Neural Networks for Biomedical Signals Classification Based on Empirical Mode Decomposition and Principal Component Analysis

Abdoul Dalibou Abdou[2], Ndeye Fatou Ngom[1(✉)], Samba Sidibé[1],
Oumar Niang[1], Abdoulaye Thioune[3], and Cheikh H. T. C. Ndiaye[3]

[1] Laboratoire Traitement de l'Inforrmation et Systémes Intelligents (LTISI),
Ecole Polytechnique de Thies (EPT), Thies, Senegal
fngom@ept.sn
[2] Univerté de Thies, LTISI-EPT, Thies, Senegal
abdouldalibou@gmail.com
[3] Univerité Cheikh Anta Diop de Dakar, LTISI-EPT, Dakar, Senegal

Abstract. The three main events presented in the electrocardiogram (ECG) signal of each heartbeat are: the P wave, the QRS complex and the T wave. Each event contains its own peak, making this important to analyze their morphology, amplitude and duration for cardiac abnormalities. In this study, we propose a system for biomedical signal analysis based on empirical mode decomposition. Mustispectral analysis is first performed to remove noise, detect QRS complex and compute the QRS wide. Then statistical features and QRS wide are after used as inputs of classifier based on neural network model. The proposed methodology is tested on real biomedical data and discussed.
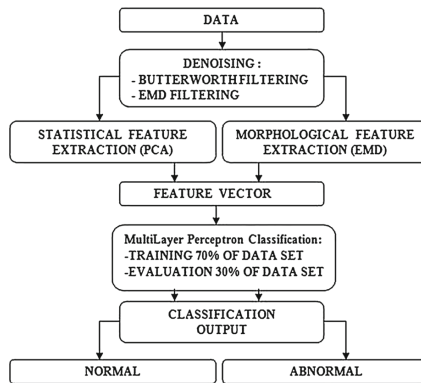
Keywords: Empirical mode decomposition · Neural network
ECG signal classification

1 Introduction

The three main events presented in the electrocardiogram (ECG) signal of each heartbeat are: the P wave, the QRS complex and the T wave. Each event contains its own peak, making this important to analyze their morphology, amplitude and duration for cardiac abnormalities. In order to provide tools to contribute to more accurate diagnosis, several signal processing algorithms have been developed to facilitate the continuous follow up and customized care. Al-Ashkar [1] used non linear filtering scheme for edge detection according a transition slope sign. Rodriguez et al. [11] proposed feature extraction based on Hilbert transform, adaptative threshold and principal component analysis. Empirical mode decomposition (EMD) has been also widely used for source separation and noise elimination. Indeed, EMD is adaptive, depends on the position of the extrema

of the signal, is non-linear and non-stationary [4,6,7,12]. Many methods have been developed for classification based on non linear features of ECG signals [3], machine vector support [9], Markov chain models [2] and artificial neural network (ANN) models [8,10,13]. Compared to the other methods, Neural networks offer a number of advantages: less formal statistical training, ability to implicitly detect complex nonlinear relationships between dependent and independent variables, ability to detect all possible interactions between predictor variables, and the availability of multiple training algorithms [14]. However, NN is often described as a black box learning approach. In this study, we propose a system for biomedical signal analysis based on Empirical Mode Decomposition (EMD) and multilayer perceptron neural network model. The Butterworth filtering is first used for signal noise elimination and then QRS detection is carried out using empirical Mode Decomposition (EMD). Statistical features and morphological feature such as QRS wide are used as inputs of a Multi Layer Perceptron network model. The schematic diagram of the processing steps is illustrated by Fig. 1.

The paper is organized as follows. Section 2 recalls related works and some basics concepts. Section 3 deals with the proposed noise elimination method, QRS wide detection and the proposed. Section 4 illustrates the developed method on real biomedical data and discusses the obtained results. The proposed method is applied to PhysioNet ECG database for classification of normal and abnormal ECG signals. Section 5 gives conclusion and perspectives of this study.


Fig. 1. Schematic diagram of the processing steps: noise elimination process is first performed before a classification based on features extracted using empirical mode decomposition (EMD) and principal component analysis (PCA)

2   Tools for Biomedical Signal Analysis

This section recalls some basics tools for ECG arrhythmia analysis such as feature extraction, empirical mode decomposition and the classifier construction basis.

