

HYBRID DISCRETE WAVELET TRANSFORM AND TEXTURE ANALYSIS METHODS FOR FEATURE EXTRACTION AND CLASSIFICATION OF BREAST DYNAMIC THERMOGRAM SEQUENCES

Khaleel Al-Rababah¹, Mas Rina Mustaffa^{2}, Shyamala C. Doraisamy³, Fatimah Khalid⁴,
Luís Filipe de Pina Júnior⁵*

^{1,2,3,4}Multimedia Department, Faculty of Computer Science and Information Technology, Universiti Putra Malaysia, 43400 UPM Serdang, Selangor Darul Ehsan, Malaysia

⁵Computer Engineering Department, Faculty of Engineering, Universidade Lúrio, Bairro Eduardo Mondlane, 3200, Pemba, Cabo Delgado, Mozambique

Email: gs46118@student.upm.edu.my¹, MasRina@upm.edu.my^{2*}(corresponding author), shyamala@upm.edu.my³, fatimahk@upm.edu.my⁴, luis.pina@unilurio.ac.mz⁵

DOI: <https://doi.org/10.22452/mjcs.sp2021no2.8>

ABSTRACT

Breast cancer is a common cancer that hits women causing thousands of casualties every year. A cancerous tumor causes an increase of temperature near the region of the tumor. The heat generated by the temperature transferred to the skin surface. The temperature in the tumor area is warmer than in the healthy area. Detecting breast cancer in early stages can save women's lives and lower the burden on the cost. Thermography is an imaging technique used for breast cancer detection. A dynamic thermography technique which is used to generate infrared images over a fixed time measured in minutes to detect the difference between the normal and cancerous areas in images. In this research, we propose a methodology to deal with the changes of temperature in patient's breasts by defining a set of efficient features resulted from extraction and reduction of coefficients obtained from breast thermogram images followed by classification. Texture feature methods (Histogram of Oriented Gradients (HOG) and Discrete Curvelet transform) are applied separately using the HH (high-high) and HL (high-low) sub band images of Discrete Wavelet transform (DWT). HOG-based features and Curvelet features are extracted by reducing coefficients' vectors returned by the two methods. Finally, Support Vector Machine (SVM) binary classifier is used to classify the images to either normal or abnormal. The proposed work has successfully achieved an Accuracy of 98.2%, Sensitivity of 97.7%, and Specificity of 98.2% through empirical studies using dynamic breast thermogram dataset.

Keywords: *Breast cancer, Curvelet wrapping, Discrete Wavelet transform, Gradient methods, Histograms, Support Vector Machine.*

1.0 INTRODUCTION

Breast cancer took the lives of thousands of women every year. If diagnosed early, breast cancer can be cured. Thermography is among many tools which are used for early breast cancer diagnosis. The Infrared (IR) thermography diagnosis is based on temperature of the skin which can be measured without direct contact with the patient. This method provides safety and comfort to the patients [1]. Dynamic Infrared Thermography (DIT) quantitatively measures the temperature changes on the breast surface in a given period of time as opposed to static infrared thermography (SIT) which measures the temperature distribution of a breast at a given time. The existence of cancer in a breast tissue shows a higher temperature than healthy breast tissue as more blood flow is needed for tumor's growth [2]. Processing the image for breast cancer detection starts with segmentation to get one or more regions of interest (ROI). Then, the diagonal detail sub band (HH) resulted from Discrete Wavelet transform (DWT) decomposition is obtained [3]. Wavelets can capture point singularities and reveal image features across edges [4]. Wavelets are being used in many research works. Authors in [5] mentioned the use of approximation and details terms in Wavelet analysis where the approximation is the high scale low frequency component while the details are the low scale high frequency component. Lahmiri and Boukadoum in [6] used DWT to extract high frequency image and then applied Gabor bank to this image. Faye and Samir in [7] used

Wavelet to classify mammograms into either healthy or cancerous. Dalal and Triggs in [8] stated that HOG method is a powerful texture analysis method as it works on local cells and it is invariant to geometric transformations, except for object orientation. Authors in [9] used HOG descriptor to classify thermogram images while authors in [10] used HOG to classify mammograms.

The goal of this work is to model the changes of temperature in patient's breasts over a period of time by introducing a methodology to define efficient features through extraction and reductions of vectors of coefficients extracted from thermogram images. The objectives in this work are to utilise the dynamic breast thermography and to increase the accuracy of classifying the patients as having normal or abnormal breasts.

Results of experiments show that extracting features from more than one image per patient's case described the changes in patient's breast temperature and improved the accuracy. Also using HH+HL of DWT improved the results over when using the original images. The proposed methodology was able to describe the changes in breasts over a period of time and to assist in the diagnosis of abnormalities in the patient's breast.

The performance of the methodology is assessed using two texture analysis methods: Histogram of Oriented Gradients and Discrete Curvelet transform. The remaining of the paper is organised as follows. Section 2 provides the related work. Section 3 introduces texture analysis methods (HOG and Curvelet transform) and classification method. Section 4 introduces the methodology. Section 5 introduces the experimental results and discussion. Section 6 is the conclusion and future work.

2.0 LITERATURE REVIEW

Several techniques have been developed and presented to detect breast cancer from mammograms and thermograms. Lahmiri and Boukadoum in [6] used Gabor filter bank based on HH DWT sub band which are applied to three datasets related to mammograms, retina, and brain magnetic resonance (MR). On mammogram dataset, authors reported 96.67% accuracy with DWT, Gabor filter bank, and SVM classifier.

Abdel-Nasser, Moreno, and Puig in [9] used rank-based technique to model the changes on temperature in dynamic thermal images and texture-based methods to extract features. The authors reported 95.8%, 97.1%, and 94.6% for accuracy, recall, and precision, respectively. Santana et al. in [11] used different features which are based on texture and shape from Haralick and Zernike by using the grey-level temperature matrices. The images are to be classified into three classes: malignant, benign, and cyst. Extreme Learning Machines (ELM) and Multilayer Perception networks (MLP) used to classify those images. Authors reported 76.01% of accuracy.

Francis, Sasikala, and Saranya [12] used statistical and textural features extracted from Curvelet transform and Support Vector Machine to detect abnormal thermograms. Authors used 22 images (11 normal and 11 abnormal). Authors reported a 90.91% of accuracy for detecting abnormal thermograms. Sathish, Kamath, Prasad, Kadavigere, and Martis in [13] used features extracted from Histogram and Grey -Level Co-occurrence Matrix (GLCM) using a set of 80 thermogram images acquired from PROENG (40 normal and 40 abnormal). Authors reported the accuracy (90.0%), sensitivity (87.50%), and specificity (92.50%) results using SVM Radial Basis Function (RBF) classifier.

Acharya et al. in [14] presented a work based on 50 breast images (25 normal and 25 cancerous) where features extracted using Co-occurrence Matrix and Run Length Matrix were forwarded to SVM to differentiate between normal and malignant. Authors reported 88.10%, 85.71%, and 90.48% for accuracy, sensitivity, and specificity, respectively.

