Comparison of Spherical Wavelet Transform (SWT) and Discrete Wavelet Transform (DWT) features on Mammographic Images

Sushma S[1], Latha KC[2], Balasubramanian S[3], Sridhar R[4]

Abstract One of the most widely used technology to detect breast cancers used in the primary diagnosing stage is mammograms. Detection of cancer in initial stages can increase the probability of prolong patient life and better recovery. Thus, there is high demand for early identification and diagnosis of breast cancer with the help of mammograms. To increase the accuracy of diagnosis and image interpretation consistency of healthcare professionals, Cad is introduced in the field of radiology. Texture based feature extraction strategies are commonly used for analysis of mammograms. To be particular, wavelets are a favorable choice to texturally analyze the image. For this purpose, previously discrete wavelets have been utilized, but spherical wavelets have hardly been utilized for Computer-Aided Diagnosis (CAD) of breast cancer with the help of mammographic images. In this study, a comparative analysis of the performance between the features of Spherical Wavelet Transform (SWT) and Discrete Wavelet Transform (DWT) on the basis of three classification results of malignant, benign and normal stage was studied. Classification was done with the help of Parzen Classifier (ParzenC), Support Vector Machines (SVM), Nearest Mean Classifier (NMC), Quadratic Discriminant Classifier (QDC), and Linear Discriminant Classifier (LDC). The maximum achieved classification accuracy is 89.90% for SWT and 82.85% for DWT features with combination of SVM classifier.

Keywords Breast cancer · SWT feature · DWT feature · Mammographic images

1. Introduction

One of the most prevalent non-skin cancer and the leading fatality among women and is breast cancer (DeSantis et al. 2013). In fact, among all cancer deaths in women, breast cancer mortality is second highest (Hashemi et al. 2014). At present time, there is no medication technology that have the potential to cure cancer. However, it is a well-known fact that detection of cancer in initial stages can increase the probability of prolong patient life and better recovery (Scharl et al. 2015). That is the reason why doctors, oncologists, and related health care professionals want to detect breast cancer as early as possible. Nevertheless, there are numerous issues associated with early detection of the disease. The first tool that is used to detect breast cancer is mammogram. There is a unique issue associated with this diagnosing tool because of inter-observer variations that occur when diagnosing breast cancer via mammograms (Masroor et al. 2016). To overcome this problem, Computer Aided Diagnosis (CAD) is introduced in the field of radiology. The purpose of introducing CAD in the field of radiography is to increase the accuracy of diagnosis and image interpretation consistency of healthcare professionals with the help of computer output as a guidance (Litjens et al. 2015). This is quite possible because a radiologist interprets the mammogram on the basis of judgments that are subjective to the radiologist. Moreover, it also helps in diagnosing the masses and microcalcifications that are usually missed by radiologists because they are not clear in mammograms. Furthermore, it has been established that inter-observer and intra-observer variability are significant factors in determining the accuracy of diagnosis (Masroor et al. 2016). Various studies have proven improvement and positive influence in diagnostic accuracy of radiologists when they used CAD system (Litjens et al. 2015).

In this study, the classification that is used is illustrated in Figure 1. Although CAD system is in its rudimentary, evidence have supported the use of CAD system by radiologists with noticeable success rate (Abbas 2016). Numerous studies have been conducted in the past to investigate the effectiveness of combination of various techniques. In the work of Murakami, the overall accuracy rate was over 90 percent, it can be said that the CAD systems are at advanced stage for breast cancer (Murakami et al. 2013). Nonetheless, there is still room for improvement because of several factors. For example, the accuracy rates were inspiring in the study of Murakami et al. (Murakami et al. 2013), but it must be noted also that high accuracy was achieved in the experiments which used Full-Field Digital Mammograms (FFDM). The highest accuracy rate was 100% for microcalcification manifestations. However, the accuracy declined with other types of mammographic appearances and the density of breasts (Murakami et al. 2013). Likewise, Uppal (2016) also conducted a study by using more complicated strategy including Genetic Programming (GP) based filter, with the combination of DWT features and Discrete Cosine Transform (DCT) to be used in a CAD system. This technique achieved the accuracy rate of 96.97%. Hence, we can see that there are various techniques ranging from very simple to very complex (Uppal 2016).

