Classification of Arrhythmia using Multi-Class Support Vector Machine

Tonghui Li, Jieming Ma, Xinyu Pan, Yujia Zhai, and Ka Lok Man

Abstract—**Arrhythmia has become the most common disease in the medical field. Manual diagnosis of arrhythmia beats is very tedious owing to its nonlinear and complex nature of electrocardiogram (ECG). In this article, a multi-class support vector machine (MSVM) based approach is proposed to solve ECG multi-classification problem. Based on the characteristics of the R-R interval, it has the capability of detecting normal heart rate (NOR), left bundle branch block (LBBB), right bundle branch block (RBBB), atrial premature complex (APC) and Ventricular premature beat (VPC) was mainly discussed. Using ECG MIT-BIH database, simulation results show the proposed method achieves a very high classification accuracy.**

Index Terms—Arrhythmia, MSVM, ECG, classification, kernel function.

I. INTRODUCTION

IN clinical medicine, the arrhythmia, referred to an abnormal rhythm [1], usually affects the rate of heart beats. Clinically, according to the heart rate of arrhythmia, the rate of heart rate is divided into two categories: fast and slow arrhythmia [2]. Detection of arrhythmia classification is of importance in clinical medicine, especially in the real-time implementation of cardiology.

There are a variety of methodologies for automatic detection of the classification of cardiac arrhythmias have been proposed in recent years. [3] - [11]. Khazaee et al. [3] used support vector machines (SVM) and genetic algorithms (GA) to detect arrhythmia, which is referred to as identification of ventricular contraction (PVC). However, this approach can only distinguish three types of arrhythmia, including NOR, PVC and others, and it did not clearly express what kind of kernel function was used in SVM. Lin [4] presented a new method of grey correlation analysis of electrocardiogram (ECG) heartbeat discrimination (GRA). In [5], Zeraatkar et al. used T wave of electrocardiogram (ECG) on the heart repolarization phase, since the T wave can be explained in the correct operation of the electrical activity of the heart and the classification of ECG detection. Ebrahimzadeh [6] proposed a three-stage technique for detection of premature ventricular contraction (PVC) to distinguish between normal beats and other heart diseases. This method takes a denoising module, a feature extraction module, and a classification module to analyze them. In [7], Ince presented a feature extraction process that utilizes morphological wavelet transform features. It can be projected onto a lower dimensional feature space which uses principal component analysis and temporal features from the ECG data. Inan et. al [8] were aimed at coupling morphological information with timing information, which is more constant among patients, and they combine wavelet-transformed ECG waves with timing information as our feature set for classification. They used different feature values to classify the ECG data, but the classification is not much. Shyu et. al [9] presented a novel method for detecting ventricular premature contraction (VPC) from the Holter system is proposed using wavelet transform (WT) and fuzzy neural network (FNN), which detect QRS wave. In [10], ECG signals can efficiently and automatically detect QRS wave, based on the inverse double orthogonal wavelet decomposition and nonlinear filter. The resulting coefficients are filtered using non-linear filters (averaging and median filter). In [11], Yazdani et. al study is aimed at extracting Q-wave, R-wave, S-wave (QRS) complexes and their fiducial points by mathematical morphology (MM) with an adaptive structuring element, on a beat-to-beat basis. The structuring element is based on the characteristics of the previously detected QRS complexes for a more robust and precise detection. The process of extracting QRS wave needs wavelet change, and there will be data error, more cumbersome. Here MIT-BIH Library in the existing parameters can be used.

The multi−class support vector machine (MSVM) method is based on the Vapnik-Chervonenkis (VC) dimension theory and structural risk minimization principle of statistical learning theory. In this paper, we use SVM to distinguish normal heartbeat from left bundle branch block (LBBB), right bundle branch block (RBBB), premature atrial contraction (APC) and premature ventricular premature beat (VPC). Besides, the different eigenvalues of their ECG as classification can be grabbed and different kernel functions are made use of to classify them.

The paper is organized as follows. In Section 2, the specific working principle of SVM and the different kernel functions are introduced in detail. In Section 3 Kernel function of MSVM is applied to classification of arrhythmia and simulation results are discussed in detail. We draw conclusions in the last section.

.

II. ECG HEARTBEAT CLASSIFICATION

The MSVM method is a kind of machine learning method that is based on the structural risk minimization principle and theoretical basis, by selecting the appropriate subset function

Tonghui Li, Jieming Ma and Xinyu Pan are with School of Electronic and Information Engineering Suzhou University of Sicence and Technology Suzhou, P.R. China. (e-mail: Jieming.ma@gmail.com).

Yujia Zhai is with Electrical Electronic Engineering Department, Xi'an Jiaotong-Liverpool University, P.R. China.

Ka Lok Man is with Computer Science and Software Engineering Department, Xi'an Jiaotong-Liverpool University, P.R. China.