2.1   Feature Extraction

In general, the ECG signal of a single cardiac cycle lies on the P, T and QRS complex waves as depicted in Fig. 2. The ECG analysis is related to the detection of QRS because of the presence of low amplitudes, negative polarities and noise. The QRS wide is one of the most feature selected for signal recognition and classification. In this study, the QRS detection is performed using the empirical mode decomposition (EMD).
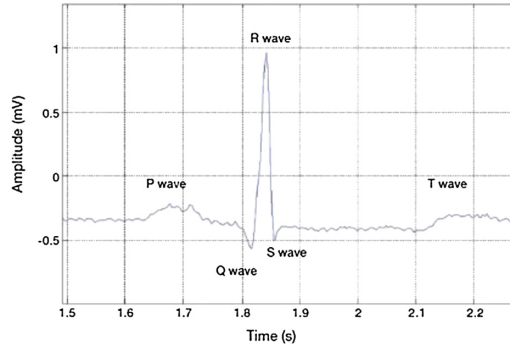
Fig. 2. Electrocardiogram QRS complex [11]: cardiac cycle lies on the P, T and QRS complex waves.

2.2   Empirical Mode Decomposition

EMD algorithms decompose iteratively a complex signal $s(n)$ into elementary AM-FM type components called Intrinsic Mode Functions (IMF) [6]:

$$s(n) = r_k(n) + \sum_{k=1}^{K} imf_k(n) \tag{1}$$

where imf_k is the k-th mode or IMF of the signal and r_k stands for residual trend. Sifting procedure generates a finite number of IMFs. Indeed, the underlying principle of the EMD decomposition is to locally identify in the signal, the most rapid oscillations defined as the waveform interpolating local maxima and minima. To do so, these later points are interpolated with cubic spline to yield the upper and lower envelopes. The mean envelope is then subtracted from the initial signal and the same interpolation scheme is reiterated. In this study, we use the algorithm presented in [6] for the empirical mode decomposition.

2.3   Principal Component Analysis

Principal component analysis is carried out to select the most important features of the ECG data, among computed statistical properties (mean, variance, covariance, correlation, energy, power...), to reduce their number and at the same time

retain as possible of their class discriminatory information. The PCA involves three main phases

1. Computation of variance matrix from data X

$$V = \left(X - \tilde{X}\right)\left(X - \tilde{X}\right)^T \tag{2}$$

   where X is the data matrix and $tildeX$ the mean vector of X.
2. Calculate the array of eigenvectors E and diagonal matrix of eigenvalues

$$E^{-1}VE = D \tag{3}$$

3. Sort the eigenvectors in E in descending order of eigenvalues in D and project the data on these eigenvector directions by taking the inner product between the data matrix and the sorted eigenvector matrix

$$U = \left[E^T\left(X - \tilde{X}\right)^T\right]^T \tag{4}$$

2.4   Classifier Construction

ECG beat recognition and classification depend on various features (morphological, temporal and statistical). These features can be used as input of a neural network classifier. The most used neural networks are Multi Layer Perceptron neural networks (MLP-NN). A MLP consist of an input layer, several hidden layer and an output layer. A node also called a neuron includes a summer and nonlinear activation function g as illustrated by Fig. 3.
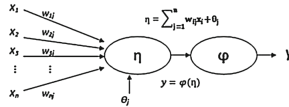
Fig. 3. Schematic illustration of a node (neuron): the inputs of a neuron are multiplied by weights summed up with bias terms.

The inputs $(x_1, ..., x_K)$ to the neuron are multiplied by weights w_{ki} and summed up together with the constant bias term θ_i. The resulting $n_i = \sum_{j=1}^{K} w_{ji}x_j + \theta_i$ is the input to the activation function g.

3   The Proposed Biomedical Analysis Methods

This section describes the proposed methods for noise removal, QRS wide extraction and classification. The proposed classification system adopts different methods following morphological feature extraction through empirical mode decomposition, QRS complex detection and the most discriminant statistical features extraction using principal component analysis. For each signal, the noise is first removed before performing a QRS complex detection and classification.

3.1   Noise Removal

We are interested in noise elimination method for QRS wide detection. Zhang et al. [15] listed many advantages of low Butterworth low pass filter of order 6 with frequencies from 5 to 15 for ECG signal noise removal. Other methods are based on empirical mode decomposition. Indeed, the first IMF contains mostly high frequency noise and some QRS information [7]. However, if the first IMF are removed and others retained, the resulting output may contains considerable level of noise. In this study we propose noise removal method based on Butterworth filtering and the first IMF removal method. The principle consist at first perform Butterworth filtering for a smooth preprocessing and then remove the first IMF which contains high frequency noise.