Authors in [15] applied the process of dilation and erosion on images and minimum variance quantisation. A total of 20 features were extracted based on GLCM and *K*-Nearest Neighbours (KNN) classifier from 40 images (26 normal and 14 abnormal). An accuracy of 92.50% was reported. Ali, Sayed, Gaber, and Hassanien in [16] used texture and statistical features as a new segmentation method where they successfully obtained an accuracy of 85% using SVM on DMR-IR dataset.

Pramanik, Bhattacharjee, and Nasipuri in [17] used DWT's detail coefficients of breast thermogram to detect breast cancer and used Principal Component Analysis (PCA) for feature selection. Based on Artificial Neural Network (ANN) to classify normal and abnormal images, the results reported were 90.48%, 87.60%, and 89.73% for accuracy, sensitivity, and specificity, respectively. EtehadTavakol et al. in [18] segmented images using Radon projection, then extracted bispectral invariant features using Adaboost classifier to classify malignant, benign, and normal cases. Based on 32 images (9 are malignant, 12 are benign and 11 are normal), 95% of accuracy was obtained for detecting malignant class.

Authors in [19] segmented 28 images and extracted few statistical features which include range, mean, and standard deviation, and perform quantisation for higher tone in an eight level posterisation. SVM was used as the classification method. The proposed framework achieved an average result of 85.71% for accuracy, 95.83% for sensitivity, and 25% for specificity.

Cruz-Ramírez et al. [20] reported 71.88%, 82% and 37% for accuracy, sensitivity, and specificity respectively using Temperature features and Naïve Bayes classifier to classify 98 (77 malignant, 21 healthy) images. Wakankar and Suresh in [21] segmented 34 private thermogram images using fuzzy *C*-means and *K*-means, where skewness, kurtosis, and entropy features are extracted for image representation. They managed to obtain 91.60%, 88.80%, and 100% for accuracy, sensitivity, and specificity. Silva et al. in [22] used 22 samples of DMR-IR dynamic thermogram database and statistical features and reported 90.1% for accuracy and 93% for precision.

Authors in [23] used 50 private thermograms and obtained 86% for recall using thermal minutia points. Etehadtavakol et al. [24] used mean, standard deviation, entropy, skewness, and kurtosis features from histogram and Haralick statistical features to classify breast thermograms. Forty breast thermal images from Department of Diagnostic Radiology, Singapore General Hospital were used. Twelve benign and eight malignant cases were used for training.

Gogoi et al. [25] used Mann-Whitney-Wilcoxon statistical test to select the most distinguish features from a collection of 24 features. Six classifiers are used to diagnose breast thermogram irregularity. Reported results showed that ANN and SVM Linear classifier provided the highest accuracy of 87.5% over DMR database. Jeyanathan et al. [26] extract features by applying Wavelet, Curvelet and Contourlet transform on thermograms. One sample *t*-test and independent *t*-test were used to measure the features. A neural network with multilayer and back-propagation classifier showed that Gabor wavelet features can be used to classify the thermograms into normal and abnormal.

De Vasconcelos, dos santos, and de Lima in [27] used left and right breasts to compute twenty characteristics based on maximum and minimum temperatures using 233 thermograms for binary class (cancer and healthy) and multi-class (malignant, benign, cyst, and normal) classification. 93.42% accuracy and 94.73% sensitivity for binary classification were obtained whereas 80.77% sensitivity for multi-class was reported. These results were achieved by applying the Sequential Minimal Optimisation (SMO) classifier. Rodrigues et al. [28] proposed a feature selection approach using genetic algorithms (GA) and particle swarm optimization (PSO) in breast thermographic images to optimise the identification and classification of breast lesions. Authors reported an accuracy of 91.12% using 169 features whereas when using 57 features, the GA achieved an accuracy of 87.08%. An accuracy of 86.16% is achieved when using 60 features generated from PSO.

Araújo et al. in [29] developed a methodology to classify breast thermogram images which starts with extracting region of interest using 244 images followed by the extraction of textural features. Using SVM, an accuracy of over 90% is achieved. Gogoi et al. in [30] used Database of Mastology Research (DMR) and singular value decomposition to detect abnormality in breast thermogram images and they reported an accuracy of 98% when using SVM. Karthiga and Narasimhan in [32] used Frontal images from the visual laboratory database. The authors used morphological operation to segment hot regions and enhance ROIs. Curvelet transform with wrapping method is applied to the ROIs while 20 features based on GLCM are then extracted. Based on hypothesis testing, 16 features out of 20 are selected. Authors reported an accuracy of 93.3% when using SVM.

Based on the literature review mentioned above, it has been found that most of the previous works utilised statistical features such as GLCM, skewness, kurtosis, mean, and standard deviation for feature representation. Still, the reported results need to be improved. Here in this research, we propose a technique to reduce the resulted coefficient matrices based on HOG and Curvelet transform. Most of the present techniques extract features over the entire ROIs which might cause many information losses that can be important to differentiate between healthy and unhealthy ROIs. In our

research, we divide the resulted coefficients into equal consecutive intervals and extract features based on those intervals. Our approach showed an improvement in sensitivity, specificity, and accuracy.

3.0 TEXTURE ANALYSIS AND CLASSIFICATION METHODS

Effective feature extraction and classification methods are crucial in helping to determine whether a breast thermogram image is cancerous or healthy. Below is the explanation on the texture-based Histogram of Oriented Gradients, Discrete Curvelet transform, and Support Vector Machine binary classifier.

3.1 Discrete Curvelet Transform

Discrete Curvelet transform is a multi-scale approach which was proposed by Candes and Donoho [33]. The first generation of Curvelet transform presented in [33] was too complicated and have a lot of redundancy. A second generation of the Curvelet transform was presented which is characterised as being simpler and faster [34]. Curvelets represent anisotropic features efficiently as the anisotropic elements of the curvelets have high directional sensitivity [35]. Curvelet transform extends the Ridgelet transform to include multiple scale analysis. Ridgelet transform was introduced by Candes and Donoho [36].

Both ridgelets and wavelets provide multi-resolution texture information. Ridgelet transform provides multi-directional capabilities which has better texture discrimination compared to Wavelet transform [37]. While ridgelets are effective in detecting linear radial, it is not effective to detect image activity along curves as in Curvelet transform [38]. Fig. 1 below shows the radial grid of the Ridgelet transform.

