Classification of Ventricular Premature and Ischemic Beats in Animal Electrograms

Marina Ronzhina[1,2], Lucie Marsanova[1], Radovan Smisek[1], Veronika Olejnickova[3], Oto Janousek[1,2], Petr Vesely[2], Jana Kolarova[1,2], Marie Novakova[2,3], Ivo Provaznik[1,2]

[1]Department of Biomedical Engineering, Brno University of Technology, Brno, Czech Republic
[2]International Clinical Research Center, St. Anne's University Hospital Brno, Czech Republic
[3]Department of Physiology, Faculty of Medicine, Masaryk University, Brno, Czech Republic

Abstract

Many approaches have been proposed for automatic classification of different pathological events in ECG signals. Present study is focused on analysis and classification of non-ischemic, moderate and severe ischemic, and VPB segments in data obtained in rabbit isolated hearts. It is shown, that use of low number of morphological parameters calculated from electrograms combined with even simple classification method allows achieving of accurate results (with overall accuracy up to 0.99) for four types of the segments.

1. Introduction

There are many approaches for automatic classification of different pathological events, including ventricular premature beats (VPBs) and other arrhythmias in human ECG. Other interesting issue is the detection of ischemic changes in ECG. Most classification approaches are based on use of rhythm- and morphology-related ECG parameters, such as duration and amplitude of different parts of ECG [1]-[4]. In recent studies, time-frequency methods, such as wavelet analysis and higher-order statistical approaches, are also proposed for representation of different changes in ECG [1],[3]-[7]. Besides the calculation of such ECG parameters, the appropriate classification method should be chosen to achieve accurate results. Artificial neural network, support vector machines and discriminant function (linear or quadratic) are the most popular in this area [1],[3]-[7].

In experimental data, VPBs may appear as a result of some changes of experimental conditions, such as pharmacological intervention or ischemia induction. In case of ischemia, VPBs detection may be complicated due to the similarity between morphology of VPBs and ischemic beats. In this work, classification of four types of heart beats is presented.

2. Methods

2.1. Experimental data

Data were recorded during experiments performed in accordance with the guidelines for animal treatment approved by local authorities and conformed to the EU law. Isolated hearts of ten New Zealand rabbits perfused according to Langendorff were used in this study. Three orthogonal electrograms (EGs) (see Fig.1) were recorded during the whole experiment (20 min stabilization and three repetitions of 10 min ischemia and reperfusion) with sampling frequency of 2kHz [8].

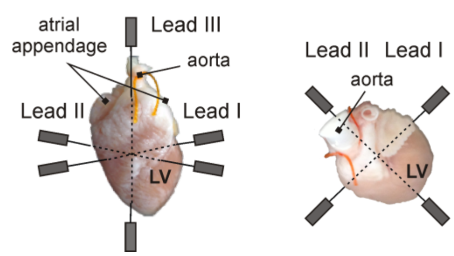

Figure 1. Orthogonal leads used for electrograms recording: front (left) and top (right) view. LV - left ventricle.

2.2. Electrograms processing and analysis

The parts of EGs with any artifacts were excluded from the study. Lynn's filter with cut-off frequency of 0.5Hz was used to eliminate the low-frequency baseline wandering. After that, in EGs recorded during stabilization and ischemic periods, QRS complexes were detected automatically and selected QRS-T segments were then manually delineated, i.e. the beginning and the end of QRS and the end of T wave were localized regarding three-lead EGs.

For further analysis, VPBs were found with human expert among selected QRS-T segments. Non-ischemic

(NOR), VPB, moderate (ISM), and severe (ISE) ischemic segments (see Fig.2) were then chosen (172 segments for VPB and 220 segments for other types, respectively) from data set for feature extraction. VPBs were found in ischemic periods only. Last two types were selected from data recorded in about the 5th and the 10th minute of ischemia, which can be characterized with slight and significant changes in EG morphology, respectively. As can be seen, ISE and VPB segments are of quite a similar morphology.

