Independent Component Analysis and Complex Wavelet Decomposition for Classifying Medical Data

Corina Săraru, Luminiţa State, Maria Miroiu
Faculty of Mathematics and Computer Science
University of Piteşti
Piteşti, Romania
corina.sararu@upit.ro, lstate@clicknet.ro, maria.miroiu@upit.ro

Abstract

In this article, we describe a new classification methodology based on the use of Independent Component Analysis and Wavelet decomposition (ICAW) techniques.

An ensemble system of classifiers is built such that each classifier independently decides the assignation of the test examples on several representations resulted by taking projections computed by wavelets and Independent Component Analysis (ICA).

The representations used by the individual classifiers are obtained by taking the real and imaginary part of the wavelet decompositions, as well as the magnitude and phase.

The decision of the ensemble system is based on several types of voting rules (such as the majority voting rule or a weighted voting rule).

The experimental results presented in the paper show that the proposed ensemble systems of classifiers provide higher accuracy in the particular problem of classifying biomedical data.

Keywords: : independent components analysis, wavelet decomposition, pattern recognition, signal processing.

I. Preliminaries

Processing biomedical data becomes more and more important nowadays, because of its relevance and support for the decisions of the specialists. Due to its nature, the process of acquisition for medical information about a certain patient usually supplies data that contains a significant amount of noise.

In the particular problem of classifying certain types of biomedical signals, many techniques that extract relevant information from the data have been used.

In order to detect abnormalities in informational data, Principal Component Analysis (PCA) is used. The idea behind extracting the principal components is to find the spatial directions of the data set, that have the maximum data variance ([1], [2]).

PCA is a very efficient technique used widely to obtain relevant statistical description of the data. It is extensively applied in preprocessing and classification steps of the information in several domains. Its popularity comes from the fact that it uses first and second order statistics in order to characterize the data sets, thus giving a high level of confidence in the obtained results [3].

Although PCA is very useful in the case of extracting relevant information from some data sets, there are other techniques that use higher-order statistics in characterizing the data.

Independent Component Analysis (ICA) may be regarded as a particular case of blind source separation problem ([4], [5], [6]) that attempts to separate all underlying sources contributing to the data without knowing these sources or the way that they are mixed.

In order to achieve the goal of separating independent components from mixed signals, the ICA model does not need any prior knowledge about each source.

The most important assumption that is associated with the ICA model is the independence of the sources to be estimated.

Another assumption used in the ICA model is the following: the signal from each sensor (that is, an observable variable) has different mixing ratio of the independent components. ICA has been introduced to mechanical dynamic signal analysis in the last few years ([7], [8]).

ICA is more suitable when the purpose is to find a component from a mixture of many independent sources. However, in some circumstances, there is limitation to install too many sensors to satisfy the requisition for ICA.

In our application, the signals received from the sensors are formally represented by the examples in the data set that is taken into consideration.

The sources are, in fact, the independent components estimated by the ICA model or the latent variables depending on which the recorded signals can be expressed.

In biomedical data, noise is almost always present, because of the residuals in the signals, coming from other body activities [9].

The main disadvantage of the various types of approaches in the problem of noise reduction is that they are time consuming and computationally expensive. That is why the idea of preprocessing the data is much more popular nowadays.The noise reduction may be done by applying suitable low pass filters in time domain before implementing ICA algorithm [10].

A wavelet transform can focus on localized signal structures with a zooming procedure that progressively reduces the scale parameter. Singularities and irregular structures often carry essential information in a signal. For example, discontinuities in images may correspond to occlusion contours of objects in a scene.

Taking into account a wavelet transform of the observed data may be regarded as a filtering procedure. It is performed on localized signal parts, a scaling procedure allowing to change the values of the corresponding parameters. It is generally accepted that singularities and irregular parts of an signal often contain essential information about the objects in the signal.

Singularities and edges are detected from wavelet transform local maxima at multiple scales. These maxima define a geometric scale/space support from which signal or image approximations are recovered. Non-isolated singularities appear in highly irregular signals such as multifractals [11].