3.2   QRS Complex Detection

Detection of QRS complex is the entry point of almost all ECG analysis technique. For a signal, we propose the step by step process described in Algorithm 1 for the QRS wide computation.

Algorithm 1. Pseudocode to compute the QRS wide

 1: function larger(x)
 2: x : ecg records
 3: m, im : the maxima of x and it's index
 4: iim : the extremum maxima index
 5: $l1, l2$: previous and next index from iim
 6: w : the QRS with
 7: $ex \leftarrow extrema(x)$ extrema
 8: $m \leftarrow max(x)$
 9: $im, iim \leftarrow find(x = max(x)), find(ex = im)$
10: $l1, l2 \leftarrow ex(iim - 1), ex(iim + 1)$
11: $w \leftarrow (l2 - l1)$
12: return w

Our method for QRS detection is based on empirical mode decomposition (EMD). Indeed, It has been shown that the first IMFs will have the QRS information as QRS regions is high frequency component [4,7,12]. The two first IMFs (Eq. 1) contain only the most low frequency and waves like P and T are filtered out from consideration [7]. Thus, in this study we suppose that the first IMF contains the QRS complex. The proposed method involves the computation of the first IMF before the QRS computation as described in Algorithm 1.

3.3   The Neural Network Model

The input data for the Neural Network are: the QRS width detection and the most discriminant statistical properties among mean, variance, root mean

squared (RMS), energy and power. Each input data is associated a weight and is computed from each record. If for example, the RMS is not a discriminant properties then it's weight is equal to zero.

For each node, the input is multiplied by weights (pm, en, pen, pp, pl, r): $sp = m * pm + v * pv + en * pen + p * pp + l * pl + s * r$ where m, pv, e, en, p, l, s are respectively the mean, the variance, the energy, the power, the QRS width and the RMS. Then, the difference $s = sp - w$ between the neural network threshold w and sp is used as input of the activation (sigmoid) function defined as follows

$$f(x) = \frac{1}{(1 + \exp(-x))} \tag{5}$$

Algorithm 2 gives detailed description of the step processing of a node.

Algorithm 2. Pseudocode to compute a node output

1: function neurone(x,w)
2: x; ecg records
3: w : the node's threshold
4: $n : length(x)$
5: $s \leftarrow 0$
6: **for** $i \leftarrow 1$ to n **do**
7:    $s \leftarrow s + x(1, i) * x(2, i)$
8:    $i \leftarrow i + 1$
9: **end for**
10: $s \leftarrow s - w$
11: $y \leftarrow 1/(1 + exp(-s))$
12: return y

Let us suppose that, the number of input of the classifier is six (the mean, the variance, the energy, the power, the QRS width and the RMS). If a statistical parameter is not considered after ACP analysis, then it's associated weight is equal to zero. We propose to compute the 1st node with all the six parameters, the six other nodes ($2^{nd}, 3^{rd}, 4^{th}, 5^{th}, 6^{th}, 7^{th}$) with five parameters (the variance, the energy, the power, the QRS width and the RMS) and the 8th node with the output of the node 1, the node 2, the node 3, the node 4, the node 5, the node 6 and the node 7. The computed nodes are then assembled using the step processing described in Algorithm 3.

The training provide the thresholds of the network nodes. It is implemented with a function called *train* which take as input the feature vector from the training set database and gives as output a threshold. The threshold defines the neural network quality and is computed from the nodes of the neural network. Algorithm 4 describes the training evaluation process.

The error rate and the performance rate are used as validation and test parameters during the classification process (see Algorithm 5). The network performance is evaluated with the percentage of correct ecg classification results.

