A Wavelet-Wavelet Based Processing Approach for Microcalcifications Detection in Mammograms

Salim Lahmiri
University of Quebec at Montreal/Department of Computer Science, Montreal, Canada
Email: lahmiri.salim@courrier.uqam.ca

***Abstract*—A new methodology for the detection of microcalcifications (MCs) in mammograms is presented. Since MCs correspond to the high frequency components of the mammograms, a further multiresolution analysis is applied to these components. In particular, we seek to capture better high frequency features of a mammogram by performing a second analysis only to its high frequency components. For instance, in the first step, discrete wavelet transform (DWT) is applied to the mammograms and HH image is extracted. In the second step, the DWT is applied to the previous HH image. Then six statistical features are computed. Finally, principal component analysis is employed to reduce the number of features. The *k*-Nearest Neighborhood (*k*-NN) algorithm is employed for the classification task using cross-validation technique. A similar approach is adopted with use of the discrete Fourier transform (FT). The experimental results show strong evidence of the proposed methodology for MCs detection in digital mammograms.**

***Index Terms*—MCs, discrete Fourier transform, discrete wavelet transform, *k*-NN.**

I. INTRODUCTION

In the Western World, breast cancer is the most common form of cancer in the female population and early detection of it is one of the most important factors helping to recover from the disease. Indeed, the causes of breast cancer remain unknown, and early detection is the key to reduce the death rate [1]. This can be achieved through mammography screening programs performed by computer-assisted diagnosis (CAD) systems to eliminate the operator dependency and improve the diagnostic accuracy. Therefore, it is a valuable and beneficial tool for breast cancer detection and classification. Symptoms of breast cancer include clustered MCs, speculated lesions, circumscribed masses, ill defined masses, and architectural distortions [2]. The MCs can be divided into benign (passive) and malignant (active). Pathologists consider benign cell not dangerous. But malignant (cancer) cell can become larger and affect the around area and causes dead. The detection of MCs has received a large attention in the literature. In particular, many multi-resolution techniques have been employed to process the mammograms in order to detect the clustered MCs. For example, the discrete wavelet transform (DWT) was used to characterize digital mammograms and a specific threshold based on standard deviation of the decomposed images coefficients is employed to select a subset of representative features [3]. The employed wavelet bases were the Haar, Daubechies4, Biorthogonal2.4, Coiflets2 and Symlets2. Then distance metrics are used to measure the similarity between unknown mammogram and classes signatures. The validation and the testing process are respectively performed with 75% and 25% of the images. The experiments performed show that successful accuracy varies from 62.50% to 100% for the classification problem between normal, MCs, spiculated, and circumscribed areas. The authors conclude that generally the distance metrics used for the classification purpose present similar results. In addition, the Haar wavelet achieves better results in all the tested classes. Furthermore, the selection of features by threshold helps reducing the number features used to form the signatures of classes. In [4], different types of wavelet packets including Daubechies, Symlet, Coiflet, and Biorthogonal were employed, all at two-level expansion. Then, 18 features were computed from the high frequency sub-band, and principal component analysis (PCA) was performed to eliminate the features that contribute less than 3% to the total variation of the data set. As a result, only 7 features formed the inputs set to be fed to the backpropagation neural networks for the classification purpose. The performance was evaluated in terms of the receiver operating characteristic (ROC). It shows that best performance was achieved by the Coiflet wavelet with areas under ROC curve ranging from 0.90 to 0.97. A multiresolution approach to automated classification of mammograms using Gabor filters was proposed in [5]. First, Gabor filters of different frequencies and orientations have been applied on mammograms to produce filtered outputs. Then, for each filtered output, the mean and the standard deviation of the coefficient magnitude are used as image features for classification. Second, *t*-test statistic is performed on each feature to select significant features. Finally, the selected features are applied in mammogram classification using *k*-NN algorithm. The obtained classification rate with 14 selected features is 80% significantly higher than 75% performance obtained with 48 features. The authors concluded that Gabor filter is able to extract textural

