Computers in Biology and Medicine 40 (2010) 384–391

Contents lists available at ScienceDirect

Computers in Biology and Medicine

journal homepage: www.elsevier.com/locate/cbm

A comparison of wavelet and curvelet for breast cancer diagnosis in digital mammogram

Mohamed Meselhy Eltoukhy [a,*], Ibrahima Faye [b], Brahim Belhaouari Samir [b]

[a] Electrical and Electronic Engineering Department, Universiti Teknologi PETRONAS, Bandar Seri Iskandar, 31750 Tronoh, Perak, Malaysia
[b] Fundamental and Applied Sciences Department, Universiti Teknologi PETRONAS, Bandar Seri Iskandar, 31750 Tronoh, Perak, Malaysia

ARTICLE INFO

Article history:
Received 24 February 2009
Accepted 1 February 2010

Keywords:
Wavelet
Curvelet
Breast cancer
Digital mammogram
Multiresolution
Feature extraction

ABSTRACT

This paper presents a comparative study between wavelet and curvelet transform for breast cancer diagnosis in digital mammogram. Using multiresolution analysis, mammogram images are decomposed into different resolution levels, which are sensitive to different frequency bands. A set of the biggest coefficients from each decomposition level is extracted. Then a supervised classifier system based on Euclidian distance is constructed. The performance of the classifier is evaluated using a 2×5-fold cross validation followed by a statistical analysis. The experimental results suggest that curvelet transform outperforms wavelet transform and the difference is statistically significant.

1. Introduction

Breast cancer is one of the most dangerous types of cancer among women all over the world. It happens to over 11% women during their life time. The world health organization's International Agency for Research on Cancer (IARC) estimates that more than a million cases of breast cancer will occur worldwide annually and more than 400,000 women die each year from this disease [1]. Early detection of breast cancer is essential in reducing life fatalities [2]. Digital mammography has been used in attempts to reduce the negative biopsy ratio and the cost to society by improving feature analysis and refining criteria for recommendation for biopsy. Digital mammography is a convenient and easy tool in classifying tumors and many applications in the literature proved its effective use in breast cancer diagnosis [3].

Image features extraction is an important step in image processing techniques. The features of digital images can be extracted directly from the spatial data or from a different space. Using a different space by special data transforms such wavelet or curvelet can be helpful to extract specific characteristics from a data. Detecting the features of image texture is a difficult process since these features are mostly variable and scale dependent [4–7].

Wavelet provides an efficient representation for images. In recent years, several schemes for mammogram analysis using wavelet were introduced. Liu et al. [8] proved that the use of multiresolution analysis of mammograms improves the effectiveness of any diagnosis system based on wavelets coefficients. In their mammogram analysis study, they used a set of statistical features with binary tree classifier. Ferreira and Borges [7] indicated that, the biggest wavelet coefficients in the low frequency of wavelet transform could be used as a signature vector for the corresponding mammogram. The obtained results showed that the biggest wavelet coefficients gave good classification accuracy rates. Rashed et al. [4] used a multiresolution mammogram analysis in multilevel decomposition to extract a fraction of the biggest coefficients. They used Daubechies-4, -8 and -16 wavelets with four level decompositions. They showed that the biggest coefficients in multilevel decomposition have a remarkably high efficiency.

Angelini et al. [9] presented a study to classify between two classes, masses versus non-masses. In order to find the optimal solution to this two class classification problem, three image representations were tested; a pixel based, discrete wavelet transform (DWT) representation and an overcomplete wavelet transform (OWT) representation. Support vector machine (SVM) was used as a classifier. The best possible results were achieved by DWT and OWT, both sophisticated image representations. Mousa et al. [6] proposed a system based on wavelet analysis and used the adaptive neuro-fuzzy inference system (ANFIS) for building the classifier to distinguish normal from abnormal and to determine whether the type of abnormality is mass or

* Corresponding author. Tel.: +60 17 517 86 34.
E-mail addresses: tokhy2478@yahoo.com, mtokhy@gmail.com (M. Meselhy Eltoukhy).

0010-4825/$ - see front matter © 2010 Elsevier Ltd. All rights reserved.
doi:10.1016/j.compbiomed.2010.02.002

microcalcification. The features were extracted by summing a predefined number of energy values together. The given results showed a successful classification rate.

Sakka et al. [10] carried out a comparative study on some wavelet functions which are widely used in microcalcification detection and feature extraction. The detection of microcalcification was achieved by decomposing the mammograms into different frequency sub-bands, and reconstructing the mammogram from the sub-bands containing only high frequencies, due to the fact that microcalcification correspond to high frequencies in the frequency domain of the image. Experimental results showed that the (sym8) wavelet achieved the best detecting result.

Yang et al. [11] did a comparative study for microcalcification detection in digital mammogram using wavelet. They decomposed the image into different resolution levels which are sensitive to different frequency bands. Several normal wavelet family functions were studied comparably, and for each wavelet function, different resolution levels were explored for detecting the microcalcification. Experimental results showed that the Daubechies wavelet with the 4th level decomposition achieved the best detecting result.

Moayedi et al. [12] presented a study of contourlet-based mammography mass classification using SVM. In their study, a set of statistical properties of contourlet coefficients from 4 decomposition levels, co-occurrence matrix features and geometrical features were used as feature vector of region of interest (ROI). Genetic algorithm was used for feature selection based on neural network pattern classification. They conclude that the contourlet offers an improvement of the classification process.

Curvelet efficiently represents discontinuities along edges or curves in images or objects [13]. Some studies using curvelet transform in image processing have been carried out. Ali et al. [14] presented a curvelet approach for the fusion of magnetic resonance (MR) and computed tomography (CT) images. They found that curvelet transform achieved good results in their fusion. Bind and Tahan [15] presented a method for object detection of speckle image based on curvelet transform. They constructed a segmentation method that provides a sparse expansion for typical images having smooth contours. Murtagh and Stark [16] used second, third, and fourth order moment of Multiresolution transform (wavelet and curvelet) coefficients as features, and *K*-nearest neighbors supervised classifier for image classification process.

Curvelet has also showed its efficiency for mammogram analysis. Eltoukhy et al. [17] presented a study of mammogram classification based on curvelet transform. The feature extraction depends on a percentage of the biggest coefficients from each decomposition level.

This paper introduces a comparative study between wavelet and curvelet transforms. The motivation is to determine through experimental work, which method is more efficient for representation, analysis and classification of breast cancer in digital mammograms.