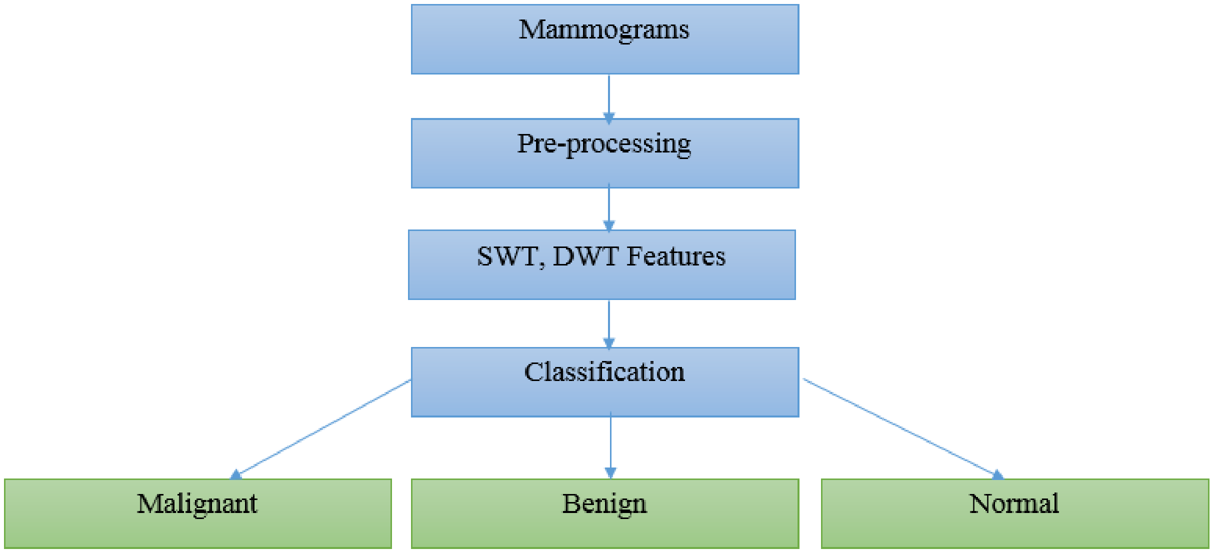


Figure 1: Framework used in this study

Since, it is clear that there is a scope for further progression in the primary classification system, particularly in the phase of feature extraction, we put forward the application of a new methodology related to feature extraction to be use in analysis of mammographic images. Initially, Spherical Wavelet Transforms (SWT) was introduced for data analysis and astronomical image, their potential to retrieve minute information make it appropriate, specifically for our application. In popular applications, the common use of SWT is yet to be achieved, particularly in the analysis of medical image. In this field, they might be very beneficial knowing that they are extremely useful in sharpening the images and filtering off the noises in addition to the provision of vital information about details that might not be accessible by using Discrete Wavelet Transforms (DWT).

This study elaborated the advantages of SWT in contrast to DWT features with the help of mammographic images. Three classes of mammograms are used; malignant, benign, and normal. In this paper, we consider the features that are extracted by SWT and DWT. For both these sets, the classification was done distinctly for assessing the performance of SWT and DWT. The classification was done with the help of Parzen Classifier (ParzenC), Support Vector Machines (SVM), Nearest Mean Classifier (NMC), Quadratic Discriminant Classifier (QDC), and Linear Discriminant Classifier (LDC).

2. Materials and Methods

During the experiment, 67 images or mammograms were used. The images were of patients aged between 46-75 years. There were 37 normal images, 14 benign images, and 16 malignant images. Each image was normalized and preprocessed to counter disparities in imaging conditions before any processing. For achieving quantitative results with greater accuracy, the pectoral muscles were removed from all mammograms manually so that only masses of breast tissues are visible. The mammographic images processed at 1024 X 1680 resolution. Every single image was processed at Mediolateral-oblique (MLO) view and CranioCaudal (CC) view. Figure 2 presents some samples of mammographic images utilized in this research.

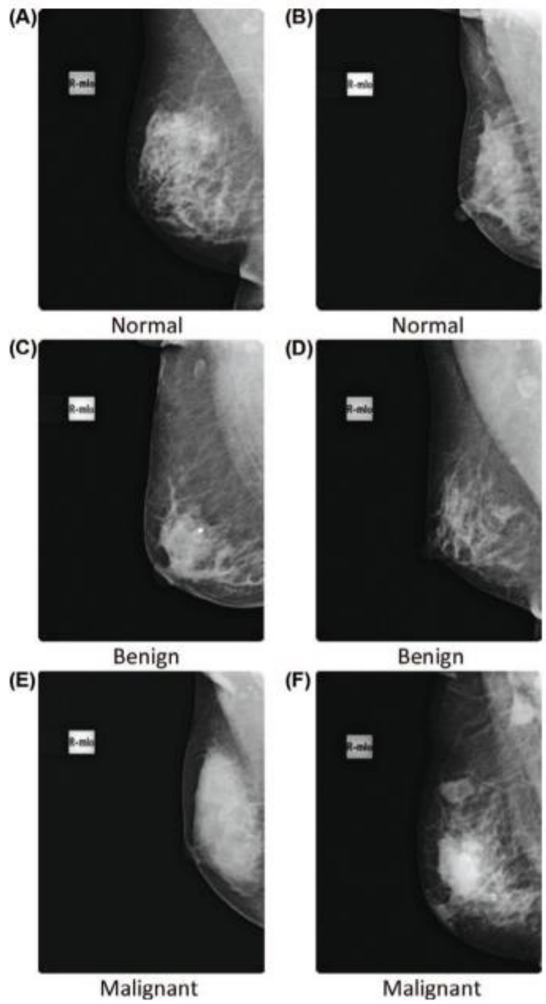


Figure 2: Sample mammograms

2.1 Feature Extraction

To extract texture features from the mammographic images to be studied, two feature extraction techniques are utilized in this study, namely Spherical Wavelet Transform (SWT) and Discrete Wavelet Transform (DWT).