doi:10.4304/jait.3.3.162-167

patterns of mammograms, and that statistical *t*-test and its *p*-value can be used to reduce feature space and speed up the classification process while providing good classification rates. More recently, dual-tree complex wavelet, curvelet and contourlet were also successfully applied to mammograms to extract features. For instance, the authors in [6] used curvelet transform and extracted a set of coefficient from each level of decomposition. These coefficients are used to classify the mammogram images into normal or cancer classes. The Euclidean distance was used to design the classifier. At scale 3, a 100% successful classification rate was obtained in benign class when a set 10%, 20%, 60%, 70%, and 80% of coefficients are used. The malignant class achieved 100% classification rate at 30%, 40%, 50%, 70%, 80%, and 90% percentages of coefficients used. The authors in [7] used dual-tree complex wavelet transform (DT-CWT) as feature extraction technique and support vector machine (SVM) to classify two classes of MCs: benign and malignant. Using this methodology, the experimental result achieved 88.64% classification accuracy. In [8], the contourlet transform was employed for feature extraction. Using the contourlet coefficients, the classification is performed based on successive enhancement learning (SEL) weighted SVM, support vector-based fuzzy neural network (SVFNN), and kernel SVM. The obtained correct classification accuracies are 96.6%, 91.5% and 82.1% respectively.

In an attempt to capture directional features in mammograms, the authors in [9] combined the discrete wavelet transform and the Gabor filter. In particular, the two-dimensional discrete wavelet transform is employed to process the mammogram and obtain its high-high (HH) frequency sub-band image. Then, a Gabor filter bank is applied to the latter at different frequencies and spatial orientations to obtain new Gabor images from which the average and standard deviation are computed. Finally, these statistics are fed to a support vector machine with polynomial kernel to classify normal versus cancer mammograms. The obtained classification results using ten cross-validation technique showed the superiority of their approach to the standard approach, which only uses the discrete wavelet transform to extract features from mammograms. Therefore, the authors concluded that high frequency directional features are important to improve the correct classification rate of MCs in mammograms.

Although the curvelet, contourlet, DT-CWT, and Gabor transforms were employed for features extraction, the discrete wavelet transform (DWT) remains the most employed tool for processing mammograms [4]since it is able to perform signal analysis at different time and frequency scales. In addition, DWT offers a low computational cost in comparison with cuvelet, contourlet, DT-CWT and Gabor transform. The DWT decomposes an image into four orthogonal sub-bands: low-low (LL), high-low (HL), low-high (LH), and high-high (HH). The sub-bands (sub-images) LL, HL, LH, and HH contain respectively approximation, horizontal, vertical, and diagonal information. In the next octave, the LL sub-band is further decomposed in the same manner.

The purpose of this paper is to propose a simple methodology for features extraction from mammograms based on a further analysis of high frequency components of the mammograms. We rely on high frequency components of the mammograms since MCs are usually found in dense biological tissue, which corresponds to high frequencies in the frequency domain of the image [4][2][10]. In particular, we aim to apply DWT to HH sub-bands to extract further accurate high frequency information. In particular, DWT is applied to the mammogram and its high-high (HH) image is obtained. Then, a second DWT is applied to the HH image obtained in the previous step. The purpose of applying a second DWT uniquely to HH image is to accurately capture high frequency information from high frequency image. A similar approach with use of discrete Fourier Transform (FT) is proposed. For instance, FT is applied to the mammogram to obtain a Fourier image. Then, a second FT is applied to the Fourier image obtained in the previous step. As in the first approach, the purpose is to accurately capture high frequency information from high frequency image. Therefore, a second FT is applied to the Fourier image obtained in the previous step. We aim to adopt Fourier transform in the second approach to check the effectiveness of high frequency information in the detection of MCs. To the best of our knowledge, no such methodologies have been adopted in the literature to extract better high frequency features from mammograms. Therefore, we suggest examining the effectiveness of a further analysis of high frequency components of a mammogram to better characterize MCs.

