

## APPLICATION OF STATISTICAL FEATURES AND MULTILAYER NEURAL NETWORK TO AUTOMATIC DIAGNOSIS OF ARRHYTHMIA BY ECG SIGNALS

Amine B. Slama<sup>1,3)</sup>, Łukasz Lentka<sup>2)</sup>, Aymen Mouelhi<sup>1)</sup>, Mohamed F. Diouani<sup>4)</sup>,  
Mounir Sayadi<sup>1)</sup>, Janusz Smulko<sup>2)</sup>

- 1) University of Tunis, ENSIT, LR13ES03 SIME, 1008, Montfleury, Tunisia (aymen.mouelhi@yahoo.fr, mounir.sayadi@esstt.rnu.tn)
- 2) Gdańsk University of Technology, Faculty of Electronics, Telecommunications and Informatics, G. Narutowicza 11/12, 80-233 Gdańsk, Poland (✉ lukasz.lentka@pg.edu.pl, +4858 3471737, jsmulko@eti.pg.edu.pl)
- 3) University Tunis ElManar, ISTMT, LR13ES07, LRBTM, Tunis, Tunisia (amineslama1@gmail.com)
- 4) Laboratory of Epidemiology and Veterinary Microbiology, Pasteur Institute, Tunis, Tunisia (fethi.diouani@pasteur.rns.tn)

### Abstract

Abnormal electrical activity of heart can produce a cardiac arrhythmia. The electrocardiogram (ECG) is a non-invasive technique which is used as a diagnostic tool for cardiac diseases. Non-stationarity and irregularity of heartbeat signal imposes many difficulties to clinicians (e.g., in the case of myocardial infarction arrhythmia). Fortunately, signal processing algorithms can expose hidden information within ECG signal contaminated by additive noise components. This paper explores a method of de-noising ECG signal by the discrete wavelet transform (DWT) and further detecting arrhythmia by estimated statistical parameters. Parameters of the de-noised ECG signals were used to form an input data vector determining whether the examined patient suffers from a cardiac arrhythmia or not. Input data were transformed using selected linear methods in order to reduce dimension of the input vector. A neural network was used to detect illness. Compared with the results of recent studies, the proposed method provides more accurate diagnosis based on the examined ECG signal data.

Keywords: Multilayer Neural Network, Arrhythmia diagnosis, ECG signal processing, Principal Component Analysis, Fisher's Linear Discriminant.

© 2018 Polish Academy of Sciences. All rights reserved

## 1. Introduction

Each year, millions of people die because of a variety of cardiovascular diseases. In fact, cardiac arrhythmia is an object of interest due to its increased frequency and incidence. In daily clinical analyses, electrocardiography (ECG) is mostly limited to the quantitative expression of electrical activity in heart muscles produced by depolarization and repolarization of atria and ventricle. An ECG cycle consists of P-QRS-T waves, as shown in Fig. 1. The ECG signal is applied to evaluate stability of heartbeats, effects of drugs or devices used to control the heart, such as a pacemaker. The major information about the signal is included in a frequency range of  $0.5 \div 45$  Hz [1–4]. Different parameters of ECG waveform can be extracted from the recorded

data, such as the P wave which indicates atrial depolarization. Its amplitude is usually between 0.1 and 0.2 mV and its duration – between 0.06 and 0.08 s. The QRS complex is a ventricular depolarization of an amplitude about 1 mV and duration time between 0.06 and 0.12 s.

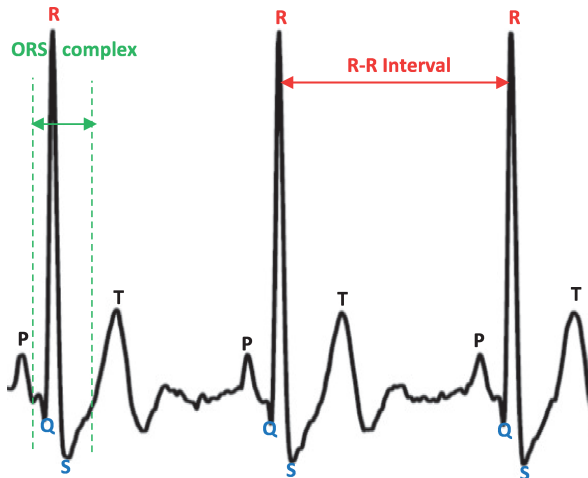


Fig. 1. An example of ECG signal.

Information about heart state in an ECG is contained in its amplitude and duration of characteristic wave peaks, as was mentioned above. However, evaluation of the ECG signal status is often based on subjective assessment of different factors at the presence of noise and irregularities of heartbeat. Furthermore, development of automatic methods, used to divide into sections and classify ECG signal according to different levels of pathological states, can assure a serious improvement in cardiovascular diseases diagnostics (CVD). The main aim of such algorithms is to recognize quality of different CVDs by precise measurements assisting clinicians in the diagnosis process [1–3]. In fact, the accuracy of a classification method depends on the nature of filtering and feature extraction techniques used in the ECG analysis process. In other words, the global features contribute to the applied filtering technique and must be highly accurate to assess the presence or absence of different cardiac disorders. Indeed, the feature extraction from ECG signals is a very complicated task because of large variability in their time and frequency representation (peak amplitudes' distribution, P-QRS-T complex, peak intervals). Frequently, Idioventricular Rhythm (IVR) and Ventricular Tachycardia (VT) are confused. These two rhythms present very similar ECG morphology. Also, Supraventricular Tachycardia (SVT) and Atrial Fibrillation (AF) are regularly confused with Atrial Flutter (AFL) which results in a decision of assuming atrial arrhythmias. The main distinguishing criterion for this rhythm is an irregularity of P and R waves. This difference is very subtle to be detected, especially when intense noise is present in the ECG signal or variability in wave morphologies of the examined patients takes place.

In recent years, there have been presented in the literature [4, 5] many filtering methods which can overcome the presence of noise, as well as the variability and complexity of RR intervals' series enabling to provide ECG features relevant for CVD. Several studies have been presented on processing ECG signals [6]. Procedures of analysis, filtering process [7, 8] and feature extraction [9, 10] from ECG datasets have exceeded some limitations, so that the algorithms cannot identify different cardiac diseases. In fact, due to the variety of amplitudes in an ECG signal dataset, different algorithms using time and frequency domains do not offer efficient detection based on ECG signals [11]. Therefore, additional studies are necessary to improve it.

