SVM Based ECG Classification Using Rhythm and Morphology Features, Cluster Analysis and Multilevel Noise Estimation

Radovan Smíšek[1,2], Jakub Hejč [1,3], Marina Ronzhina[1], Andrea Němcová[1], Lucie Maršánová[1], Jiří Chmelík[1], Jana Kolářová[1], Ivo Provazník[1,4], Lukáš Smital[1], Martin Vítek[1]

[1] Department of Biomedical Engineering, Brno University of Technology, Brno, Czech Republic
[2] Institute of Scientific Instruments of the Czech Academy of Sciences, Brno, Czech Republic
[3] University Hospital Brno – Children's Medical Center, Brno, Czech Republic
[4] Department of Physiology, Masaryk University, Brno, Czech Republic

Abstract

Background: Smartphone-based ECG devices comprise great potential in screening for arrhythmias. However, its feasibility is limited by poor signal quality leading to incorrect rhythm classification. In this study, advanced method for automatic classification of normal rhythm (N), atrial fibrillation (A), other rhythm (O), and noisy records (P) is introduced.

Methods: Two-step SVM approach followed by simple threshold based rules was used for data classification. In the first step, various features were derived from separate beats to represent particular events (normal as well as pathological and artefacts) in more detail. Output of the first classifier was used to calculate global features describing entire ECG. These features were then used to train the second classification model. Both classifiers were evaluated on training set via cross-validation technique, and additionally on hidden testing set.

Results: In the Phase II of challenge, total F1 score of the method is 0.81 and 0.84 within hidden challenge dataset and training set, respectively. Particular F1 scores within hidden challenge dataset are 0.90 (N), 0.81 (A), 0.72 (O), and 0.55 (P). Particular F1 scores within training set are 0.91 (N), 0.85 (A), 0.76 (O), and 0.73 (P).

1. Introduction

This paper describes an automated algorithm for classification of ECG from the 2017 PhysioNet/CinC challenge database. Challenge was focused on atrial fibrillation classification from short single lead ECG. The aim was to classify these signals into one of 4 types: normal (N), atrial fibrillation (A), other (O), and noisy (P). The testing set consisted of 8,528 signals of length from 9 s to 61 s. A detailed description of the challenge is given in [1].

In this study, ECG classification method based on supervised non-parametric technique – support vector machine (SVM) – is introduced. SVM is widely used for ECG classification due to its simplicity, robustness and efficiency [2, 3], which was confirmed in our previous study, too [4, 5].

2. Method

The block diagram of the proposed algorithm is shown in Figure 1. The method for classification of ECG into 4 required groups consists of 7 main blocks. At first, ECGs are preprocessed. Various features are then extracted from selected PQ intervals and QRS complexes and used to classify these particular parts of ECG. In the next step, so called global features describing the entire ECG are derived. The most appropriate features are further selected using genetic algorithm (GA) and they are used for training of the supervised classifier. Finally, additional threshold based rules (defined empirically) are applied on the classifier output in order to improve the overall performance of the method. Each part of the algorithm is described in following subchapters in more detail.

2.1. Preprocessing

A broad-spectrum noise and a baseline wander were removed using wavelet wiener filter and pyramid median filter, respectively. Error signals were further used for the feature extraction and noise classification. To obtain reliable QRS positions, the detection was performed using combination of 3 detectors – phasor transform, continuous wavelet transform (CWT), and S-transform. QRS onset and offset for each complex were detected by a custom-made algorithm [6]. Moreover, Teager-Kaiser Energy Operator (TKEO) was applied to the original signal for the subsequent feature extraction.