The idea behind applying wavelet decomposition before feeding the observation data to ICA is to improve the assumption of non-gaussianity distribution of sources enforced for ICA algorithm and increasing the independency of sources. The projection of data to a set of orthogonal basis function in wavelet domain produces fewer coefficients to represent the data leading to super-gaussian distribution of data.

Removing noise from signals is possible only if some prior information is available. This information is encapsulated in an operator designed to reduce the noise while preserving the signal. Ideally, the joint probability distribution of the signal and the noise is known. Bayesian calculations then derive optimal operators that minimize the average estimation error. However, such probabilistic models are often not available for complex signals such as natural images.

In order to improve the accuracy with noise present in data, the kNN algorithm introduces a parameter k so that for each new example q to be classified the classes of the k nearest neighbors of q are considered: q will be labeled with the majority class. Another alternative consists in assigning that class whose average distance is the smallest one or introducing a heuristically obtained threshold $k1 < k$ so that the assigned class will be that with a number of associated examples greater than this threshold [12].

This type of classification [13] implies the search of a group of k objects in the training set that are closest to the test object, and bases the assignment of a label on the predominance of a particular class in this neighborhood. This addresses the issue that, in many data sets, it is unlikely that one object will exactly match another, as well as the fact that conflicting information about the class of an object may be provided by the objects closest to it. There are several key elements of this approach: (i) the set of labeled objects to be used for evaluating a test objects class (ii) a distance or similarity metric that can be used to compute the closeness of objects, (iii) the value of k, the number of nearest neighbors, and (iv) the method used to determine the class of the target object based on the classes and distances of the k nearest neighbors. In its simplest form, k-NN can involve assigning an object the class of its nearest neighbor or of the majority of its nearest neighbors.

In this paper an heuristic classifier design based on ICA and wavelets (ICAW) methods is presented. In the classification process of the individual and ensemble systems, the k-NN method is also used. The particular metric used in experiments is the cosine distance.

Being given a database containing a set of observations coming from two classes such that the true provenance class is known for each individual, new representations are computed in terms of the features extracted using the wavelets described in Equation 8.

The quality of each type of features is tested by evaluating the empirical error when the k-NN classification technique is applied to the resulted representations.

The FastICA algorithm is applied to the initial database and to each collection of representations in terms of the above mentioned features.

The design of the new classifier corresponds to a weighted voting procedure that combines the decisions of the ensemble of resulted classifiers where the weights are given by the corresponding correctness scores computed for each classifier.

II. Theoretical Framework: ICA and Wavelet Decomposition

A. Description of Independent Component Analysis

ICA has become an important signal processing and data analysis technique; it is a particular case of blind source separation and it is used on a wide range of data, such as biomedical, acoustical and astrophysical signals.

Generally speaking, ICA is viewed (Jutten [14] , Cardoso [15], Jutten and Herault [16], Comon [17] , Hyvarinen et al. [6]) as a statistical signal processing technique that models a set of observations, x, with an instantaneous linear mixing of independent latent variables, s

$$x(t) = As(t) + ns(t) \tag{1}$$

where ns is additive noise.

ICA supplies a series of techniques allowing the decomposition of a random vector in linear components which are "as independent as possible", where the independence should be understood in its strong statistical sense.

The problem of recovering sources from their linear mixtures without knowledge of the mixing channel can be expressed in its simplest form as the problem of identifying the factorization of the N-dimensional observations x into a mixing channel A and M-dimensional sources s, a large body of work being devoted to the case when the statistical independency of sources is assumed.

The goal of ICA is to recover the latent components from the observations. If noise is negligible, this can be achieved by the determination of an inverse linear mapping from x to s, say

$$x = As \tag{2}$$

The equation represents a simplified ICA model, resembling the one presented in equation 1, in which the noise is considered to be negligible and the time component is implicit.

Provided the model is used to estimate the sources, or the latent variables, denoted by s, the simplified ICA model may be written in the following form [4]:

$$s = Bx \tag{3}$$

In order to determine the matrix B, usually, an intuitively justified criterion function is selected, yielding to an unconstrained optimization problem.