The Discrete Wavelet Transform (DWT) is the most generally used time-frequency filtering method [12–14] which provides good resolution in both time and frequency domains by reducing any additive noise components. A combined method based on the wavelet transform and support vector machine (SVM) algorithm was proposed by Zhao et al. [15]. Another method, proposed by Emran et al. [16], applied DWT to feature extraction from ECG signals and resolved the problem of analysing non-stationary signals with an optimum frequency resolution. In [17], a discrimination method based on the principal component analysis (PCA) and DWT was able to identify features pertinent to arrhythmia and sinus rhythm databases. Unfortunately, it was inefficient when noise was too intense [18]. The same experience was observed when higher-order spectra were considered (the detection accuracy did not exceed 94% [19]). Some other methods presented in the literature [20, 21] were combined with PCA and SVM algorithms and reached the detection accuracy of up to 95.5%.

Other promising methods were presented recently in the literature, giving higher classification accuracy of different heart diseases [22, 23]. Many proposed approaches based on machine learning, such as neural networks [24] and SVM [25], provide accurate ECG classification. These methods present many drawbacks regardless of the use of data processing without evaluating the presence of additive noise component and selection of the most optimal detection method.

In addition, an arrhythmia recognition algorithm [23] is developed which can assess irregular heart rhythms better than experienced cardiologists. The employed expert technique is a deep convolutional network used to classify 14 output classes of 64,121 ECG records from 29,163 patients. Nevertheless, all of the above mentioned algorithms have many disadvantages. It is probable that some of classification methods, such as deep learning and computationally complex algorithms, need a large amount of data. Thus, additivity and homogeneity principles are not pursued by these algorithms and do not integrate the regularity properties [26].

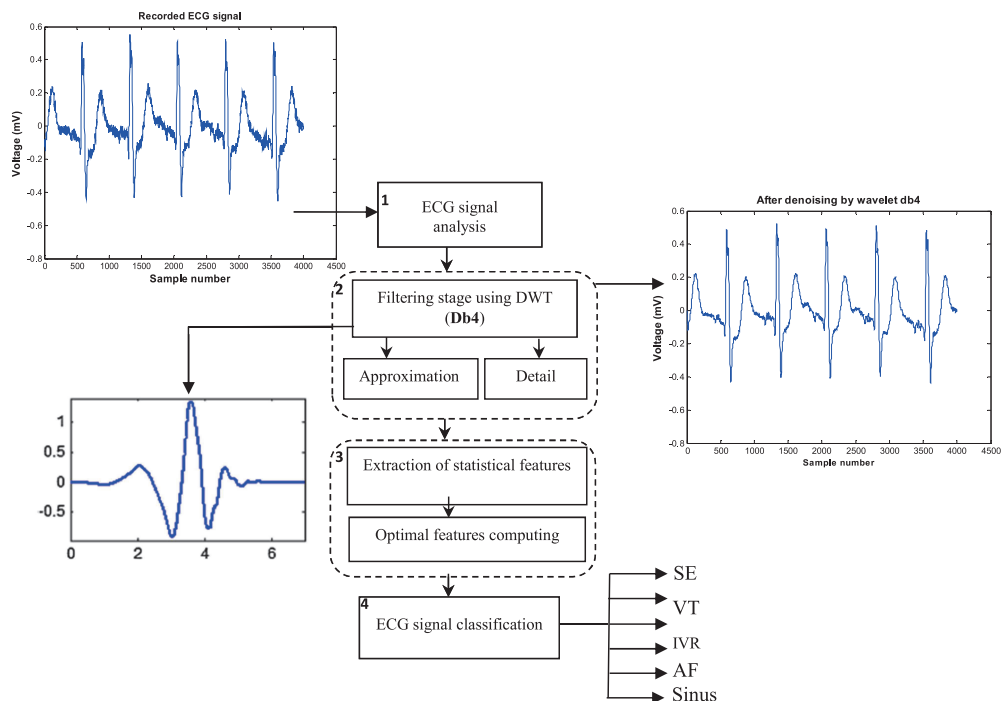


Fig. 2. An illustration of the automated arrhythmia recognition system.

Hence, a combination of features' optimization and classification method would serve as an excellent solution for overcoming these drawbacks. In our study, an advanced technique for arrhythmia recognition is developed. First, all ECG datasets are filtered using DWT (see Fig. 2). Then, a multilayer neural network (MNN) classifier is applied after the pre-processing stage using Fisher's linear discriminant (FLD) features to recognize either arrhythmia or sinus rhythm behaviour (Fig. 2, stage 3 and 4). The proposed method reduces dimensionality of ECG signals' features and its ability to cooperate with the MNN classifier in evaluating ECG signals. Thanks to the proposed procedure, the ECG classification process can be improved by about a few percent when compared with other methods, even at the presence of intense additive noise. The automatic analysis system used with artificial intelligence methods can be a reliable diagnosis tool.

## 2. Algorithmic methods of applied data processing

An ECG dataset used in this work was collected from the archives of the MIT-BIH arrhythmia dataset [27]. Five classes are used: the Supra-ventricular Ectopic beats (SE), Ventricular Tachycardia (VT), Idioventricular Rhythm (IVR), Atrial Fibrillation (AF) and Sinus Rhythm (Sinus). The ECG dataset signals are classified into the five separate classes. Four classes present different types of arrhythmia and are compared to the Sinus Rhythm class. The database includes 100 files of various time records. In this work, the process of analysing ECG signals includes the pre-processing stage followed by the feature extraction. The ECG signals are pre-processed to eliminate noise at a high frequency range. The features are extracted using DWT to locate different parameters of ECG signal like peaks (P, Q, R, S and T) and QRS interval [17, 18], which are helpful in diagnosing arrhythmia conditions.