2.2 Spherical Wavelet Transform (SWT)

The Spherical Wavelet Transform operates on the phenomenon that original data can be replicated by sum of the scales. The purpose of this technique is to decrease the redundancy that is usually present within the traditional DWT. This redundancy results in large data sets due to which averaging strategies are applied, leading to loss of data. An undecimated isotropic transform was used to develop SWT. This isotropy has the ability to statistically capture isotropic features in an isotropic field; thus, it is used in SWT development and is advantageous for creating a wavelet pyramid. In medical images' textural feature extraction, this characteristic of statistically capturing isotropic features can be put to a worthy use, as medical image analysis utilizing textures is centered on isotropic region extraction.

On a dyadic resolution, the estimates of a mammographic image (I) can be acquired with the help of scale function ϕl_c as; $c_0 = \phi l_c * f$, c1=ϕ2−1lc*f,....,cj=ϕ2−jlc*f, where ϕl_c have a cut-off frequency $2^{-j}\, l_c$ and ϕ2−jlc is its rescaled version (Ganesan et al. 2013). With the help of this function, for each scale j, a low-pass filter h_j is described by;

$$\widehat{H}_j(l,m) = \sqrt{\frac{4\pi}{2l+1}}\,\widehat{h}_j(l,m)$$

Correspondingly, on each scale j, *a* high-pass filter *gj* can be defied by;

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

From above two equations, it is well-established that the filter used for this study can be illustrated as;

$$\widehat{G}_j(l,m) = 1 - \sqrt{\frac{4\pi}{2l+1}}\,\widehat{h}_j(l,m) = 1 - \widehat{H}_j(l,m)$$

The above equation demonstrated a model of wavelet function which can be utilized in spherical domain; whereas, to find the approximate and detailed coefficients, any other wavelet function can substitute it. An implementation of SWT is shown in Algorithm 1, illustrated in Figure 3.

Algorithm 1:	Implementation of the SWT
Step 1:	Compute a multiresolution sphere.
Step 2:	Computer 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.

Figure 3: Implementation of SWT algorithm

2.3 Discrete Wavelet Transform (DWT)

It is a well-known technique used for textural feature extraction. Considering this technique, images were passed through an array of down-sampling filters. These down-sampling filters are consisted of a series of low-pass and high-pass filters. The low-pass filters generate the coefficients of approximation *A*[n], whereas high-pass filters generate the detail coefficients *D*[*n*] (Sharma and Jain 2014). Mathematically, these coefficients are presented as;

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

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

Where, *g[2n-k]* denotes low-pass filter's transfer function, *h[2n-k]* denotes high-pass filter's transfer function, and *x[k]* characterizes the under considered image.

In this experiment, biorthogonal wavelets were utilized (Sudarshan et al. 2015). In these wavelets, transformation of wavelet is invertible, however, it is not essentially orthogonal. In contrast to orthogonal wavelets, biorthogonal wavelets have more degrees of freedom. The wavelet's first level generates an approximation coefficients A_1, a vertical Dv_1, diagonal Dd_1, and horizontal Dh_1 detailed coefficients (Sharma and Jain 2014). Since the number of elements are too high in the output of these matrices; for this reason, these outputs cannot be utilized directly for computation. Accordingly, to reduce dimensionality, averaging techniques have been derived as follows;

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

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

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

Finally, the averaging method utilized not the values of intensity but averages; however, it averages the intensity values' energy which can be defined as follows;

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

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

3. Classification

3.1 Linear Discriminant Classifier (LDC)

The underlying concept of linear classifier is that each of the object present in the sequential pattern $x_1, x_2, \ldots, x_n$ is allocated to a class ω_1 or ω_2 which is based on threshold τ_0. This classifier can be mathematically presented as;

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

Where, ω^T denotes the weight vector. all linear classifiers principally work on this concept. For all the samples in a single class, when $\omega^T x + \omega_0 > 0$ then the data is regarded as linearly separable (Tharwat 2016).

3.2 Nearest Mean Classifier (NMC)

It is one of the most effective and simplest classifiers. Due to its computational efficiency, it is simple, as it requires very little effort for computing the mean of the under consideration classes. Every object present in the dataset is allocated to one of the classes, as soon as the mean of the classes are figured at random (Poudel et al. 2013). Although, when the data is well spread, NMC has found to be nominal, the data present in our research that is non-linear does not give a good result because of the overlap between the means of numerous classes under consideration.

3.3 Parzen Classifier (ParzenC)

These algorithms are based on the non-parametric estimation technique. In this technique, the total histogram of a provided feature set is divided into a number of bins and estimate the probability of a random sample associated with one specific bin. Parzen classifiers are based on this classification process modeling the data into a multidimensional scale. This means that rather than division of histogram into numerous bins, similar to the usual non-parametric techniques, the *n*-dimensional space is separated into hyperccubes with a volume h^l and side h (Lesniak 2012). In this situation, the probability $p\hat{}(x)$ of a single variable related to specific hypercube can be described by;

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

Where, $\phi(x_i) = 1$ if $x_i < ½$, whereas $\phi\,(x_i) = 0$ for any other value of x_i, and $x_i = 1, 2….l$ are the available feature vectors (Tharwat 2016).