Most algorithms, directly or indirectly, minimize the mutual information, I, between the component estimates. It can be shown (Hyvarinen [5], [18], Hyvarinen et al. [6]) that minimizing I corresponds to the maximization of the negentropy, a measure of non-Gaussianity of the components.

Most of the existing ICA algorithms can be viewed as approximating negentropy through simple measures, such as high-order cumulants (Cardoso [15], Hyvarinen [19], Hyvarinen et al. [6]).

In our work, we used the FastICA algorithm introduced by [4]. In [4] the negentropy is approximated by

$$F(y) = [EG(y) - EG(\nu)]^2 \tag{4}$$

where G is an nonquadratic function, ν and $y = w^T z$ are Gaussian variables of zero mean and unit variance, yielding to the constraint optimization problem [4]:

$$maxF(w^T z), ||w||^2 = 1 \tag{5}$$

B. Wavelet Transforms

The most popular decomposition in the domain of the frequencies is obtained using the Fourier transform. The elementary waveforms that appear in the decomposition have the same time and frequency resolution. Hence, this type of transform is efficiently characterizing the signal if it does not contain different time-frequency resolutions. The main disadvantage of the Fourier transform is that it can not characterize signal.

In the medical domain, many of the signals describing a certain activity possess the property of being very localized in time in a different way than being localized in frequency [11].

In order to have a better understanding of the nature of the signal, one needs to be able to locally analyze the information in the signal. Unlike the Fourier transform, a wavelet transform provides the means to overcome the problem of changing time-frequency resolution. The introduction of two parameters regarding time and space, denoted by u and s.

A wavelet transform performs on localized signal structures, a zooming procedure allowing to progressively reducing the value of the scale parameter. It is generally accepted that singularities and irregular structures often contain essential information in a signal. For example discontinuities in images may correspond to occlusion contours of objects in a scene. Singularities and edges are detected using the local maxima of wavelet transform at different scales. These maxima define a geometric scalespace support from which signal and image approximations are recovered [11].

A wavelet is a real valued function ψ, having the property that

$$\int_{-\infty}^{+\infty} \psi(t)dt = 0 \tag{6}$$

Using a given wavelet (referred to as the mother wavelet), ψ, to model the window shape and scaling and translation transforms, the following family of functions is obtained ([11]):

$$\psi_{u,s}(t) = \frac{1}{\sqrt{(s)}}\psi\left(\frac{t-u}{s}\right). \tag{7}$$

where the parameters s and u are used for scaling and translating the windows.

So far there have been introduced a long series of types of wavelets, the option about the shape of the window being imposed by the particular problem that is handled.

We intend to use wavelet based techniques in processing sets of observations taken on electrocardiographic data. Being given the specificity of this problem, our option is to use as the mother wavelet ψ a member of the complex Gauss wavelet family , defined by:

$$\psi(n) = f^{(n)} \tag{8}$$

where $f^{(n)}$ is the n-th derivative of the function $f(x) = C_n e^{-ix} e^{-x^2}$ and C_n is chosen such that $||f^{(n)}||^2 = 1$.

The features extracted from signals result by convoluting the signal f with the windows defined by the wavelets derived from the mother wavelet (8) ([20]):

$$Wf(u,s) = < f, \psi_{u,s} >= \tag{9}$$
$$\int_{-\infty}^{+\infty} f(t)\frac{1}{\sqrt{(s)}}\psi\left(\frac{t-u}{s}\right)dt.$$

where $f \in \mathbf{L}^2(\mathbf{R})$.

III. The ICAW-k-NN Algorithm

The scheme of the classification system is represented in Figure 1.

The Preprocessing module consists in applying denoising methods to clean the initial input data.

The process involves using a particular family of wavelets called symlets ([21]), in order to reduce the noise present in the signals.

The symlets are a particular case of the widely used Daubechies wavelets ([21]); the feature that characterizes symlets is a high degree of asymmetry, as compared to traditional Daubechies, which are symmetric wavelets.