The DWT method enables a multiresolution analysis by decomposing a discrete signal  $x(n)$  into low and high frequency components by low-band and high-band filters determined by a mother wavelet function and a corresponding scaling function. A signal  $x(n)$  is decomposed into low  $a_{j,k}$  (approximations) and high  $d_{j,k}$  (details) frequency components (Fig. 3) with three level banks, where  $n$  is the number of level:

$$\begin{cases} a_{j,k} = \sum_n h(n-2k)a_{j-1,m} \\ d_{j,k} = \sum_n g(n-2k)a_{j-1,m} \end{cases}, \quad (1)$$

where  $j$  is a parameter (scaling level) that influence the scaling of the wavelet transform and  $k$  is related to the translation within each level of the wavelet function,  $(j,k) \in \mathbb{Z}$ ;  $m$  is used in the scaling function as a translation (shifting) corresponding to the scaling level  $j$ ;  $h(n)$  and  $g(n)$  are low-pass and high-pass quadrature mirror filters, respectively (Fig. 3);  $n \in \mathbb{Z}$  is an integer scale. The wavelet transform contains different families of wavelet functions, such as Haar, Daubechies, etc. The *daubechies 4* (db4) and *symlets 6* wavelets have been selected for de-noising ECG signal and exhibiting some resemblance to QRS intervals (having the most intense part of power spectrum at a similar low frequency range).

The Fisher's linear discriminant (FLD) analysis is a method used in pattern recognition and machine learning issues [28, 29]. It is a statistical approach which is applied to create a linear combination of the original data by maximization of the ratio of between-class variation to within-class variation [29]. FLD can also be applied to analysis of ECG signal by processing a dataset to reduce its dimension. In this work, FLD is used to evaluate the training dataset and to generate independent features which are frequently more appropriate for classification into a



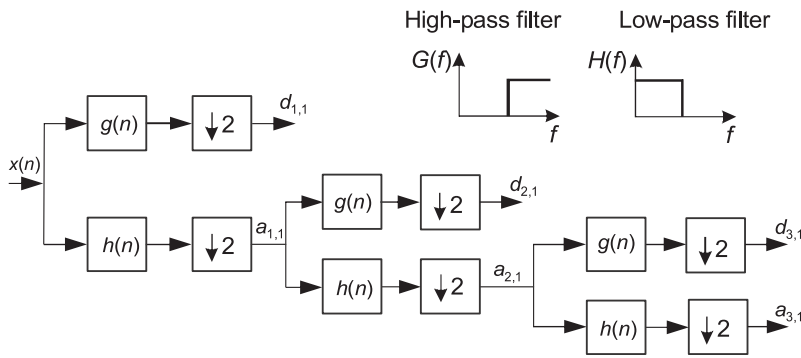


Fig. 3. An example of three-stage signal  $x(n)$  decompositions by the wavelet transform; the operator means down-sampling by 2, i.e. rejecting each second sample.

selected number of classes  $c$  than the PCA method. The goal of FLD is to create a basis  $w$  that minimizes the within-class measure while maximizing the between-class measure. This is done by maximizing the Fisher value  $J(w)$  defined by:

$$J(w) = \frac{w^T S_b w}{w^T S_w w}, \quad (2)$$

where  $S_b$  is a between-class scatter matrix:

$$S_b = \sum_{j=1}^c (\mu_j - \mu)(\mu_j - \mu)^T \quad (3)$$

and  $\mu_j$  is mean of class  $j$  while  $\mu$  is mean of all classes.  $S_w$  is a within-class scatter matrix:

$$S_w = \sum_{j=1}^c \sum_{i=1}^N (\mu_{j,i} - \mu_j)(\mu_{j,i} - \mu_j)^T, \quad (4)$$

where  $\mu_{j,i}$  is the  $i$ -th sample extracted from class  $j$  and  $N$  is the number of samples in class  $j$ . A linear transformation  $w$  is obtained by the eigenvector of  $S_w^{-1} S_b$ . The data set dimension is reduced and a new feature dataset  $F_{in} = (F_1, F_2, \dots, F_n)$  is obtained using the following transformation:

$$F_{in} = X_{in} w. \quad (5)$$

Consequently, after applying the FLD analysis, we obtain a limited dataset optimized to maximize the ratio of between-class variation to within-class variation. The procedure is linear and therefore can be easily computed.

PCA is a technique used for feature reduction [32] (similarly to the above presented FLD method). It is based on a discrimination function transforming a set of components from the original features into a lower-dimension space. As a linear method, the new variable, obtained by a fixed thresholding, is called Principal Component (PC). The first principal component provides a vector of the highest variability [26]; the second variable consists of the next highest variability, and so on. The PCA algorithm requires estimation of a covariance matrix of the dataset and eigenvectors in the decreasing order of eigenvalues. In this work, the PCA method is used to remove non-pertinent features from the recorded ECG signals.

The following algorithm presents different phases of the proposed PCA method:

1. Compute a covariance matrix  $C$  from the data as:

$$C = E \{ (x - \tilde{x}) \cdot (x - \tilde{x})^T \}, \quad (6)$$

where  $x$  is an input signal,  $\tilde{x}$  presents its mean vector and  $E$  is an operator of averaging within the set of  $N$  samples.

2. Calculate a diagonal matrix of eigenvalues  $D$  and a matrix of eigenvectors  $V$  from the covariance matrix  $C$ , as given by the equation:

$$V^{-1}CV = D. \quad (7)$$

3. Compute an eigenvectors' matrix of PCs in the descending order corresponding to the matrix of eigenvalues  $D$ . Then, the data are projected into the direction of the computed eigenvectors as:

$$p = [V^T(x - \tilde{x})^T]^T. \quad (8)$$

An output vector of variables  $p$  comprises PCs selected from the input signal.

Note that the PCA and FLD pre-processing methods increase evidently the classification process performance by reducing data dimensions. Moreover, the FLD pre-processing for further classification by the MNN classifier should give different results when compared with the PCA pre-processing because the methods assume diverse optimization conditions.

In our work, the MNN is approved to identify various patterns. There are different classifiers that can be used for arrhythmia detection. Indeed, it has been confirmed that MNN is a powerful tool for several classification tasks [30]. In order to classify the ECG dataset, we have chosen for computation simply the R–R interval, which is the main feature used to distinguish between sinus rhythm and arrhythmia case [20]. Due to correlation between features and similarities between pattern classes in the dataset, the FLD analysis is used in order to obtain new independent variables which should help to maximize the difference between classes in the input training set by considering various statistical parameters of the ECG dataset.