3.4 Quadratic Discriminant Classifier (QDC)

It is commonly assumed in a classification problem that data can be described with the help of Gaussian distribution in each class. Considering this assumption, a conclusion is drawn about the classifier's quadraticity and linearity, which depends on the data's covariance matrices. When the covariance matrices found to be different, a Quadratic Discriminant strategy is followed; whereas Linear Discriminant strategy is followed if the covariance matrices become equal (Tharwat 2016). In our research, the insinuation of this approach can be observed clearly, as LDC underperforms QDC. This represent that there is more scope for linear classifiers as the data is not linearly separable.

3.5 Support Vector Machine (SVM)

It is one of the most commonly used classifiers, while LDS are the driving force of SVMs. For SVMs, the separations by hyperplanes can be represented as;

$$\omega^T x_i + \tau_0 \geq +1, \forall\ x\ \varepsilon\ \omega_1$$
$$\omega^T x_i + \tau_0 \leq -1, \forall\ x\ \varepsilon\ \omega_2$$

Where, τ_0 represents threshold determined by above equation, ω denotes a weight vector, while ω1 and ω2 are the classes under consideration (Suykens et al. 2014). From the above equation, it can be observed that it is an extension of the linear discriminant classifier. In a way, it is different that the hyperplanes as a substitute of SVMs operate on the support vectors theory. At this point, the location of the hyperplane can be determined by the vectors closest to the separating hyperplanes. Additionally, in the case of non-linear data, the performance of DVMs is appropriate because in the original data space, these classifiers do not discrete the data. Using specific function such as kernels, it maps the original data into manageable space, which aid in transforming non-linear data into linear data with the help of kernel functions (Suykens et al. 2014). In this research, a radial basis kernel is utilized for building our classifier. The kernel used in this study is described as;

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

Where, *z* and *x* represent the mean and object of the classes respectively, while σ denotes standard deviation of the class. Binary classifiers were used in this research, and a majority vote was utilized for extending it into a structure that is multi-class classification.

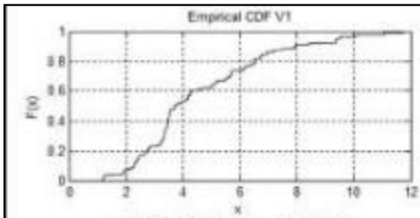
4. Results

This study provides a quantitative and comparative analysis of the utilization of SWT in contrast to DWT. 67 mammographic images were used in this study, with analogous processing performed on all the images. To extract the features, all the levels of sub-bands SWT and DWT were utilized. Feature selection or ranking was not done because it is a well-established fact that wavelet transformation‘s all sub-bands are significant, since information absent in a band is found in another band. For both SWT and DWT, this characteristic is true. Table A shows a sample of features for the mammograms three classes for SWT and DWT features.

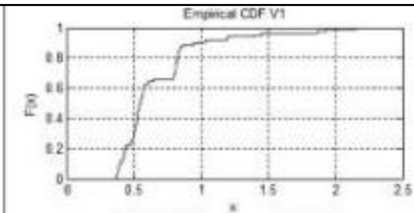
Features	Normal	Benign	Malignant
DWTA1	2.4E+02 ± 1.2E+02	1.7E+02 ± 1.04E+02	2.6E+02 ± 1.5E+02
DWTD1	1.9E-03 ± 1.3E+01	-3.8E+01 ± 1.19E+01	2.2E-06 ± 1.3E+01
DWTH1	1.0E-02 ± 3.4E+01	1.1E+00 ± 3.26E+01	9.0E-04 ± 3.5E+01
DWTH2	5.0E-02 ± 1.0E+02	1.9E+00 ± 8.23E+01	1.8E+00 ± 8.9E+02
DWTV1	-2.0E-02 ± 4.3E+01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

Table A: Extracted significant features (Mean ± SD). For all the features, the p-value is < 0.0001

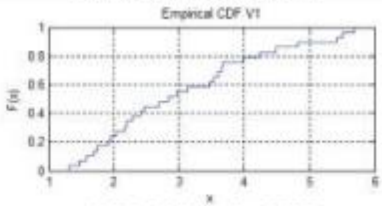
The features' data distribution that are obtained by SWT and DWT can be studied with the help of empirical Cumulative Distributive Functions (eCDF), presented in Figure 5. This technique of visualization has aided in better interpretation of variations in the features extracted by SWT and DWT.