In this study, each of mammogram images is decomposed using wavelet and curvelet separately. Then a set of the corresponding coefficients of each mammogram are extracted. Finally, a nearest neighbor classifier based on Euclidian distance is used to classify the mammogram images. The proposed system consists of two main steps: the first is to differentiate between normal tissue, benign and malignant tumors. The second step is to classify different types of abnormalities based on geometrical properties such as microcalcification clusters, circumscribed mass, spiculated mass, ill-defined mass, architectural distortion and asymmetry.

The remaining of this paper is organized as it follows. Section 2 gives a brief introduction to wavelet and curvelet transform. Section 3 discusses the experimental work. Section 4 presents the statistical analysis method. Results and discussions are introduced in Section 5, while Section 6 contains the conclusion of the work.

2. Preliminaries

2.1. Wavelet transform

2.1.1. Multiresolution and one-dimensional wavelet representation

The multiresolution approximation of one-dimensional signal $f(x) \in \mathbf{L}^2(\mathbf{R})$ at a resolution 2^j is defined as the orthogonal projection of a signal on subspace V_{2^j} of $\mathbf{L}^2(\mathbf{R})$. The approximation $A_{2^{j+1}}f(x)$ at resolution 2^{j+1} contains more information than the approximation $A_{2^j}f(x)$ at resolution 2^j. The details signal of $f(x)$ at resolution 2^j denoted by $D_{2^j}f(x)$. The details can be defined as the difference between $A_{2^{j+1}}f(x)$ and $A_{2^j}f(x)$. $D_{2^j}f(x)$ is equivalent to the orthogonal projection of $f(x)$ on the complement O_{2^j} of vector space V_{2^j} in $V_{2^{j+1}}$. According to the theory of multiresolution signal decomposition [18], there exists a unique scaling function $\varphi(x) \in \mathbf{L}^2(\mathbf{R})$ and a unique corresponding wavelet function $\psi(x) \in \mathbf{L}^2(\mathbf{R})$, where $\varphi_{2^j}(x) = 2^j\varphi(2^j x)$ and $\psi_{2^j}(x) = 2^j\psi(2^j x)$, such that $\{2^{-j/2}\varphi_{2^j}(x-2^{-j}k)\}_{k\in Z}$ and $\{2^{-j/2}\psi_{2^j}(x-2^{-j}k)\}_{k\in Z}$ are orthogonal bases of O_{2^j} and V_{2^j}, respectively. The approximation and detail signals of the original signal $f(x)$ at resolution 2^j are completely characterized by the sequence of inner products of $f(x)$ with φ_{2^j} and ψ_{2^j} as follows:

$$\{A_{2^j}f(k)\}_{k\in Z} = \{\langle f(o), \varphi_{2^j}(o-2^{-j}k)\rangle\}_{k\in Z} \tag{1}$$

$$\{D_{2^j}f(k)\}_{k\in Z} = \{\langle f(o), \psi_{2^j}(o-2^{-j}k)\rangle\}_{k\in Z} \tag{2}$$

Let H be a low-pass filter and G be a high-pass filter, where the impulse response of the filter H is $h(k) = \langle \varphi_{-1}(x), \varphi(x-k)\rangle$, and the impulse response of the filter G is $g(x) = \langle \psi_{-1}(x), \psi(x-k)\rangle$. Define $\tilde{H}$ with impulse response $\tilde{h}(k) = h(-k)$ to be the mirror filter of H, and $\tilde{G}$ with impulse response $\tilde{g}(k) = h(-k)$ to be the mirror filter of G. The multiresolution representation of $f(x)$ at any resolution 2^j can be implemented by a pyramidal algorithm as shown in Fig. 1

$$A_{2^{j-1}}f(x) = \sum_{k=-\infty}^{\infty} \tilde{h}(2x-k)A_{2^j}f(k) \quad \text{where } j = 0, -1, -2, \ldots \tag{3}$$

$$D_{2^{j-1}}f(x) = \sum_{k=-\infty}^{\infty} \tilde{g}(2x-k)A_{2^j}f(k) \quad \text{where } j = 0, -1, -2, \ldots \tag{4}$$

2.1.2. Two-dimensional wavelet representation

The wavelet model can be extended to two-dimensional signals by separable multi-resolution approximation of $\mathbf{L}^2(\mathbf{R}^2)$ with scaling function $\varphi(x,y) = \varphi(x)\varphi(y)$. And $\psi(x)$ is the one-dimensional wavelet function associated with $\varphi(x)$. There are three associated wavelet functions $\psi^1(x,y) = \varphi(x)\psi(y)$, $\psi^2(x,y) = \psi(x)\varphi(y)$ and $\psi^3(x,y) = \psi(x)\psi(y)$. With this formulation, the wavelet decomposition of a two-dimensional signal can be computed with a separable extension of the one-dimensional decomposition algorithm as shown in Fig. 2.

Fig. 3 illustrates the decomposition of the image $A_{2^{j+1}}f$ into $A_{2^j}f$, $D^h_{2^j}f$, $D^v_{2^j}f$, and $D^D_{2^j}f$ in the frequency domain. The images $A_{2^j}f$, $D^h_{2^j}f$, $D^v_{2^j}f$, and $D^d_{2^j}f$ corresponding to the lowest frequencies, the vertical high frequencies (horizontal edges), the horizontal high frequencies (vertical edges) and the high frequencies in both

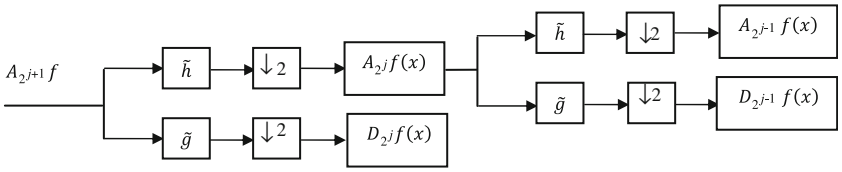

Fig. 1. A wavelet decomposition of a signal $A_{2^{j+1}}f$.