ISBN: 978-988-14047-7-0
ISSN: 2078-0958 (Print); ISSN: 2078-0966 (Online)

and the subset of the discriminant function [12,13]. The MSVM is used to distinguish NOR, LBBB, RBBB, APC, PVC. One part of the data can be used as training, and the other part of data is used as a test. The MIT-BIH library is used as a database to provide data of arrhythmia.

A. MSVM

In this section, an "one-against one" multiclass support vector machine (MSVM) algorithm for shading pattern identification is presented. The basic idea behind MSVM is to construct separating hyperplanes between classes in feature space using support vectors. Given the module voltage vector V_i ($i = 1, 2, ..., n$)and the number of shaded modules $y_i (y_i \in 1, ..., n)$, the decision function for the training data from the i^{th} and the j^{th} classes can be expressed by

$$f_{ij}(V) = w^{ij} \cdot \Phi(V) + b^{ij} \tag{1}$$

The two-class classification problem is solved by:

$$\begin{aligned} \min_{w^{ij}, b^{ij}, \xi^{ij}} \quad & \frac{1}{2}(w^{ij})^T w^{ij} + C \sum_t^m (\xi^{ij})_t \\ s.t. \quad & (w^{ij})^T \Phi(V_t) + b^{ij} \geq 1 - \xi_t^{ij}, \text{if:} y_i = i \\ & (w^{ij})^T \Phi(V_t) + b^{ij} \leq 1 - \xi_t^{ij}, \text{if:} y_i = j \\ & \xi^{ij} \geq 0 \end{aligned} \tag{2}$$

where C is the associated penalty for excessive deviation, ξ^{ij} is the non-negative slack variables, and Φ is a mapping function.

Equation (2) can be solved by introducing the Lagrange multipliers α^{ij} for its dual optimization model. After the optimal solution $(\alpha^{ij})^*$ is obtained, the optimal hyperplane parameters $(w^{ij})^*$ and $(b^{ij})^*$ can be determined, and the indicator function (classifier) can be written as:

$$\text{sign}[\sum_{t=1}^{t} y_t (\alpha^{ij})_t^* \Phi(V) \cdot \Phi(V_t) + (b^{ij})^*] \tag{3}$$

The "one-against-one" strategy is used to extend SVM to the multi-class scenario. As shown in Fig.3, there are $c_2^k = k(k-1)/2$ classifiers used in training. Each classifier is trained with two different classes. The strategy gives one vote to the j^{th} class, and the classes that receive the most votes serve as classification results. Through this pairwise comparison, the number of shaded modules of a PV string is determined.

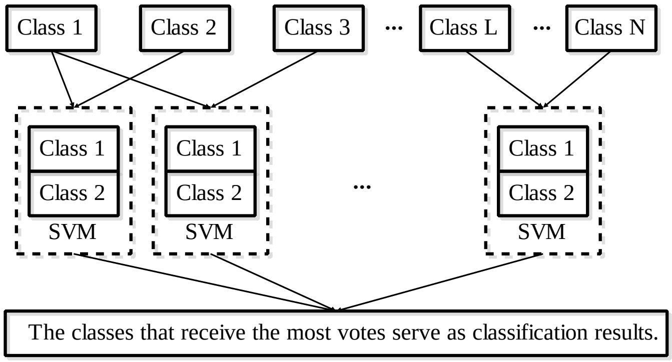

Fig. 1 Separation of two classes by SVM.

B. Kernel function

The performance of MSVM is largely dependent on the choice of kernel function, the different kernel function, the training results will be very different. There are four kinds of kernel functions of common support vector machines. Their kernel functions are as follows:

Linear: $K(x_i, x_j) = x_i \cdot x_j$ (4)

RBF: $K(x_i, x_j) = \exp(-\|x_i - x_j\|^2 / 2\delta^2)$ (5)

Polynomial: $K(x_i, x_j) = (x_i \cdot x_j + 1)^d$ (6)

Sigmoid: $K(x_i, x_j) = \tanh(\gamma x_i \cdot x_j + \eta)$ (7)

C. Proposed method

MIT-BIH library, provided by the Massachusetts Institute of Technology (MIT), is recognized as one of the three standard ECG databases and is used in this paper. In the MIT-BIH arrhythmia database, there are different types of heart disease data with 48 sets, including ECG and anteroventral leads (V1) leads. A frequency of 360 HZ is used in ECG sampling.