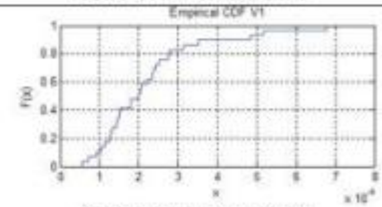
eCDF of V1 normal DWT



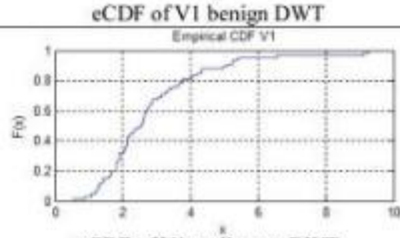
eCDF of V1 normal SWT



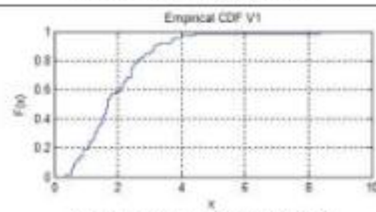
eCDF of V1 benign DWT



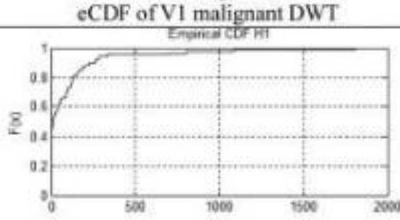
eCDF of V1 benign SWT



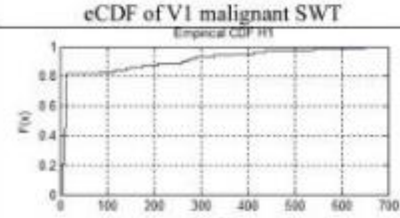
eCDF of V1 malignant DWT



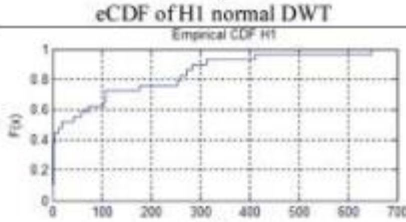
eCDF of V1 malignant SWT



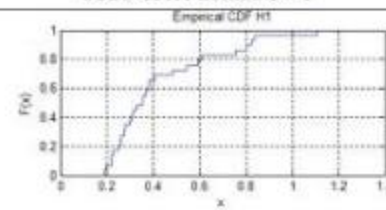
eCDF of H1 normal DWT



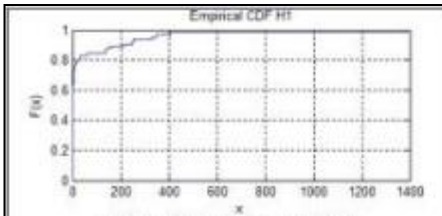
eCDF of H1 normal SWT



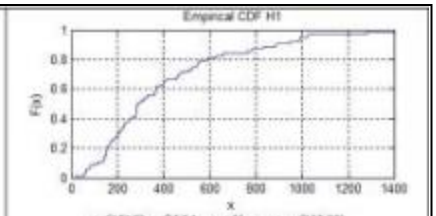
eCDF of H1 benign DWT



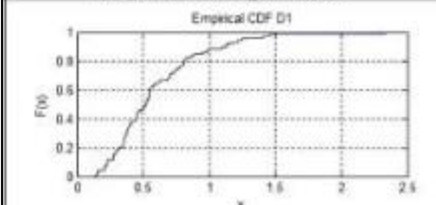
eCDF of H1 benign SWT



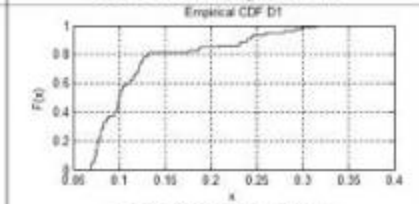
eCDF of H1 malignant DWT



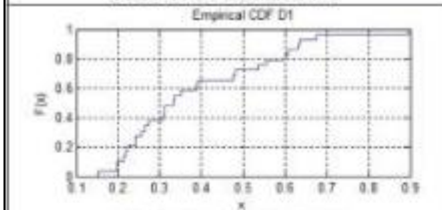
eCDF of H1 malignant SWT



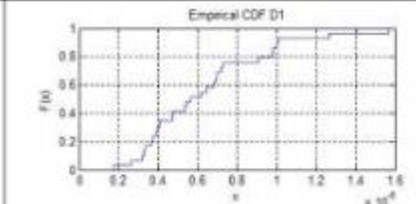
eCDF of D1 normal DWT



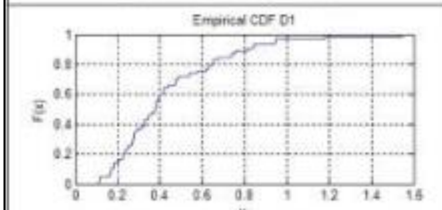
eCDF of D1 benign SWT



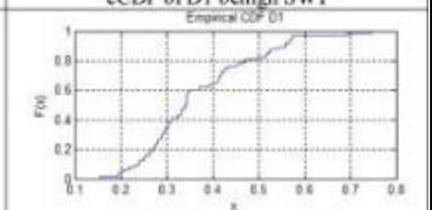
eCDF of D1 benign DWT



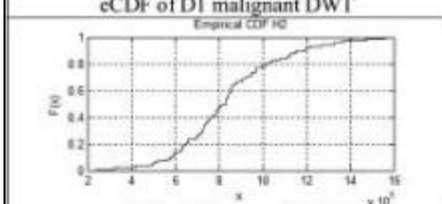
eCDF of D1 benign SWT



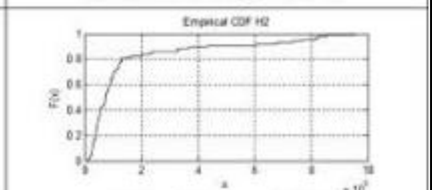
eCDF of D1 malignant DWT