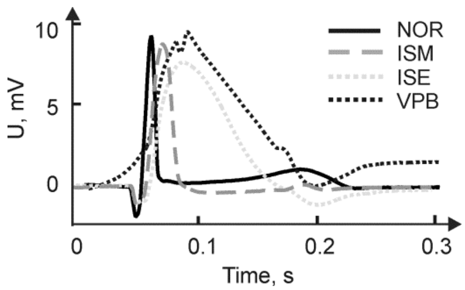

Figure 2. Types of classified QRS-T segments. NOR, ISM, ISE, VPB – non-ischemic, moderate and severe ischemic beats, and ventricular premature beats, respectively.

2.3. Features extraction

Classification features representing the morphology were calculated from EGs recorded with two horizontal leads. EG from vertical lead was excluded from the study because of poor changes in its morphology caused by ischemia. Total number of 71 morphological features was calculated from QRS-T segments. Seven features (duration of QRS, QT, ST-T, angle and length of maximum vector calculated from 2D QRS and ST-T loops separately) were calculated using EGs from both leads together. Remaining 32 features were derived from each lead separately. They include durations, amplitudes and areas under various parts (AUC) of QRS-T segment, such as QRS complex, ST-T segment and T wave. Additionally, relative values of AUC based features were calculated to study the changes in morphology of corresponding part of the segment regarding its total change.

In Fig.3, boxplots are shown for two different morphological parameters derived from four types of QRS-T.

2.4. Feature number reduction

Only the most informative features were then chosen from the whole data set using Kruskal-Wallis test accompanied with multicomparison Tukey-Kramer test (α=0.05). As a result, 10, 12 and 4 features with significantly different values among all types of data were selected for lead I, lead II and joint features, respectively. Examples of values distribution of two features selected with statistical test are shown in Fig.3.

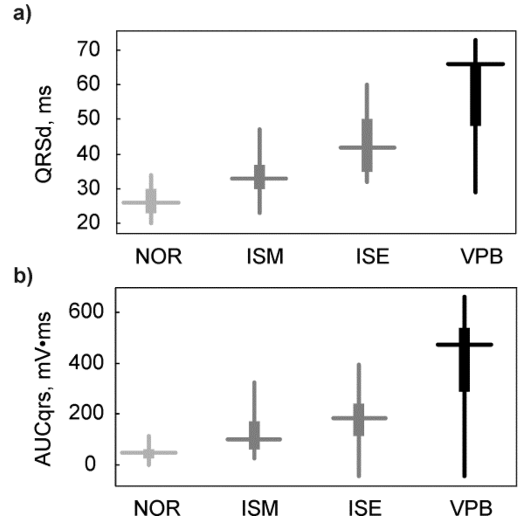

Figure 3. Boxplots for values of features calculated from different types of QRS-T segments with significant difference among all types ($p<0.05$) confirmed with statistical analysis: a) QRS complex duration (QRSd) joint for lead I and lead II, b) area under QRS (AUCqrs) for lead II.

Besides statistical test, principal component (PC) analysis was also applied on the whole data set (71 features) to reduce the number of features. In this case, only ten first PCs explaining approx. 93% of total variance of data (see Fig.4) were selected for further classification.

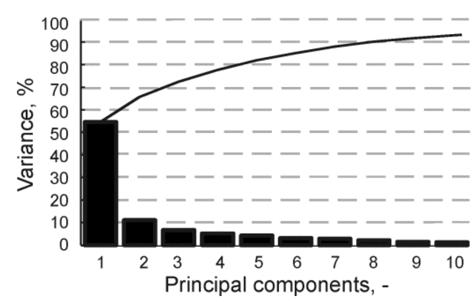

Figure 4. Dependence of explained variance on number of used principal components.

Thus, five different groups of features were proposed for classification of four types of QRS-T segments:

– 4 features selected from data joint for both leads,
– 10 features selected from data recorded with lead I,
– 12 features selected from data recorded with lead II,
– all selected features together (n=26),
– 10 features obtained with PC analysis.