MITBIH library contains all the period of the R-R interval. Five local timing features can be extracted based on the R-R interval, which can promote the ability of morphological characters' recognition. The effect of five feathers is most significant when distinguishing the similar heartbeats patterns. Five local timing features are an R-R time interval ratio (IR), an R-R time interval difference (ID), a sum of two R-R time intervals (SI) and two R-R time intervals. The IR feature reflects the heartbeat rate deviation of two adjacent R-R intervals and the ID feature reflects the deviation of non-adjacent R-R intervals of heartbeat rate. Besides, the SI shows the sum of two non-adjacent R-R intervals, which can reveal the heartbeat speed, because an R-R cannot obviously show heart rates' differences among different kinds of arrhythmia. When taking sum of two non-adjacent R-R time intervals, the heart rates' differences will be presented more obviously twice as much as that using only one R-R intervals. Not adjacent to the RR interval is to make it more universal. They are defined as:

$$IR_i = \frac{T_i - T_{i-1}}{T_{i+1} - T_i} \tag{8}$$

$$ID_i = T_{i+2} - T_{i+1} - (T_i - T_{i-1}) \tag{9}$$

$$SI_i = T_{i+2} - T_{i+1} + T_i - T_{i-1} \tag{10}$$

In the above formulas, T_i defined as the time at which the R-wave for beat i occurs. Except IR, ID and SI, two other timing features are the previous and next R-R time intervals for each kind of heartbeat. A normal heartbeat, the value of IR is approximately equal to 1 and ID is about equal to 0. SI and two other timing features are used to distinguish different arrhythmias from the heart rate.

III. RESULTS AND DISCUSSIONS

To evaluate the feasibility of the proposed method, six records (118, 207, 208, 209, 214, and 223) from the MIT-BIH database are used as data of arrhythmia. It is not hard to get the data of R-R intervals since an internal tool developed by MIT can extract R-R intervals. There are 70% samples are used to train the MSVM, and all that remained are used for testing. The classification accuracy is defined as the percentage of correct classifications.

Table I compares the accuracy of the MSVMs with linear, quadratic and RBF kernel functions in the detection of NOR, LBBB, RBBB, APC, VPC. The data used in this experiment are all selected from untrained data. In a MSVM, the associated penalty for excessive deviation C and the turning parameter δ involved in the RBF affects the classification accuracy. An optimization algorithm, such as particle swarm optimization (PSO), can be used to optimize the two tuning parameters. The PSO-MSVM therefore refers to such a MSVM with parameter optimization process. From Table I, it is observed that the MSVM with a linear kernel obtains the lowest accuracy, especially for the case of RBBB. It is because that the linear kernel can only be applied to the case of linear separable. The quadratic polynomial kernel function improves the correctness of the linear kernel function by using higher order operation, and the overall accuracy reaches 90%. However, the error rate of VPC detection is relatively large. The RBF-MSVM show better results in VPC detection. This is consistent with results obtained in many SVM classification applications, that the RBF kernel function is commonly preferred to other kernel function types. With optimum turning parameters, the accuracy of PSO-MSVM is up to 99.17%.

Fig.2 further evaluate the average classification performance of different MSVMs in 50 runs. It can been see that PSO-MSVM obtains the most steady and accurate classification performance in classification of arrhythmia.

TABLE I. COMPARISON OFTHE ACCURACY OF DIFFERENT MULTI-CLASS SUPPORT VECTOR MACHINES

ECG	Qty.	Linear	Quadratic Polynomial	RBF	PSO-RBF
NOR	77	95.83%	95.83%	94.11%	95.83%
LBBB	108	100.00%	100.00%	100.00%	100.00%
RBBB	116	0.00%	100.00%	100.00%	100.00%
APC	129	100.00%	100.00%	100.00%	100.00%
VPC	104	94.11%	64.71%	94.11%	100.00%
Overall	534	77.50 %	91.17%	98.82%	99.17%

Fig. 3 shows the classification accuracy along the number of function evaluations in the 10 runs of the PSO. The mean accuracy become stable after about 18 iterations. The optimum accuracy value converges to 99.1758%, and the parameters C and g involved in MSVM converges to 66.1902 and 30.5375, respectively. This is as expected, because PSO possesses an extraordinarily faster convergence rate. It can also be interpreted as PSO consuming short time to yield high accuracy classifications.

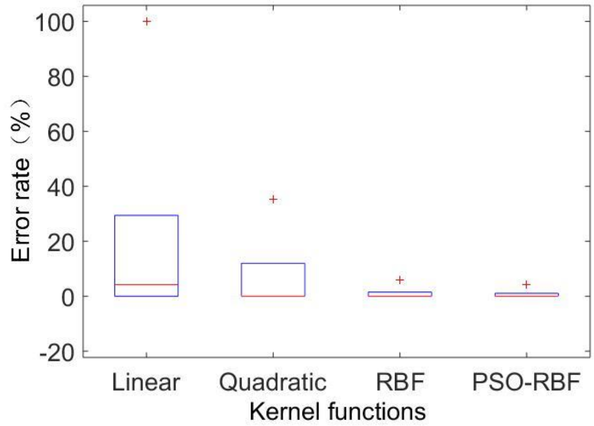

Fig. 2 Statistical analysis of the accuracy of different multi-class support vector machine.