The database is divided into two sub-sets: one for training (containing 75% of the samples), and the other for test (containing 25% of the samples). Indeed, the combination of MNN and FLD or PCA pre-processing is already used in some applications of pattern recognition, but in our work the pre-processing step is applied to ECG signals with preselected statistical variables. That approach takes into account natural fluctuations of ECG signal which are an additional source of information about arrhythmia disease. Seventy ECG signals present four types of beat arrhythmia classes and 30 seem to be normal. This technique aids the projected intelligent system to avoid making errors in the learning phase. The classification stage is composed of two phases: The ECG dataset of training cases is firstly characterized by the significant FLD features – 75 cases randomly chosen (55 abnormal classes and 20 normal) constitute the training set. Then, the validation sets represent the smallest part of database and the network test is done with 25 remaining cases (15 abnormal and 10 normal).

In the training stage, statistical parameters representing five classes (SE, VT, IVR, AF, and Sinus) are chosen and are applied to feed the MNN. A synaptic weight vector is adapted to minimize the cost function using the back-propagated error between the actual and the desired outputs [29]. The neural network structure is composed of five input classes; they are: (0.5, -0.5, -0.5, -0.5, -0.5) for (SE), (-0.5, 0.5, -0.5, -0.5, -0.5) for (VT), (-0.5, -0.5, 0.5, -0.5, -0.5) for (IVR), (-0.5, -0.5, -0.5, 0.5, -0.5) for (AF) and (-0.5, -0.5, -0.5, -0.5, 0.5) for Sinus cases. Therefore, to the input layer constituted of input features, we used a single hidden layer and one output layer. The hidden layer size is optimized using the  $k$ -fold cross-validation technique [28].



### 3. Detection efficiency

We used the wavelet transform to decompose ECG signal. This approach was examined to achieve the best accuracy regarding a baseline drift removal and noise component attenuation. The DWT technique performance was further evaluated by estimating mean square error (*MSE*) and signal-to-noise ratio (*SNR*). The *MSE* and *SNR* values were estimated by analysing the difference between the recorded ECG signal and its de-noised version after wavelet transform decomposition and thresholding of selected details [33]. We selected *daubechies 4* (db4) and *symlets 6* which performed very well giving very similar results. Therefore, we have decided to further present the detailed data for db4 wavelet only (Table 1). *MSE* between the recorded and processed ECG signals within both considered groups was also calculated. We can notice the difference of efficiency between the considered methods of drift removal and noise component attenuation.

*MSE* is a parameter used to estimate deviation between the recorded ECG and the filtered signal, as given by the following equation:

$$MSE = \sqrt{\frac{1}{N-1} \sum_{i=1}^N (x(i) - \bar{x}(i))^2}. \quad (9)$$

The noise reduction efficiency was evaluated using the square error between the original  $x(i)$  and filtered  $\bar{x}(i)$  ECG signals by estimating *SNR* defined by the following equation:

$$SNR = 10 \cdot \log \left[ \frac{\sum_{i=1}^N x(i)^2}{\sum_{i=1}^N (x(i) - \bar{x}(i))^2} \right], \quad (10)$$

where  $N$  is a length of the ECG signal,  $x(i)$  is a recorded ECG signal and  $\bar{x}(i)$  is an ECG signal after the operation of filtering.

The results of de-noising by applying DWT are compared with those obtained for Butterworth and median filters. Table 1 presents the results of signal-to-noise ratio denoted as  $SNR_{res}$  and estimated by (10) after using selected de-noising methods. The recorded ECG signal contained an additive white noise component of an intensity given by the value of signal-to-noise ratio  $SNR_{ad}$ . Five data sets were examined. The results confirmed that de-noising by DWT secured the highest  $SNR_{res}$  despite of the additive noise component intensity. It means that this method is more efficient than the considered filtering methods. The improvement is different for various conditions and the differences reach a few dB and are clearly visibly by eye inspection in the filtered time sequences (Fig. 4).

The DWT filter ranges from 2 to 10. The results showed that the order 4 gives the best performance in terms of noise reduction compared with other DWT levels (db2, db6, db8, and db10). For the Butterworth low-pass filter with a 3 dB ripple and a cut-off frequency equal to 0.92, different values of filter order are employed to evaluate the best de-noising results. Each filter order (2, 4, 6, 8, 10) held out the de-noising ECG signal where the used filtering algorithm provided the best filtering results when the order was set to 10 and 8 for the Butterworth and median filters, respectively.

To evaluate more in-depth the quality of considered filtering methods, we have identified amplitudes of R peaks. We evaluated the ability of considered filtering methods to preserve the amplitude of the peak R. Using the wavelet mother function db4 we have observed that amplitudes of peaks R after the filtering stage are very close to their original mean amplitude. When