The paper is organized as follows: The methodology is described in Section II. The experimental results are presented in Section III. Finally, conclusions are given in Section IV.

II. METHODOLOGY

A. Discrete Fourier Transform

Fourier transform (FT) is an effective tool for signal analysis. For instance, FT is more generic than power spectral density (PSD) and phase spectrum approaches and provides better recognition than the PSD approach [11]. In addition, the DFT is efficient in detecting periodicity in the frequency domain [12] and Fourier descriptors are greatly immune to the noise [13]. Moreover, DFT is translation invariant with respect to the spectrum [14]. The two-dimensional discrete Fourier Transform is defined as:

$$F(u,v)=\frac{1}{MN}\sum_{m=0}^{M-1}\sum_{n=0}^{N-1} f(m,n)\exp\left[-j2\pi\left(\frac{mu}{M}+\frac{nv}{N}\right)\right] \quad (1)$$

where $f(m,n)$ is an image with $M\times N$, $j=\sqrt{-1}$, $u = 0,1,\ldots,M-1$; $v = 0,1,\ldots,N-1$. The values of $F(u,v)$ are the Fourier coefficients of the expansion of the exponential into sine and cosine with the variables u and v. The features used to detect cancer images are extracted from the frequency image $F(u,v)$.

B. Discrete Wavelet Transform

Wavelet transform belongs to the multiresolution transformation, performing the decomposition of the signal on different levels. Then, the wavelet functions have good localization abilities in both time and frequency, enabling good representation of the local features of the patterns [13]. As mentioned in the introduction, the discrete wavelet transform (DWT) decomposes an image into several sub-bands according to a recursive process (see Figure. 1). These include LH1, HL1 and HH1 which represent detail images and LL1 which corresponds to the approximation image. The approximation and detail images are then decomposed into second-level approximation and detail images, and the process is repeated to achieve the desired level of the multi-resolution analysis. The obtained coefficients values for the approximation and detail sub-band images are useful features for texture categorization [15][16]. To obtain the set of features that characterize a given texture image, the 2D-DWT wavelet transform is used to find its spectral components. This allows transforming each texture image into a local spatial/frequency representation by convolving the image with a bank of filters. Then, the image features are extracted from the obtained 2D-DWT representation.

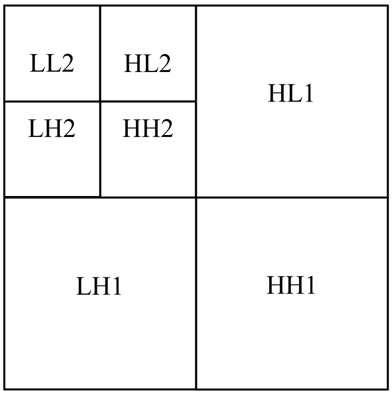

Figure 1: Two level 2D-DWT decomposition of an image.

A 1D-DWT is defined as follows:

$$f(x) = \sum_{i,j} c_{i,j} \psi_{i,j}(x) \qquad (2)$$

where $\psi_{i,j}(x)$ are the wavelet functions and $c_{i,j}$ are the DWT coefficients of $f(x)$. They are defined by:

$$c_{i,j} = \int_{-\infty}^{+\infty} f(x) \psi_{i,j}(x) \qquad (3)$$

A mother wavelet $\psi(x)$ is used to generate the wavelet basis functions by using translation and dilation operations:

$$\psi_{i,j}(x) = 2^{-i/2} \psi\left(2^{-i} x - j\right) \qquad (4)$$

where j and i are respectively the translation and dilation parameters. The one-dimensional wavelet decomposition can be extended to two-dimensional objects by separating the row and column decompositions [10][17]. For instance, the 2D-DWT is achieved by alternating row and column filtering in each level with iteration from the LL (low-pass/low-pass) subband as shown in Figure 1. For instance, the 2-D wavelet analysis process (See Figure 2) consists of filtering and down-sampling horizontally using 1-D low-pass filter to each row in the image F(x, y) to produce the coefficient matrices $F_L(x, y)$ and $F_H(x, y)$. Then, vertical filtering and down-sampling are performed using the low-pass and high-pass filters L and H to each column in $F_L(x, y)$ and $F_H(x, y)$ to produce four sub-images $F_{LL}(x, y)$, $F_{LH}(x, y)$, $F_{HL}(x, y)$ and $F_{HH}(x, y)$ for one level of decomposition. The sub-images images $F_{LL}(x, y)$, $F_{LH}(x, y)$, $F_{HL}(x, y)$ and $F_{HH}(x, y)$ correspond respectively to LL, LH, HL, and HH images in Figure 1. A second level decomposition is considered in this study. The Daubechies-4 wavelet [18] is chosen as the mother wavelet in this paper since it has the advantage of better resolution for smoothly changing signals [19]. Finally, features are extracted from HH2 sub-image since detail coefficients in level 2 and 3 contain fine breast structure and micro-calcifications [2].

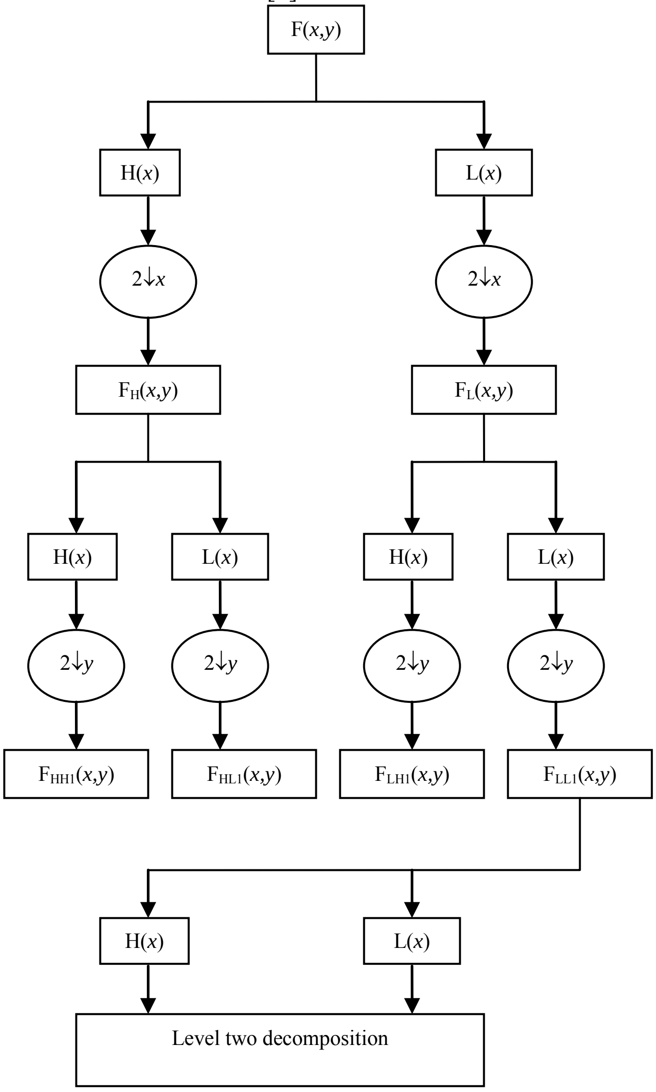

Figure 2: 2D-DWT decomposition process of an image.