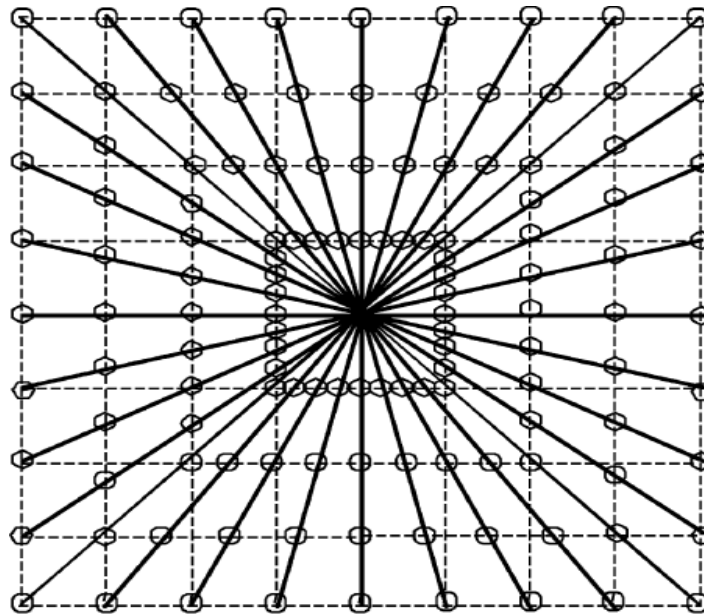


Fig. 1: Radial grid of the Ridgelet transform [38]

The new implementation of Curvelet transform is based on wrapping of Fourier samples by taking a 2D image in the form of a Cartesian array $f[m, n]$ where $0 \leq m < M$, $0 \leq n < N$ where M and N are the dimensions of the array. Equation 1 shows the output coefficients $C^D(j, l, k1, k2)$ which is indexed by scale j , l is the orientation, while $k1$ and $k2$ as spatial location parameters. Each $\Phi_{j,l,k1,k2}^D$ is a digital Curvelet waveform. 'D' stands for digital. The effective parabolic scaling law on the sub bands in the frequency domain is implemented by curvelet approach to capture curved edges within an image.

$$C^D(j, l, k1, k2) = \sum_{\substack{0 \leq m < M \\ 0 \leq n < N}} f[m, n] \phi_{j,l,k1k2}^D[m, n]. \quad (1)$$

Wrapping-based Curvelet transform consists of several sub bands at different scales with different orientations and positions in the frequency domain. Curvelets are so fine and looks like a needle shaped elements at high frequency level whereas at low frequency level they are non-directional coarse elements. Combination of frequency responses of curvelets at different scales and orientations result in a rectangular frequency tiling that covers the whole image in the spectral domain (Fig. 2). Therefore, the curvelet spectra completely covers the frequency plane with no loss of spectral information [39].

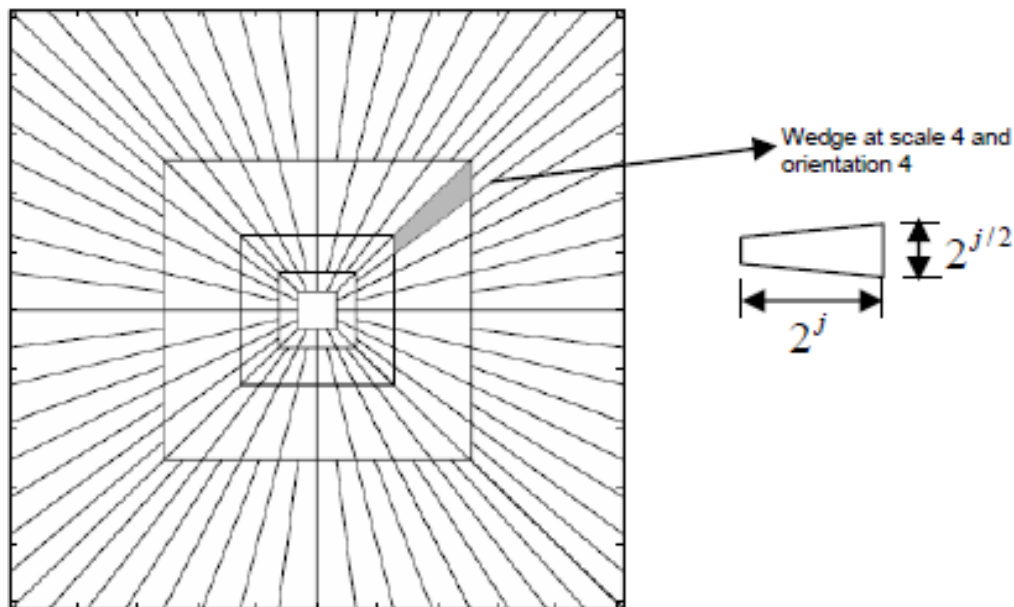


Fig. 2: Radial grid of the Ridgelet transform [39]

Curvelet transform when implemented in the frequency domain achieves a higher level of efficiency. Both the curvelet and the image are transformed by Fast Fourier transform (FFT) and then multiplied in the Fourier frequency domain. Inverse Fast Fourier transform (IFFT) is then applied on the product to obtain the curvelet coefficients. The process can be described as Curvelet transform = IFFT [FFT (Curvelet) × FFT (Image)] and the product from the multiplication is a wedge. The resulted wedge cannot fit into a rectangle of size $2^j \times 2^{j/2}$. A wrapping method was formulated by Candes et al. [34] to overcome this problem. The wrapping is done by periodic tiling of the spectrum inside the wedge and then collecting the rectangular coefficient area in the center. The center rectangle of size $2^j \times 2^{j/2}$ successfully collects all the information in that parallelogram (Fig. 3). Thus, the Discrete Curvelet coefficients are obtained by applying 2-D Inverse Fourier transform to this wrapped wedge data.

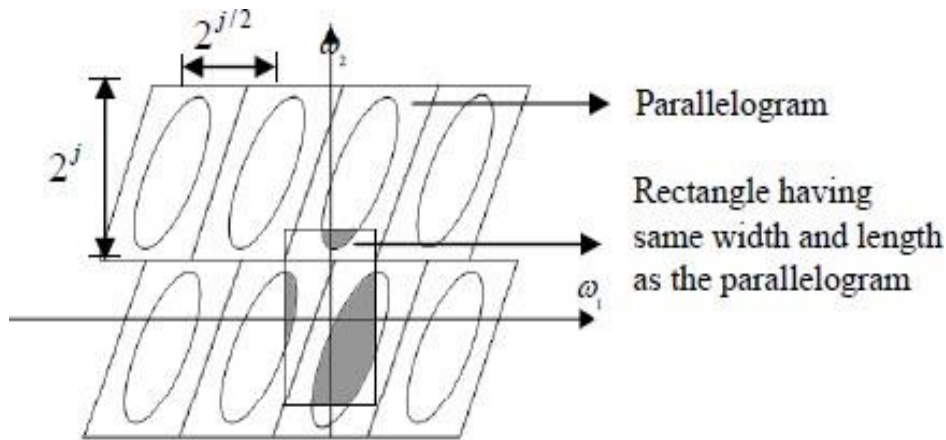


Fig. 3: Wrapping wedge procedure [39]

3.2 Histogram of Oriented Gradients(HOG)

Histogram of Oriented Gradients (HOG) is a powerful texture analysis method to detect objects [9]. HOG construction is done by dividing the image window into small spatial regions. Occurrences of gradient orientation of the HOG in local portions of an image are counted. The entries of the histogram are then combined to form the representation. The frequencies are then normalised to consider the illumination changes. Concatenation of the resulted histogram of all blocks formed the final HOG descriptor [8].

3.3 Discrete Wavelet Transform

The Wavelet transform is a multi-resolution analysis tool. It is based on decomposing the images into different resolution levels. Different physical structures in an image can be revealed at different resolution levels. The resulting wavelet coefficients contain information about the underlying signal at different t scales. The detail coefficients of Wavelet transform resemble looking a nearby object whereas the approximation coefficient exhibit a coarser resolution. High-frequency components of short duration (details) and low frequency components of long duration (approximations) exist in most images as shown in Fig. 4 [31].