2.2. Extraction of features

Local features
For more detail representation of various pathological patterns present in the analysed ECG, morphological features were first calculated from PQ intervals and QRS complexes, respectively. PQ features were based on the

```
┌─────────────────┐
│   ECG signals   │────────┐
└────────┬────────┘        │
         ▼                 │
┌─────────────────┐        │
│  Preprocessing  │        │
└────────┬────────┘        │
         ▼                 │
┌──────────────────────┐   │
│ Extraction of local  │   │
│      features        │   │
├───────────┬──────────┤   │
│PQ interval│QRS complex│  │
└─────┬─────┴─────┬────┘   │
      ▼           ▼        │
┌──────────┐ ┌──────────┐  │
│ SVM RBF 1│ │ SVM RBF 2│  │
└─────┬────┘ └────┬─────┘  │
      ▼           ▼        │
┌──────────────────────┐   │
│ Extraction of global │◄──┘
│      features         │
└──────────┬───────────┘
           ▼
┌──────────────────────┐
│   Selection of        │
│  features using GA    │
└──────────┬───────────┘
           ▼
┌──────────────────────┐
│    Linear SVM 3       │
└──────────┬───────────┘
           ▼
┌──────────────────────┐
│    Threshold rules    │
└──┬────┬────┬────┬────┘
   ▼    ▼    ▼    ▼
 ┌──┐ ┌──┐ ┌──┐ ┌──┐
 │ N│ │ A│ │ O│ │ P│
 └──┘ └──┘ └──┘ └──┘
```
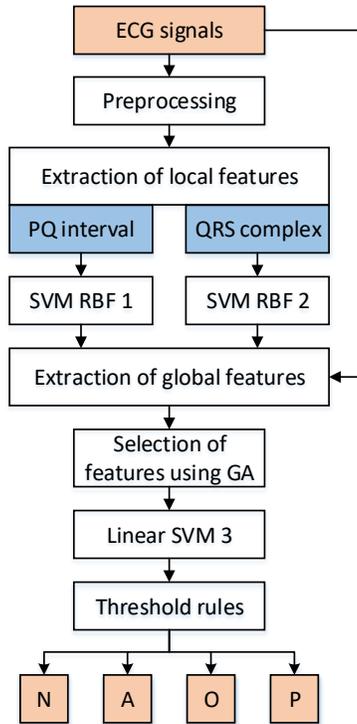
Figure 1. Block diagram of the ECG classification.

spectral analysis of signal-averaged PQ intervals and local statistics of RR dynamics, and mainly represented patterns associated with atrial fibrillation. Both, TKEO-transformed and filtered ECGs were used to extract QRS features characterizing pathological events such as ventricular extrasystoles, and myocardial ischemia. The features were calculated as a root mean square and 3rd–6th central moments within various time windows: a) starting at QRS onset and lasted 70 ms, b) starting at QRS onset and lasted 200 ms, c) started 90 ms after QRS onset and lasted 110 ms. Thus, PQ and QRS local features mainly contributed to recognize heart beats from N, A and O groups.

The set of features was used to classify the individual PQ and QRS intervals into 4 (N, A, O, and P) and 3 (N, O, and P) classes, respectively. To learn the classifier, PQ and QRS were derived from particular beats (P-QRS-T randomly selected through the entire ECGs, n = 31,404) and scored by the visual analysis.

Global features

To represent the entire signal, altogether 227 global features were extracted.

The first global feature group was based on the results of PQ and QRS evaluation (see Classification) via above local features. This group consisted of features such as relative frequency, the length of the largest continuous segment, average length of all segments, and the number of segments with only single ECG beat.

Other groups were directly derived from the entire ECGs. One of them included common linear and nonlinear RR interval parameters: a) percentage of RR intervals 50 ms and 100 ms longer than the previous ones, b) percentage of NN intervals 20 ms and 50 ms longer than the previous ones (NN is used for a normal beat), c) standard deviation of RR intervals.

Relative wavelet energy (RWE) [7] global features were then calculated as a mean and median of the local RWE. These features were based on the stationary wavelet transform (SWT) applied on PQ interval.

Other group was based on the evaluation of inter-QRS similarity within the signal using cluster analysis and consisted of number of clusters, percentage rate of QRS in the largest cluster, number of single-QRS clusters, etc.

One of the most interesting global features was an estimate of the noise level contained in ECG. For this purpose, an algorithm based on signal-to-noise ratio (SNR) computation was used. The first step of this procedure is to gain an estimate of noise-free signal using the wavelet wiener filter [8] and separate the noise and signal component. Once these two parts are separated, SNR computation can be applied. This calculation is performed in a time-limited sliding window to receive time dependency of SNR.

2.3. Feature selection

The global features were evaluated in terms of their ability to distinguish between 4 ECG groups. The most successful features were selected for further classification via the genetic algorithm (GA) with a population of 30 individuals. Each individual contained 227 binary values indicating whether the appropriate parameter is suitable for the classification or not. To create the new populations, cross-over with probability of 0.8 and mutation with probability of 0.015 were applied. The selection of individuals into the new population was made by the tournament selection. The principle of elitism was applied. This means that the best individual of the old population was always selected into a new population.

