



## Automated Diagnosis of Mammogram Images of Breast Cancer Using Discrete Wavelet Transform and Spherical Wavelet Transform Features: A Comparative Study

www.tcr.org  
DOI: 10.7785/tcr.2013.600262

Mammograms are one of the most widely used techniques for preliminary screening of breast cancers. There is great demand for early detection and diagnosis of breast cancer using mammograms. Texture based feature extraction techniques are widely used for mammographic image analysis. In specific, wavelets are a popular choice for texture analysis of these images. Though discrete wavelets have been used extensively for this purpose, spherical wavelets have rarely been used for Computer-Aided Diagnosis (CAD) of breast cancer using mammograms. In this work, a comparison of the performance between the features of Discrete Wavelet Transform (DWT) and Spherical Wavelet Transform (SWT) based on the classification results of normal, benign and malignant stage was studied. Classification was performed using Linear Discriminant Classifier (LDC), Quadratic Discriminant Classifier (QDC), Nearest Mean Classifier (NMC), Support Vector Machines (SVM) and Parzen Classifier (ParzenC). We have obtained a maximum classification accuracy of 81.73% for DWT and 88.80% for SWT features using SVM classifier.

### Introduction

Breast cancer is a leading cause of fatality among women and is also the most prevalent non-skin cancer in women (1, 2). This is evident from the fact that mortality of breast cancer is the second highest among all cancer deaths in women (2). It is well known that there is no technology at present, which is capable of curing cancer. But, it is well known that early detection of cancer can aid in good recovery and prolong patient life (2). For this reason, radiologists want to detect breast cancer at an early stage. But there are several problems with early detection of breast cancer. One problem is that, most often, the first tool for breast cancer screening is a mammogram. This in itself poses a unique problem since there is a lot of inter-observer variations which occur while diagnosing breast cancer through mammograms (3). According to Dio *et al.*, the purpose of Computer-Aided Diagnosis (CAD) in radiology, is “to improve the diagnostic accuracy as well as the consistency of radiologists’ image interpretation by using the computer output as a guide” (3). This is very true, since a radiologist makes his diagnostic interpretations based on judgements which are subjective to the radiologist and also because of the fact that radiologists tend to miss microcalcifications and masses which are not visible clearly in mammograms. This, along with the fact that intra-observer and inter-observer variability play an important role in diagnostic accuracy, has been statistically proven by Balleyguier *et al.* in (4). The influence and improvement in diagnostic accuracy of radiologists in the presence of a CAD system has been studied and proven by Fenton *et al.* in (5).

**Karthikeyan Ganesan, M.Sc.<sup>1\*</sup>**  
**U. Rajendra Acharya, Ph.D.,**  
**D.Eng.<sup>1,3</sup>**  
**Chua Kuang Chua, Ph.D.<sup>1</sup>**  
**Lim Choo Min, Ph.D.<sup>1</sup>**  
**Thomas K. Abraham, Ph.D.<sup>2</sup>**

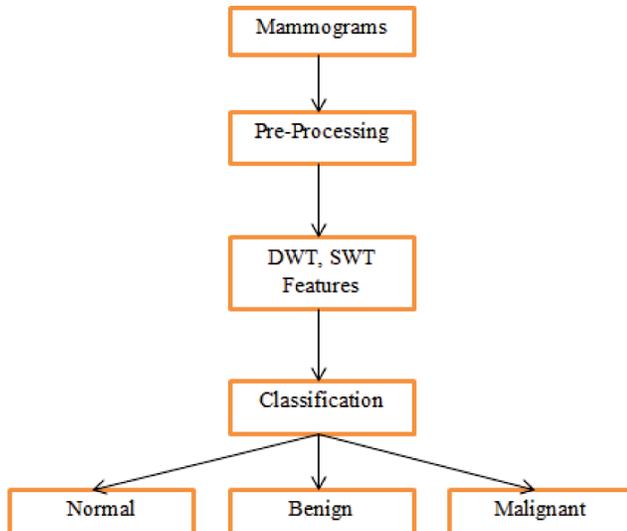
<sup>1</sup>Department of ECE, Ngee Ann Polytechnic, Clementi Road, Singapore-599489

<sup>2</sup>SATACommHealth Singapore

<sup>3</sup>Department of Biomedical Engineering, Faculty of Engineering, University of Malaya, Malaysia

\*Corresponding author:  
Karthikeyan Ganesan, M.Sc.  
Phone: (+65)-90017046  
E-mail: g.karthikeya@gmail.com

The classification pipeline used in this work is shown in Figure 1. Though CAD systems are thought to be in a nascent stage, evidence in (5) proves quite the contrary with high success rates for radiologists using CAD systems. There are numerous works which have been done in the past with a



**Figure 1:** Classification framework used in this work.

combination of several techniques (9) With accuracy rates reaching up to a maximum of 99.5% in the work of Sadaf *et al.* (6), we see that the field of CAD systems for breast cancer is at a very advanced stage. Yet, there is scope for improvement. This is due to various factors. For instance, in the work reported in (6), though the accuracy rates seem to be impressive, it must also be noted that the experiments which gave a high accuracy rate were done on Full-Field Digital Mammograms (FFDM) and only on images which manifest as microcalcifications. In images which manifest as other mammographic appearances, the accuracy rate was mentioned to be 86%. Similarly, a more complicated technique for usage in a CAD system was proposed by Szekeley *et al.* in (7). In this study, texture features were used with combined classifiers decision tree and multiresolution markov random models, which gave an accuracy rate of 94%. Hence we see that there are a range of wide array of techniques ranging from the seemingly simple ones (6) to more complicated ones (7).

Since there is clearly scope for further improvement in the basic classification framework, especially in the feature extraction step, we propose the application of a new feature extraction methodology for usage in mammographic image analysis. Though Spherical Wavelet Transforms (SWT) was initially proposed for astronomical image and data analysis (9), their inherent properties and ability to extract minute information

make it particularly suitable for our application. Spherical Wavelet Transform (SWT) is yet to be widely used in popular applications, especially in the field of medical image analysis where they might prove extremely useful given the fact that they prove to be very useful in filtering off noise and sharpening the images (10) apart from providing vital information regarding details which might otherwise not be available in the case of Discrete Wavelet Transforms (DWT).