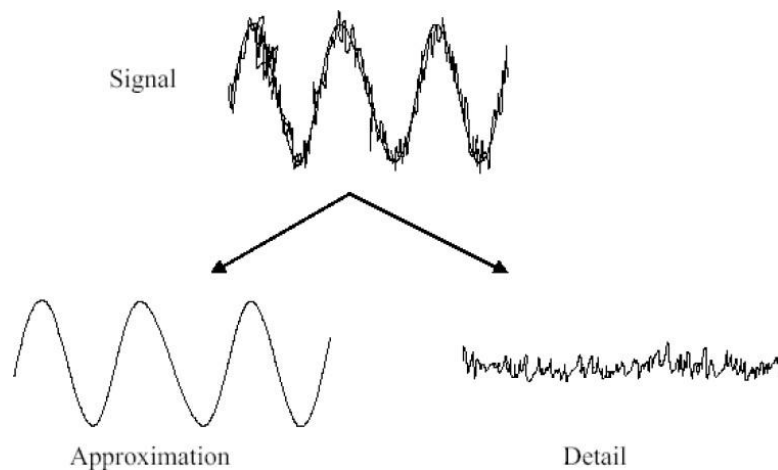


Fig. 4: Details and approximation of a signal

Wavelets decompose an image $f(t)$ into four sub band components: Low–low (LL) which corresponds to approximation whereas low–high (LH), high–low (HL), and high–high (HH) correspond to detail components. Further, the LL component can be decomposed into another four sub bands; and the (LLL) component which is approximation component can then be decomposed again and the process could continue.

3.4 Classification Method

Support Vector Machine (SVM) is introduced by Vapnik [40]. It is a statistical based classifier. It is proven to work well on small data as well as on big data and tolerates high-dimensional or incomplete data. SVM has been used in medical applications such as in breast cancer detection. SVM is an algorithm which looks to create a line or a hyperplane to separate the data into different classes (data are the images in our case). The support vectors are data points of the training set that are closest to the hyperplane that can influence the position of the hyperplane. Binary classification is the original use of SVM to solve the binary classification problem. Establishing the optimal hyperplane for two types of samples lead to maximising their separation.

4.0 METHODOLOGY

DWT is applied to images to obtain HH and HL sub band images. This is to obtain its high frequency components in an image which represents the desired information about the image tissues. The abnormal biological tissues found in an image can be indicated by the sudden change in the image tissues. An example of processing of normal and abnormal images is illustrated in Fig. 5 and Fig. 6.

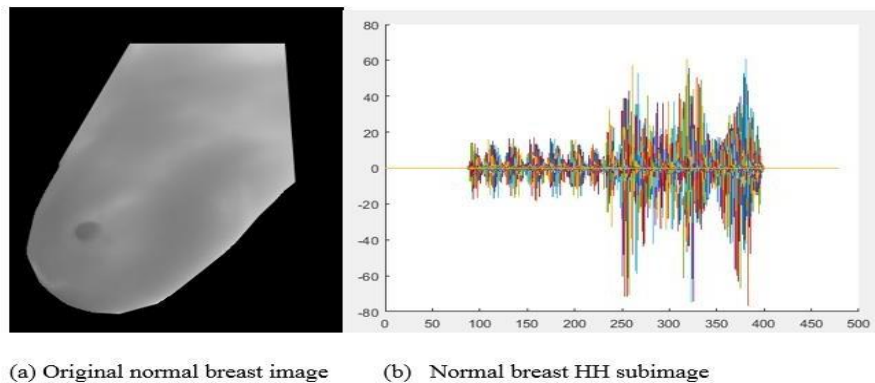


Fig. 5: Extraction of an HH image from a normal breast

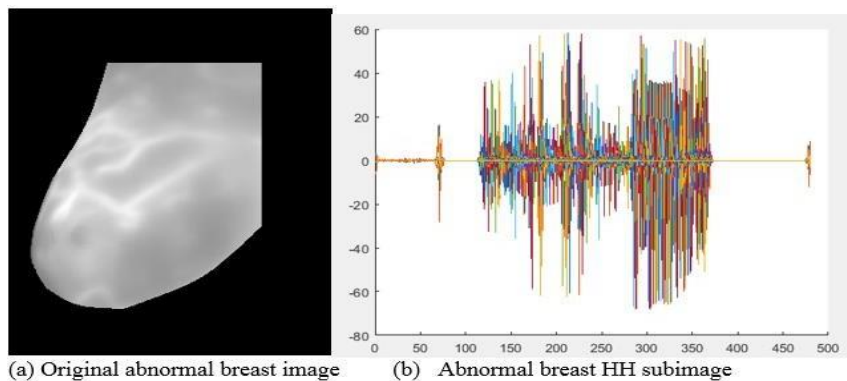


Fig. 6: Extraction of an HH image from an abnormal breast

4.1 Processing

1040 images (52 breast images, each with 20 sequential images) used in the learning phase and 240 images (12 breast images, each with 20 sequential images) used in the test stage. The 64 images are divided evenly as normal and abnormal. Temporal aspects of the image dataset are utilised to improve the classification rates as we conducted experiments using just one image per case and two images per case. For two images per case, we used 46 different combinations of pair of images. The following are five groups of the combinations where t is the time followed by a number which indicates the time (at time 0 for example):

$(t_0, t_1), (t_1, t_2), \dots, (t_{18}, t_{19})$
 $(t_2, t_4), (t_4, t_6), \dots, (t_{16}, t_{18})$
 $(t_1, t_3), (t_3, t_5), \dots, (t_{17}, t_{19})$
 $(t_1, t_4), (t_4, t_7), \dots, (t_{16}, t_{19})$
 $(t_1, t_5), (t_5, t_9), \dots, (t_{13}, t_{17})$

All images of all cases are covered in the experiments. The idea behind the combinations is to take images per case at different time stamp. The first group of two images is arranged by fixing the first image of each case at time stamp 0 and take the second image of the same case at time stamps 1:19. The second and third images were taken at odd and even time stamps. The fourth and fifth groups have 3 and 4 differences between the time stamp of two images of the same case. The intuition behind forming those group combinations is to find out if we need to extract features from all images of each case or extracting features from some images can already be used to describe the temperature changes in breasts.

Daubechies 4 and 5 (db4 and db5) of DWT are applied on the pair of images to obtain HH and HL images.

4.2 Feature Extraction

It is crucial step to extract appropriate information from images which can lead to accurate classification of the images.

4.2.1 Feature Extraction Based on Discrete Curvelet Transform

Curvelet decomposition of a 2D image is defined as the inner product of the image with the mother curvelet function's translated, scaled, and rotated. The result of the Curvelet decomposition of an image is a sparse representation of curve-like features which have different lengths and orientations at different resolution scales. The resulted coefficients of curvelet of an image make the texture descriptor of that image. We adopted the curvelet implementation of the Discrete Curvelet transform which can be found at <http://www.curvelet.org/software.html>. The number of scales used is set to 4 and the number of angles is 16. The output of using the implementation is a collection of coefficients indexed by a scale j , an orientation l , and spatial location parameters k_1 and k_2 . Out of these settings, 24 matrices are obtained which are explained in Table 1 below.