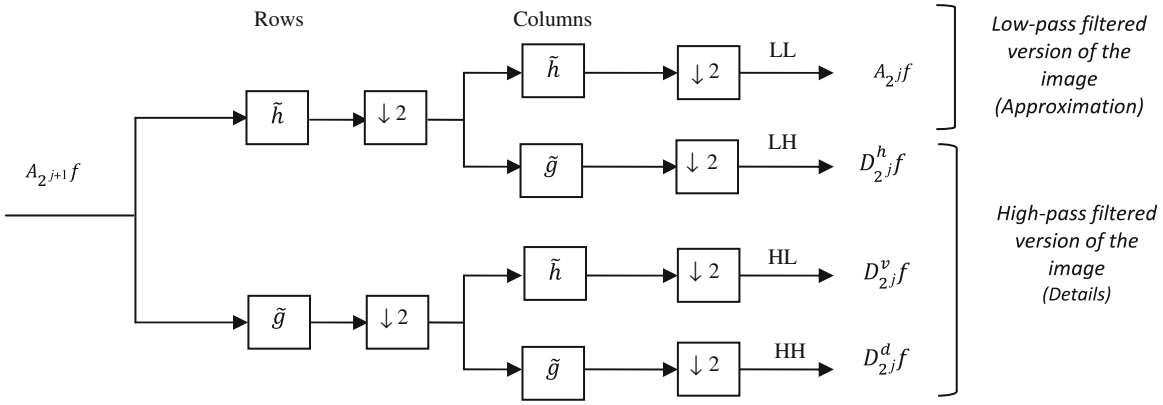

Fig. 2. A wavelet decomposition of an image, the outputs are approximation and details.

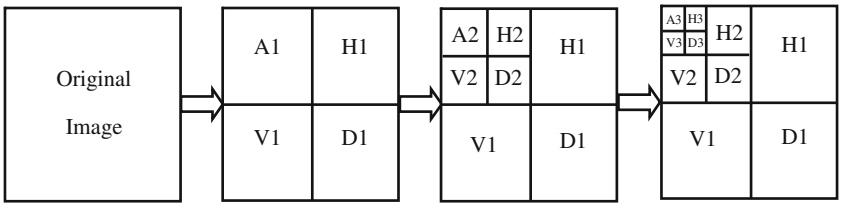

Fig. 3. Wavelet multiresolution decomposition for three levels.

directions (diagonal), respectively. i.e., the image $A_{2^{j+1}}f = A_{2^j}f + D^h_{2^j}f + D^v_{2^j}f + D^d_{2^j}f$. This set of images is called an orthogonal wavelet representation in two dimensions [18]. The image $A_{2^j}f$ is the coarse approximation at the resolution 2^j and the images $D^h_{2^j}f$, $D^v_{2^j}f$ and $D^d_{2^j}f$ give the detail signals for different orientations and resolutions. If the original image has N pixels, then each of the images $D^h_{2^j}f$, $D^v_{2^j}f$ and $D^d_{2^j}f$ will have 2^jN pixels ($j<0$), so that the total number of pixels in this new representation is equal to the number of pixels of the original image, to keep the volume of data maintained. This process can be summarized as, wavelet decompose an image into orthogonal sub-bands with low–low (LL), low–high (LH), high–low (HL), and high–high (HH) components which correspond to approximation, horizontal, vertical and diagonal, respectively. The LL sub-band is further decomposed into another four sub-bands low–low–low–low (LLLL) component, which represents the image approximation at this level, and then it is decomposed once again and so on [19].

2.2. Curvelet transform

The discrete curvelet transform is a new image representation approach. It was proposed by Candes and Donoh [20], from the idea of representing a curve as superposition of functions of various length and width obeying the curvelet scaling law $width \approx length^2$ [20]. Fig. 4 presents the curvelet analysis method.

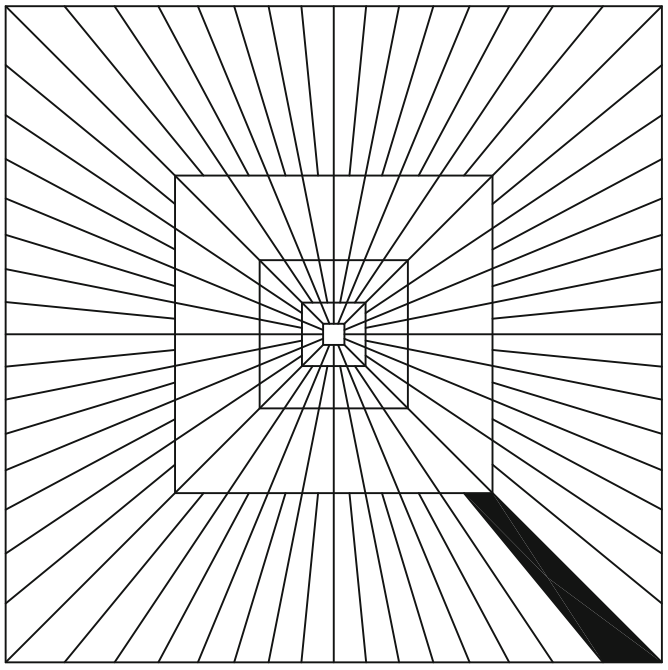

Fig. 4. Curvelet basic digital tiling. The shaded region represents one such typical wedge.

The work is done throughout in two dimensions, i.e., $\boldsymbol{R}^2$ with x as spatial variable, ω as frequency domain variable, r and θ as polar coordinates in the frequency domain. A pair of windows $W(r)$ and $V(t)$ are defined as the radial window and angular window, respectively. These are smooth, nonnegative and real-valued, with W taking positive real arguments and is supported on $r \in (1/2, 2)$ and V taking real arguments and is supported on $t \in [-1, 1]$. These windows will always obey the admissibility conditions [21]

$$\sum_{j=-\infty}^{\infty} W^2(2^j r) = 1, \quad r \in \left(\frac{3}{4}, \frac{3}{2}\right) \tag{5}$$

$$\sum_{l=-\infty}^{\infty} V^2(t-l) = 1, \quad t \in \left(\frac{-1}{2}, \frac{1}{2}\right) \tag{6}$$

For each $j \geq j_0$, a frequency window U_j is defined in the Fourier domain by

$$U_j(r,\theta) = 2^{-\frac{3}{4}j} W(2^{-j} r) V\left(\frac{2^{\lfloor j/2 \rfloor}\theta}{2\pi}\right) \tag{7}$$

where $\lfloor j/2 \rfloor$ is the integer part of $j/2$. Thus the support of U_j is a polar wedge defined by the support of W and V, applied with scale-dependent window widths in radial and angular directions. The symmetrized version of (7), namely, $U_j(r,\theta)+U_j(r,\theta+\pi)$ is used to obtain real-valued curvelet.