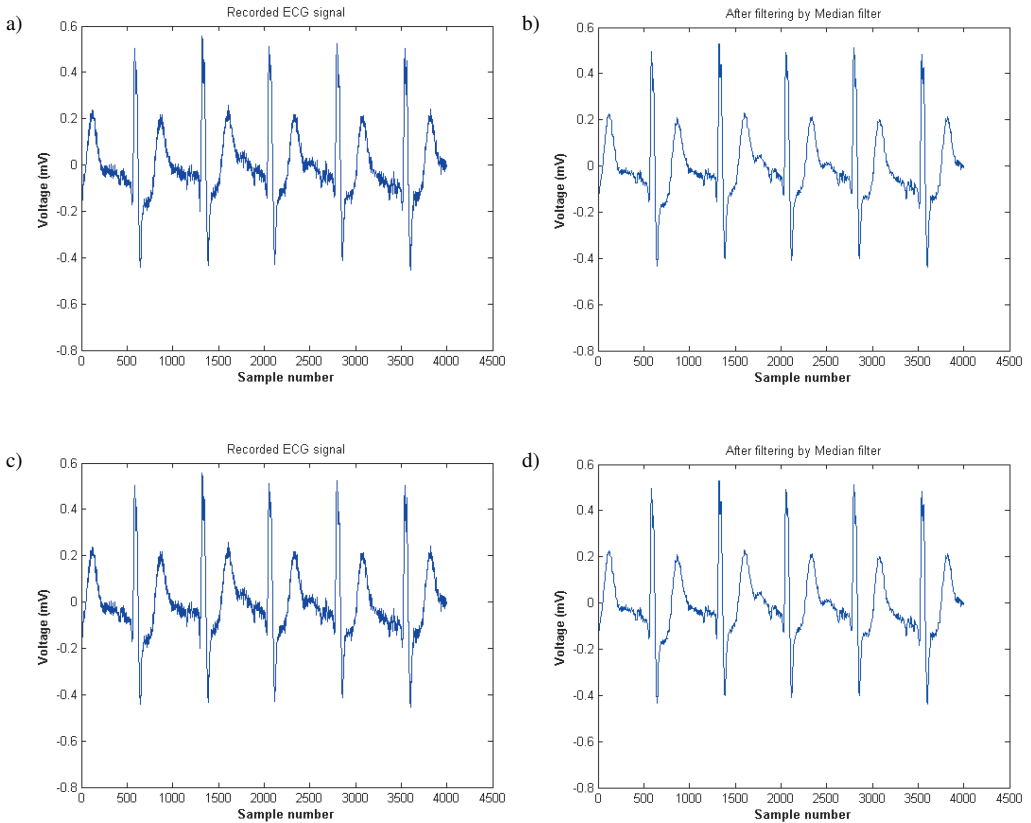


Fig. 4. De-noising results of ECG signals: a) recorded, b) de-noised using the median filter, c) de-noised using db4 wavelet, d) de-noised using the Butterworth filter.

Butterworth or median filtering was used the results were much worse – we have recovered only part of the original amplitude: 48% for the Butterworth and 53% for median filtering in the worst case of the analysed data sets (ECG signals 1÷10, Table 1). Thus, the DWT method again gave better detection results than other applied filtering methods.

After the filtering stage, the data sets can be further processed to reduce the ECG signals' dimension. Eventually, separated features are more suitable for classification tasks [23]. Various statistical parameters can be extracted from ECG signals as proposed for analysis of other random data series, comprising Gaussian noise [34], random transient pulses [35] or train of random pulse spikes [36]. The benefits of these features are simplicity of implementation and reduced computational time. Thus, we have selected a few popular statistical parameters characterizing random time series, including non-Gaussian random signals. The equations defining the selected parameters are gathered in Table 2.

All estimated statistical parameters (Table 2) are correlated in an unpredictable way and therefore further reduction of data dimension was required. Thus, in the next step the FLD analysis was applied to reduce the dimensionality of the dataset to establish the most informative parameters that maximize the ratio of between-class variation to within-class variation. The results enabled the reduction of the set of estimated parameters from eight (Table 2) to five only





Table 1. Values of  $SNR_{res}$  [dB] estimated for the applied filtering algorithms when the ten recorded ECG signals (cases 1 ÷ 10) contain an additive white noise component of one of selected intensities  $SNR_{ad} = 5$  dB, 10 dB or 15 dB.

$SNR_{ad}$ [dB]	DWT (wavelet db4)			Median filter			Butterworth filter		
	5	10	15	5	10	15	5	10	15
$SNR_{res}$ [dB] ECG signal 1	15.34	18.23	25.28	11.13	11.87	22.16	10.29	11.87	18.12
ECG signal 2	12.55	20.48	25.54	10.36	13.32	21.41	8.34	11.42	19.67
ECG signal 3	16.45	21.54	27.56	12.23	14.98	22.34	7.55	10.65	19.12
ECG signal 4	14.09	19.76	24.45	11.21	14.63	20.45	9.14	11.85	18.54
ECG signal 5	16.56	20.44	25.98	12.86	15.58	23.98	9.54	12.64	17.43
ECG signal 6	14.78	19.86	25.09	11.44	14.89	22.64	9.22	12.06	17.56
ECG signal 7	13.53	18.21	23.65	12.55	13.63	20.07	8.16	10.77	17.89
ECG signal 8	12.74	18.07	24.29	11.39	13.20	20.98	10.47	11.57	18.65
ECG signal 9	14.89	19.54	25.04	12.49	14.75	23.08	8.75	11.16	18.54
ECG signal 10	16.46	20.87	26.73	12.69	13.94	21.59	10.83	12.27	17.61

Table 2. Statistical parameters applied to reducing random data dimension by using the equations below.

Parameter	Definition
Root mean square	$\left(\frac{1}{N} \sum_{i=1}^N x_i^2\right)^{1/2}$
Skewness	$\frac{1}{N} \sum_{i=1}^N \frac{(x_i - \bar{x})^3}{\sigma^3}$
Peak to peak (P_P)	$x_{\max} - x_{\min}$
Crest factor	$\frac{\max  x_i }{RMS}$
Margin factor	$\frac{\max  x_i }{\left(\frac{1}{N} \sum_{i=1}^N  X_i ^{1/2}\right)^2}$
Shape factor	$\frac{RMS}{\frac{1}{N} \sum_{i=1}^N  X_i }$
Kurtosis	$\frac{1}{N} \sum_{i=1}^N \frac{(x_i - \bar{x})^4}{\sigma^4}$
Impulse factor	$\frac{\max  x_i }{\frac{1}{N} \sum_{i=1}^N  X_i }$

Note:  $\bar{x}$  is mean of  $x$  vector.

(Table 3), representing the main signal features: root mean square, skewness, peak to peak (P-P), crest factor and margin factor due to those having the highest values of Fisher’s linear discriminant (FLD) (exceeding 1).

Table 3. Evaluation of ECG signal features using Fisher’s linear discriminant (FLD).

Features	Fisher’s linear discriminant	Highest Priority
Root mean square	1.3964	*
Skewness	1.8336	*
Peak to peak (P-P)	2.9512	*
Crest factor	1.5120	*
Margin factor	1.8675	*
Shape factor	0.8837	–
Kurtosis	0.6792	–
Impulse factor	0.2394	–

To demonstrate effectiveness of the presented classification approach, we compared our proposed method with methods published in the literature [37–41]. In fact, using the FLD-MNN technique we obtain a classification of ECG datasets into five separate classes. The FLD-MNN method combines data reduction by FLD and further processing by the detection based on the MNN classifier. The results are then compared with those obtained with the support vector machine (SVM) non-linear classifier, as published elsewhere [17, 19]. The main advantage of employing these algorithms is having different measures which link various signal features.