In the module Complex Wavelet Decomposition, the cleaned data are processed in order to extract significant features using (9).

In order to execute the feature extraction, the convolution $R^i = \psi \star X^{\ i}$ is computed for each sample from the design set X^i and $R = (R^1, R^2, ..., R_1^N)$.

The mixing matrices are estimated using the FastICA algorithm [4] and the computation of

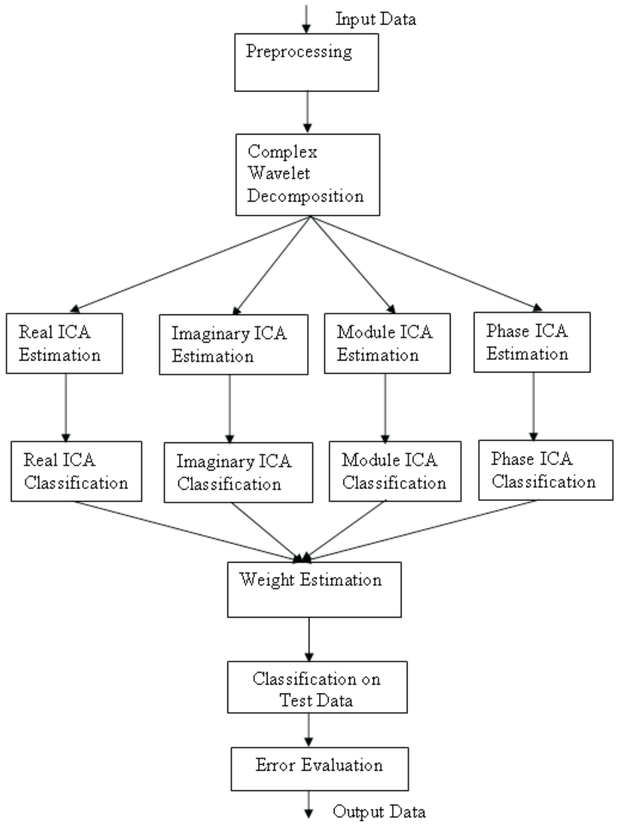

Fig. 1. The classifier design

$$\begin{aligned} real(R) &= A_1 S_1 \\ imag(R) &= A_2 S_2 \\ abs(R) &= A_3 S_3 \\ phase(R) &= A_4 S_4 \end{aligned} \tag{10}$$

is carried out.

The classification of each of the classifiers Real ICA, Imaginary ICA, Module ICA and Phase ICA is performed using the corresponding representations. The FastICA algorithm is applied in order to identify its latent independent structure.

Each of the resulted databases is classified using the kNN method and the performance is evaluated in terms of the empirical error against the true known classifications.

The resulted correctness scores are fed in the Weight Estimation module and combined in order to determine a set of classification weights used in the voting procedure that combines the classification decisions corresponding to each classifier into a final decision.

The performance of the resulted ensemble classifier is evaluated in terms of the empirical error against the true known classifications.

In case the classification error is not acceptable the system can ask for further data and/or to initiate more refined de-noising techniques.

The classification task is solved using three classifiers referred in Table 1 as Ensemble1+ICA, Ensemble2+ICA and Ensemble3+ICA. These classifiers are essentially modified versions of the 2-NN classification rule by including a voting mechanism that involves the decisions of individual classifiers. In Ensemble1+ICA the voting mechanism is a simple majority rule.

The Ensemble2+ICA classifier uses a more refined classification rule, involving the connectivity of individual classification. The decisions of individual classifiers are combined additively using the weights computed in the Weight Estimation module. The logistic function is applied to the resulted value and the class is computed by applying the threshold θ_1.

In our tests the data contains equal sized samples coming from the normal/abnormal classes and for this reason we set $\theta_1 = 0.5$.

In cases when the data contains $N = N' + N"$ samples, where N' and $N"$ are the sizes of data coming from the class labeled $-1/1$ and $N' \neq N"$, the value of the threshold is set by taking into account the sizes of data coming from each class in the design phase and test phase respectively.