Table 1: Distribution of curvelet sub bands at each scale

Curvelet transform (4 level decomposition)				
Scale	1	2	3	4
Total no. sub bands	1	16	32	1
Sub bands for feature calculation	1	8	16	1

As shown in the above table, only half of sub bands at a resolution level are used as features and this is because curvelet at angle θ produces the same coefficients at $(\theta+\pi)$ in the frequency domain.

The obtained matrices are of different sizes: $\{60 \times 106, 66 \times 108, 81 \times 80, 81 \times 87, 120 \times 108, 120 \times 107, 126 \times 108, 81 \times 160, \text{ and } 81 \times 167\}$. All these coefficient matrices need to be transformed into a vector first. Then, sum and skewness are computed

on every 2000 consecutive coefficients in every vector as one feature. The corresponding features for the pair of images per case for all images are added as the final feature. Finally, the feature vector is fed into the SVM for classification as shown in Fig. 7. Skewness identifies the deformation of the symmetric curve in the set of a complete data. The shortfall of symmetry can identify the data distributions.

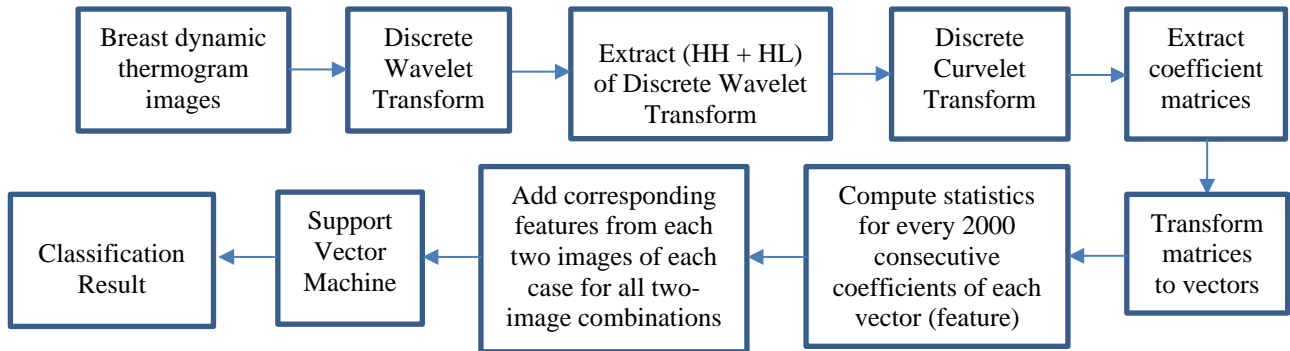


Fig 7: Schematic diagram of DWT-Discrete Curvelet Transform approach

4.2.2 Feature Extraction from HOG

The process starts by applying DWT to extract HH or (HL+HH) for each image of each case of the cases. Then, HOG is applied on the transformed image. The resulted coefficients vector size is 167000 coefficients. One feature is obtained by taking every (1000 or 2000) consecutive coefficients of the returned coefficients vector and taking their sum.

We choose HOG to extract features from each image. Each image has a resolution of 640×480 pixels. Each image returns a coefficient vector of size 167000. Most research take the mean, skewness, or standard deviation over this vector. Our method is to form smaller vectors by taking 1000 or 2000 consecutive coefficients and compute the mean on the smaller vector. As we considering two images per case, we compute the average on corresponding smaller vectors and the average makes one feature.

In case of taking two images per case, corresponding features are added to make the final feature. This applies to all combinations of two images at a time per case. The final constructed feature vector will be utilised by SVM for classification as shown in Fig. 8.

Experiments were done by taking 25%, 56%, 70%, 97%, and 100% of coefficients from the returned vector, which is equivalent to around 41000, 93000, 118000, 162000, and 167000 coefficients, respectively.

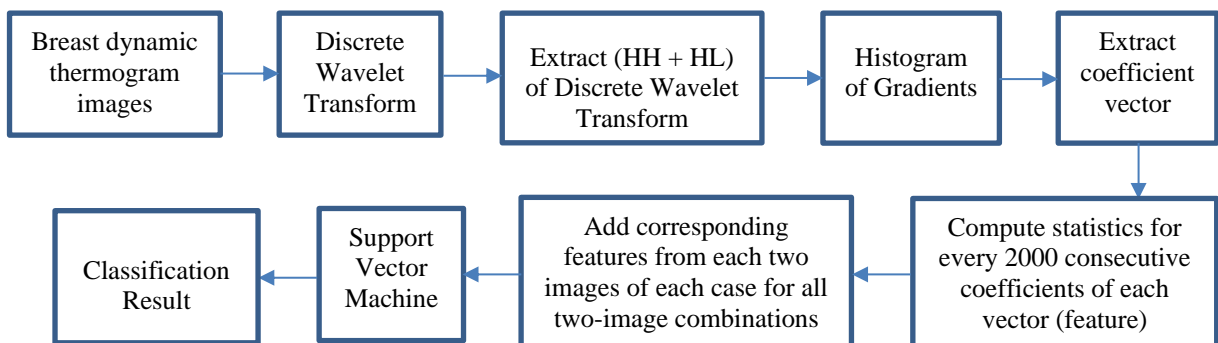


Fig 8: Schematic diagram of DWT-Discrete HOG approach

4.3 Evaluation Measurement

Specificity, sensitivity, and accuracy are used to evaluate the approach. Equation for each of them is given below:

$$\text{Specificity} = \text{TN} / (\text{TN} + \text{FP}) \quad (2)$$

$$\text{Sensitivity} = \text{TP} / (\text{TP} + \text{FN}) \quad (3)$$

$$\text{Accuracy} = (\text{TP} + \text{TN}) / (\text{TP} + \text{TN} + \text{FP} + \text{FN}) \quad (4)$$

where TP: count of true positive cases, TN: count of true negative cases, FN: count of false negative cases, and FP: count of false positive cases. The same measurements were used in related work as in [14] and [21].

A publicly available dataset that is available at <http://visual.ic.uff.br/dmi/> was used. It consists of 47 segmented dynamic breast thermogram images where there are 31 sick cases and 16 health cases. Each case of the 47 cases has 20 shots taken over five minutes which represents the change in breast temperature. Samples of images are shown in Fig. 9. Image resolution is 640×480 pixels.