eCDF of D1 malignant SWT



eCDF of H2 normal DWT



eCDF of H2 normal SWT

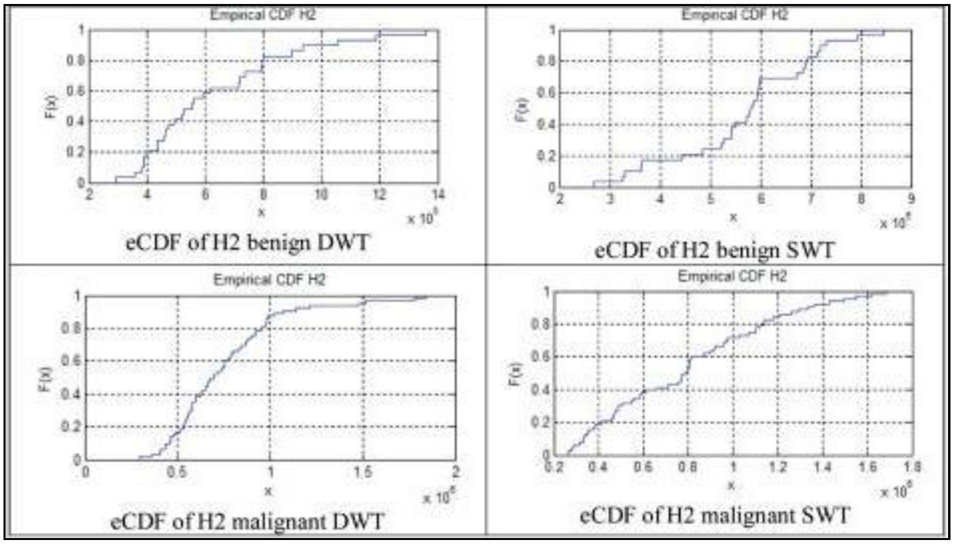


Figure 4: Plots of Empirical Cumulative Distributive Functions (eCDF) for different images

As this research emphasizes on a comparison of performance between SWT and DWT; thus, no other techniques were utilized. By using DWT features in combination with SVM radial basis classifier, 82.85% optimal classification accuracy was achieved. Likewise, with the help of SWT features in combination with SVM radial basis classifier, an optimal accuracy of 89.90% was attained. Using ten-fold cross validation, the performance of all the other classifiers in average can be observed ion Table B.

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]

In Figure 6, the application of SWT's decomposition sub-bands on a sample image can be observed. Furthermore, the results of all the ten-fold associated with the cross validation scheme are presented in Figure 7 for DWT and in Figure 8 for SWT.

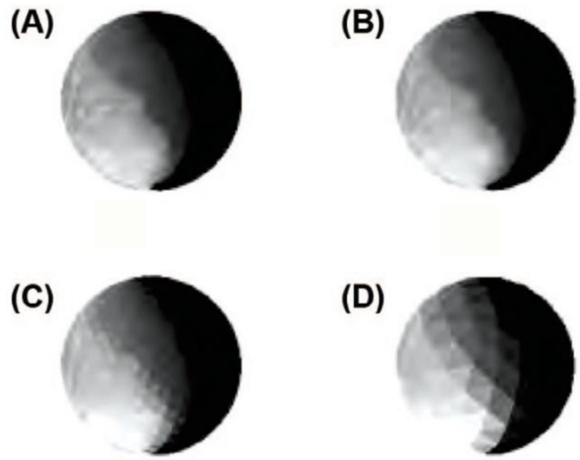


Figure 5: SWT decomposition levels, (A) actual image (B, C, D) successive 3 decomposition levels