2.4. Classification

Two-step approach based on the SVM (one-vs-all approach) was used for the multi-class classification. First, PQ and QRS were classified into 3 or 4 classes (see above) via two separate SVM with radial basis function (SVM RBF 1 and SVM RBF 2 in Figure 1) using local features. RBF kernel was chosen due to better predictive performance of SVM RBF as compared to linear SVM (not shown here). In this step, targets were defined via visual analysis of particular PQ and QRS.

To reduce the negative effect of data imbalance (lower number of samples in O and P as compared to other

groups) on classification performance, misclassification cost arrays were defined according to obtained classification results (represented by confusion matrix, see Table 1). Particularly, higher cost values were set in the cost arrays on the positions corresponding to that from confusion matrix with high number of incorrectly assigned samples. For example, value '2' for the true class O (row) and the predicted class N (slope) in cost arrays from Table 1 means that we care twice as much for the correct classification of the class O (small group, where each incorrectly assigned sample could be crucial for overall classification performance) than for class N. Cost arrays were mainly adjusted to avoid the misclassifications between following group pairs: N – O, N – P, and P – A.

Table 1. PQ (top) and QRS (bottom) local classification: left - confusion matrix for one cross-validation fold; right - corresponding misclassification cost array. Rows - real output, slopes - predicted output.

PQ	N	A	O	P		N	A	O	P
N	1027	18	4	56	N	0	1	1	2
A	12	1039	8	18	A	1	0	1	2
O	10	5	170	7	O	2	1	0	1
P	76	40	6	644	P	2	2	1	0

QRS	N	O	P		N	O	P
N	2214	3	89	N	0	1	1
O	35	177	14	O	2	0	1
P	60	5	543	P	2	1	0

Second, the global features selected by the GA were used to train SVM with linear kernel function (linear SVM 3 in Figure 1), and classify ECGs into 4 classes (N, A, O, and P). Linear SVM is fast and allows achieving more reliable predictive performance as compared to other kernel functions.

The performance of local approaches was evaluated by 10-fold cross-validation technique which is preferable for large data sets. In the case of global approaches, an evaluation on validation set (300 signals out of the whole data set) was chosen instead of k-fold cross-validation due to low number of data in A and P groups.

Overall success of the ECG classification was additionally increased by two threshold based rules applied on the global classifier output:

a) If the analyzed signal was classified by SVM as group N and the heart rate was higher than 100 bpm or lower than 50 bpm (indicating tachycardia and bradycardia, respectively), the signal was assigned to group O.

b) If the analyzed signal was classified by SVM as N and the ST segment depression is greater than 0.3 mV

(indicating myocardial ischemia or myocardial infarction) then the signal is assigned to the group O.

3. Results and discussion

Resulting test F1-measures for the first classification step are summarized in Table 2. According to the results, the recognition of both PQ and QRS from O and P groups was the most difficult.

Table 2. Performance of the first classification step: classification of individual PQ and QRS. Mean ± STD; evaluated by 10-fold cross-validation.

F1	Local PQ	Local QRS
N	0.92±0.01	0.96±0.00
A	0.95±0.00	-
O	0.89±0.02	0.88±0.02
P	0.87±0.01	0.88±0.01

F1-measures for the global classifier are shown in Table 3. It is evident that recognizing ECGs from O group is a difficult task. According to the confusion matrix calculated for the global classifier with threshold (see Table 4, Global & GA & Threshold), quite low F1-measure for O group was obtained. It is due to ECGs from this group were misclassified as N (and vice versa).

Table 3. Performance of the second classification step: the global classifier using all features (Global) or GA selected features (Global & GA) and above classifier supplemented with threshold rules (Global & GA & Threshold). Evaluated on the validation set (300 ECGs).

F1	Global (all features)	Global & GA	Global & GA & Threshold
N	0.89	0.90	0.91
A	0.84	0.85	0.85
O	0.70	0.72	0.76
P	0.75	0.75	0.75
Overall*	0.81	0.82	0.84

* - mean from F1 values calculated for N, A, and O

Table 4. Confusion matrix for the global classifier (Global & GA & Threshold) for the whole public dataset.