C. Features Extraction

The statistics used to describe the processed images are the mean, standard deviation, smoothness, third moment, uniformity, and entropy. They are chosen since they are widely used in patter recognition [20]. The statistics are expressed as follows:

$$Mean = m = \sum_{i=0}^{L-1} z_i\, p(z_i) \tag{5}$$

$$St.Dev = \delta = \sqrt{\mu_2(z)} = \sqrt{\delta^2} \tag{6}$$

$$Smoothness = R = 1 - \frac{1}{1/(1+\delta^2)} \tag{7}$$

$$3th.Moment = \mu_3 = \sum_{i=0}^{L-1} (z_i - m)^3\, p(z_i) \tag{8}$$

$$Uniformity = U = \sum_{i=0}^{L-1} p^2(z_i) \tag{9}$$

$$Entropy = e = -\sum_{i=0}^{L-1} p(z_i) \log_2 p(z_i) \tag{10}$$

where z is a random variable indicating intensity, p is the probability density of the ith pixel in the histogram, and L is the total number of intensity levels. Finally, principal component analysis (PCA) is applied to the features set to reduce the number of characteristics to be fed to the classifier. For instance, *k*-NN is employed in this study for classification.

D. Design of Experiments

As mentioned in the introduction, the following experiments are conducted:

(1) **Fourier**: Apply Fourier transform to mammogram. Extract high frequency image. Then, compute features. Finally, apply PCA. Employ *k*-NN for classification task.

(2) **Fourier-Fourier**: Apply Fourier transform to mammogram. Extract Fourier processed image. Then, apply Fourier transform to the previous Fourier processed image to obtain the Fourier-Fourier image. Compute features. Finally, apply PCA. Employ *k*-NN for classification task.

(3) **DWT**: Apply wavelet transform to mammogram. Extract high frequency image (HH2). Then, compute features. Finally, apply PCA. Employ *k*-NN for classification task.

(4) **DWT-DWT**: Apply DWT to the mammogram. Extract high frequency image (HH2). Then, apply another DWT to the previous high frequency image (HH2) to obtain HH2*. Compute features. Finally, apply PCA. Employ *k*-NN for classification task.

The experiments (1) to (4) are shown in Figures 3,4,5, and 6 respectively.

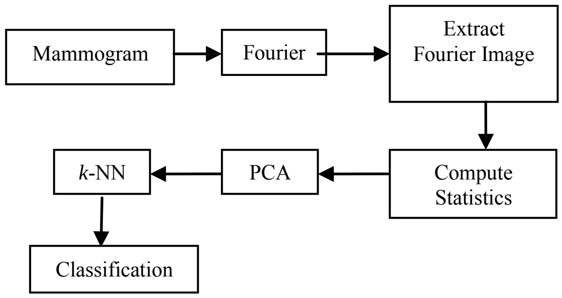

Figure 3. Experiment (1): Fourier approach.

E. The Classifier

The *k*-nearest neighbor algorithm (*k*-NN) [21] is employed for classification task in this study. It is a nonparametric method that assigns query data to the class that the majority of its *k*-nearest neighbors belong to. For instance, the *k*-NN algorithm uses the data directly for classification without the need of an explicit model.

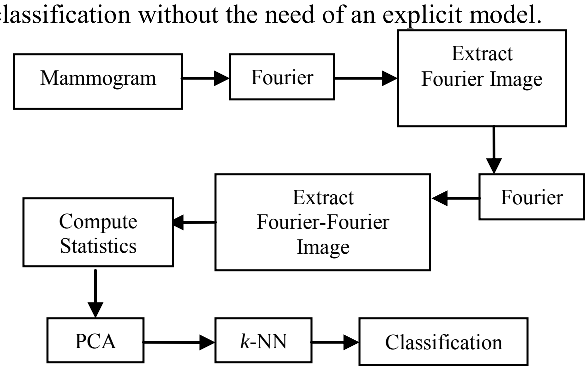

Figure 4. Experiment (2): Fourier-Fourier approach.