The waveform $\varphi_j(x)$ is defined by means of its Fourier transform $\hat{\varphi}_j(\omega) = U_j(\omega)$. Let $U_j(\omega 1, \omega 2)$ be the window defined in the polar coordinate system by (7). φ_j is the mother curvelet in the sense that all curvelets at scale 2^{-j} are obtained by rotations and translations of φ_j. Rotation angles $\theta_l = 2\pi 2^{-\lfloor j/2 \rfloor} l$ are introduced, with l=0,1,… such that $0 \leqslant \theta \leqslant 2\pi$, (the spacing between consecutive angles is scale dependent), sequence of translation parameters $k = (k_1, k_2) \in \boldsymbol{Z}^2$. The curvelets are defined as a function of $x = (x_1, x_2)$ at scale 2^{-j}, orientation angle θ_l and position $x_k^{(j,l)} = R_{\theta_l}^{-1}(k_1 \cdot 2^{-j}, k_2 \cdot 2^{-j/2})$ by

$$\varphi_{j,l,k}(x) = \varphi_j(R_{\theta_l}(x - x_k^{(j,l)})) \tag{8}$$

where R_θ is the rotation by θ radians and R_θ^{-1} is its inverse,

$$R_\theta = \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix}, \quad R_\theta^{-1} = R_\theta^T = R_{-\theta}.$$

A curvelet coefficient is the inner product of an element $f \in L^2(\boldsymbol{R}^2)$ and a curvelet $\varphi_{j,l,k}$,

$$c(j,l,k) := \int_{R^2} f(x)\overline{\varphi_{j,l,k}(x)}\, dx \tag{9}$$

where $\boldsymbol{R}$ denotes the real line. Curvelet transform obeys an anisotropy scaling relation, $length \approx 2^{-j/2}$, $width = 2^{-j}$, such that $width \approx length^2$. Fast digital curvelet transform can be implemented via two methods, using unequispaced FFTs or using wrapping [21]. In this paper, the method of unequispaced FFTs is used.

Curvelet has indeed more advantages in geometric features than wavelet making it superior over the later, like in the following cases [21]:

1. Optimally sparse representation of objects with edges.
2. Optimal image reconstruction in severely ill-posed problems.
3. Optimal sparse representation of wave propagators.

Suppose we have a function f which has a discontinuity across a curve, and which is smooth otherwise, and consider approximating f from the best m-terms in the expansion. For wavelet transform, the squared error of such an m-term expansion obeys [13]

$$\|f - f_{\tilde{W}}\|^2 \propto \frac{1}{m}, m \to +\infty \tag{10}$$

where $f_{\tilde{W}}$ is the approximation from m best wavelet coefficients.

While for curvelet transform it is

$$\|f - f_{\tilde{C}}\|^2 \propto \frac{1}{m^2}(\log m)^3, m \to +\infty \tag{11}$$

where $f_{\tilde{C}}$ is the approximation from m best curvelet coefficients.

The mean squared error is reduced in curvelet as compared to wavelet. As shown in Eqs. (10) and (11).

3. Experimental work

The proposed system is built based on multiresolution representation of the mammogram images by applying wavelet and curvelet. The largest 100 coefficients from each decomposition level are used for classification process. A nearest neighbor classifier based on Euclidian distance is used to classify the images by calculating the distances between the feature vectors and class core vectors. This section describes the dataset, feature extraction and classification method.

3.1. Dataset

In the present study, a set of images provided by the Mammographic Image Analysis Society (MIAS) [22] is used in applying the proposed technique. These images were previously investigated and labeled by an expert radiologist based on technical experience and biopsy. The dataset is selected due to the various cases it includes. It is also widely used in similar research work [4–8,12]. The dataset is composed of 322 mammograms of right and left breast, from 161 patients, where 51 were diagnosed as malignant, 64 as benign and 207 as normal. The abnormalities are classified into microcalcification, circumscribed mass, spiculated mass, ill-defined mass, architectural distortion, and asymmetry. In this study 142 mammogram images were selected as described in Table 1.

The original mammograms are 1024×1024 pixels, and almost 50% of the whole image comprised of the background with a lot of noise. Therefore a cropping operation is applied to the images to cut off the unwanted portions of the images. Regions of Interest (ROI's) 128×128 are cropped. The cropping process was performed manually, where the given center of the abnormality area is selected to be the center of ROI. Thus, almost all the background information and most of the noise are eliminated. By this method we are sure that no abnormality was suppressed with the background. An example of cropping that eliminates the label on the image and the black background is given in Fig. 5. Some examples to the ROI's are presented in Fig. 6.

3.2. Feature extraction

Once the images are cropped as described, both wavelet and curvelet transform methods are applied separately, and the features vectors are extracted. Features are extracted from the

Table 1
The distribution of selected cases from the MIAS dataset.

Class	Benign	Malignant	Total
Microcalcification	12	13	25
Circumscribed mass	19	4	23
Spiculated mass	11	8	19
Ill-defined mass	7	7	14
Architectural distortion	9	10	19
Asymmetry	6	9	15
Normal	–	–	27
Total	64	51	**142**

ROI based on a multiresolution transform. For wavelet, four different decomposition levels based on three different wavelet functions, Daubechies-8 (db8), symlet (sym8) and bi-orthogonal (bior3.7) are used. The used levels of decomposition and wavelet functions are selected based on previous work [4,10,11], the number of decomposition levels used for curvelet transform is 4 based on the work of candes [23].

In each decomposition level, the obtained coefficients are sorted in descending order. Then, the biggest 100 coefficients are extracted to represent the corresponding mammogram (i.e., feature vector). This means that each mammogram image is represented by 400 coefficients. Then these coefficients are passed to classification step.

3.3. Classification method

The Euclidian distance is used to design the nearest neighbor classifier. The dataset is divided into 5 samples two times, i.e. 2×5-fold cross validation. Then 10 experiments are performed. In each experiment a single sample is used to build the classes' core vectors and the remaining samples are used to test the model. The evaluation statistics for each method is then assessed via an average of 10 experiments followed by t-test statistical significant method.