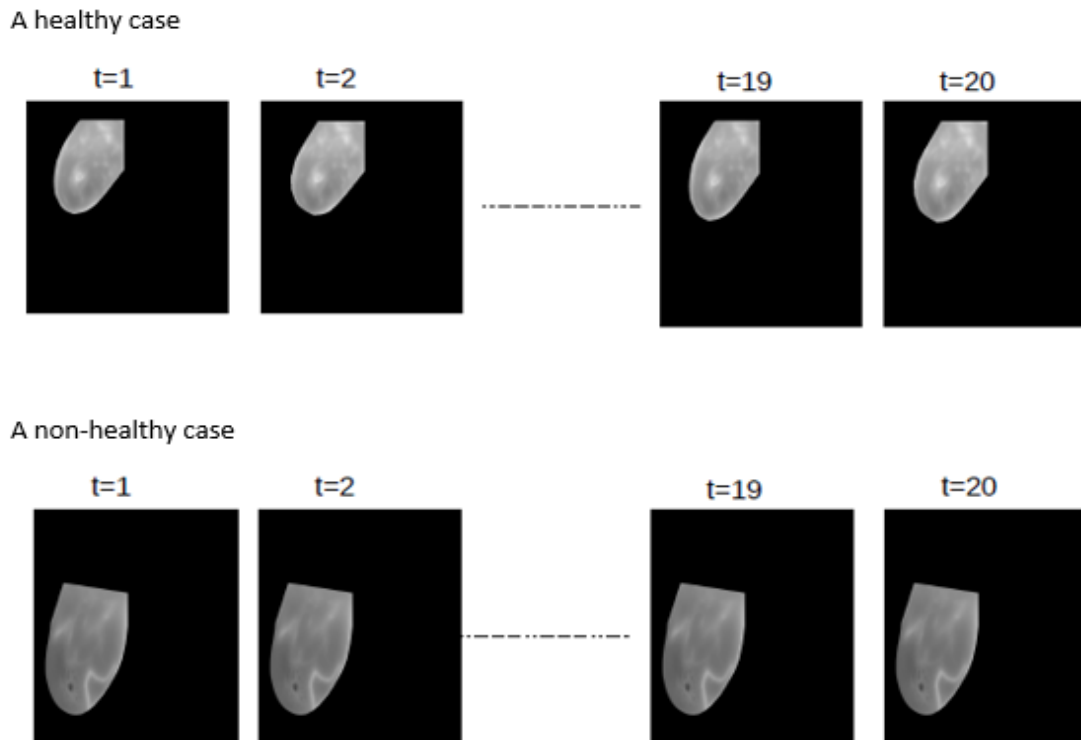


Fig. 9: Healthy and unhealthy cases at different time stamps

5.0 EXPERIMENTAL RESULTS AND DISCUSSION

We conducted several experiments using one image per case and two images per case based on HOG and Discrete Curvelet transform. Table 1 shows the results using one image per case while Table 2 shows the results of using two images per case, taking 70% of HOG coefficients. Table 3 shows the results of combining 2000 HOG coefficients using HH+HL sub bands and original images. While Table 4 shows results of using 97% coefficients on HH sub band and yielded very promising results. Table 5 shows the performance of our methodology per number of features.

Table 1: Average Accuracy, Sensitivity, and Specificity for one image per case (original, HH sub band, 167 features)

	Accuracy	Sensitivity	Specificity
Original	78.56%	90.10%	55.20%
HH Sub band	73.60%	88.40%	58.50%

Table 2: Average Accuracy, Sensitivity, and Specificity for HH high frequency sub band of DWT and 70% of HOG coefficients (118000 Coefficients, 118 features)

Accuracy	Sensitivity	Specificity
96%	98%	91%

Table 3: Average Accuracy, Sensitivity, and Specificity for original image and (HL+HH) frequency sub bands of DWT and 97% of HOG coefficients (162000 Coefficients, taking every 2000 consecutive coefficients, 81 features)

	Accuracy	Sensitivity	Specificity
Original	98.2%	97.7%	98.2%
HH +HL	98.0%	97.7%	98.7%

Table 4: Average Accuracy, Sensitivity, and Specificity for HH high frequency sub band of DWT and 97% of HOG coefficients (162000 Coefficients, 162 features)

Accuracy	Sensitivity	Specificity
98%	97%	100%

Table 5: Accuracy, Sensitivity, and Specificity for number of HOG-based features related to our research

No. of features	Accuracy	Sensitivity	Specificity
162000	87.8%	84.4%	94.3%
162	98.0%	97.0%	100.0%
118	96.0%	98.0%	91.0%
81	98.2%	97.7%	98.2%
41	89.0%	89.0%	88.0%

5.1 Discrete Wavelet Transform Results

We used 64 images (32 normal and 32 abnormal) for experimentation with Discrete Wavelet transform (DWT) based on the same dataset. Two-dimensional DWT and Daubechies wavelet (db5) were used. The following Table 6 show the results of using different settings for the image selection as well as on the number of features. When considering 188 features (using the sum and skewness), it is clear that using (HL+HH) sub bands yielded better results than using original images. When using only the sum feature which yielded 94 features, utilising original images also produced lower results than when using (HL+HH) sub-bands. Results are better when using DWT sub bands over using original images. Previous related publications support the claim that DWT sub bands are more efficient in characterising images other than LL or image approximation as authors in [6] found that using LH and HL coefficients are effective as they characterise the changes in the biological tissues and help in differentiating the normal from abnormal image textures. When using one image per case based on the original image and (HH+HL) sub bands using sum as a feature with a total of 94 features, results obtained are much lower compared with when using two images per case.

Table 6: Average Specificity, Sensitivity, and Accuracy for different image setting and number of features

Image type and features	Specificity	Sensitivity	Accuracy
Original images + 188 features	57.9%	78.6%	72.4%
HL + HH sub bands + 188 features	90.6%	98.2%	95.6%
Original images + 94 features	52.9%	77.5%	69.3%
HL + HH sub bands + 94 features	92.7%	97.1%	95.6%
HL + HH sub bands + one image/case + 94 features	60.8%	60.5%	60.6%
One original image/case + 94 features	62.3%	44.9%	50.5%

5.2 HOG and Curvelet Transform:

HOG and Discrete Wavelet both showed comparable results when using the original image with DWT sub bands. Both methods show better results when using sub band images over using the original images. Also, both show much better results when using two images per case over using one image per case. Coefficient matrix reduction for both the HOG and curvelet shows effectiveness and improvement in results. Results show better sensitivity when using the Discrete Wavelet over HOG.

The experimental results show that the proposed methodology described the changes in temperature of patient's breasts by determining a set of features formed by reducing the huge coefficients vectors returned from thermogram images. Those vectors are then reduced to smaller vectors and the mean is computed over those vectors. Then, each feature is formed by taking the average for every corresponding reduced vectors of each pair of images per patient's case.

5.3 Comparison with related work:

Table 7 shows the results of related work compared to our proposed work. Our methodology aims to describe the changes in temperatures in patient's breasts over a period of time. In addition to reducing HOG vector into smaller vectors, we transformed different matrix sizes resulted from Curvelet transform into a vector before processing. Table 7 shows that classification results using dynamic thermogram images are better than using static thermogram images. The results of most works presented in the table were based on extracting texture and statistical features over the whole image. Authors in [9] used a different approach than ours to model temperature changes in breasts. While authors in [9] accumulate the occurrences of the temperature matrices along different orientations, we reduced the HOG vectors into smaller matrices to reduce features set. Also, we used different percentages of the HOG vector. Further, while [9] used original images, we used HH and (HL+HH) sub band images of DWT besides the original images. We achieved 98% of accuracy using HOG and (HL+HH) of DWT. Authors in [22] achieved 90% of accuracy when using only 22 samples of the dataset. Authors in [32] used hypothesis testing to determine statistically significant features, reported an accuracy of 93.3% using 60 images. Our methodology is to reduce the resulted coefficients vectors by combining every 2000 coefficients and apply the mean on those vectors and then average the corresponding vectors to form features. We achieved 95% of accuracy which is better than the one reported in [32].