This paper highlights the advantages of SWT over DWT features using mammogram images. We do this by studying the classification accuracy of a three class classification problem. The database was obtained from SATACommHealth, Singapore. The three classes are normal, benign and malignant mammograms. We consider only features extracted by DWT and SWT. The classification was performed for both these feature sets separately, to evaluate the performance of SWT and DWT. Classification was performed using Linear Discriminant Classifier (LDC), Quadratic Discriminant Classifier (QDC), Nearest Mean Classifier (NMC), Support Vector Machines (SVM) and Parzen Classifier (ParzenC).

### Materials and Methods

The experiment was performed using 282 mammograms provided by the SATACommHealth, Singapore. All the images were acquired from patients between 45-70 years of age. There were 65 malignant images, 60 benign images and 157 normal images. Before further processing, each image was preprocessed and normalized to counter variations in imaging conditions. In order to obtain accurate quantitative results, the pectoral muscles from all images were manually removed in order to work only on the breast tissue masses. The images were received and processed at a resolution of  $1024 \times 1680$ . Each image were processed at two views: the Cranio-Caudal (CC) view and the Mediolateral-oblique (MLO) view. Figure 2 shows samples of mammograms which were used in this study.

### Feature Extraction

In this study, two feature extraction techniques; Discrete Wavelet Transform (DWT) and Spherical Wavelet Transform (SWT) were used to extract texture features from the images in consideration.

### Discrete Wavelet Transform (DWT)

Discret wavelet transforms are well known textural feature extraction techniques. In this method, the images were sent through a series of down-sampling filters which are composed up of a sequence of high-pass and low-pass filters. The high-pass filters produce the detail coefficients  $D[n]$  while the

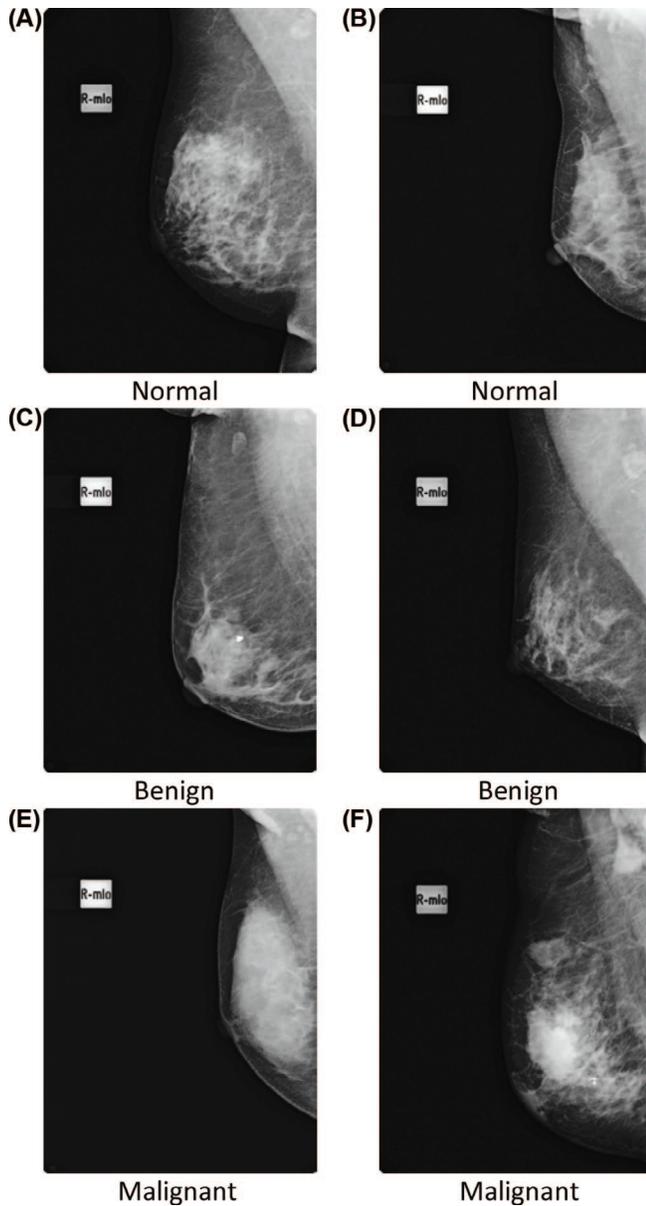


Figure 2: Sample mammograms.

low-pass filters produce the approximation coefficients  $A[n]$  (11). These coefficients are mathematically represented by:

$$D[n] = \sum_{k=-\infty}^{\infty} x[k]h[2n-k] \quad [1]$$

$$A[n] = \sum_{k=-\infty}^{\infty} x[k]g[2n-k] \quad [2]$$

Where,  $x[k]$  represents the image in consideration and  $h[2n-k]$  represents the transfer function of a high-pass filter,

while  $g[2n-k]$  represents the transfer function of a low-pass filter.

Biorthogonal wavelets were used in this study (12). Biorthogonal wavelets are those in which the wavelet transform is invertible, but not necessarily orthogonal. The advantage of biorthogonal wavelets over orthogonal wavelets is that biorthogonal wavelets allow more degrees of freedom compared to orthogonal wavelets. The first level wavelet produces a horizontal  $Dh_1$ , diagonal  $Dd_1$  and vertical  $Dv_1$  detailed coefficients and one approximation coefficient  $A_1$ (11). The output of these matrices, though provide intensity values, cannot be directly used due to the fact that the number of elements are simply too high for computation. So, averaging methods for dimensionality reduction have been developed as follows:

$$\text{Average } Dh_1(Ah) = \frac{1}{NxM} \sum_{x=N} \sum_{y=M} |Dh_1(x,y)| \quad [3]$$

$$\text{Average } Dv_1(Av) = \frac{1}{NxM} \sum_{x=N} \sum_{y=M} |Dv_1(x,y)| \quad [4]$$

$$\text{Average } Dd_1(Ad) = \frac{1}{NxM} \sum_{x=N} \sum_{y=M} |Dd_1(x,y)| \quad [5]$$

The final averaging method used averages not the intensity values as such; but averages the energy of the intensity values.

$$\text{Energy } (Ed) = \frac{1}{N^2 x M^2} \sum_{x=N} \sum_{y=M} |Dd_1(x,y)|^2 \quad [6]$$

$$\text{Energy } (Ev) = \frac{1}{N^2 x M^2} \sum_{x=N} \sum_{y=M} |Dv_1(x,y)|^2 \quad [7]$$

A more detailed description of the technique used can be seen in (11, 12).