For each class, the class core vector is calculated as the mean of set of the class vectors using Eq. (12). For a new mammogram image to be classified, the feature vector is extracted as discussed above (the biggest 100 coefficients from each level of decomposition), and then the distances between this feature vector and the class core vectors are calculated using Eq. (13). The system automatically classifies the feature vector in the class for which the distance obtained is the smallest.

$$V_{core}^{i} = \frac{1}{N}\sum_{j=1}^{j=N} V_{j}^{i} \tag{12}$$

$$Dist = \sqrt{\sum_{i=1}^{k}(V_{core}^{i} - V_{test}^{i})^{2}} \tag{13}$$

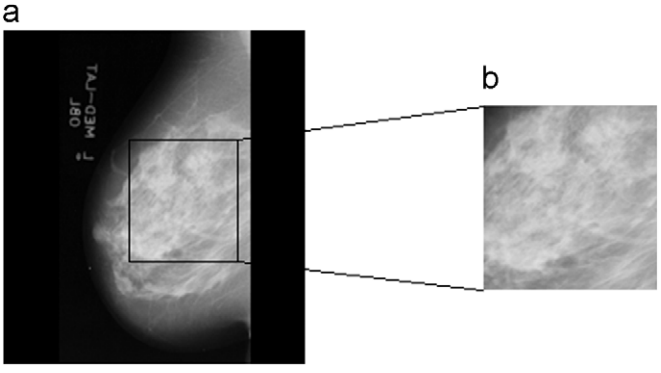

Fig. 5. (a) Original image (1024×1024), (b) Cropped image (128×128).

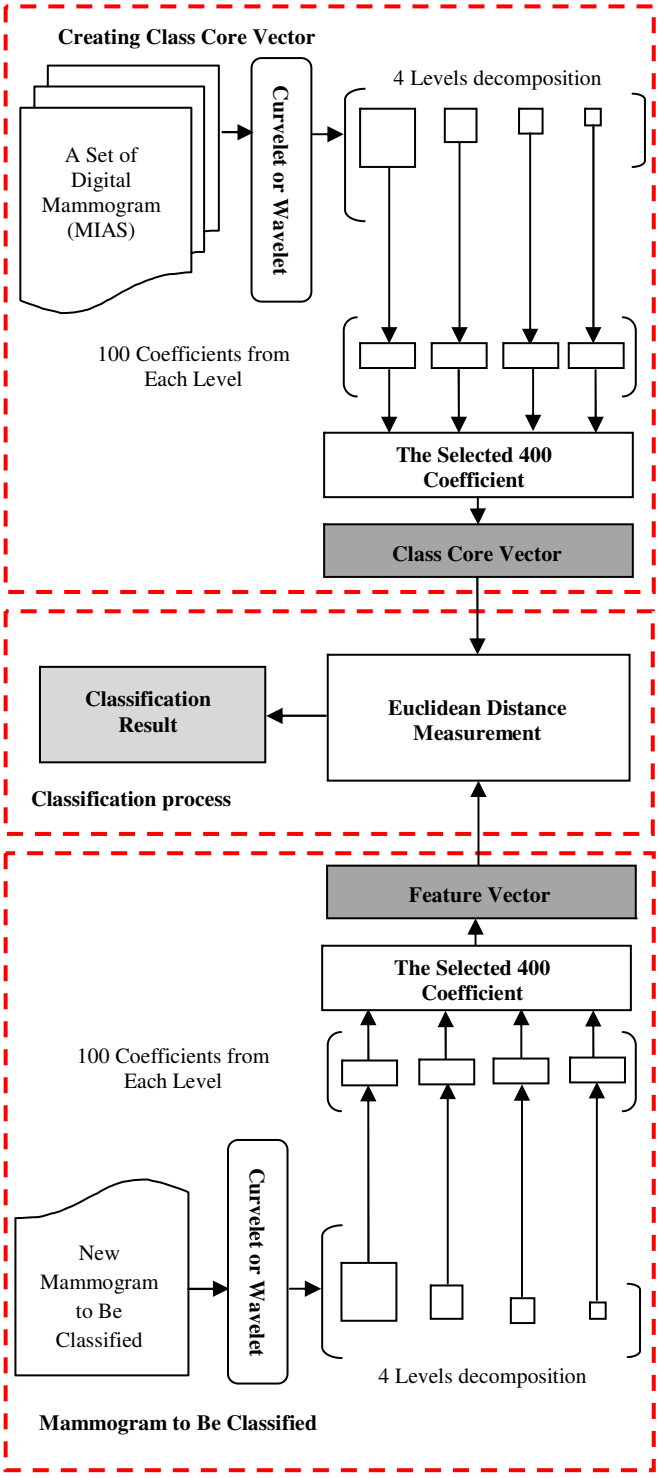

Fig. 7. The proposed multiresolution mammogram diagnosis system.

Fig. 6. Sample of the used mammogram images.

where V^i_{core} is the vector core of the corresponding class, V_j is the coefficient vector for ROI's of the corresponding class, i is the index of the vector, N is the number of images used to produce the class core vector, $Dist$ is the calculated distance between the tested image and the class core vector, k is the length of vector, and V^i_{test} is the feature vector of mammogram to be classified. The procedure of the feature extraction and classification method is summarized in Fig. 7.

The work in the experimental level consists of two different main functions. The first is to distinguish between three classes normal tissue, benign and malignant tumors. The second is to classify between six classes, microcalcification, circumscribed mass, spiculated mass, ill-defined (misclassified) mass, architectural distortion and asymmetry.

4. Statistical analysis

In this study, to test whether the classification rates for curvelet are significantly higher than those of wavelet, a paired t-test is performed on the results of the 2×5-fold cross validation. Let μ_c and μ_w be respectively the mean accuracies of curvelet and wavelet. The null hypothesis is that the difference between the means of the two techniques is zero ($H_0 : \mu_c - \mu_w = 0$), and the alternative hypothesis is that the difference is positive ($H_a : \mu_c - \mu_w > 0$). For a trial i of the cross validation, let P^i_w and P^i_c be respectively the results obtained by using wavelet and curvelet. The test statistic is computed as follows:

$$t = \overline{P}.\frac{\sqrt{n}}{\sqrt{\frac{\sum_i^n (P^i - \overline{P})^2}{n-1}}}$$

where $\overline{P} = 1/n \sum_1^n P^i$, $P^i = P^i_c - P^i_w$ and n is the number of runs (10 for this experiment) [24]. The P-value is obtained from a t-distribution table at the degree of freedom $(n-1)$, and is compared to the critical value 0.05 (i.e. 5% significance level). If the P-value is smaller than 0.05, the null hypothesis is rejected at 0.05 significance level.