2.5. Automatic classification

Four classification models were trained and used for automatic classification of QRS-T segments based on selected features mentioned above:

- discriminant function analysis (DFA) with linear and quadratic function,
- naive Bayes (NB) classifier with Gaussian density function and kernel density function estimation,
- support vector machine (SVM) (one-vs-all approach for multiclass classification) of general type and with radial basis function (RBF),
- k nearest neighbours (k-NN) with different *k* value (k=1, 5, 10, 15, and 20).

Training and testing of classifiers were performed using 10-fold cross validation approach. Besides initial feature values, their normalized modification was also used as input for training and testing of different classification models. Statistical characteristics, such as mean and standard deviation values, were calculated from training data and then used for normalization of training and corresponding testing feature values within each fold separately.

Mean overall accuracy (ACC) was calculated among ten folds of testing of each classifier approach. ACC was defined as a number of correctly classified segments of all groups related to their total number. Mean ACCs separately for each type of the segments were also calculated for more detailed interpretation of the results.

3. Classification results

Mean overall ACC values for different classification approaches are summarized in Table 1. It is evident, that the lowest ACC was obtained in case of DFA classification. The best results were generally obtained with k-NN (within wide range of k-values) and NB with kernel density estimation (ACC in range of 0.83-0.99 and 0.83-0.92, respectively). ACC of Gaussian NB is of about 0.08-0.15 lower than that of kernel NB. SVM is probably the most sensitive to data normalization with similar results for both common and RBF model.

Results obtained using feature sets from lead I and lead II are very similar with slightly better results for lead I data in case of SVM classification. Use of features joint for both leads allows obtaining of similar results regardless of lower number of features in this case. For all classifiers, use of all selected features together improves the accuracy; especially it is true in case of NB and SVM (increasing of ACC is at least 0.05 and 0.09, respectively). It can be also seen from Table 1, that using of PCA-reduced feature set allows reaching of similar results regardless the classification model. In this case, reached ACC values are in general higher than that obtained using other feature sets.

Mean ACCs for classification with linear discriminant function and selected features calculated from lead II data are: 0.93, 0.77, 0.52, and 0.62 for NOR, ISM, ISE, and VPB group, respectively.

Table 1. Mean overall accuracies (among 10 folds of validation) of different classification approaches.

Classifier	Group of used features								
	Joint for leads		lead I		lead II		All together		PCA
	Init.	Norm.	Init.	Norm.	Init.	Norm.	Init.	Norm.	Norm.
DFA									
Linear	0.72	0.72	0.77	0.77	0.72	0.72	0.83	0.83	0.84
Quadratic	0.73	0.73	0.70	0.70	0.71	0.71	0.79	0.79	0.87
NB									
Gaussian	0.75	0.75	0.71	0.71	0.71	0.71	0.79	0.79	0.87
Kernel	0.83	0.83	0.84	0.84	0.86	0.86	0.95	0.95	0.92
SVM									
Common	0.76	0.75	0.58	0.81	0.54	0.76	0.78	0.91	0.89
RBF	0.98	0.76	0.32	0.80	0.27	0.76	0.26	0.87	0.92
k-NN									
k=1	0.99	0.99	0.98	0.99	0.99	0.99	0.99	0.99	0.99
k=5	0.98	0.98	0.95	0.97	0.98	0.98	0.99	0.99	0.99
k=10	0.90	0.92	0.89	0.92	0.91	0.95	0.93	0.96	0.97
k=15	0.87	0.88	0.85	0.89	0.89	0.90	0.91	0.94	0.94
k=20	0.86	0.87	0.83	0.87	0.87	0.89	0.89	0.93	0.92

Corresponding confusion matrix of the results (example within one fold of testing) is shown in Table 2. It is evident, that the classification of ISE and VPB segments is the most difficult. ISE segments were misclassified as ISM and VPB (5 and 6 segments from 22, respectively). On the other hand, 10 VPB segments (from 17) were misclassified as ISE.