Table 7: Comparison with related work

Author	Method / Feature	Result	Dataset
Abdel-Nasser et al. [9]	HOG and learning to rank	Accuracy: 95.8% Recall: 97.1% Precision: 94.6%	All sample images of DMR-IR database
Santana et al. [11]	Texture and shape from Haralick and Zernike features	Accuracy: 76.01%	825 images from Private data
Acharya et al. [14]	Texture features based on histogram and GLCM	Accuracy: 90.0% sensitivity: 87.50% specificity: 92.50%	80 images PROENG: database
Wakankar et al. [21]	Skewness, kurtosis, entropy, K-means, and fuzzy C-means	Accuracy: 91.60% Sensitivity: 88.8% Specificity: 100%	34 images from private data
Silva et al. [22]	Statistical features	Accuracy: 90.1%	A sample of 22 images of DMR-IR database
Saniei. et al. [23]	Thermal minutia points	specificity: 61% Recall: 86%	50 thermograms of private dataset
Rodrigues et al. [28]	Reduced features to optimise classification time	Accuracy: 91.12%	1052 images drawn from University Hospital of the Federal University of Pernambuco
Karthiga and Narasimhan [32]	GLCM	Accuracy: 93.3%	60 images from Visual laboratory, the Fluminense Federal University in Rio de Janeiro, Brazil
Proposed work	DWT (HL+HH) and Curvelet	Accuracy: 95.6% Sensitivity: 97.1% Specificity: 92.7%	All samples of DMR-IR database
Proposed work	DWT (HL+HH) and HOG	Accuracy: 98.2% Sensitivity: 97.7% Specificity: 98.2%	All samples of DMR-IR database

6.0 CONCLUSION AND FUTURE WORK

In this work, we presented a methodology to extract features from two texture extraction methods (HOG and Curvelet transform) by reducing the resulted vectors. We utilised the temporal aspect of the images for each case in the database by processing two images per case. We compared results when using one and two images per case, also when using original images and DWT transformed images. Results using Discrete Curvelet transform showed better results than the ones presented in [32]. Our results using HOG showed comparable results with the ones in [9]. Our future work is to apply the methodology on different datasets and to use time series to extract features from all images of cases.

7.0 ACKNOWLEDGEMENT

This work was supported / funded by the Ministry of Higher Education Malaysia under the Fundamental Research Grant Scheme (FRGS/1/2016/ICT04/UPM/02/6).

REFERENCES

- [1] G. Schaefer, T. Nakashima, M. Zavisek, "Analysis of Breast Thermograms Based on Statistical Image Features and Hybrid Fuzzy Classification", *Bebis (eds) Advances in Visual Computing: Lecture Notes in Computer Science*, 2008, Vol. 5358. Springer, Las Vegas, pp. 753–762.

- [2] S. Pramanik, D. Bhattacharjee, M. Nasipuri, "Wavelet Based Thermogram Analysis for Breast Cancer Detection", in *Proceedings of the Advanced Computing and Communication (ISACC)*, 2015, pp. 205-212, 10.1109/ISACC.2015.7377343.
- [3] S. Beura, B. Majhi, R. Dash, "Mammogram Classification Using Two-Dimensional Discrete Wavelet Transform and Gray-Level Co-Occurrence Matrix for Detection of Breast Cancer", *Neurocomputing*, Vol. 154, 2015, pp. 1-14.
- [4] M. Etehadtavakol, E.Y.K. Ng, V. Chandran, H. Rabbani. "Separable and Non-Separable Discrete Wavelet Transform Based Texture Features and Image Classification of Breast Thermograms ", *Infrared Physics & Technology*, Vol. 61, 2013, pp. 274–286.
- [5] M.K. Abu Mahmoud, A. Al-Jumaily, M. Takruri, "Wavelet and Curvelet Analysis for Automatic Identification of Melanoma Based on Neural Network Classification", *International Journal of Computer Information Systems and Industrial Management (IJCISIM)*, Vol. 5, 2013, pp. 606-614.
- [6] S. Lahmiri, M. Boukadoum, "Hybrid Discrete Wavelet Transform and Gabor Filter Banks Processing for Features Extraction from Biomedical Images", *Hindawi Publishing Corporation, Journal of Medical Engineering*, 2013, 13 pages. DOI- 10.1109/NEWCAS.2011.5981217.
- [7] I. Faye, B. Samir, "Digital Mammograms Classification Using a Wavelet Based Feature Extraction Method", in *Proceedings of the Second International Conference on Computer and Electrical Engineering*, 2009, Vol. 2, pp. 318-322, DOI: 10.1109/ICCEE.2009.39.
- [8] N. Dalal, B. Triggs, "Histograms of Oriented Gradients for Human Detection", in *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05)*, San Diego, CA, USA, 2005, Vol. 1, pp. 886-893, DOI: 10.1109/CVPR.2005.177.
- [9] M. Abdel-Nasser, A. Moreno, D. Puig, "Breast Cancer Detection in Thermal Infrared Images Using Representation Learning and Texture Analysis Methods", *Electronics*, Vol. 8 No. 1, pp. 18 pages; <https://doi.org/10.3390/electronics8010100>, 2019.
- [10] Z. Alhakeem, Se-In. Jang, "A Convolution-Free LBP-HOG Descriptor for Mammogram Classification". *arXiv*, 2019.
- [11] M.A.D Santana, J.M.S .Pereira, F.L.D. Silva , N.M.D. Lima, F.N.D. Sousa, G.M.S.D. Arruda, de R.C.F. Lima, W.W.A.D. Silva, W.P.D. Santos, "Breast Cancer Diagnosis Based on Mammary Thermography and Extreme Learning Machines", *Res. Biomed. Eng.*, Vol. 34, pp. 45–53, <https://doi.org/10.1590/2446-4740.05217>, 2018.
- [12] S.V.M. Francis, M. Sasikala, S. Saranya, "Detection of Breast Abnormality from Thermograms Using Curvelet Transform Based Feature Extraction", *J Med Syst*, Vol. 38 No. 23, DOI 10.1007/ s10916-014-0023-3, 2014.
- [13] D. Sathish, S. Kamath, K. Prasad, R. Kadavigere, R. Martis, "Asymmetry Analysis of Breast Thermograms Using Automated Segmentation and Texture Features", *Springer- Verlag London*, pp. 745-752, 2016, 10.1007/s11760016-1018-y.
- [14] U.R. Acharya, E.Y. Ng, JH. Tan, SV. Sree, "Thermography Based Breast Cancer Detection Using Texture Features and Support Vector Machine", *J Med Syst.*, 2012, Vol. 36, pp. 1503-15010, <https://doi.org/10.1007/s10916-010-9611-z>.