The rule of the collective decision is obtained by computing the outcome of the error function, with the threshold $\theta_2 = 0$.

The training phase aims to obtain the weights used to classify new data. We use N_1 samples from the total amount of N to estimate the weights and N_2 samples to evaluate the performance.

The N_1 samples are submitted to be classified by nine classifiers, (see Table 1) and the empirical errors are computed for each classifier.

In the Weight Estimation module the empirical errors are used to compute the weight of each classifier, the weight being the ratio of the success rate and the overall sum of success rates.

The testing procedure consists of submitting the remaining N_2 samples for being classified by Ensemble1+ICA, Ensemble2+ICA and Ensemble3+ICA.

The results of our tests on the MIT-BIH database are presented in Table 1.

IV. Experimental Results

The electrocardiogram (ECG) signal is the electrical interpretation of the heart activity; it consists of a set of well defined, successive waves denoted: P, Q, R, S, and T waves.

Of a particular interest is the adequate and accurate analysis in order to identify possibly cardiac anomalies [22].

Unfortunately, very often the ECG signals are not accurate enough because the measurements are affected by uncontrollable noise.

Consequently, the ECG signals have to be preprocessed in order to clean them up, that is to remove, at least partially, the noise.

Different developments have been proposed to design filtering algorithms aiming to improve the signal to noise ratio and to recover the ECG waves in the framework of different noisy environments [23].

The denoising based on wavelet theory [24] has been extensively exploited in filtering noisy ECG. [25].

The analysis of the different DWT levels shows that the first level detail sequence of the noisy ECG signal is highly dominated by the wavelet gaussian noise (WGN) energy [25].

When the biomedical signals are corrupted by some artifact, a preprocessing step is needed in order to extract relevant clinical information from the data.

For this reason the artifact cancelation is a key topic in biomedical data processing [22], [23]. In particular, the artifact removal is often necessary for the clinical study of the electrocardiographic (ECG) signal [22].

The ECG signal looks like a repeating and almost periodic pattern. This characteristic of physiological signals was explored in order to synchronize the parameters of the filter with the period of the signal.

For instance, Liang and Lin [26] proposed an efficient method based on discrete wavelet transform (DWT) in order to perform the cancelation of stimulus artifact in the serosal recordings of gastric myoelectric activity, but it works well only when there is no interference between the filter and the ECG signal.

Aiming to remove the artifacts in biomedical signals, even in the presence of interference with the ECG signals, several methods based on Independent Component Analysis (ICA) have been also proposed [27], [28].

The methodology that was previously described in Section 3 was applied on the MIT-BIH database, http://www.physionet.org/physiobank/database/, containing samples of ECG signals taken from patients with/without supraventricular arrhythmia, for each sample being provided the correct diagnosis.

In our tests, we considered 35 records from which 25 records came from patients with supraventricular arrhythmia.

Each record represents the signal measured using a 128Hz sampling frequency for 10 seconds.

The training data consisted of all records coming from patients without supraventricular arrhythmia and 10 records selected from the remaining 25 signals.

The classifier denoted by Std implements the standard 2-NN classification rule on the unprocessed samples, while the ICA classifier uses the same classification rule applied to the representations computed by the FastICA algorithm.

The Real+ICA, Imag+ICA, Magn+ICA, Phase+ICA perform the classifications on the representations in terms of the features extracted by combined wavelet and ICA method, $real(R), imag(R), abs(R), phase(R)$, respectively.

The results of the individual classification are then used in three different ensemble classification systems.

Let cls_i^1, cls_i^2, cls_i^3 be the resulted classification for the i^{th} test sample, obtained by using the system Ensemble1+ICA, Ensemble2+ICA and Ensemble3+ICA, respectively.

For simplicity, let us denote the individual weights obtained by training the individual classifiers by:

1) p_1 = weight for classification on Real+ICA;
2) p_2 = weight for classification on Imag+ICA;
3) p_3 = weight for classification on Magn+ICA;
4) p_4 = weight for classification on Phase+ICA;

where $p_i \in [0, 1]$.