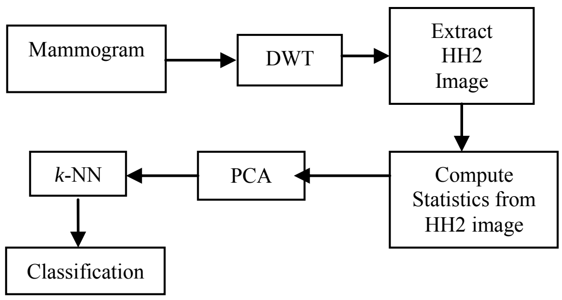

Figure 5. Experiment (3): DWT approach.

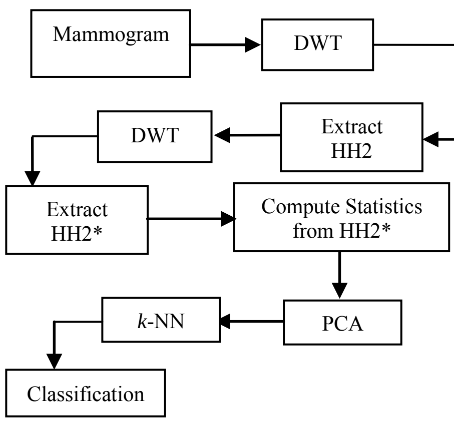

Figure 6. Experiment (4): DWT-DWT approach.

The performance of k-NN depends on the number of the nearest neighbor k. In general, there is no solution to find the optimal k. However, trial and error approach is usually used to find its optimal value. The main objective is to find the value of k that maximizes the classification accuracy. The main advantage of k-NN algorithm is the ability to explain the classification results. On the other hand, its major drawback is the need to find the optimal k and to define the appropriate metric to measure the distance between the query instance and the training samples. In this paper, the distance metric chosen is the Euclidean distance. The standard algorithm of k-NN is given as follows:

(1) Calculate Euclidean distances between an unknown object (o) and all the objects in the training set;

(2) Select k objects from the training set most similar to object (o), according to the calculated distances;

(3) Classify object (o) with the group to which a majority of the K objects belongs.

III. Data and Results

In order to test the proposed methodology, one hundred digital mammograms were taken from The Digital Database for Screening Mammography (DDSM) [22]. They consisted of fifty normal images and fifty cancer images. An example of a digital mammogram is shown in Figure 7. The DWT and DWT-DWT images of a normal mammogram are shown in Figure 8, and its Fourier and Fourier-Fourier images are shown in Figure 9.

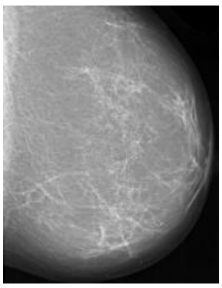

Figure 7: Example of a normal mammogram.

Figure 8: Left: DWT. Right: DWT-DWT.

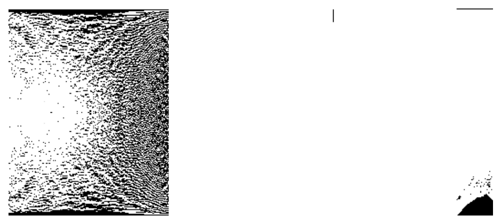

Figure 9: Left: Fourier. Right: Fourier-Fourier.

The features selected by PCA and their respective cumulative variance proportion (CVP) are shown in Table 1. It shows that the most significant features are smoothness and uniformity for DWT-DWT images and uniformity and entropy for Fourier-Fourier images. The value of k used to perform the k-NN algorithm was varied from 2 to 10. The optimal k that maximizes the average recognition rate was found to be 2. The experiments were conducted with 10-fold cross-validation. For each fold, the correct classification rate (hit ratio) and standard deviation are computed. Figure 10 provides the obtained results.