Algorithm 3. Pseudo code for the neural network

```
 1: function network(x,w)
 2: x : ecg records
 3: w : the threshold
 4: v1 ← vector(x): six parameters
 5: v2 ← [v1(1), v1(2), v1(3), v1(4), v1(5)]
 6: v3 ← [v1(1), v1(2), v1(3), v1(4), v1(6)]
 7: v4 ← [v1(1), v1(2), v1(3), v1(5), v1(6)]
 8: v5 ← [v1(1), v1(2), v1(4), v1(5), v1(6)]
 9: v6 ← [v1(1), v1(3), v1(4), v1(5), v1(6)]
10: v7 ← [v1(2), v1(3), v1(4), v1(5), v1(6)]
11: s1 ← neurone(v1, w)
12: s2 ← neurone(v2, w)
13: s3 ← neurone(v3, w)
14: s4 ← neurone(v4, w)
15: s5 ← neurone(v5, w)
16: s6 ← neurone(v6, w)
17: s7 ← neurone(v7, w)
18: v8 ← [s1, s2, s3, s4, s5, s6, s7]
19: s ← neurone(v8, w): the output of the network
20: return s
```

Algorithm 4. Pseudo code for the network performance evaluation

```
1: function evaluate(t,s)
2: t : test db
3: s : desired output
4: w ← train(t) threshold
5: r ← network(t, w) computed output
6: e ← error(s, r)
7: p ← performance(s, r)
8: v ← [e, p]
9: return v
```

## 4   Application to Biomedical Signal Analysis

The proposed methods for biomedical signal analysis are applied to PhysioNet ECG database using Matlab 2013 software. The ECG records used in this study is the MIT-BIH arrhythmia database [5]. Example of computed features on the 10 first ecg records are resumed in Table 1.

4.1   MIT BIH Databases Records

The MIT-BIH arrhythmia database contains two channels ambulatory recordings obtaining from 48 subjects studied by BIH Arrhythmia Laboratory between 1975 and 1979. The recordings were digitized at 360 samples per second per channel with 11 bit resolution over a 10 mV range. In this study, the second channel recording is studied. The records (100, 101, 102, 103, 104, 105, 106, 107, 108,

Table 1. Computed characteristics on a sample of ECG records

N^o	MIT-MIH patient	Mean (mm)	Variance (mm^2)	Root mean square (mm)	Energy (mm^2)	Power (mm^2)	QRS wide (ms)
1	100	−0.078	0.028	0.168	0.034	4.116	0.0500
2	101	0.152	0.015	0.122	0.038	4.59	0.1170
3	102	0.087	0.100	0.316	0.107	4.116	0.3000
4	103	0.182	0.002	0.041	0.035	4.212	0.0550
5	104	0.138	0.001	0.033	0.020	2.408	0.2050
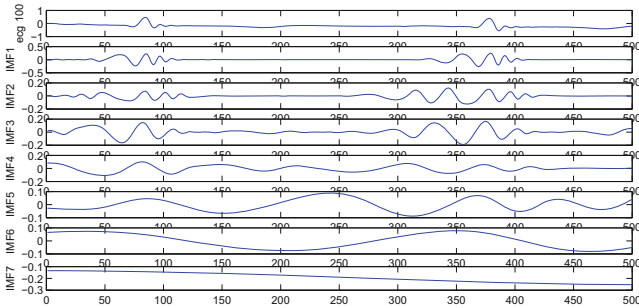
Fig. 4. MIT BIH patient 100 ECG empirical mode decomposition

109, 111, 112, 113, 114, 115, 116, 117, 118, 119, 121, 122, 123, 124, 200, 201, 202, 203, 205, 207, 208, 209, 210, 212, 213, 214, 215, 217, 219, 220, 221, 222, 223, 228, 230, 231, 232, 233, 234) are numbered respectively from 1 to 48.
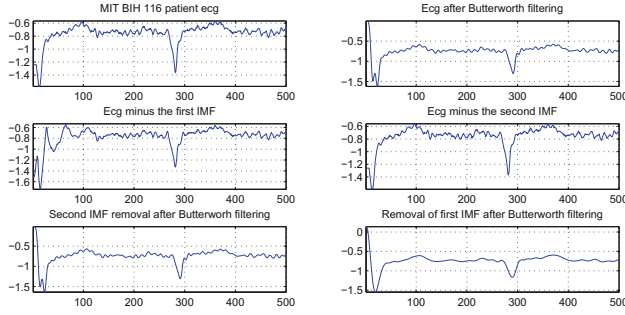
4.2 Noise Elimination

For each ecg record of the MIT BIH sample, a noise elimination procedure is first performed before the QRS wide computation. The proposed method is based on empirical mode decomposition (EMD). Figure 4 shows an electrocardiogram record and the resulting IMFs after the EMD decomposition.

Figure 5 illustrates the noise removal method proposed on two sets of ECG samples.