In addition, we denote by $res_i = (res_i^j)_{j\in\{1,2,3,4\}}$, the results of the individual classifications, of the i^{th} test sample, $0 \leq i \leq 15$.

Each of the four individual classifications are represented; each of the representations (Real+ICA, Imag+ICA, Magn+ICA, Phase+ICA) is numbered with 1, 2, 3 and 4, respectively.

Given the above considered notations, the ensemble decisions of the three systems are defined as follows:

1) Ensemble1+ICA performs the classification by a majority voting procedure using the decisions taken by the classifiers Real+ICA, Imag+ICA, Magn+ICA, Phase+ICA:

$$cls_i^1 = \begin{cases} 1, & \text{if } \sum_{j=1}^4 p_j * res_i^j \geq 0 \\ \text{-1}, & \text{otherwise.} \end{cases} \quad (11)$$

2) The classification computed by Ensemble2+ICA results by combining the decisions of Real+ICA, Imag+ICA, Magn+ICA, Phase+ICA, using the logistic function and the threshold $\theta_1 = 0.5$:

$$cls_i^2 = \begin{cases} 1, & \text{if } \dfrac{1}{1+e^{-\sum_{j=1}^4 p_j * res_i^j}} \geq \theta_1 \\ \text{-1}, & \text{otherwise.} \end{cases} \quad (12)$$

3) The Ensemble3+ICA system uses the error function in order to asses the global decision. The error function is defined as follows[29]:

$$erf(x) = \frac{2}{\sqrt{\pi}} \int_0^x e^{-t^2} dt \quad (13)$$

Based on the above definition, the collective decision of the ensemble system is computed by the following formula:

$$cls_i^3 = \begin{cases} 1, & \text{if } erf\left(\dfrac{\sqrt{\pi}}{2} * (\sum_{j=1}^4 p_j * res_i^j)\right) \geq \theta_2 \\ \text{-1}, & \text{otherwise.} \end{cases}$$

The results of our tests on MIT-BIH database are summarized in Table 1 and Figure 2.

While the performance of the classifier Std and ICA expressed in terms of the success rate is pretty poor, substantial improvements are obtained in case of the classifiers Real+ICA, Imag+ICA, Magn+ICA, Phase+ICA. The best performance is obtained by the ensemble classifiers, the highest success rate corresponding to Ensemble2+ICA and Ensemble3+ICA.

Type of Classifier	Success Rate(Percent)
Std	26.67
ICA	53.33
Real+ICA	73.33
Imag+ICA	66.67
Magn+ICA	46.67
Phase+ICA	66.67
Ensemble1+ICA	80.00
Ensemble2+ICA	86.67
Ensemble3+ICA	86.67

TABLE I. Success Rate for Several Types of Classifiers

V. Conclusive remarks and suggestions for further work

The research aimed to propose a new classification technique based on wavelet and ICA methods.

The novelty is represented by the Ensemble1+ICA, Ensemble2+ICA and Ensemble3+ICA, that use voting procedures for classification purposes, each of the voting classifiers being of the 2-NN classification rule but applied to different sets of features extracted from the input signals.

The tests performed on the MIT-BIH database point out significant improvements in case of the proposed ensemble type classifiers and encourage further work on one hand, in refining the classification scheme and on the other hand, in identifying classes of wavelets better adapted to this particular problem.