#### Spherical Wavelet Transform (SWT)

The Spherical Wavelet Transform (SWT) proposed by Starck *et al.* (9) works on the principle that the sum of the scales reproduces the original data. The goal of this method is to reduce redundancy which occurs in the conventional DWT, which results in huge data sets, in turn requiring averaging techniques, which result in loss of data as seen from discussions in the previous section. The development of a SWT by (9) was carried out using an undecimated isotropic transform. Isotropy utilized in this development is favorable for construction of a wavelet pyramid since it can capture isotropic features in a statistically isotropic field. This property can be

exploited and put to good use in textural feature extraction of medical images since isotropic region extraction is the basis of medical image analysis using textures.

The approximations of an image  $I$  on a dyadic resolution scale can be obtained using the scaling function  $\phi_{l_c}$  as:  $c_0 = \phi_{l_c} * f$ ,  $c_1 = \phi_{2^{-1}l_c} * f, \dots, c_j = \phi_{2^{-j}l_c} * f$ , where  $\phi_{2^{-j}l_c}$  is a rescaled version of  $\phi_{l_c}$  with cut-off frequency  $2^{-j}l_c$  (8). From this, a low-pass filter  $h_j$  for each scale  $j$  is defined by:

$$\hat{H}_j(l, m) = \sqrt{\frac{4\pi}{2l+1}} \hat{h}_j(l, m) \quad [8]$$

Similarly, a high-pass filter  $g_j$  on each scale  $j$  can be derived as:

$$\hat{G}_j(l, m) = \sqrt{\frac{4\pi}{2l+1}} \hat{g}_j(l, m) \quad [9]$$

From Equations 8 and 9, it is evident that the high-pass filter used in this specific case can be expressed as:

$$\hat{G}_j(l, m) = 1 - \sqrt{\frac{4\pi}{2l+1}} \hat{h}_j(l, m) = 1 - \hat{H}_j(l, m) \quad [10]$$

Equation 10 shows an example of a wavelet function that can be used in the spherical domain, while any other wavelet function can be used in its place to determine the detailed and approximate coefficients.

Algorithm 1 shows an implementation of SWT (9).

---

#### Algorithm 1: Implementation of the SWT

---

- Step 1: Compute a multiresolution sphere.
  - Step 2: Compute the center of each face of the sphere.
  - Step 3: Load the image and precompute the local wavelet matrix using Eq.8-Eq.10.
  - Step 4: Initialize the forward transform for extracting the low-pass components from the orthogonal direction details in Step 3.
  - Step 5: Extract these low-pass components and create a matrix.
  - Step 6: Back store the coefficients.
  - Step 7: Calculate the Spherical Wavelet Coefficients.
- 

### Classification

#### Linear Discriminant Classifier (LDC)

The idea of linear classifiers is that, in a training pattern  $x_1, x_2, \dots, x_n$  each of the object in the pattern is assigned to a

class  $\omega_1$  or  $\omega_2$  depending on a threshold  $\tau_0$  determined by a mathematical function defined as:

$$\mathbf{w}^T x + \tau_0 \begin{cases} > 0 \\ < 0 \end{cases} \Rightarrow x \in \begin{cases} \omega_1 \\ \omega_2 \end{cases} \quad [11]$$

Where,  $\mathbf{w}^T$  is the weight vector. The data is said to be linearly separable when  $\mathbf{w}^T x + \omega_0 > 0$  for all samples in a single class (21). All linear classifiers function on this logic.

#### Quadratic Discriminant Classifier (QDC)

In a classification problem, it is common to assume that data in each class is adequately described by a Gaussian distribution. From this assumption, a conclusion is brought about the linearity or quadraticity of the classifier, depending on the covariance matrices of the data (21). If the covariance matrices turn out to be equal, a Linear Discriminant approach is followed, while a Quadratic Discriminant approach is followed when the covariance matrices turn out to be different (22). The implication of this can very well be seen in our study where QDC clearly outperforms LDC. This shows that the data is not linearly separable giving more scope for non-linear classifiers.

#### Nearest Mean Classifier (NMC)

Nearest Mean Classifier is one of the simplest, yet most effective classifier. It is simple, because of its computational efficiency where it takes very little effort to compute the mean of the classes in consideration. Once the mean of the classes are computed at random, every object in the dataset is assigned to one of the classes, depending upon a distance measure which calculates the distance between the mean of the class and the object in consideration. The procedure is followed for every object, until all objects in the dataset are assigned to one of the classes (21). Though NMC has found to be effective in cases where the data is well spread out, highly non-linear data like the one in our current study does not provide a good result due to the high overlap between the means of various classes in consideration. This will result in poor performance as seen from the results in our study.

#### Parzen Classifier (ParzenC)

Parzen classifiers are non-parametric classifications algorithms. The idea of a non-parametric estimation technique is to divide the total histogram of a given feature set into a number of bins and calculate the probability of a random sample belonging to one particular bin in the histogram. Parzen classifiers are based on this non-parametric classification technique which models data in a multidimensional scale. That is, instead of dividing the histogram into several bins,

like usual non-parametric techniques, the  $n$ -dimensional space is divided into hypercubes with a length of side  $h$  and volume  $h^l$  (20). In this case, the probability  $\hat{p}(x)$  of a variable belonging to a particular hypercube can be given by:

$$\hat{p}(x) = \frac{1}{h^l} \left( \frac{1}{N} \sum_{i=1}^N \phi \left( \frac{x_i - x}{h} \right) \right) \quad [12]$$

Where  $x_i = 1, 2, \dots, l$  are the available feature vectors and  $\phi(x_i) = 1$  for  $x_i < 1/2$  and  $\phi(x_i) = 0$  otherwise.