If the first IMF is removed, the signal shape is more smooth but there are appearance of non desired artifacts. With the removal of the second IMF, the shape is conserved but the noise are not removed. The Butterworth filter gives good results but it changes sometimes the shape of signal. Indeed, for the MIT BIH patient 116 the shape is locally changed whenever the Butterworth filter is performed. We obtained better results with the proposed method which consist at first to perform Butterworth filtering and after remove the first IMF. The amplitude of the QRS decreases but the computation of the extrema is more easy. And as we are concerned in this study with the computation of QRS wide from the extrema, the proposed method gives better results.

(a) MIT BIH patient 116
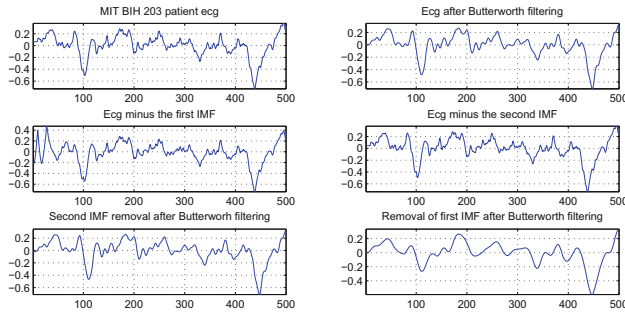
(b) MIT BIH patient 203

Fig. 5. EMD based noise elimination for QRS wide detection: the proposed method consist at first to perform Butterworth filtering and then remove the first imf

4.3   Features Extraction and Analysis

After noise removal, the ECG QRS width is computed according the step by step process described in Algorithm 1. The higher QRS width (0.3 ms) is obtained with ECG record of MIT BIH patient 102. The lower QRS width (0.038 ms) is computed from the ECG record of MIT BIH patient 124. The mean and the dispersion computed from the ecg records are respectively 0.13 ms and 0.12.

From the considered statistical features, the most discriminant are extracted. The evolution of the selected characteristics according to the ecg record number is illustrated in Fig. 6.

To study the influence of the selected feature, we perform ecg record shape similarity analysis. From the selected statistical properties and the QRS wide, a clustering based on the minimal linkage distance is performed. The result is illustrated in Fig. 7.

Among the more similar ecg records are: ecg 12 (MITBIH patient 112) and ecg 22 (MITBIH patient 123), ecg 5 (MITBIH patient 104) and ecg 9 (MITBIH patient 108), ecg 2 (MITBIH patient 101) and ecg 4 (MITBIH patient 103). Among the most dissimilar ecg records are: ecg 12 (MITBIH patient 112) and
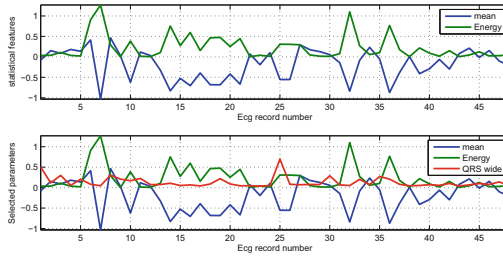
Fig. 6. Evolution of the selected feature characteristics according to the ecg record number
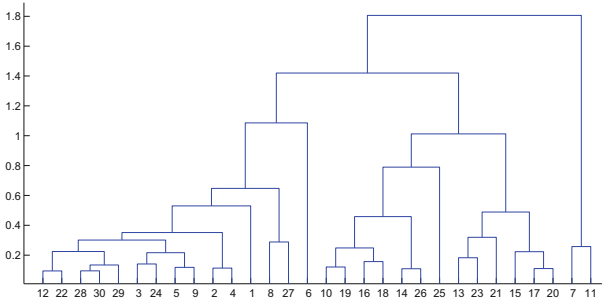
Fig. 7. ECG clustering from the feature vector

ecg 11 (MITBIH patient 111), ecg 22 (MITBIH patient 123) and ecg 11 (MITBIH patient 111). Ecg records of the MIT BIH patients 116 and 203 are also very dissimilar (Fig. 4).

4.4   Classification

The output coding is defining as follows: 0 for abnormal signal and 1 normal signal. The output of the neural network (Algorithm 3) computed from all the nodes are then used for the ecg records classification. Algorithm 5 gives the main processing of the classifier.

The network is created with 70% training set, 15% validation set and 15% test set. The results are compared with clinical experimental results available on the physio-net web set. The performance of the classification is 82%. The samples that were not correctly classified are MIT BIH patient 213, 214, 217, 220, 222, 231, 232 and 233.