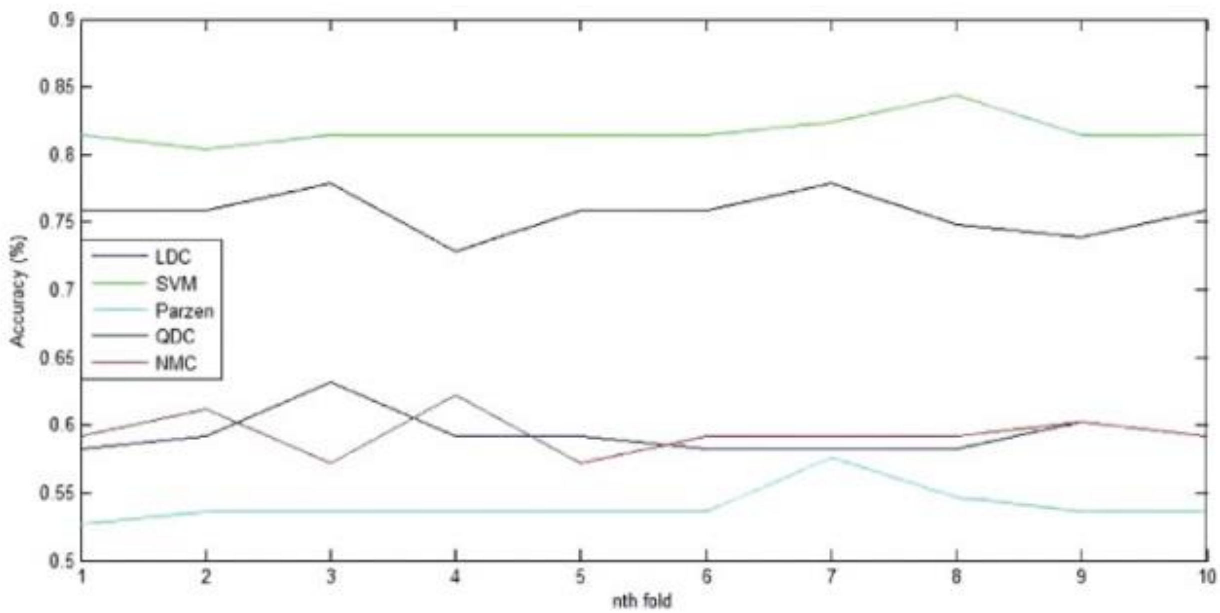


Figure 6: DWT's classification accuracy by ten-fold

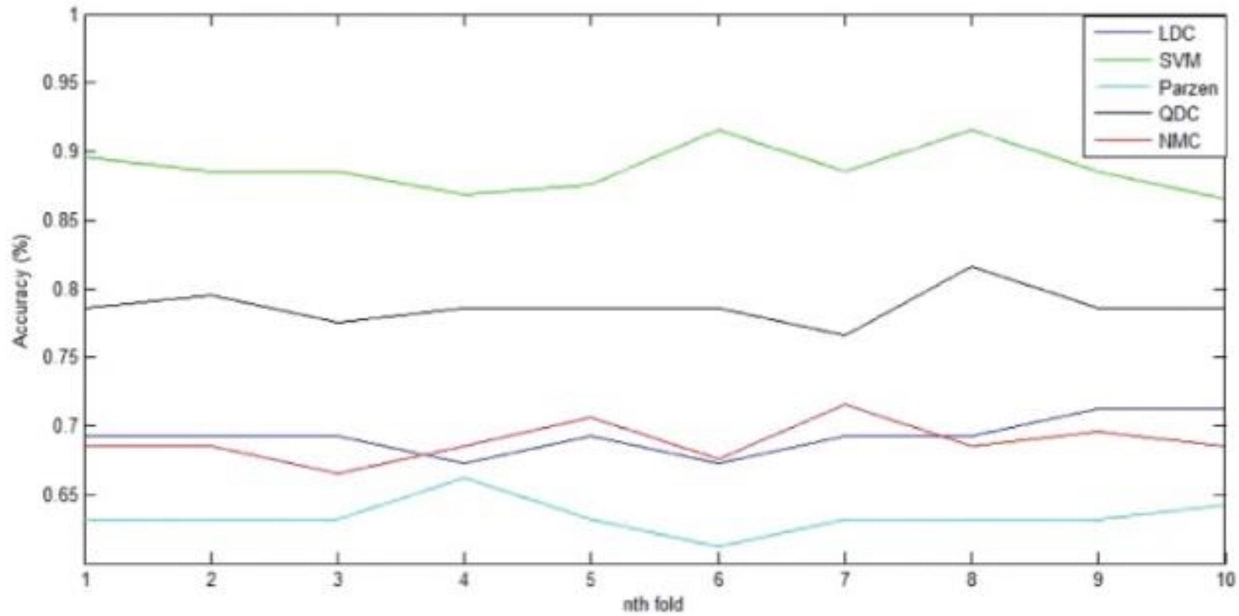


Figure 7: SWT's classification accuracy by ten-fold

5. Discussion

Given the results regarding classification, it is evident that the performance and accuracy level of SWT is better in contrast to DWT features. Since the main purpose of the study is to determine the efficiency and effectiveness of the features overall, rather than the targeted and specific features; therefore, we did not select either feature selection of feature ranking for the available features, whose results are presented in the earlier sections. It is significant to recognize that the pots of eCDF of the feature sets extracted utilizing both SWT and DWT, illustrated a good sign about the nature of the features acquired using each technique. We discuss with respect to the data distribution for benign and normal images, for easy analysis of the results, attained by using both techniques. At first glance, no difference can be observed in the distributions, but it is proved that the differentiation provided by SWT is better than the DWT features. This is also observed in figures, from Table A. in the table, the SWT features offered a better differentiation range in contrast to DWT features. It is also observed that the classification accuracy of SWT features is higher in comparison to the DWT features. SWT in combination with other techniques of other textural feature extraction might aid in developing an innovative set of texture features that would be good step for advancement of the CAD system.

Moreover, the data analysis in a different coordinate system pertaining to the actual reference plane would assist in identifying the minute modifications and information that might not be observable otherwise in a CAD system. In our view, this is a significant contribution of the present work. In addition, the algorithm's computational time is a vital point to consider. The time required for computation of SWT and DWT is practically same. It a well-known fact that DWT is an efficient algorithm with regard to real-time implementations due to the accurate execution and speed in real-time applications.