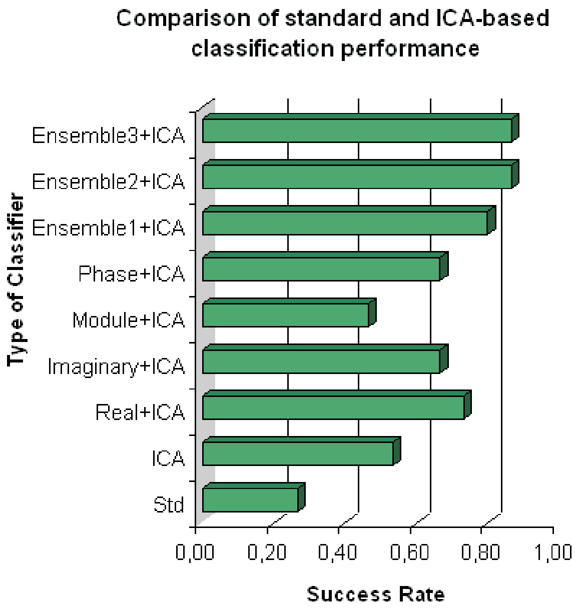

Fig. 2. Success Rates of the Classifiers

Complex wavelets and in particular, the Gauss wavelet family defined in Equation 8 have higher accuracy in describing the characteristics of the signal processed.

The decomposition and hence the use of specialized functions in describing the projected data, using time and space windows is proved to be much more efficient in the medical signals classification problems.

References

[1] I. Rosca, L. State, and C. L. Cocianu, "Investigations on the potential of pca based neural implementation attempts in solving specific tasks in image processing," *Proceedings of the 9th WSEAS International Conference on Mathematics and Computers in Business and Economics*, 2008.

[2] ——, "Learning schemes in using pca neural networks for image restoration purposes," *WSEAS Transactions on Information Science and Applications*, 2008.

[3] D. Constantin, "Principal directions for local independent components analysis," *Proceedings of the 9th WSEAS International Conference on Neural Networks*, 2008.

[4] A. Hyvarinen, J. Karhunen, and E. Oja, *Independent Component Analysis*, 2001.

[5] A. Hyvarinen, "Fast and robust fixed-point algorithms for independent component analysis," *IEEE Transactions on Neural Networks*, vol. 10(3), 626–634, 1999.

[6] A. Hyvarinen, P. Hoyer, and M. Inki, "Topographic independent component analysis," *Neural Computation*, vol. 13(7), 1525–1558, 2001.

[7] G. Gelle, M. Colas, and G. Delaunay, "Blind sources separation applied to rotating machines monitoring by acoustical and vibrations analysis," *Mechanical Systems Signal Processing*, 2000.

[8] W. Li, F. Gu, A. Ball, A. Leung, and E. Phipps, "A study of the noise from diesel engines using the independent component analysis," *Mechanical Systems Signal Processing*, 2001.

[9] A. Tarczynski, "Reliability of signal reconstruction from arbitrarily distributed noisy samples," *WSEAS Transactions on Signal Processing*, vol. 2, pp. 925–932, 2006.

[10] A. Boroomand, M. A. Ahmadian, A. Oghabian, J. Alirezaie, and C. Beckman, "An efficient hybrid wavelet-ica algorithm for analyzing simulated fmri data in noisy environment," *The 7th IEEE International Symposium on Signal Processing and Information Technology*, 2007.

[11] S. Mallat, *A Wavelet Tour of Signal Processing, 3rd Edition*. New York: Academic Press, 2009.

[12] P.-N. Tan, M. Steinbach, and V. Kumar, *Introduction to Data Mining*. Addison-Wesley, 2006.

[13] T. H. Cover and J. A. Thomas, *Elements of Information Theory*. John Wiley And Sons, 2006.

[14] C. Jutten, "Calcul neuromimétique et traitement du signal, analyse en composantes indépendentes," *PhD Thesis, Univ. Grenoble*, 1987.

[15] J. F. Cardoso, "Source separation using higher order moments," *Proc. ICASSP*, pp. 2109–2112, 1989.

[16] C. Jutten and J. Herault, "Blind separation of sources, part i: An adaptive algorithm based on neuromimetic architecture," *Signal Processing, 24:1–10*, 1991.

[17] P. Comon, "Independent component analysis. a new concept?" *Signal Processingss*, 1994.

[18] A. Hyvarinen, "Complexity pursuit: Separating interesting components from time-series," *Neural Computation*, pp. 883–898, 2001.