- [15] M. Milosevic, D. Jankovic, A. Peulic, "Thermography Based Breast Cancer Detection Using Texture Features and Minimum Variance Quantization", *EXCLI Journal*, Vol. 13, pp. 1204-1215, 2014.
- [16] M. Ali, G. Sayed, T. Gaber, AE. Hassanien, "Detection of Breast Abnormalities of Thermograms based on a New Segmentation Method", *Computer Science and Information Systems*, Vol. 5, pp.255–261, 2015, DOI: 10.15439/2015F318.
- [17] S. Pramanik, D. Bhattacharjee, M. Nasipuri, "Wavelet Based Thermogram Analysis for Breast Cancer Detection", in *Proceedings of the Advanced Computing and Communication (ISACC), International Symposium on IEEE*, pp. 205-212, 2015, doi: 10.1109/ISACC.2015.7377343.
- [18] M. EtehadTavakol, V. Chandran, E.Y.K. Ng, R. Kafieh, "Breast Cancer Detection from Thermal Images Using Bispectral Invariant Features", *International Journal of Thermal Sciences*, 2013, Vol. 69, pp. 21-36, <https://doi.org/10.1016/j.ijthermalsci.2013.03.001>.
- [19] T.B. Borchardt, R. Resmini, A. Conci, "Thermal Feature Analysis to Aid on Breast Disease Diagnosis", in *Proceedings of 21st Brazilian Congress of Mechanical Engineering—COBEM2011*, pp. 24–28, 2011.
- [20] N. Cruz-Ramírez, E. Mezura-Montes, Y. Maria, M. Enrique, H. Gabriel, AM. Nancy, G. Alejandro, G. Hoyos-Rivera, R. Barrientos-Martínez, "Evaluation of The Diagnostic Power of Thermography in Breast Cancer Using Bayesian Network Classifiers", *Computational and Mathematical Methods in Medicine, Hindawi Publishing Corporation*, 264246, 10.1155/2013/264246.
- [21] A. Wakankar, G.R. Suresh, "Automatic Diagnosis of Breast Cancer using Thermographic Color Analysis and SVM Classifier", in *Proceedings of the International Symposium on Intelligent Systems Technologies and Applications*, Vol. 530, 2016, https://doi.org/10.1007/978-3-319-47952-1_2.
- [22] L.F. Silva, G.O. Sequeiros, M.L.O. Santos, C.A. Fontes, D.C. Muchaluat-Saade, A. Conci, "Thermal Signal Analysis for Breast Cancer Risk Verification", in *Proceedings of the World Congress on Medical and Health Informatics (MEDINFO2015)*, São Paulo, Brazil, Vol. 216, 2015, pp. 746–750, DOI: 10.3233/978-1-61499-564-7-746.
- [23] E. Saniei, S. Setayeshi, M.E. Akbari, M. Navid, "A Vascular Network Matching in Dynamic Thermography for Breast Cancer Detection", *Quant. InfraRed Thermogr. J.*, Vol. 12, No. 1, 2015, pp. 24-36, <https://doi.org/10.1080/17686733.2015.1005398>.
- [24] M. Etehadtavakol, E.Y.K. Ng, C. Vinod, H. Rabbani, "Separable and Non-Separable Discrete Wavelet Transform Based Texture Features and Image Classification of Breast Thermograms", *Infrared Physics & Technology*, Vol. 61, pp. 274–286, 2013.
- [25] U.R. Gogoi, M.K. Bhowmik, A.K. Ghosh, D. Bhattacharjee, G. Majumdar, "Discriminative Feature Selection for Breast Abnormality Detection and Accurate Classification of Thermograms", in *Proceedings of the International Conference on Innovations in Electronics, Signal Processing and Communication, IESC*, pp. 39–44, 2017.
- [26] J. Jeyanathan, P. Jeyashree, A. Shenbagavalli, "Transform Based Classification of Breast Thermograms Using Multilayer Perceptron Back Propagation Neural Network", *Int. J. Pure Appl. Math*, pp. 1955–1961, 2018.
- [27] J.H. de Vasconcelos, W.P. dos Santos, R.C.F. de Lima, "Analysis of Methods of Classification of Breast Thermographic Images to Determine Their Viability in The Early Breast Cancer Detection", *IEEE Latin America Transactions*, Vol. 16, No. 6, pp. 1631-1637, 2018.

- [28] A.L. Rodrigues, M.A. de Santana, W.W. Azevedo, W.W. et al., "Identification of Mammary Lesions in Thermographic Images: Feature Selection Study Using Genetic Algorithms and Particle Swarm Optimization", *Research on Biomedical Engineering*, Vol. 35, No. 3–4, pp. 213–222, 2019, <https://doi.org/10.1007/s42600-019-00024-z>.
- [29] A.D.S. Araújo, A. Conci, R. Resmini, A. Montenegro, C. Araujo, F. Lebon, "Computer Aided Diagnosis for Breast Diseases Based on Infrared Images", in *Proceedings of the IEEE/ACS 14th International Conference on Computer Systems and Applications (AICCSA), Hammamet, Tunisia*, pp. 172-177, 2017.
- [30] U.R. Gogoi, M.K. Bhowmik, D. Bhattacharjee, A.K. Ghosh, "Singular Value-Based Characterization and Analysis of Thermal Patches for Early Breast Abnormality Detection", *Australas Phys Eng Sci Med.*, Vol. 41, pp. 861–879, 2018, doi: 10.1007/s13246-018-0681-4.
- [31] N. Hamad, K. Taouil, M.S. Bouhlel, "Mammographic Microcalcifications Detection using Discrete Wavelet Transform", *International Journal of Computer Applications*, 2013, DOI: 10.5120/10758- 5695.
- [32] R. Karthiga, K. Narasimhan, "Medical Imaging Technique Using Curvelet Transform and Machine Learning for The Automated Diagnosis of Breast Cancer from Thermal Image", *Pattern Analysis and Applications*, 2021.
- [33] E J. Candès, D L Donoho, "Curvelets: A Surprisingly Effective Nonadaptive Representation of Objects with Edges", In: A. Cohen, C. Rabut, L.L. Schumaker, eds. *Curve and Surface Fitting: Saint-Malo 1999. Nashville: Vanderbilt University Press*, pp. 1–10.
- [34] E.J. Candès, D.L. Donoho, "New Tight Frames of Curvelets and Optimal Representations of Objects with Piecewise C2 Singularities", *Communications on Pure and Applied Mathematics*, Vol. 57 No. 2, pp. 219–266, 2004.
- [35] E.J. Candès, L. Demanet, D.L. Donoho, "Fast Discrete Curvelet Transforms", *Multiscale Modeling & Simulation*, Vol. 5 No. 3, pp. 861–899, 2006.
- [36] D. Donoho, "Ridge Functions and Orthonormal Ridgelets", *J. Approx. Theory*, Vol. 111 No. 2, 2001, pp. 143–179.
- [37] G. Zha, Q. He, C. Guan, J. Sun, M. He, "Speckle Suppression in Synthetic Aperture Radar Ocean Internal Solitary Wave Images with Curvelet Transform", *Acta Oceanologica Sinica*, Vol. 35 No. 9, pp. 13–21, 2016, doi: 10.1007/s13131-016-0929-3.
- [38] L. Dettori, L. Semler, "A Comparison of Wavelet, Ridgelet, and Curvelet-Based Texture Classification Algorithms in Computed Tomography", *Comput. Biol Med.* 2007 Apr, Vol. 37 No. 4, pp. 486-98. doi: 10.1016/j.combiomed.2006.08.002. Epub 2006 Oct 18. PMID: 17054933.
- [39] I.J Sumana, "Image Retrieval Using Discrete Curvelet Transform", Monash University, thesis, 2008.
- [40] V.N. Vapnik, "The Nature of Statistical Learning Theory", *Springer*, 1995.