In order to provide an assessment of the classifiers by the proposed FLD-MNN and SVM algorithms published earlier [17, 19], a statistical analysis in terms of accuracy *AC*, sensitivity *SE*, and specificity *SP* was applied. These parameters are given by the following equations:

$$SE = \frac{TP}{TP + FN}, \tag{11}$$

$$SP = \frac{TN}{FP + TN}, \tag{12}$$

$$AC = \frac{TP + TN}{TP + FN + TN + FP}, \tag{13}$$

where *TP* is the number of true positives, *TN* is the number of true negatives, *FN* is the number of false negatives and *FP* is the number of false positives.

In this section, we present a comparison study of classification results using the FLD-MNN method with those obtained by the PCA-MNN and MNN only (without reducing the number of eight considered statistical parameters). PCA-MNN is a combination of a multivariate statistical method (PCA) and a multilayer neural network (MNN). Combined with MNN, PCA is used only to obtain new uncorrelated parameters from the input data in the same subspace dimension, and then MNN is used to classify new components [31]. MNN is a simple feed-forward neural network trained with selected components extracted from the training ECG signals, as presented elsewhere [39].

In order to obtain a reasonable comparison of these three neural network-based methods, the same neural network structures were trained and validated during different experiments using the features appropriate for each method. It is known that the network structure has a great impact on its capability of generalization [42]. A neural network’s ability of data separation increases with the number of hidden layers. In addition to the input layer, we used a single hidden layer

and one output layer only. In practice, the applied MNN may rely on sophisticated and growing algorithms to define optimal structures of the network [38, 39]. The cross-validation technique was used for training and testing the datasets [28] in order to select the pertinent MNN structures for FLD and PCA methods of extracting ECG signals' features. The training set of feature vectors was divided into five equally-sized sub-datasets (five folds). Furthermore, cross-validation experiments of the five folds were performed to select optimal numbers of hidden layers and nodes in each layer. Then, five iterations of training and validation were performed, such that within each iteration a different fold of the data set was held-out for testing, and the remaining four folds – for training. In each fold, the performance measures were computed based on the accurate classification of cases in the testing ECG dataset. Besides, cross-validation experiments of the five folds were completed to validate the classification accuracy of each dataset randomly selected. A size of the hidden layer was being successively changed every training session from 1 to 15. Then, the average classification accuracy was computed for each original fold. The best MNN structure used for ECG classification selected also the hidden layer size that directly led to a minimum validation error. Other fast training algorithms may be used as well [37, 38]. The training patterns were chosen randomly from our database which contains ECG signals belonging to five separate classes. Note that the PCA and FLD pre-processing methods improve appreciably the MNN classifier performance. However, as shown in Table 4, the FLD-MNN classifier is superior to the PCA-MNN one for the analysed data sets. A classification accuracy of 97.38% was achieved with a standard deviation less than 0.82% for the ECG dataset. Thus, we conclude that the FLD-MNN method assures better classification accuracy for the training and validation datasets because, in some sense, it applies a signal-to-noise ratio to class labelling and therefore can be more robust against any additive noise components.

Table 4. Accuracy *AC*, sensitivity *SE*, and specificity *SP* of data classes detection achieved by applying the MNN classifiers and data reduction by PCA or FLD methods.

Data record	Cross validation	MNN			PCA-MNN			FLD-MNN		
		<i>SE</i> (%)	<i>SP</i> (%)	<i>AC</i> (%)	<i>SE</i> (%)	<i>SP</i> (%)	<i>AC</i> (%)	<i>SE</i> (%)	<i>SP</i> (%)	<i>AC</i> (%)
<b>Training</b> <b>ECG</b> <b>Dataset</b>	Fold1	88.1±2.1	90.6±1.6	89.1±1.4	94.2±1.5	95.2±1.1	93.3±1.4	95.8±1.1	0.4±1.2	<b>98.2±0.9</b>
	Fold2	82.3±2.6	87.7±2.1	86.5±2.4	91.3±1.8	93.6±1.3	91.5±1.8	94.2±1.3	98.3±0.5	<b>97.2±0.6</b>
	Fold3	84.5±2.1	85.5±2.5	84.4±2.5	92.5±1.7	92.9±1.5	90.9±1.9	95.5±1.2	97.9±0.7	<b>97.3±0.8</b>
	Fold4	86.3±2.3	89.2±1.9	87.2±2.3	91.4±1.7	93.6±1.8	90.4±1.9	95.2±1.2	98.3±0.5	<b>98.1±0.5</b>
	Fold5	84.5±2.4	86.9±2.1	86.2±2.4	92.5±1.6	93.8±1.5	92.5±1.5	96.5±1.1	98.5±0.4	<b>97.9±0.7</b>
<b>Validation</b> <b>ECG</b> <b>Dataset</b>	Fold1	85.4±2.1	88.4±1.8	87.5±1.9	91.5±1.6	93.6±1.3	92.7±1.5	93.5±1.2	96.3±1.1	<b>97.9±0.7</b>
	Fold2	80.6±2.2	86.1±2.4	85.3±2.1	89.3±1.9	91.2±1.5	89.3±1.9	93.1±1.4	95.1±1.1	<b>96.8±0.9</b>
	Fold3	83.1±2.4	83.9±2.5	83.7±2.5	90.4±1.6	90.7±1.5	89.4±1.6	92.4±1.6	96.4±0.8	<b>97.4±0.8</b>
	Fold4	85.8±2.5	87.6±2.1	85.3±2.3	87.8±1.9	92.5±1.3	88.3±2.1	94.5±1.4	96.5±0.9	<b>97.6±0.9</b>
	Fold5	83.2±2.4	82.4±2.6	84.9±2.6	90.5±1.9	91.5±1.7	90.3±1.6	93.8±1.3	97.2±0.7	<b>97.2±0.8</b>

In their work [19], the authors proposed a classification method that can resolve the problem of five types of beats using a feed-forward NN and least-square support vector machine (SVM); the combined HOS bi-spectrum and PCA were used for extracting bi-spectrum features and reduction of dimensionality, respectively. The average classification accuracy obtained was

93.48%. Compared with previous classification method, Ince et al. [41] proposed an arrhythmia identification system to classify two MIT-BIH database classes: (S) and (V), using DWT coefficients and PCA as the feature extraction and reduction process, followed by employing NN in the classification stage. The classification results were optimized by the multidimensional particle swarm optimization (MDPSO) method. A classification accuracy of 95.58% was obtained using this approach. In another study of Martis et al. [17] five classes were used from the MIT-BIH arrhythmia database and classified using the DWT and PCA dimensionality reduction technique. The average classification accuracy obtained was 96.92% using SVM.