#### Support Vector Machine (SVM)

SVMs are one of the most widely used classifiers. The idea of SVMs are derived from the LDC. The separating hyperplanes for SVMs are defined as:

$$\begin{aligned} \omega^T x_i + \tau_0 &\geq +1, \forall x \in \omega_1 \\ \omega^T x_i + \tau_0 &\leq -1, \forall x \in \omega_2 \end{aligned} \quad [13]$$

where,  $\omega_1$  and  $\omega_2$ , are the classes in consideration,  $\omega$  a weight vector, and a threshold  $\tau_0$  determined with the above equation (23). From Equation 13, it is evident that, it is just an extension of the generic LDC. It is difference in a way that the hyperplanes in place for SVMs function on the theory of support vectors. Here, the vectors closest to the separating hyperplanes determine the location of the hyperplane rather than the hyperplane determining the location of the objects in the classes. Also, SVMs perform well in the case of non-linear data due to the fact that these classifiers do not separate data in the original data space, but rather map the original data into a more manageable space using specified functions referred to as kernels, which help in converting a non-linear data into a linear

data through the kernel functions (23, 24). In the current study, a radial basis kernel is used to build our classifier. This kernel is defined by:

$$K(x, z) = \exp \left( -\frac{\|x - z\|^2}{\sigma^2} \right) \quad [14]$$

where,  $\sigma$  provides the standard deviation of the class in consideration and  $x$  and  $z$  provide the object and mean of the classes respectively.

In this study, where binary classifiers have been used, a simple majority vote was used to extend it into a multi-class classification structure.

#### Results

The current study provides a comparative and quantitative analysis of the usage of SWT with respect to DWT. 282 images were used for this study, with similar processing done on all the images. All the levels of sub-bands of both DWT and SWT were used to extract the features. Feature ranking or selection was not performed due to the fact that it is quite evident that all sub-bands of a wavelet transformation scheme are important as information missing in one sub-band is always found in another. This holds true for both DWT and SWT. Table I shows a sample window of features for the three classes of mammograms for DWT and SWT features considered in this study.

To study the data distribution of the features extracted by both DWT and SWT, the features were plotted as their empirical Cumulative Distributive Functions (eCDF), which is seen in Figure 3. This visualization technique helps in better inference of variations in feature spread obtained using DWT and SWT.

**Table I**

Significant features extracted (Mean  $\pm$  SD) using the DWT and SWT. The  $p$ -values were  $< 0.0001$  for all features.

Features	Normal	Benign	Malignant
DWTA1	2.4E+02 $\pm$ 1.2E+02	1.7E+02 $\pm$ 1.04E+02	2.6E+02 $\pm$ 1.5E+02
DWTD1	1.9E-03 $\pm$ 1.3E+01	-3.8E+01 $\pm$ 1.19E+01	2.2E-06 $\pm$ 1.3E+01
DWTH1	1.0E-02 $\pm$ 3.4E+01	1.1E+00 $\pm$ 3.26E+01	9.0E-04 $\pm$ 3.5E+01
DWTH2	5.0E-02 $\pm$ 1.0E+02	1.9E+00 $\pm$ 8.23E+01	1.8E+00 $\pm$ 8.9E+02
DWTV1	-2.0E-02 $\pm$ 4.3E+01	-1.3E+01 $\pm$ 3.15E+01	-1.2E+01 $\pm$ 2.2E+01
SWTA1	5.2E+01 $\pm$ 3.0E-04	1.4E+01 $\pm$ 9.0E-04	1.5E+02 $\pm$ 1.1E-03
SWTD1	8.2E+00 $\pm$ 5.0E-03	6.6E+00 $\pm$ 1.5E-04	1.7E+01 $\pm$ 7.0E-04
SWTH1	-8.4E+00 $\pm$ 1.3E+00	5.9E+00 $\pm$ 1.7E+01	1.5E+01 $\pm$ 1.0E+02
SWTH2	5.7E+02 $\pm$ 1.34E-03	1.0E+01 $\pm$ 3.6E-03	6.8E+02 $\pm$ 1.9E-01
SWTV1	1.02E+02 $\pm$ 5.0E-04	9.1E+02 $\pm$ 1.7E-02	2.0E+02 $\pm$ 1.0E-01