Table 1: Features selected by PCA

	Features	**CVP**
DWT image	smoothness, uniformity	99%
DWT-DWT image	smoothness, uniformity	98%
Fourier image	mean, entropy	92%
Fourier-Fourier image	uniformity, entropy	98%

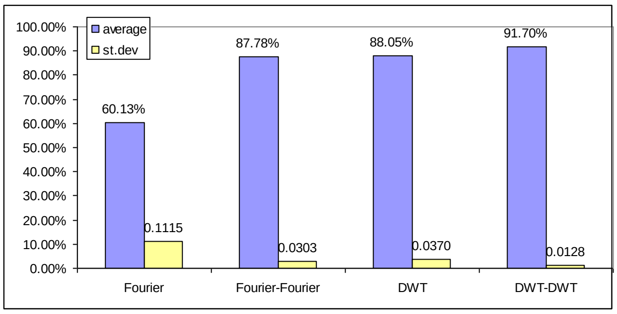

Figure 10. Classification results using k-NN.

It shows that the feature extraction approach based on DWT-DWT clearly improved the classification accuracy with respect to using DWT only. For instance, when using only DWT feature extraction, the achieved average accuracy was 88.05% (±0.037), and when using DWT-DWT based feature extraction, it was 91.70.13% (±0.012). On the other, the feature extraction approach based on Fourier-Fourier approach clearly improved the classification accuracy with respect to using Fourier only. For instance, when using only Fourier feature extraction, the achieved average accuracy was 60.13.05% (±0.1115), and when using Fourier-Fourier based feature extraction, it was 87.78% (±0.0303). Then, the results show strong evidence that the two-step based approaches are suitable to extract high frequency features to better detect MCs in digital mammograms.

In sum, the DWT-DWT approach provides higher accuracy than the Fourier-Fourier approach. In addition, both proposed approaches allow obtaining high frequency characteristics needed to detect MCs. In other words, our methodology based on a further analysis of high frequency images using DWT or Fourier transform helps extracting suitable features to detect MCs.

IV. CONCLUSION

The discrete wavelet transform (DWT) is the most employed tool for processing mammograms to detect microcalcifications (MCs) since it is able to perform signal analysis at different time and frequency scales. Indeed, MCs are usually found in dense biological tissue, which corresponds to high frequencies in the frequency domain of the image. Therefore, we seek to capture better high frequency features of a mammogram by performing a further analysis of its high frequency components. In particular, the DWT is applied to the mammogram to obtain its high-high (HH) image at level two. Then, a further decomposition by DWT at level two is applied to the HH image of the previous step to obtain HH*. Finally, statistical features are computed from HH* and PCA is performed to select the most significant features. The *k*-NN algorithm is employed to classify normal versus cancer images using cross-validation technique. Recall this approach DWT-DWT. A similar approach with use of discrete Fourier Transform (FT) is proposed. For instance, FT is applied to the mammogram to obtain a Fourier image. Then, a second FT is applied to the Fourier image obtained in the previous step. Recall this approach Fourier-Fourier.

The simulation results show strong evidence of the effectiveness of our two methodologies since the average recognition rate improves and its standard deviation decreases.

REFERENCES

[1]H.D Cheng, Juan Shan, Wen Ju, Yanhui Guo and LingZhang, ''Automated breast cancer detection and classification using ultrasound images: A survey'', Pattern Recognition, Vol. 43, pp.299 – 317, 2010.

[2]Bouyahia S., J. Mbainaibeye, and N. Ellouze, ''Wavelet Based Microcalcifications Detection in Digitized Mammograms'', ICGST-GVIP Journal, Vol.8, pp.23-31, 2009.

[3]Ferreira Cristiane Bastos Rocha and Dibio Leandro Borges, ''An Evaluation of Wavelet Features Subsets for Mammogram Classification'', Lecture Notes in Computer Science, Vol.3773, pp.620-630, 2005.

[4]Sepehr M.H. Jamarani, Gholamali Rezai-Rad and Hamid Behnam, ''A Novel Method for Breast Cancer Prognosis Using Wavelet Packet Based Neural Network'', Proceedings of the IEEE Engineering in Medicine and Biology, Shanghai, China, September 1-4, pp.3414- 3417, 2005.