Table 5. Comparison of classification efficiency obtained by selected methods presented in the literature.

Literature	MNN		
	Feature reduction	Classifier	AC (%)
(Ince et al., 2009) [41]	DWT+PCA	MDPSO	95.58%
(Martis et al., 2013b) [19]	Bispectrum+PCA	SVM	93.48%
(Martis et al., 2013c) [17]	DWT+PCA	SVM	96.92%
Proposed approach	DWT+FLD	MNN	<b>97.38%</b>

In this paper, we have established that a combined FLD-MNN method as a feature extraction technique is able to extract the pertinent information from a non-stationary ECG signal. The method applied discriminant statistical parameters to the experimental MIT-BIH data [27]. The technique provided effective extraction of the most relevant features and clearly distinguished arrhythmia from the Sinus cases. As shown in Table 5, the proposed method is able to classify datasets with a 97.38% accuracy using the FLD-MNN classifier. It is the highest detection accuracy compared with other considered methods. The accuracies of the proposed system and different existing methods are presented in Table 5.

#### 4. Conclusions

In our experimental data studies, selected methods of ECG data sets' classifications have been considered. The signals were filtered using the DWT algorithm to reduce additive noise and non-stationary drifts. Next, two pre-processing methods (PCA and FLD) and a multilayer neural network (MNN) as detection algorithms were applied to assign the data to one of five classes only. These methods (PCA or FLD) were used to determine features of ECG signals in a more efficient way by considering selected statistical parameters. This combination of statistical features and a classifier is a promising way of high-accuracy medical diagnosis (arrhythmia detection) based on non-invasive ECG signals. Compared with the previously published data, our results demonstrate better detection, up to 97.38%. It means that the proposed methodology is more efficient than other considered methods, using even non-linear classifiers. We suppose that such results were observed thanks to the shape of analysed signals exhibiting strong spikes with some variation in amplitude and time occurrence. The proposed methods, especially the pre-processing FLD, take into account such variations and therefore can assure better classification results.



## References

- [1] Mathers, C., Lopez, A., Stein, C., Fat, D., Rao, C. (2005). *Deaths and disease burden by cause: global burden of disease estimates for 2001 by World Bank country groups*.
- [2] Rai, H.M., Trivedi, A. (2012). De-noising of ECG waveforms using multiresolution wavelet transform. *International Journal of Computer Application*, 45, 25–30.
- [3] Martis, R.J., Acharya, U.R., Mina, L.C. (2013). ECG beat classification using PCA, LDA, ICA and Discrete Wavelet Transform. *Biomedical Signal Processing and Control*, 8, 437–448.
- [4] Goldberger, A.L. (2012). *Clinical Electrocardiography: A Simplified Approach*. Elsevier Health Sciences.
- [5] Shiyovich, A., Wolak, A., Yacobovich, L., Grosbard, A., Katz, A. (2010). Accuracy of Diagnosing Atrial Flutter and Atrial Fibrillation from a Surface Electrocardiogram by Hospital Physicians: Analysis of Data from Internal Medicine Departments. *The American Journal of the Medical Sciences*, 340(4), 271–275.
- [6] Bakul, G., Tiwary, U.S. (2010). Automated risk identification of myocardial infarction using relative frequency band coefficient (RFBC) features from ECG. *Biomedical Engineering Journal*, 4, 217–222.
- [7] Mahmoodabadi, S., Ahmadian, A., Abolhasani, M. (2005). ECG feature extraction based on multi-resolution wavelet transform. *Conf. Proc. IEEE Eng. Med. Biol. Soc.*, 4, 3902–3905.
- [8] Pal, S., Mitra, M. (2010). Detection of ECG characteristic points using multi resolution wavelet analysis based selective coefficient method. *Measurement*, 43, 255–261.
- [9] Josko, A. (2007). Discrete Wavelet Transform in automatic ECG Signal Analysis. *Conf. Proc. IEEE Instrumentation and Measurement Technology*, 1–3.
- [10] Banerjee, S., Gupta, R., Mitra, M. (2012). Delineation of ECG characteristic features using multi-resolution wavelet analysis method. *Measurement*, 45, 474–487.
- [11] Lin, C.H. (2008). Frequency-Domain Features for Ecg Beat Discrimination Using Grey Relational Analysis-Based Classifier. *Computers & Mathematics with Applications*, 55(4), 680–690.
- [12] Arif, M. (2008). Robust Electrocardiogram (Ecg) Beat Classification Using Discrete Wavelet Transform. *Physiological Measurement*, 29(5), 555.
- [13] Zhao, Q.B., Zhang, L.Q. (2005). ECG Feature Extraction and Classification Using Wavelet Transform and Support Vector Machines. *Proc. of the 2005 International Conference on Neural Networks and Brain*, 1–3, 1089–1092.
- [14] Li, C., Zheng, C., Tai, C. (1995). Detection of Ecg Characteristic Points Using Wavelet Transforms. *IEEE Transactions on Biomedical Engineering*, 42(1), 21–28.
- [15] Zhao, Q.B., Zhang, L.Q. (2005). ECG Feature Extraction and Classification Using Wavelet Transform and Support Vector Machines. *Proc. of the 2005 International Conference on Neural Networks and Brain*, 1–3, 1089–1092.
- [16] Tamil, E.M., Kamarudin, N.H., Salleh, R., Idris, Noorzaily, M., Y.I., Noor, M., Tamil, A.M. (2008). *Heartbeat Electrocardiogram (ECG) Signal Feature Extraction Using Discrete Wavelet Transforms (DWT)*.
- [17] Martis, R.J., Acharya, U.R., Min, L.C. (2013). Ecg Beat Classification Using Pca, Lda, Ica and Discrete Wavelet Transform. *Biomedical Signal Processing and Control*, 8(5), 437–448.
- [18] Martis, R.J., Acharya, U.R., Ray, A.K., Chakraborty, C. (2011). Application of Higher Order Cumulants to Ecg Signals for the Cardiac Health Diagnosis. *Engineering in Medicine and Biology Society. EMBC, Annual International Conference of the IEEE*, 1697–1700.
- [19] Martis, R.J., Acharya, U.R., Mandana, K., Ray, A., Chakraborty, C. (2013b). Cardiac Decision Making Using Higher Order Spectra. *Biomedical Signal Processing and Control*, 8(2), 193–203.