Consequently, with SWT having superior classification accuracy in contrast to DWT, and with time taken for calculation being almost same for SWT and DWT, SWT appears to be a feasible substitute to DWT. Even it can be said that SWT is a better option in comparison to DWT, particularly when wavelets have to be used. It can also be observed from Table A that the change in standard deviation and mean values between malignant, benign, and normal mammograms are much more differentiable in SWT features in contrast to DWT features. This illustrated that small changes are better identified by utilizing SWT features.

Considering the classifiers, it is very considerable to note that classifiers have a major role in drawing our conclusion, as QDC and SVM are the only classifiers that achieved good results while other classifiers are relatively poor in performance. However, this inconsistency in classifiers can be described through observing the data under consideration, which was non-linear. That is the reason why ParzenC, NMC, and LMC were not effective because they are utilized or linear data, whereas QDC and SVM perform much better because of their properties to adjust to non-linear data. Furthermore, SVMs do not have the ability to classify non0-liniear data, rather this technique selects a mapping function that is kernel, so that it can map itself into a linear space in which the classifier can carry out its function better (Suykens et al. 2014). A radial basis kernel function was applied to the data in this study.

6. Conclusion

A comparable research has been conducted for finding the functional differences and variations in the effectiveness between SWT and DWT techniques. In contrast to DWT, SWT has been found to perform better in terms of identifying or capturing delicate differences in mammographic images, as proved from the plots of eCDF of Feature sets. We have also proposed that combination of SWT and SVM radial basis kernel provided a maximum accuracy of 89.90%, whereas the combination of SVM radial basis kernel and DWT features gave a maximum accuracy of 82.85%, in comparison to other classifiers. The utilization SWT for analysis of medical images is in tis ini9tial stages and yet to be documented in literature with restricted application in sharpening of the image. Considering the fact that SWT has performed better when compared with DWT for classification of medical images as illustrated in the study, there is a possibility of extending this technology to other modalities of medical images for application in CAD.

References

DeSantis C, Ma J, Bryan L, Jemal A. Breast cancer statistics, 2013. CA: a cancer journal for clinicians. 2014 Jan 1;64(1):52-62.

Hashemi SH, Karimi S, Mahboobi H. Lifestyle changes for prevention of breast cancer. Electronic physician. 2014 Jul;6(3):894.

Scharl A, Kühn T, Papathemelis T, Salterberg A. The right treatment for the right patient– personalised treatment of breast cancer. Geburtshilfe und Frauenheilkunde. 2015 Jul;75(07):683-91.

Masroor I, Rasool M, Saeed SA, Sohail S. To Assess Inter And Intra-Observer Variability For Breast Density And Birads Assessment Categories In Mammographic Reporting. JPMA. 2016;66(194).

Litjens GJ, Barentsz JO, Karssemeijer N, Huisman HJ. Clinical evaluation of a computer-aided diagnosis system for determining cancer aggressiveness in prostate MRI. European radiology. 2015 Nov 1;25(11):3187-99.

Abbas Q. DeepCAD: A Computer-Aided Diagnosis System for Mammographic Masses Using Deep Invariant Features. Computers. 2016 Oct 31;5(4):28.

Murakami R, Kumita S, Tani H, Yoshida T, Sugizaki K, Kuwako T, Kiriyama T, Hakozaki K, Okazaki E, Yanagihara K, Iida S. Detection of breast cancer with a computer-aided detection applied to full-field digital mammography. Journal of digital imaging. 2013 Aug 1;26(4):768-73.

Uppal MT. Classification of mammograms for breast cancer detection using fusion of discrete cosine transform and discrete wavelet transform features. Biomedical Research. 2016;27(2).

Ganesan K, Acharya UR, Chua CK, Min LC, Abraham KT, Ng KH. Computer-aided breast cancer detection using mammograms: a review. IEEE Reviews in Biomedical Engineering. 2013;6:77-98.

Sharma P, Jain T. Robust digital watermarking for coloured images using SVD and DWT technique. In Advance Computing Conference (IACC), 2014 IEEE International 2014 Feb 21 (pp. 1024-1027). IEEE.

Sudarshan V, Acharya UR, Ng EY, Meng CS, San Tan R, Ghista DN. Automated identification of infarcted myocardium tissue characterization using ultrasound images: a review. IEEE reviews in biomedical engineering. 2015;8:86-97.

Tharwat A. Linear vs. quadratic discriminant analysis classifier: a tutorial. International Journal of Applied Pattern Recognition. 2016;3(2):145-80.

Poudel A, Kanneganti R, Gupta L, Chu TP. Nearest mean classifier for defect classification in CFRP panels. InASNT 22nd Research Symposium 2013. 2013 Mar 18 (pp. 55-60).

Lesniak JM, Hupse R, Blanc R, Karssemeijer N, Székely G. Comparative evaluation of support vector machine classification for computer aided detection of breast masses in mammography. Physics in medicine and biology. 2012 Aug 1;57(16):5295.

Suykens JA, Signoretto M, Argyriou A, editors. Regularization, optimization, kernels, and support vector machines. Chapman and Hall/CRC; 2014 Oct 30.