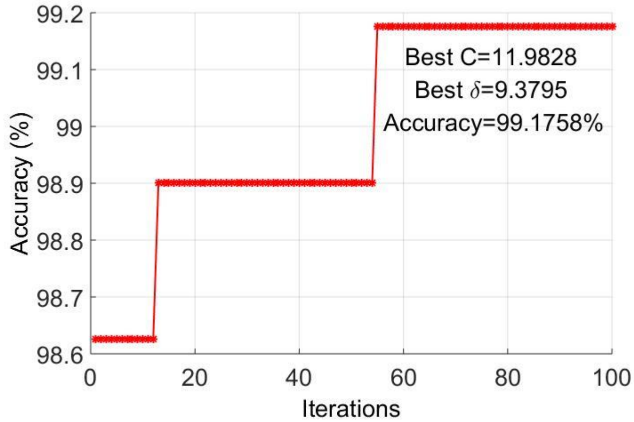

Fig. 3 Classification accuracy along the number of function evaluations in the 10 runs of the PSO.

IV. CONCLUSION

In this paper, five features of ECG have been grabbed from MIT-BIH database and a MSVM based classification method has been proposed to distinguish normal heartbeat from LBBB, RBBB, APC, VPC. Besides, the four M-SMVs with different kernel functions have been evaluated. Simulation results show the PSO-MSVM with RBF kernel function has made the highest accuracy, achieving the overall average accuracy of rate 99.1758%. The proposed method could be efficient tool which has intensive applications in early diagnosis and mass screening of cardiac health.

ACKNOWLEDGMENT

The authors would like to thank Suzhou Electrical Apparatus Science Research Institute Co., Ltd (China) for providing the materials and computer facilities necessary in achieving the results reported. This research is supported by the Natural Science Foundation of Jiangsu Province (Grant No. BK20160355), the Natural Science Research Project of Higher Education of Jiangsu (Grant No. 15KJB480002), and the Science and Technology Project of Ministry of Housing and Urban-Rural Development (Grant No. 2016-K1-019).

ISBN: 978-988-14047-7-0
ISSN: 2078-0958 (Print); ISSN: 2078-0966 (Online)

REFERENCES

[1] G D Clifford, F Azuaje, P McSharry. Advanced Methods and Tools for ECG Data Analysis. Artech House London, 2006.

[2] E Sandoe, B Sigurd. Arrhythmia: A Guide to Clinical Electrocar Diology.Publishing Partners, 1991.

[3] A Khazaee, A Ebrahimzadeh. Heart arrhythmia detection using support

vector machines. Intelligent Automation & Soft Computing, 2013, 19(1): 1-9.

[4] C H Lin. Frequency-domain features for ECG beat discrimination using

grey relational analysis-based classifier. Computers &Mathematics with Applications, 2008, 55(4): 680-690.

[5] E Zeraatkar, S Kermani, A Mehridehnavi, A Aminzadeh, E Zeraatkar, H Sanei. Arrhythmia detection based on morphological and timefrequency features of t-wave in electrocardiogram. Journal of Medical Signals and Sensors, 2011, 1(2): 99-106.

[6] A Ebrahimzadeh, A Khazaee. Detection of premature ventricular contractionsusing MLP neural networks: a comparative study. Measurement,2010, 43(1): 103-112.

[7] T Ince, S Kiranyaz, M Gabbouj. A generic and robust system for automated patient-specific classification of ECG signals. IEEE Transactions on Biomedical Engineering, 2009. 56(5): 1415–1426.

[8] O T Inan, L Giovangrandi, G T Kovacs. Robust neural-network-based

classification of premature ventricular contractions using wavelet transformand timing interval features. IEEE Transactions on BiomedicalEngineering, 2006, 53(12): 2507-2515.

[9] L Y Shyu, Y H Wu, W Hu. Using wavelet transform and fuzzy neuralnetwork for VPC detection from the holter ECG. IEEE Transactions on Biomedical Engineering, 2004, 51(7): 1269–1273.

[10] M Kholkhal, FB Reguig. Efficient automatic detection of QRS complexes in ECG signal based on reverse biorthogonal wavelet decomposition and nonlinear filtering. Measurement, 2016, 94:663-670.

[11] S Yazdani, JM Vesin. Extraction of QRS fiducial points from the ECG using adaptive mathematical morphology. Digital Signal Processing, 2016, 56:100-109.

[12] C J Burges. A tutorial on support vector machines for pattern recognition. Data Mining and Knowledge Discovery, 1998, 2(2): 121–167.

[13] C Cortes, V Vapnik. Support-vector networks. Machine Learning,1995, 20(3): 273–297.

ISBN: 978-988-14047-7-0
ISSN: 2078-0958 (Print); ISSN: 2078-0966 (Online)