine-learning methods for classification of malignant and benign clustered microcalcifications, IEEE Transactions on Medical Imaging, **24**(3), pp.371-380.

[3] Karahaliou,A.N.; Boniatis,I.S.; Skiadopoulos,S.G.; Sakellaropoulos,F.N.; Arikidis,N.S.; Likaki,E.A.; Panayiotakis,G.S.; Costaridou.L.I. (2008): Breast cancer diagnosis: analyzing texture of tissue surrounding microcalcifications, IEEE Transactions on Information Technology in Biomedicine, **12**(6), pp.731-738.

[4] Yang,T.; Guo,S.; Wu,X. (2007): An Approach Based on Immune Algorithm and SVM for Detection and Classification of Microcalcifications, The 1st International Conference on IEEE Bioinformatics and Biomedical Engineering, pp. 588-591.

[5] [Geetha,K.; Thanushkodi,K. (2008): New particle swarm optimization for feature selection and classification of microcalcifications in mammograms, International Conference on IEEE Signal Processing, Communications and Networking,. ICSCN'08, pp. 458-463.

[6] Cheran,S.C.; Cataldo,R.; Cerello,P.; De Carlo,F; Fauci,F.; Fomi,G.; Golosio,B.; Lauria, ,A.; Torres,L.E.; De Mitri,I.; Masala,G.; Raso,G.; Retico,A; Tata,A. (2004): Detection and classification of microcalcifications clusters in digitized mammograms, IEEE Nuclear Science Symposium Conference Record,**7**, pp. 4136-4140.

[7] Xu,W.; Li,L.; Zou,S. (2007): Detection and Classification of Microcalcifications Based on DWT and ANFIS., The 1st International Conference on Bioinformatics and Biomedical Engineering,. ICBBE 2007, pp. 547-550.

[8] Karnan,M.; Thangavel,K.; Sivakuar,R.;Geetha,K (2006): Ant Colony Optimization for Feature selection and Classification of Microcalcifications in Digital Mammograms, International Conference on IEEE Advanced Computing and Communications ADCOM, pp. 298-303.

[9] Gorgel,P.; Sertbas,.; Kilic,N, Ucan,O.N.; Osman,O.(2009): Mammographic Mass Classification Using Wavelet Based Support Vector Machine, Journal of Electrical & Electronics Engineering, **9**(1), pp.867-875.

[10] Chenl,Y.; Chang,C.I.(2004): 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, **1**, pp.1275 – 1278.

[11] Mohammed,J.I,; Ahmad.M.; Sid-Ahmed,M.A.(2010): An Efficient Automatic Mass Classification Method in Digitized Mammograms Using Artificial Neural Network, International Journal of Artificial Intelligence & Applications (IJAIA), July **, 1**(3).

[12] Eltoukhy,M.M.; Faye.I.(2010): Curvelet Based Feature Extraction Method for Breast Cancer Diagnosis in Digital Mammogram, IEEE International Conference on Intelligent and Advanced Systems, June, pp 1-5.

[13] Sahiner,B, Petrick,N, Chan,H.P.(2001): Computer-aided characterization of mammographic masses: accuracy of mass segmentation and its effects on characterization, IEEE Transactions on Medical Imaging, **20**(12), pp.1275–1284.

[14] Viton, J.L.; Rasigni,M.; Rasigni,G.; Liebaria,A. (1996) : Method for characterizing masses in digital mammograms, Journal of Optical Engineering, **35**(12), pp.3453–3459.

[15] Gopinath,R.A., Odegard,J.E.; Burrus,C.S.(1994): Optimal wavelet representation of signals and the wavelet sampling theorem, IEEE Transactions on Circuits and Systems II: Analog and Digital Signal Processing, **41** (4), pp.262-277.

[16] Mallat,S.G.(1989): A theory for multiresolution signal decomposition: the wavelet representation, IEEE Transactions on Pattern Analysis and Machine Intelligence, **11**(7) ,pp.674-693.

[17] Steffen,P.; Heller,P.N.; Gopinath,R.A.; Burrus,C.S.(1993): Theory of regular M-band wavelet bases, IEEE Transactions on Signal Processing, **41**(12),pp.3497-3511.

[18] Reynolds,D.A.; Rose,R.C (1995): Robust text-independent speaker identification using gaussian mixture speaker models, IEEE Transactions on Speech and audio processing, **3**(1), 1995, pp.72–83.

[19] Heath,M.; Bowyer,K.W.; D.K.,et al., Current status of the digital database for screening mammography , in: Karssemeijer,N.; Thijssen.M.; Hendriks,J.L,vanErning (Eds.) ,Digital Mammograpy, Kluwer Academic Publishers, Boston, MA, USA, 1998, pp.457–460, (Online).Available from: http://marathon.csee.usf.edu.