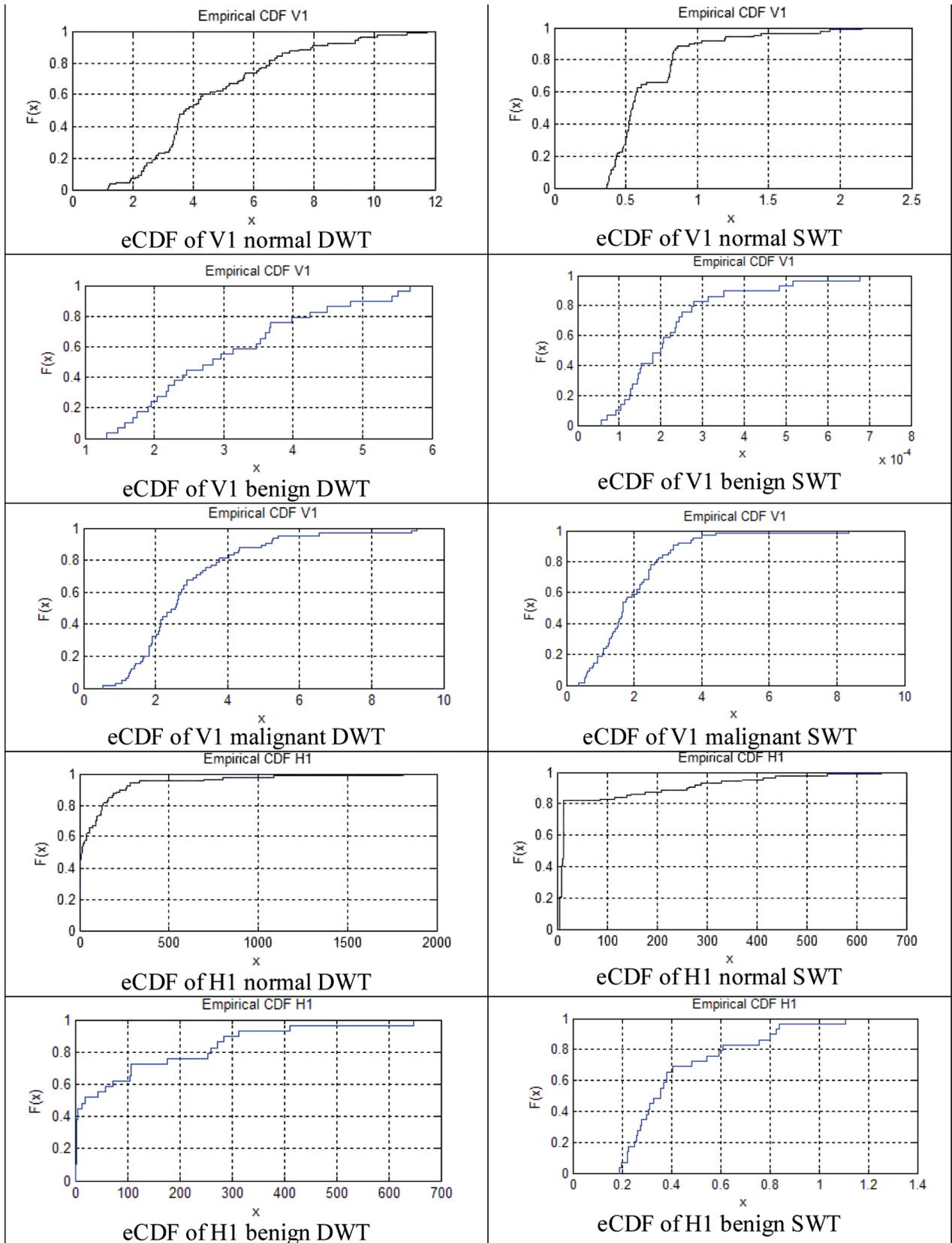


Figure 3: (Continued)

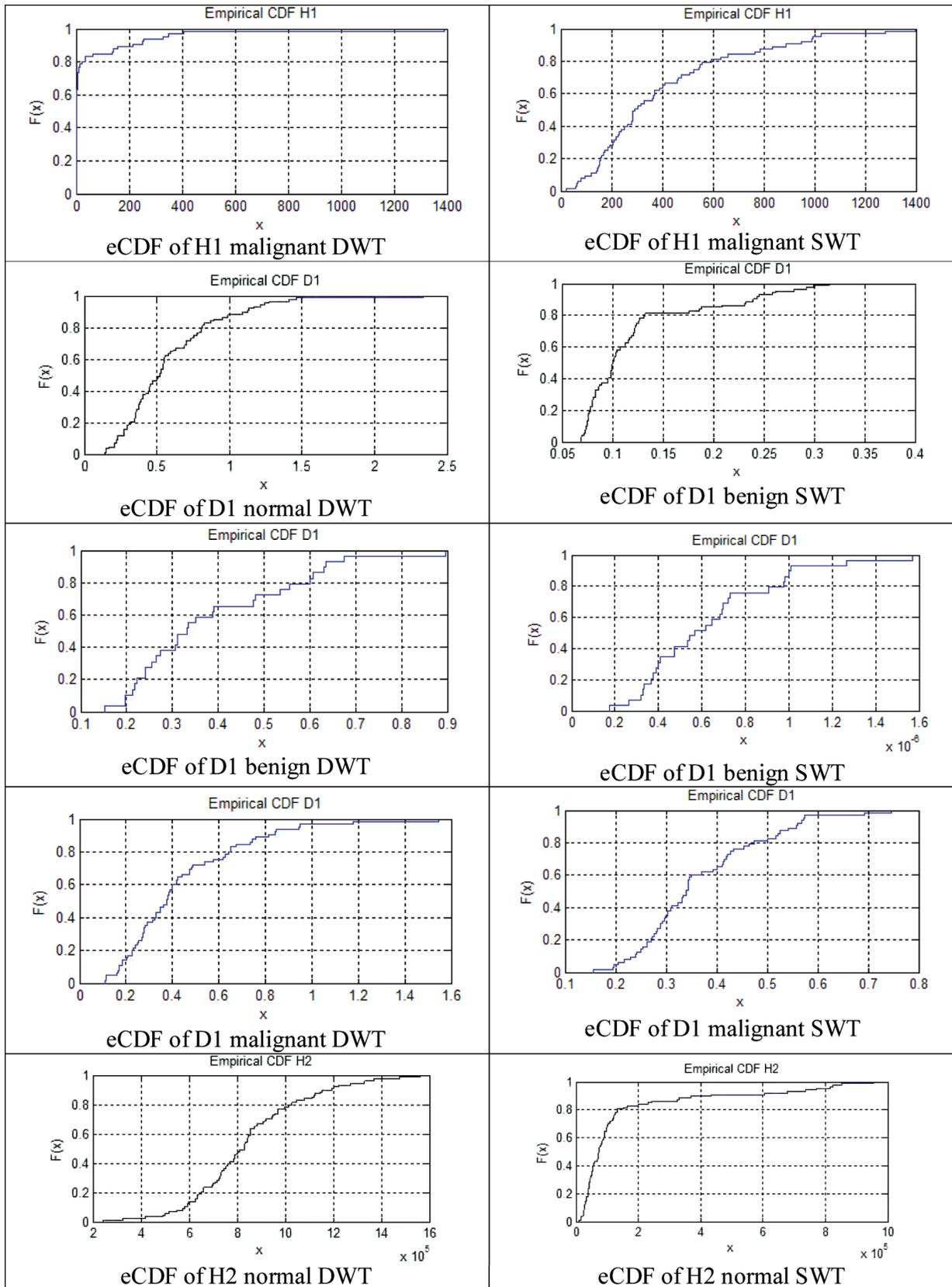
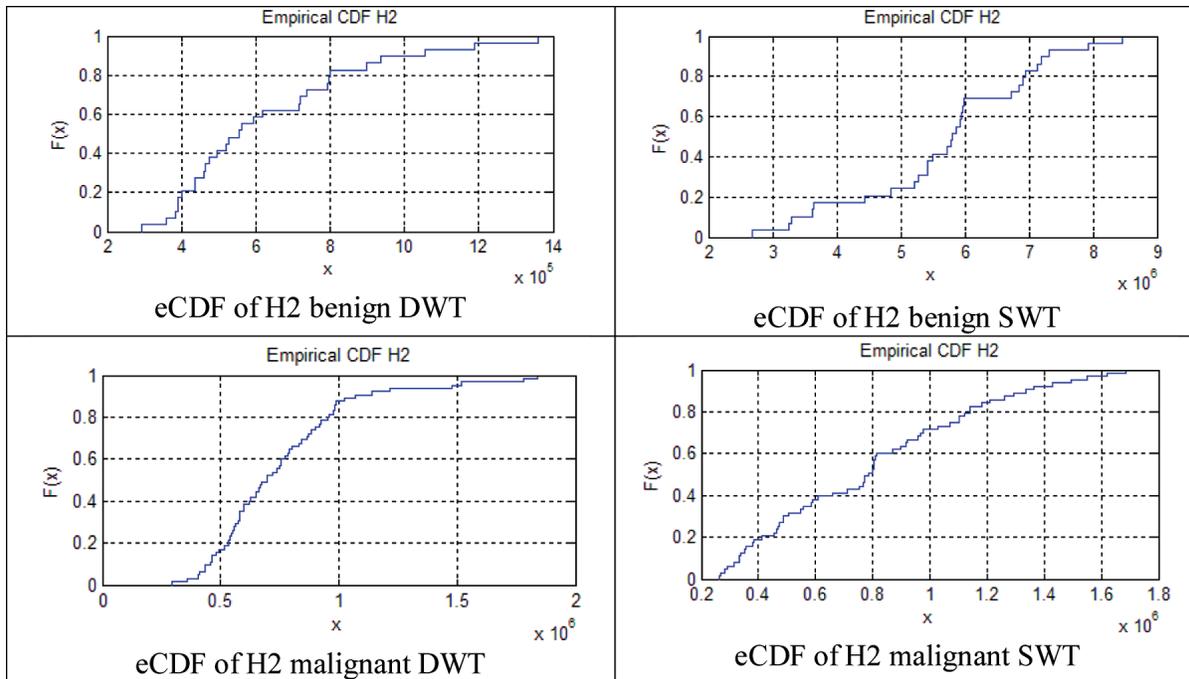
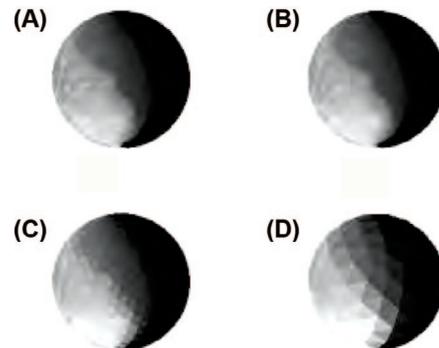