Future studies will aim to improve the neural network architecture by more straightening the choice of the number of nodes, the training set, the test set and the validation set. The proposed classifier and the performance evaluation model will also be improved.

Algorithm 5. Pseudo code for the classifier

1: $function\ classify(x, w)$
2: x : ecg records
3: $w \leftarrow train(t)$: the threshold
4: $fs \leftarrow filter(x)$ the filtered signal
5: $v \leftarrow vector(fs)$: the vector of the filtered signal attributes
6: $rn \leftarrow network(v, w)$ the neural network output
7: **if** $rn \leq 0.5$ **then**
8:      $c \leftarrow 0$ for abnormal class
9: **else**
10:      $c \leftarrow 1$ for normal class
11: **end if**
12: return c

5   Conclusion

In this work, tools for biomedical signal analysis are proposed. We first propose noise removal process based on empirical mode decomposition (EMD) and the Butterworth filtering. Then, QRS complex detection method from EMD is proposed and an algorithm for QRS wide computation is given and tested on ECG records. This latter set was used to perform data analysis and study ECG records similarities. A neural network classifier taking as input statistical characteristics and the QRS wide is also proposed and applied on MIT BIH ECG records. In future works, the following improvements will be considered: improve the proposed classification method, use more geometrical ECG features, extends the classification method to biomedical signals of higher dimensions, performs validation tests and add security supplement for data privacy management.

References

1. Al-Ashkar, T.: Medical multi signal signature recognition applied to cardiac diagnostic (2012)
2. Chang, P.C., Lin, J.J., Hsieh, J.C., Weng, J.: Myocardial infarcation classification with multi-lead ECG using hidden Markov models and Gaussian mixture models. Appl. Soft Comput. **12**, 3165–3175 (2012)
3. Fatin, A.E., Salim, N., Harris, A.R., Swee, T.T., Ahmed, T.: Arrhythmia recognition and classification using combined linear and nonlinear feature of ECG signals. Comput. Methods Programs Biomed. **127**, 52–63 (2016)
4. Jin, F., Sugavaneswaran, L., Krishnan, S., Chauhan, V.S.: Quantification of fragmented QRS complex using intrinsic time-scale decomposition. Biomed. Signal Process. Control **31**, 513–523 (2017)
5. Moody, G.B., Mark, R.G.: The impact of the MIT-BIH arrhythmia database. IEE Eng. Med. Biol. **20**, 45–50 (2001)
6. Niang, O., Thioune, A., Delechelle, E., Lemoine, J.: Spectral intrinsic decomposition method for adaptative signal representation. ISRN Sig. Process. **9**, 3 (2012)
7. Pal, S., Mitra, M.: Empirical mode decomposition based on ECG enhencement and QRS detection. Comput. Biol. Med. **42**, 83–92 (2012)

8. Pasa, L., Testolin, A., Sperdui, A.: Neural networks for sequential data: a pre-training approach based on hidden Markov models. Neurocomputing **169**, 323–333 (2015)
9. Polat, K., Gunes, S.: Detection of ECG arrythmia using a differential expert system approach based on principal component analysis and least square support vector machine. Appl. Math. Comput. **186**, 898–906 (2007)
10. Ripoll, V.J.R., Wojdel, A., Romero, E., Ramos, P., Brugada, J.: ECG assessment based on neural networks with pretraining. Appl. Soft Comput. **49**, 399–406 (2016)
11. Rodriguez, R., Mexicano, A., Bila, J., Cervantes, S., Ponce, R.: Feature extraction of electrocardiogram signals by aplying adaptative threshold and principal component analysis. J. Appl. Res. Technol. **13**, 261–269 (2015)
12. Slimane, Z.H., Nait Ali, A.: QRS complex detection using empirical mode decomposition. Digit. Signal Proc. **20**, 1221–1228 (2010)
13. Tjolleng, A., Jung, K., Hong, W., Lee, W., Lee, B., You, H., Son, J., Park, S.: Classification of a driver's cognitive workload levels using artificial neural network on ECG signals. Appl. Ergon. **59**, 326–332 (2017)
14. Tu, J.V.: Advantages and disadvantages of using artificial neural networks versus logistic regression for predicting medical outcomes. J. Clin. Epidemiol. **45**, 1225–1231 (1996)
15. Zhang, H.: An improved QRS wave group detection algorithm and matlab implementation. Phys. Procedia **25**, 1010–1016 (2012)