5. Results and discussions

In the first step of the work, a tested image has to be classified as normal tissue, benign or malignant tumor. Table 2 shows the successful classification rate of mammogram images with the overall classification accuracy based on 2×5-fold cross validation. The average rate for each fold is calculated then the average for each representation method is calculated. Table 2 illustrated that the average classification rate achieved for the 3 classes (benign, malignant and normal) is 94.07% with curvelet coefficients, while the highest average rate achieved by wavelet functions (db8, Biro3.7 and sym8) is 90.05%. The average of the accuracy rates

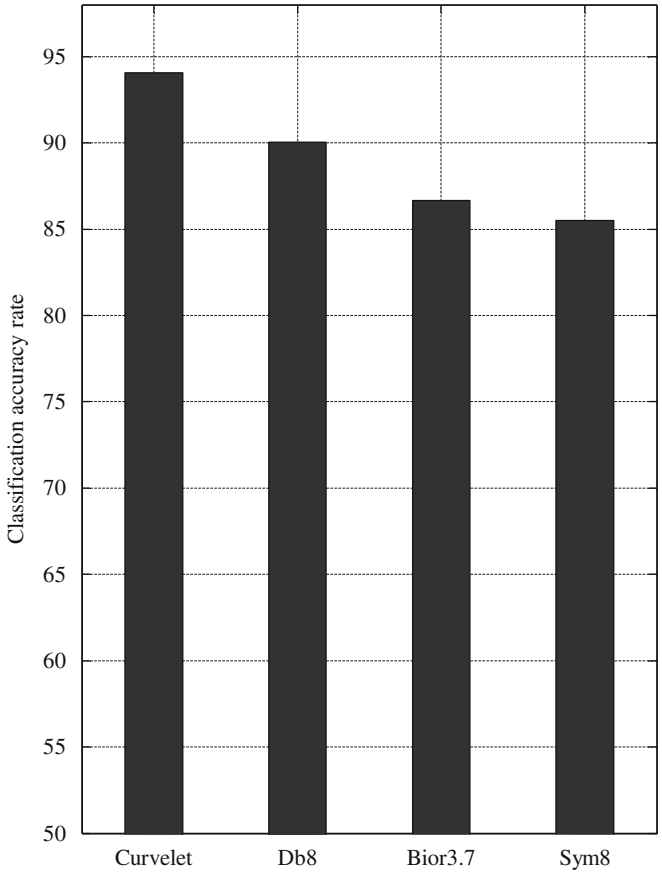

Fig. 8. The average of the results obtained for classification of normal, benign and malignant classes.

Table 2
Classification accuracy rates of normal, benign and malignant classes over the 2×5-fold.

Method	Class	Partitions 1					Partitions 2					Average
		Fold 1	Fold 2	Fold 3	Fold 4	Fold 5	Fold 1	Fold 2	Fold 3	Fold 4	Fold 5	
Curvelet	Benign	100.00	86.54	94.23	86.54	100.00	73.08	84.62	96.15	88.46	100.00	90.96
	Malignant	85.37	85.37	87.80	100.00	90.24	100.00	100.00	100.00	100.00	100.00	94.88
	Normal	100.00	100.00	100.00	100.00	90.91	77.27	100.00	100.00	100.00	95.45	96.36
	Average	95.12	90.64	94.01	95.51	93.72	83.45	94.87	98.72	96.15	98.48	**94.07**
Db8	Benign	100.00	90.38	84.62	78.85	100.00	61.54	65.38	80.77	96.15	100.00	85.77
	Malignant	85.37	92.68	100.00	87.80	70.73	100.00	85.37	102.44	82.93	100.00	90.73
	Normal	81.82	86.36	100.00	90.91	100.00	100.00	100.00	100.00	100.00	77.27	93.64
	Average	89.06	89.81	94.87	85.85	90.24	87.18	83.58	94.40	93.03	92.42	**90.05**
Bior3.7	Benign	65.38	84.62	76.92	76.92	53.85	65.38	76.92	76.92	80.77	73.08	73.08
	Malignant	100.00	100.00	82.93	100.00	100.00	85.37	85.37	100.00	100.00	100.00	95.37
	Normal	81.82	90.91	100.00	100.00	100.00	68.18	100.00	100.00	100.00	77.27	91.82
	Average	82.40	91.84	86.62	92.31	84.62	72.98	87.43	92.31	93.59	83.45	**86.75**
Sym8	Benign	100.00	88.46	50.00	76.92	61.54	63.46	80.77	84.62	100.00	100.00	80.58
	Malignant	100.00	100.00	60.98	85.37	100.00	85.37	60.98	60.98	82.93	100.00	83.66
	Normal	68.18	100.00	100.00	100.00	100.00	81.82	100.00	100.00	100.00	72.73	92.27
	Average	89.39	96.15	70.33	87.43	87.18	76.88	80.58	81.86	94.31	90.91	**85.50**

Table 3
Classification accuracy rates of abnormal classes over the 2×5 -fold.

Method	Class	Partitions 1					Partitions 2					Average
		Fold 1	Fold 2	Fold 3	Fold 4	Fold 5	Fold 1	Fold 2	Fold 3	Fold 4	Fold 5	
Curvelet	Calc	95.24	85.71	85.71	100.00	100.00	100.00	100.00	100.00	95.24	85.71	94.76
	Circ	95.00	85.00	90.00	95.00	95.00	90.00	100.00	95.00	95.00	95.00	93.50
	Spic	93.75	93.75	100.00	93.75	87.50	100.00	93.75	93.75	93.75	81.25	93.13
	Ill-def	100.00	100.00	100.00	100.00	90.91	100.00	100.00	100.00	100.00	100.00	99.09
	Arch	87.50	87.50	93.75	100.00	87.50	81.25	81.25	81.25	93.75	100.00	89.38
	Asym	100.00	91.67	100.00	100.00	91.67	75.00	100.00	100.00	100.00	100.00	95.83
	Average	95.25	90.61	94.91	98.13	92.10	91.04	95.83	95.00	96.29	93.66	**94.28**
Db8	Calc	100.00	78.95	100.00	84.21	89.47	100.00	100.00	100.00	73.68	52.63	87.89
	Circ	66.67	72.22	66.67	100.00	83.33	83.33	83.33	83.33	100.00	83.33	82.22
	Spic	73.33	80.00	66.67	93.33	86.67	60.00	93.33	73.33	73.33	73.33	77.33
	Ill-def	90.91	100.00	100.00	90.91	90.91	100.00	100.00	100.00	100.00	100.00	97.27
	Arch	73.33	66.67	73.33	100.00	66.67	86.67	100.00	100.00	100.00	73.33	84.00
	Asym	100.00	100.00	83.33	83.33	75.00	83.33	83.33	100.00	100.00	100.00	90.83
	Average	84.04	82.97	81.67	91.96	82.01	85.56	93.33	92.78	91.17	80.44	**86.59**
Bior3.7	Calc	63.16	73.68	78.95	78.95	68.42	57.89	84.21	73.68	73.68	73.68	72.63
	Circ	77.78	66.67	83.33	72.22	88.89	100.00	100.00	94.44	94.44	94.44	87.22
	Spic	73.33	73.33	73.33	66.67	80.00	100.00	53.33	100.00	100.00	100.00	82.00
	Ill-def	81.82	81.82	81.82	72.73	90.91	90.91	100.00	100.00	100.00	100.00	90.00
	Arch	80.00	53.33	73.33	73.33	66.67	100.00	100.00	100.00	66.67	100.00	81.33
	Asym	91.67	100.00	91.67	91.67	83.33	75.00	75.00	75.00	100.00	100.00	88.33
	Average	77.96	74.81	80.41	75.93	79.70	87.30	85.42	90.52	89.13	94.69	**83.59**
Sym8	Calc	100.00	73.68	100.00	100.00	94.74	100.00	100.00	100.00	73.68	52.63	89.47
	Circ	100.00	72.22	83.33	72.22	77.78	100.00	83.33	83.33	83.33	83.33	83.89
	Spic	66.67	60.00	93.33	66.67	80.00	66.67	73.33	73.33	100.00	100.00	78.00
	Ill-def	100.00	81.82	100.00	100.00	90.91	100.00	100.00	100.00	100.00	100.00	97.27
	Arch	100.00	80.00	100.00	100.00	73.33	86.67	100.00	100.00	66.67	60.00	86.67
	Asym	91.67	100.00	91.67	100.00	83.33	83.33	83.33	83.33	100.00	100.00	91.67
	Average	93.06	77.95	94.72	89.82	83.35	89.44	90.00	90.00	87.28	82.66	**87.83**