[19] ——, "New approximations of differential entropy for independent component analysis and projection pursuit," *Advances in Neural Information Processing Systems 10*, pp. 273–279, 1998.

[20] C. Liu and H. Wechsler, "Gabor-feature based classification using the enhanced fisher linear discriminant model for face recognition," *IEEE Transactions on Image Processing*, vol. 11(4), pp. 467–476, 2002.

[21] I. Daubechies, *Ten lectures on wavelets*. Philadelphia, PA, USA: Society for Industrial and Applied Mathematics, 1992.

[22] M. Sivannarayana and D. Redy, "Biorthogonal wavelet transform for ecg parameters estimation," *Medical Engineering and Physics*, 1999.

[23] M. Unser and A. Aldroubi, "A review of wavelets in biomedical applications," *Proceedings of the IEEE*, 1996.

[24] D. Donoho, "De-noising by soft-thresholding," *IEEE Trans. Information Theory*, 1995.

[25] A. G. Chouakri, F. Bereksi-Reguig, S. Ahmaidi, and O. Fokapu, "Wavelet denoising of the electrocardiogram signal based on the corrupted noise estimation," *Computers in Cardiology*, vol. 32, pp. 1021–1024, 2005.

[26] H. Liang and Z. Lin, "Stimulus artifact cancellation in the serosal recordings of gastric myoelectric activity using wavelet transform," *Biomedical Engineering*, 2002.

[27] A. K. Barros, A. Mansour, and N. Ohnishi, "Removing artifacs from electrocardiographic signals using independent component analysis," *Neurocomputing*, 1998.

[28] T. Jung, C. Humphries, T. W. Lee, S. Makeig, M. J. McKeown, V. Iragui, and T. J. Sejnowski, "Removing electroencephalographic artifacts: Comparison between ica and pca," *Neural Network for Signal Processing VIII*, 1998.

[29] W. J. Cody, "Rational chebyshev approximations for the error function," *Mathematics of Computation*, 1969.

[30] L. State, C. Cocianu, P. Vlamos, and D. Constantin, "Principal directions - based algorithm for classification tasks," *Symbolic and Numeric Algorithms for Scientific Computing*, vol. 9, no. September, pp. 137–143, January 2007.

[31] J. Kittler, M. Hatef, R. P. Duin, and J. Matas, "On combining classifiers," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 20, pp. 226–239, 1998.

[32] K. Tumer and J. Ghosh, "Error correlation and error reduction in ensemble classifiers," 1996.

[33] A. Guvenir, B. Acar, G. Demiroz, and A. Cekin, "A supervised machine learning algorithm for arrhythmia analysis," in *Proc. Computers in Cardiology Conference*, 1998, pp. 433–436.
[34] Y. Y. Tang, L. H. Yang, J. Liu, and H. Ma, *Wavelet Theory and Its Application to Pattern Recognition*. World Scientific, 2000.
[35] D. Radunovic, *Wavelets from Math to Practice*. Springer, 2009.
[36] M. Misiti, Y. Misiti, G. Oppenheim, and J.-M. Poggi, *Wavelets and their Applications*. ISTE, 2007.

Corina SĂRARU, Assistant, PhD, currently working with University of Piteşti, Department of Mathematics and Computer Science. Competence areas: digital image processing, pattern recognition, dynamical systems. Research in the fields of image processing and pattern recognition. Author of 8 papers published in national and international journals.

Luminiţa STATE, Professor, PhD, currently working with University of Piteşti, Department of Mathematics and Computer Science. Competence areas: artificial intelligence, machine learning, statistical pattern recognition, digital image processing. Research in the fields of machine learning, pattern recognition, neural computation. Author of 15 books and more than 120 papers published in national and international journals.

Maria MIROIU is currently an Assistant professor in the Computer Science Department at the University of Piteşti, Romania. Her competence areas include statistical pattern recognition, computer graphics, image processing. Dr. Miroiu's current research in the fields includes object recognition and classification, boosting algorithms, and 3D reconstruction. She has coauthored 3 books, and is the author or coauthor of 14 journal or conference papers.