Figure 3: (Continued)



**Figure 3:** Empirical Cumulative Distributions (eCDF) plots for different classes of images using DWT and SWT features.

Since this study focuses on a comparative analysis of performance between DWT and SWT, no other feature extraction techniques were used. A maximum classification accuracy of 81.73% was obtained using DWT features coupled with SVM radial basis classifier. Similarly, a maximum classification accuracy of 88.80% was obtained using SWT features with SVM radial basis classifier. The average performance of all the classifiers using ten-fold cross validation can be seen in Table II. The decomposition sub-bands of SWT applied to a sample mammograph can be seen in Figure 4. The results of all ten-folds of the cross-validation scheme can be seen in Figure 5 and Figure 6, for DWT and SWT respectively.

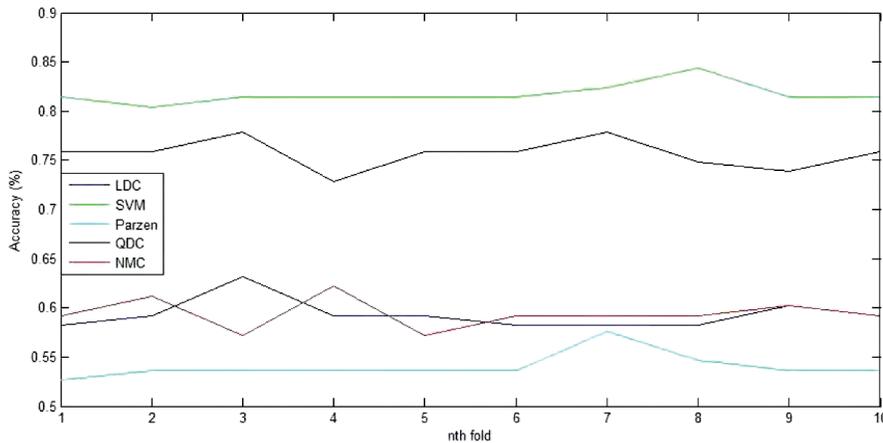


**Figure 4:** Sample images of SWT decomposition levels of a mammogram: (A) the original image (B-D) consecutive three levels of decomposition.