[5]Dong A. and Wang B., ''Feature Selection and Analysis on Mammogram Classificatio'', Proceedings of IEEE Pacific Rim Conference on Communications, Computers and Signal Processing, Victoria, B.C., Canada, August 23-26, 2009.

[6]Eltoukhy M.M.M, I. Faye, and B.B Samir, ''Using Curvelet Transform to Detect Breast Cancer in Digital Mammogram''. IEEE International Colloquium on Signal Processing & Its Applications (CSPA), pp.340-345, 2009.

[7]Tirtajaya A. and D.D. Santika, ''Classification of Microcalcification Using Dual-Tree Complex Wavelet Transform and Support Vector Machine'', IEEE International Conference on Advances in Computing, Control, and Telecommunication Technologies, pp.164-166, 2010.

[8]Moayedi F., Azimifar Z., and Boostani R., Katebi S.,'' Contourlet-based mammography mass classification using the SVM family'', Computers in Biology and Medicine, Vol.40, pp.373–383, 2010.

[9]Lahmiri S. and Boukadoum M., "Hybrid discrete wavelet transform and Gabor filter banks processing for mammogram features extraction Classifier'', IEEE New Circuits and Systems (NEWCAS) International Conference, June 26-29, Bordeaux, France, pp.53-56, 2011.

[10]Sakka E., Prentza A., Lamprinos I.E. and Koutsouris D., "Microcalcification Detection using Multiresolution Analysis Based on Wavelet Transform", IEEE International Special Topic Conference on Information Technology in Biomedicine, Ioannina, Epirus, Greece, October 26-28, 2006.

[11]Gelman L., and Braun S., ''the optimal usage of the Fourier transform for pattern recognition'', Mech. Systems Signal Process., Vol.15, pp.641–645,2001.

[12]Liu F. and Picard R.W., ''Periodicity, directionality, and randomness: Wold features for image modeling and retrieval'', IEEE Trans. Pattern Anal. Mach. Intell., Vol.18,pp.722–733, 1996.

[13]Osowski Stanislaw and Dinh Nghia Do, ''Fourier and wavelet descriptors for shape recognition using neural networks—a comparative study'', Pattern Recognition, Vol.35 pp.1949–1957, 2002.

[14]Chen Guangyi and Bui T.D, ''Invariant Fourier-wavelet descriptor for pattern recognition'', Pattern Recognition, Vol. 32, pp.1083-1088, 1999.

[15]Sengur Abdulkadir and Turkoglu Ibrahim, and M. Cevdet Ince., ''Wavelet Packet Neural Networks for Texture Classification'', Expert Systems with Applications, Vol. 32, pp.527-533, 2007.

[16]Wang Z-Z and Yong J-H., ''Texture Analysis and Classification with Linear Regression Model Based on Wavelet Transform'', IEEE Transactions on Image Processing, Vol.17, pp.1421-,14302008.

[17]Heil C.E., and Walnut D.F, "Continuous and Discrete Wavelet Transforms", SIAM Review, vol. 31, no. 4,pp.628-666, 1989.

[18]Daubechies I., ''Ten Lectures on Wavelets'', SIAM, Philadelphia, PA, 1992.

[19]Algorri M-E and Fernando F-M., ''Classification of Anatomical Structures in MR Brain Images Using Fuzzy Parameters'', IEEE Transactions on biomedical engineering, Vol. 51, No.9, pp.1599-1608, 2004.

[20]Sheshadri H.S and Kandaswamy A., ''Detection of breast cancer by mammogram image segmentation'', Journal of Cancer Research and Therapeutics, Vol.1, pp :232-234, 2005.

[21]Dasarthy, B.V., ''Nearest Neighbor Classification Techniques'', IEEE Press, Hoboken(NJ), 1990.

[22]http://marathon.csee.usf.edu/Mammography/Database.html