- [20] Kutlu, Y., Kuntalp, D. (2011). A Multi-Stage Automatic Arrhythmia Recognition and Classification System. *Computers in Biology and Medicine*, 41(1), 37–45.
- [21] De Chazal, P., Reilly, R.B. (2006). A Patient-Adapting Heartbeat Classifier Using Ecg Morphology and Heartbeat Interval Features. *IEEE Transactions on Biomedical Engineering*, 53(12), 2535–2543.
- [22] Rajpurkar, P., Hannun, A.Y., Haghpanahi, M., Bourn, C., Y. Ng, A. (2017). *arXiv preprint arXiv:1707.01836v1*. 53(12).
- [23] Ioffe, S., Christian, S. (2017) Batch normalization. Cardiologist-Level Arrhythmia Detection with Convolutional Neural Networks. *arXiv preprint arXiv:1502.03167*.
- [24] Kim, J., Min, S.D./ Lee, M. (2011). An arrhythmia classification algorithm using a dedicated wavelet adapted to different subjects. *Biomedical engineering online*, 10(1), 56, PMID:21707989. <http://dx.doi.org/10.1186/1475-925X-10-56>.
- [25] Zhang, Z., Dong, J., Luo, X., Choi, K., Wu, X. (2014). Heartbeat classification using disease-specific feature selection. *Computers in Biology and Medicine*, 46, 79–89, PMID: 24529208. <http://dx.doi.org/10.1016/j.combiomed.2013.11.019>.
- [26] Martis, R.J., Acharya, U.R., Adeli, H. (2014). Current Methods in NElectrocardiogram Characterization. *Computers in Biology and Medicine*, 48, 133–149.
- [27] Moody, G.B., Mark, R.G. (2001). The Impact of the Mit-Bih Arrhythmia Database. *Engineering in Medicine and Biology Magazine, IEEE*, 20(3), 45–50.
- [28] Mouelhi, A., Sayadi, M., Fnaiech, F., Mrad, K. (2013). A new automatic image analysis method for assessing estrogen receptors' status in breast tissue specimens. *Computers in Biology and Medicine*, 2263–2277.
- [29] Mouelhi, A., Sayadi, M., Fnaiech, F., Mrad, K., Ben Romdhane, K. (2013). Automatic Image Segmentation of Nuclear Stained Breast Tissue Sections Using Color Active Contour Model and an Improved Watershed Method. *Biomedical Signal Processing and Control*, 421–436.
- [30] Martinez, A.M., Kak, A.C. (2001). PCA versus LDA. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 23(2), 228–233.
- [31] Martis, R.J., Acharya, U.R., Lim, C.M., Mandana, K., Ray, A.K., Chakraborty, C. (2013). Application of Higher Order Cumulant Features for Cardiac Health Diagnosis Using ECG Signals. *International journal of neural systems*, 23(04).
- [32] Inan, O.T., Giovangrandi, L., Kovacs, G.T. (2006). Robust Neural-Network- Based Classification of Premature Ventricular Contractions Using Wavelet Transform and Timing Interval Features. *IEEE Transactions on Biomedical Engineering*, 53(12), 2507–2515.
- [33] Zhao, Q.B., Zhang, L.Q. (2005). ECG Feature Extraction and Classification Using Wavelet Transform and Support Vector Machines. *Proc. of the 2005 International Conference on Neural Networks and Brain*, 1–3, 1089–1092.
- [34] Smulko, J. (2006). Methods of electrochemical noise analysis for investigation of corrosion processes. *Fluctuation and Noise Letters*, 6(02), R1–R9.
- [35] Kiwiłszo, M., Smulko, J. (2009). Pitting corrosion characterization by electrochemical noise measurements on asymmetric electrodes. *Journal of Solid State Electrochemistry*, 13(11), 1681–1686.
- [36] Lentka, Ł., Kotarski, M., Smulko, J., Cindemir, U., Topalian, Z., Granqvist, C.G., Ionescu, R. (2016). Fluctuation-enhanced sensing with organically functionalized gold nanoparticle gas sensors targeting biomedical applications. *Talanta*, 160, 9–14.
- [37] Übeyli, E.D. (2009). Combining Recurrent Neural Networks with Eigenvector Methods for Classification of ECG beats. *Digital Signal Processing*, 19(2), 320–329.
- [38] Kim, J., Shin, H.S., Shin, K., Lee, M. (2009). Robust Algorithm for Arrhythmia Classification in Ecg Using Extreme Learning Machine. *Biomed. Eng. Online*, 8, 31.



- [39] Derya Übeyli, E. (2010). Recurrent Neural Networks Employing Lyapunov Exponents for Analysis of Ecg Signals. *Expert systems with applications*, 37(2), 1192–1199.
- [40] Lentka, Ł., Smulko, J.M., Ionescu, R., Granqvist, C.G., Kish, L.B. (2015). Determination of gas mixture components using fluctuation enhanced sensing and the LS-SVM regression algorithm. *Metrol. Meas. Syst.*, 22(3), 341–350.
- [41] Ince, T., Kiranyaz, S., Gabbouj, M. (2009). A Generic and Robust System for Automated Patient-Specific Classification of Ecg Signals. *IEEE Transactions on Biomedical Engineering.*, 56(5), 1415–1426.
- [42] Ben Slama, A., Mouelhi, A., Sahli, H., Manoubi, S., Mbarek, C., Trabelsi, H. Fnaiech, F., Sayadi, M. (2017). A New Preprocessing Parameter Estimation based on Geodesic Active Contour Model for Automatic Vestibular Neuritis Diagnosis. *Artificial Intelligence in Medicine*, 80, 48–62.