Global classifier		Predicted output			
		N	A	O	P
Real output	N	4728	16	289	17
	A	23	632	76	7
	O	570	103	1747	36
	P	47	5	26	206

Other challenging task is a detection of noisy beats, which were mostly assigned as N, whereas relative high number of beats from group O was recognized by the classifier as noisy (see Table 4).

Combining of the SVM with threshold based rules resulted in an improvement of the overall classification performance (F1-measure) by 0.02 (see Table 3). Threshold based rules affect N (improving F1-measure by 0.01) and O (improving F1-measure by 0.04) group.

The generalization capability of the proposed classification approach was further evaluated on hidden dataset. Resulting F1-measure obtained on training set, public testing set and hidden testing set are shown in Table 5. As seen from Table 3 and Table 5, results obtained on hidden testing set are similar to those obtained on training set, which confirms good generalization ability of the proposed classifier. Throughout the challenge, the results obtained on test data and hidden data differed by about three percent.

Table 5. F1-measure for training dataset, public testing dataset, and hidden testing dataset of 2017 PhysioNet/CinC Challenge.

F1	Phase II training	Phase II testing	Hidden dataset
N	0.90	0.91	0.90
A	0.92	0.85	0.81
O	0.76	0.76	0.72
P	0.85	0.73	0.55
Overall*	0.86	0.84	0.81

* - mean from F1 values calculated for N, A and O

4.   Conclusion

Overall F1-measure of the proposed method reaches 0.81 on hidden dataset. Use of two-step classification approach seems to be reasonable. Combination of the SVM with simple threshold based rules allows additional improvement of the classification results.

Quite high performance was achieved for group A, despite its low number in training set. Thus, proposed PQ based features seem to be appropriate for representing the patterns associated with atrial fibrillation.

As expected, reduced overall performance was mainly related to misclassifications between N and O. This is due to the similar morphology of ECGs from these groups, high variety in data within both groups, and quite low number of ECGs in group O (approx. half the size of N).

The detection of various types of noise and artefacts (including short-term events and the wideband noise distorting entire ECG) present in data recorded by a mobile device is also difficult and requires collecting more data into group P to train the classifier more accurately.

References

[1] Clifford G, Liu C, Moody B, Silva I, Li Q, Johnson A, Mark R. Classification from a short single lead ECG recording: the PhysioNet Computing in Cardiology Challenge 2017. Computing in Cardiology. Rennes; 2017.

[2] Doquire G, de Lannoy G, François D, Verleysen M, Bencherif MA. Feature selection for interpatient supervised heart beat classification. Physiological Measurement 2010;31:903-920.

[3] Shen CP, Kao WC, Yang YY, Hsu MC, Wu YT, et al. Detection of cardiac arrhythmia in electrocardiograms using adaptive feature extraction and modified support vector machines. Expert Systems with Applications 2012;39:7845-7852.

[4] Ronzhina M, Maršánová L, Smíšek R, Olejníčková V, Janoušek O, et al. Classification of ventricular premature and ischemic beats in animal electrograms. Computing in Cardiology 2015;39:1137-1140.

[5] Maršánová L, Ronzhina M, Smíšek R, Vítek M, Němcová A, et al. ECG features and methods for automatic classification of ventricular premature and ischemic heartbeats: A comprehensive experimental study. Scientific Reports 2017;7:11239.

[6] Hejc J, Vitek M, Ronzhina M, Novakova M, Kolarova J. A Wavelet-Based ECG Delineation Method: Adaptation to an Experimental Electrograms with Manifested Global Ischemia. Cardiovascular Engineering and Technology 2015;6:364-75.

[7] García M, Ródenas J, Alcaraz R, Rieta JJ. Application of the relative wavelet energy to heart rate independent detection of atrial fibrillation. Computer Methods and Programs in Biomedicine. 2016;131:157-168.

[8] Smital L, Vítek M, Kozumplík J, Provazník I. Adaptive wavelet Wiener filtering of ECG signals. IEEE Transactions on Biomedical Engineering. 2013;60(2):437-445.

Address for correspondence.

Radovan Smíšek
Department of Biomedical Engineering, Faculty of Electrical Engineering and Communication, Brno University of Technology
Technická 3058/12, 61600 Brno, Czech Republic
E-mail address: xsmise00@stud.feec.vutbr.cz