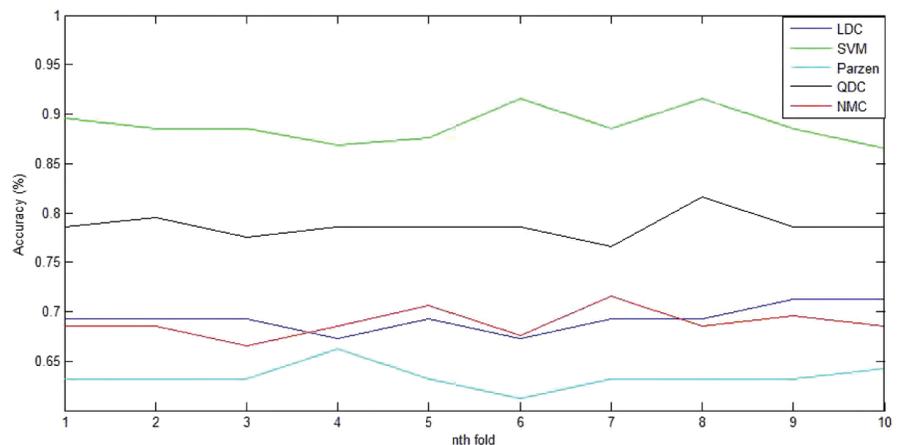
**Table II**

Classification accuracy for DWT and SWT using ten-fold cross validation. Sensitivity and specificity for the three class problem was found using the one-vs. all confusion matrix approach.

Classifier	DWT-accuracy (%) [Sensitivity (%), Specificity (%)]	SWT-accuracy (%) [Sensitivity (%), Specificity (%)]
LDC	59.31 [60.94, 57.63]	69.26 [67.36, 71.16]
QDC	75.67 [74.31, 77.03]	78.68 [77.12, 80.24]
NMC	59.41 [62.10, 56.72]	68.88 [68.40, 69.36]
SVM	81.73 [81.32, 82.14]	88.80 [89.69, 87.91]
ParzenC	54.05 [53.01, 55.10]	63.40 [62.10, 64.70]



**Figure 5:** Classification accuracy for DWT using a ten-fold cross validation scheme for various classifiers.



**Figure 6:** Classification accuracy for SWT using a ten-fold cross validation scheme for various classifiers.

### Discussion

From the classification results, it can be observed that the performance of SWT features is comparatively better than the DWT features. As the main goal of the study is to find the effectiveness and efficiency of the features as a whole, instead of specific and targeted features, we did not opt for either feature ranking or feature selection of available features, the results of which have been present in the previous sections. It is also important to note that, the eCDF plots of our feature sets obtained using both DWT and SWT, presented a good indication of the nature of features obtained using each of the techniques. For easy analysis of the provided results, we argue with respect to the distribution of data for normal and benign images obtained using both techniques. Though at first glance the distributions look quite similar, it is evident that SWT provided a much better differentiation with respect to DWT features. This is also seen numerically, from Table I, where the features of SWT provided a much better range of differentiation when compared to DWT features. It is seen that SWT features provided much higher classification accuracy when compared to DWT features. The combination

of SWT features with other textural feature extraction techniques might help in opening up a whole new set of texture features which would provide a good tool for advancing state of the art CAD systems.

Also, the analysis of data in a different coordinate system with respect to the original plane of reference in itself would help in observing minute changes and information, which might otherwise be invisible in a CAD pipeline. This, in our opinion is a very important contribution of the current work. Another vital point to consider is the computational time of the algorithm. The time taken for computation of DWT and SWT is almost the same. It is common knowledge that DWT is one of the most efficient algorithms in terms of real-time implementations, owing to its speed and accurate execution in real-time applications. So, with the time taken for computation being similar for both DWT and SWT, and with SWT having a better classification accuracy compared to DWT, SWT seems to be a viable alternative to DWT or even a better alternative, in cases where wavelets have to be used. From Table I, it can also be seen that the change in mean and standard deviation values between normal, benign and malignant

Table III

A listing of classification methods close to the current study.

Authors	Method used	Accuracy (%)
Kimme et al. [15]	Normalized statistics and texture features	74
Petrosian et al. [16]	Spatial gray level dependence and textural features with a decision tree classifier	76-89
Wei et al. [17]	Statistical features in a multiple view mammogram with SVM and KFD	85
Mudigonda et al. [18]	Gray level co-occurrence matrices, polygonal modeling with jack-knife classification	83
Wei et al. [19]	Statistical features in a multiple view mammogram with SVM and KFD	85
Szekely et al. [7]	Texture features and a combining classifier of decision trees and multiresolution Markov random models	88-94
Alolfe et al. [20]	Forward stepwise linear regression method with a combined classifier of SVM and LDA	82.5-90
Present study	SWT features and SVM	88.80

images are much more pronounced in SWT features compared to DWT features. This shows that subtle changes are captured in a better way using SWT features.

With respect to classifiers, it is very important to note that classifiers play a vital role in our conclusions since SVM and QDC are the only classifiers which provided good results with other classifiers being quite poor in performance. But this discrepancy in classifiers can easily be explained by looking at the data in consideration, which is highly non-linear. With the other classifiers such as LDC, NMC and ParzenC being not effective in classifying non-linear data, SVM and QDC perform reasonably well due to their inherent properties to adapt to non-linearity. Moreover, SVMs do not classify non-linear data as such, but rather chooses a mapping function in the form of a kernel, in order to map itself into a linear space where the classifier can perform better (14). In this case, a radial basis kernel function was used.

Some key works which are closely related to our current study and classification accuracies obtained is shown in Table III.

### Conclusion

A comparative study has been conducted to find the functional variations and differences in the efficiency between SWT and DWT methods. SWT has been found to perform better than DWT in capturing subtle differences in mammograms, as evident from the eCDF of the feature sets

presented. We have also shown that SWT features in combination with SVM radial basis kernel gave a maximum accuracy of 88.80%, while DWT features in combination with SVM radial basis kernel yielded a maximum accuracy of 81.73%, compared to the other classifiers. The usage of SWT for medical image analysis is yet to be reported in literature with very limited application being seen in image sharpening. Given the fact that SWT performed better than DWT for medical image classification as seen from our study, there is potential for extension of this technique to other imaging modalities for application in CAD.

### Acknowledgements

Authors thank SATA CommHealth, Singapore for providing grant for this research.