achieved for classification between benign, malignant and normal is shown in Fig. 8.

For the second step, the classification rates of the abnormalities of 2×5-fold cross validation are listed in Table 3. The average rate for each fold is calculated then the average for each representation method is calculated. Table 3 shows that, the average successful classification rate for all classes is 94.28% by using curvelet transform coefficients. For wavelet functions (db8, Biro3.7 and sym8) the highest average rate obtained is 87.83%. Fig. 9 presents average of classification accuracy rates achieved for this step of the work.

For both problems considered in this study, the obtained results suggest that curvelet-based features performs better that wavelet-based features. This goes in the same line with the expectations since the curvelet transform is able to capture multidimensional features in wedges as opposed to points in wavelet transform. A hypothesis test is performed to evaluate the significance of the difference between the performances of two techniques. The results at 5% significance level are summarized in Table 4. The null hypothesis is rejected at 0.05 significance level for the two problems and the different wavelets considered. It means that the classification rates obtained using curvelet are higher than those obtained using wavelet and the difference is statistically significant. It is noted that, except for the first result, all *P*-values are smaller than 0.01, which means that the differences are actually statistically highly significant.

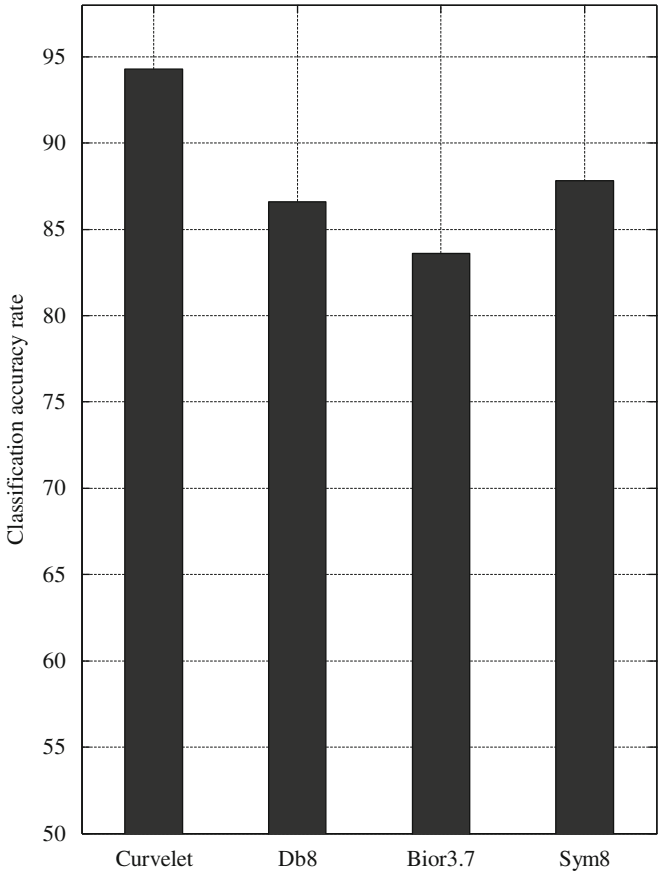

Fig. 9. The average of the results obtained for classification of abnormal classes.

6. Conclusion

In this paper, the differences between wavelet- and curvelet-based methods in digital mammogram analysis and classification

Table 4
The results of t-test at significance level α=5%.

Function	Method	T-value	P-value	Null hypothesis (H_0)
Benign, Malignant, Normal	Curvelet vs. Db8	2.7806	0.0107	Reject
	Curvelet vs. Bior3.7	4.7309	5.3626×10^{-4}	Reject
	Curvelet vs. Sym8	3.3420	0.0043	Reject
Abnormal classes	Curvelet vs. Db8	5.9521	1.0740×10^{-4}	Reject
	Curvelet vs. Bior3.7	4.7495	5.2245×10^{-4}	Reject
	Curvelet vs. Sym8	4.8725	4.4021×10^{-4}	Reject

are discussed. Firstly, each mammogram image is decomposed using wavelet and curvelet transforms separately. The 100 biggest coefficients are extracted from each decomposition level. Then, Euclidian distance is used to construct a nearest neighbor classifier. The performance of the classifier is evaluated using a 2×5-fold cross-validation test. The experimental results show that the extracted features based on curvelet give a better performance as compared to wavelet. For classification of normal, benign and malignant, based on curvelet transform coefficients, the classifier achieves 94.07%, while the highest rate achieved by wavelet coefficients is 90.05%. For abnormal classes, the curvelet transform coefficients make the classifier achieves classification rate of 94.28% as compared to wavelet functions (db8, Biro3.7 and sym8), for which the highest rate obtained is 87.83%.

Finally a hypothesis test is performed on the accuracies obtained from the 2×5-fold cross validation. The results show that the successful classification rates obtained with curvelet coefficients are higher than those obtained with wavelets and the differences are statistically significant.