In case of classification with RBF SVM method using PCA-reduced features, the mean classification ACCs for four types of QRS-T segments are: 0.93, 0.84, 0.96, and 0.94, respectively. Example of corresponding confusion matrix is shown in Table 2. Only 3 VPB segments were inaccurately recognized as ISE. Performed classification is more accurate comparing with the previous approach.

Table 2. Example of confusion matrix for linear discriminant function classification using lead II features and RBF SVM classification using PCA-reduced features.

			Classifier output			
			NOR	ISM	ISE	VPB
Linear DFA lead II	Real output	NOR	20	2	0	0
		ISM	1	18	3	0
		ISE	0	5	11	6
		VPB	0	0	10	7
RBF SVM PCA-reduced	Real output	NOR	20	2	0	0
		ISM	1	18	3	0
		ISE	0	0	22	0
		VPB	0	0	3	14

4. Discussion and conclusions

According to the results, use of low number of morphological parameters calculated from one- or two-lead electrogram in combination with simple classification methods (such as k-NN or DFA) allows achieving of accurate results for four types of QRS-T segments. One of the best achieved classification result (ACC=1.00, 0.93, 0.97, and 0.96 for NOR, ISM, ISE, and VPB segments, respectively, in case of 10-NN classification using PCA-reduced features) is similar or even better comparing with that obtained with other methods. For example, overall ACC of binary classification (NOR/VPB) obtained with SVM and Gaussian RBF classifier using morphological, wavelet and higher order statistical features are approx. 0.90 and 0.92, respectively [1]. ACC for VPB is about 0.96 and NOR about 0.93 in case of classification with multilayer perceptron (MLP) using morphological and rhythm-based features [3] and overall ACC of binary classification (ischemic/normal) with MLP using higher order statistical and spectral features is 0.96 [5].

As expected, reduced ACC was mainly related to misclassifications between ISE and VPB groups with similar character of electrogram morphology. Selection of more informative features using statistical test or PCA and use of features derived from one-lead data only reduce time and PC memory requirements that can be quite high in case of multiclass classification.

Acknowledgements

This work was supported by European Regional Development Fund - Project FNUSA-ICRC (No. CZ.1.05/1.1.00/02.0123), grant project of the Grant Agency of the Czech Republic GAP 102/12/2034 and Specific University Research Grant of Masaryk university no. MUNI/A/1326/2014, as provided by the Ministry of Education, Youth and Sports of the Czech Republic in the year 2015.

References

[1] Alajlan N, Bazi Y, Melgani F, Malek S, Bencherif MA. Detection of premature ventricular contraction atthythmias in electrocardiogram signals with kernel methods. SIViP 2014;8:931-942.
[2] Park SM, Lee JH, Choi SW. Detection of premature ventricular contractions on a ventricular electrocardiogram for patients with left ventricular assist devices. Artif Organs 2014;38:1040-1046.
[3] Das MK, Ari S. ECG beats classification using mixture of features. ISRN 2014; 2014:1-12.
[4] Llamedo M, Martínez JP. Heartbeat classification using feature selection driven by database generalization criteria. IEEE Trans Biomed Eng 2011;58:616-625.
[5] Ronzhina M, Potočňák T, Janoušek O, Kolářová J, Nováková M, Provazník I. Spectral and higher-order statistics analysis of ECG: application to study of ischemia in rabbit isolated hearts. Computing in Cardiology 2012; 645-648.
[6] Sarma P, Nirmala SR, Sarma KK. ECG classification using wavelet subband energy based features. SPIN 2014; 785-790.
[7] Correa R, Arini PD, Correa L, Valentinuzzi ME, Laciar E. Acute myocardial ischemia monitoring before and during angioplasty by a novel vectorcardiographic parameter set. J Electrocardiol 2013;46:635-643.
[8] Kolářová J, Nováková M, Ronzhina M, Janoušek O, Veselý P, Olejníčková V, Provazník I. Isolated rabbit hearts – databases of EGs and MAP signals. Computing in Cardiology 2013; 551-554.

Address for correspondence.

Marina Ronzhina
Department of Biomedical Engineering
Technická 12
616 00 Brno
Czech Republic
ronzhina@feec.vutbr.cz