### References

1. Christoyianni I, A Koutras A, Dermatas E & Kokkinakis, G. Computer aided diagnosis of breast cancer in digitized mammograms. *Computerized Medical Imaging and Graphics* 26(5), 309-319 (2002). DOI: 10.1016/S0895-6111(02)00031-9
2. Heang-Ping C, Sahiner B, Helvie MA, Petrick N, Roubidoux MA, Wilson TE, Adler DD, Paramagul C, Newman JS & Sanjay-Gopal, S. Improvement of radiologists' characterization of mammographic masses by using computer-aided diagnosis: an ROC study. *Radiology* 212, 817-827 (1999).
3. Doi K, MacMahon H, Katsuragawa S, Nishikawa RM & Jiang, Y. Computer-aided diagnosis in radiology: potential and pitfalls. *European Journal of Radiology* 31, 97-109 (1999). DOI: 10.1016/S0720-048X(99)00016-9.
4. Balleyguier C, Kinkel K, Feranian J, Malan S, Djen G, Taourel P & Helenon O. Computer-aided detection (CAD) in mammography: does it help the junior or the senior radiologist? *European Journal of Radiology* 54, 90-96 (2005). DOI: 10.1016/j.ejrad.2004.11.021
5. Fenton JJ, Taplin SH, Carney PA, Abraham L, Sickles EA, D'Orsi C, Berns EA, Cutter G, Hendrick RE, Barlow WE & Elmore JG. Influence of computer-aided detection on performance of screening mammography. *N Engl J Med* 356, 1399-1409 (2007). DOI: 10.1056/NEJMoa066099
6. Sadaf A, Crystal P, Scaranelo A & Helbich T. Performance of computer-aided detection applied to full-field digital mammography in detection of breast cancers. *Eur J Radiol* 77, 457-461 (2009). DOI: 10.1016/j.ejrad.2009.08.024
7. Szekely N, Toth N & Pataki B. A hybrid system for detecting masses in mammographic images. *Instrumentation and Measurement. IEEE Transactions* 55, 944-952 (2006). DOI: 10.1109/TIM.2006.870104
8. Ganesan K, Acharya UR, Chua CK, Lim CM, Abraham KT & Ng K-H. Computer-aided breast cancer detection using mammograms: a review. *Biomedical Engineering, IEEE Reviews* 6, 77-98 (2013). DOI: 10.1109/RBME.2012.2232289
9. Starck JL, Moudden Y, Abrial P & Nguyen M. Wavelets, ridgelets and curvelets on the sphere. *Astronomy and Astrophysics* 446 1191-1204 (2006). DOI: 10.1051/0004-6361:20053246
10. Kezele I, Descoteaux M, Poupon C, Poupon F & Mangin JF. Spherical wavelet transform for ODF sharpening. *Med Image Anal* 14(3), 332-342 (2010). DOI: 10.1016/j.media.2010.01.002; Epub 2010 Feb 1.
11. Chih-Chin L & Cheng-Chih T. Digital image watermarking using discrete wavelet transform and singular value decomposition. *IEEE Trans Instrum Meas* 59(11), 3060-3063 (2010). DOI: 10.1109/TIM.2010.2066770

12. Ganesan K, Acharya RU, Chua CK, Lim CM & Abraham TK. Decision support system for breast cancer detection using mammograms. *Proceedings of the Institution of Mechanical Engineers, Part H: Journal of Engineering in Medicine* (2013). DOI: 10.1177/0954411913480669
13. Hui Y, Jie G, Hua GZ & Chuan L. An efficient method to process the quantized acoustoelectric current: Wavelet transform. *IEEE Trans Instrum Meas* 60(3), 696-702 (2011). DOI: 10.1109/TIM.2010.2045549
14. Duda R & Hart P. *Pattern Classification and Scene Analysis*. John Wiley & Sons, New York (1973).
15. Kimme C, O'Loughlin BJ & Sklansky J. Automatic Detection of Suspicious Abnormalities in Breast Radiographs, in *Data Structures, Computer Graphics, and Pattern Recognition*, Klinger A, Fu KS, Kunii TL (Eds.), New York: Academic Press, 427-447 (1977).
16. Petrosian A, Chan HP, Helvie MA, Goodsitt MM & Adler DD. Computer-aided diagnosis in mammography: classification of mass and normal tissue by texture analysis. *Phys Med Biol* 39, 2273-2288 (1994). DOI: 10.1088/0031-9155/39/12/010
17. Wei L, Yang Y, Nishikawa RM & Jiang Y. A study on several Machine-learning methods for classification of malignant and benign clustered microcalcifications. *Medical Imaging, IEEE Transactions on* 24(3), 371-380 (2005). DOI: 10.1109/TMI.2004.842457
18. Mudigonda NR, Rangayyan RS & Desautels JEL. Gradient and texture analysis for the classification of mammographic masses. *Medical Imaging IEEE Transactions on* 19(10), 1032-1043 (2000). DOI: 10.1109/42.887618
19. Wei L, Yang Y, Nishikawa RM & Jiang Y. A study on several machine-learning methods for classification of malignant and benign clustered microcalcifications. *Medical Imaging, IEEE Transactions on* 24(3), 371-380 (2004). DOI: 10.1109/TMI.2004.842457
20. Alolfe MA, Mohamed WA, Youssef ABM, Mohamed AS & Kadam YM. Computer aided diagnosis in digital mammography using support vector machine and linear discriminant analysis classification. *Image Processing (ICIP), 2009 16th IEEE International Conference on*, 2609-2612 (2009). DOI: 10.1109/ICIP.2009.5413992
21. Theodoris S & Koutroumbas K. *Pattern Recognition*, 3rd Edition, Academic Press (2006). ISBN: 13:978-0-12-369531-4
22. Devroye L, Györfi L & Lugosi G. *A Probabilistic Theory of Pattern Recognition*. Springer-Verlag (1996). ISBN-13: 978-0387946184
23. Webb AR. *Statistical Pattern Recognition*. 2nd Edition, Wiley & Sons (2002). ISBN: 0470845147
24. David Sanchez VA. Advanced support vector machines and kernel methods. *Neurocomputing* 55(1-2), 5-20 (2003). DOI: 10.1016/S0925-2312(03)00373-4

Received: March 14, 2013; Revised: May 2, 2013;

Accepted: May 6, 2013