Conflict of interest statement

None declared.

Acknowledgment

The authors wish to thank the reviewers for their invaluable and constructive comments.

References

[1] N.R. Pal, B. Bhowmick, S.K. Patel, S. Pal, J. Das, A multistage neural network aided system for detection of microcalcifications in digitized mammograms, Neurocomputing 71 (2008) 2625–2634.
[2] J.C. Fu, S.K. Lee, S.T.C. Wong, J.Y. Yeh, A.H. Wang, H.K. Wu, Image segmentation, feature selection and pattern classification for mammographic microcalcifications, Computerized Medical Imaging and Graphics 29 (2005) 419–429.
[3] H.D. Cheng, X. Cia, X. Chen, L.H. Lou, Computer aided detection and classification of microcalcification in mammogram: a survey, Pattern Recognition Letters 36 (2003) 2967–2991.
[4] E.A. Rashed, I.A. Ismail, S.I. Zaki, Multiresolution mammogram analysis in multilevel decomposition, Pattern Recognition Letters 28 (2007) 286–292.
[5] S.N. Yu, K.Y. Li, Y.K. Huang, Detection of microcalcifications in digital mammograms using wavelet filter and markov random field model, Computerized Medical Imaging and Graphics 30 (2006) 163–173.
[6] R. Mousa, Q. Munib, A. Moussa, Breast cancer diagnosis system based on wavelet analysis and fuzzy-neural, Expert Systems with Applications 28 (2005) 713–723.
[7] C.B.R. Ferreira, D.L. Borges, Analyses of mammogram classification using a wavelet transform decomposition, Pattern Recognition Letters 24 (2003) 973–982.
[8] S. Liu, C.F. Babbs, E.J. Delp, Multiresolution detection of spiculated lesions in digital mammograms, IEEE Transactions on Image Processing 10 (6) (2001) 874–884.
[9] E. Angelini, R. Campanini, E. Iampieri, N. Lanconelli, M. Masotti, M. Roffilli, Testing the performance of image representations for mass classification in digital mammograms, International Journal of Modern Physics C 17 (2006) 113–131.
[10] E. Sakka, A. Prentza, I.E. Lamprinos, D. Koutsouris, Microcalcification detection using multiresolution analysis based on wavelet transform, Proc.in: Proceeding of the International Special Topic Conference on Information Technology in Biomedicine (IEEE-ITAB2006), Ioannina, Epirus, Greece, October 26–28, 2006.
[11] J.C. Yang, J.W. Shin, D.S. Park, Comparing study for detecting microcalcifications in digital mammogram using wavelets, Lecture Notes in Computer Science, vol. 3177, 2004, pp. 409–415.
[12] F. Moayedi, Z. Azimifar, R. Boostani, S. Katebi, Contourlet based mammography mass classification, ICIAR 2007, Lecture Notes in Computer Science, vol. 4633, 2007, pp. 923–934.
[13] K.P. Soman, K.I. Ramachandran, Insight into Wavelets: From Theory to Practice, second ed., Prentice-Hall, 2006.
[14] F.E. Ali, I.M. Eldokany, A.A. Saad, F.E. Abdelsamie, Curvelet fusion of MR and CT images, Progress in Electromagnitics Research C 3 (2008) 215–224.
[15] N.T. Binh, N.C. Thanh, Object detection of speckle image base on curvelet transform, ARPN Journal of Engineering and Applied Sciences 2 (3) (2007) 14–16.
[16] F. Murtagh, J. Starck, Wavelet and curvelet moments for image classification: application to aggregate mixture grading, Pattern Recognition Letters 29 (2008) 1557–1564.
[17] M.M. Eltoukhy, I. Faye, B.B. Samir, Using curvelet transform to detect breast cancer in digital mammogram, proc.in: proceeding of the fifth International Colloquium on Signal Processing and its Application (CSPA), Kuala Lumpur, 2009.
[18] S.G. Mallat, A theory for multiresolution signal decomposition: the wavelet representation, IEEE Transactions on Pattern Analysis and Machine Intelligence 7 (11) (1989) 674–693.
[19] M. Al-Qdaha, A. Ramlib, R. Mahmud, A system of microcalcifications detection and evaluation of the radiologist: comparative study of the three main races in Malaysia, Computers in Biology and Medicine 35 (2005) 905–914.
[20] E. Candes, D. Donoho, Curvelets: multiresolution representation, and scaling laws, in: A. Aldroubi, A.F. Laine, M.A. Unser (Eds.), Wavelet Applications in Signal and Image Processing VIII, Proc.eeding of the SPIE 4119, 2000.
[21] E. Candes, L. Demanet, D. Donoho, L. Ying, Fast discrete curvelet transforms, Multiscale Model. Simul.ing and Simulation 5 (2006) 861–899.
[22] 〈http://peipa.essex.ac.uk/ipa/pix/mias〉.
[23] E. Candes, 〈http://www.curvelet.org〉, CurveLab-2.1.2 Toolbox, 2004.
[24] T.G. Dietterich, Approximate statistical tests for comparing supervised classification learning algorithms, Neural Computation 10 (7) (1998) 1895–1923.

Mohamed Meselhy Eltoukhy received the BSc degree from the Department of Computer Science, in 1999, and the Master of Computer Science degree from the Minufiya university, Egypt, in 2006. He is working as lecturer in Akhbar Elyoum Academy, Egypt. Now he studies Ph.D. program at the Department of Electrical & Electronic Engineering, Universiti Teknologi PETRONAS, Tronoh, Perak, Malaysia. His research interests are image processing and artificial intelligence.

Ibrahima Faye is a senior lecturer at the Department of Fundamental and Applied Sciences Department, Universiti Teknologi PETRONAS, Tronoh, Perak, Malaysia. He received his Bachelor, Master and Ph.D. in Mathematics from University of Toulouse and his specialized Master in medical imaging from Ecole Centrale, Paris.

Brahim Belhaouari Samir received the engineer diploma in electrical engineering from the Telecommunication Institute of Oran, Algeria, in 1998, Master degree from INP/EENSET Toulouse, France 2000. Ph.D. in Stochastic Processes from EPFL, Lausanne Switzerland in 2006. After completing the Ph.D. degree, he held a Postdoctoral at Ecole Polytechnique Federale de Lausanne, Lausanne, Switzerland. Now, he is a senior lecturer at the Department of Fundamental and Applied Sciences Department, Universiti Teknologi PETRONAS, Malaysia. His main research interests are in data analysis, statistical signal processing, and pattern recognition.