Biomed Eng Lett (2014) 4:388-395
DOI 10.1007/s13534-014-0158-7

ORIGINAL ARTICLE

Heartbeat Classification Using Decision Level Fusion

Zhancheng Zhang and Xiaoqing Luo

Received: 13 January 2014 / Revised: 21 October 2014 / Accepted: 28 October 2014
© The Korean Society of Medical & Biological Engineering and Springer 2014

Abstract

Purpose Automatic heartbeat classification is an important technique to assist doctors to identify ectopic heartbeats in long-term ECG recording. In this paper, we employed a multi-lead fused classification schema to improve the performance of heartbeat classification.

Methods In this paper, we introduce a multi-lead fused classification schema, in which a multi-class heartbeat classification task is decomposed into a serials of one-versus-one (OvO) support vector machine (SVM) binary classifiers, then the corresponding OvO binary classifiers of all leads are fused based on the decision score of each binary classifier, the final label is predicted by voting the fused OvO classifiers. The ECG features adopted include inter-beat and intra-beat intervals, amplitude morphology, area morphology, morphological distance and wavelet coefficients. The electrocardiograms (ECG) from the MIT-BIH arrhythmia database (MIT-BIH-AR) are used to evaluate the proposed fusion method. Following the recommendation of the Advancement of Medical Instrumentation (AAMI), all the heartbeat samples of MIT-BIH-AR are grouped into four classes, namely, normal or bundle branch block (N), supraventricular ectopic (S), ventricular ectopic (V) and fusion of ventricular and normal (F). The division of training and testing data complies with the inter-patient schema.

Results Experimental results show that the average classification accuracy of the proposed feature selection method is 87.88%, the sensitivities for the classes N, S, V and F are 88.63%, 74.23%, 88.06% and 73.45% respectively, and the corresponding positive predictive values are 98.54%, 59.76%, 82.33% and 6.96% respectively.

Conclusions The proposed method demonstrates better performance than the existing fusion methods.

Zhancheng Zhang (✉)
College of Electronic and Information Engineering, Suzhou University of Science and Technology, Suzhou, 215009, China
Tel : +86-18952474219
E-mail : cimszhang@163.com

Xiaoqing Luo
School of Internet of Things, Jiangnan University, Wuxi 214122, China

Keywords Heartbeat classification, Support vector machine, Decision fusion, Multi-lead

INTRODUCTION

Electrocardiogram (ECG) is a noninvasive, inexpensive and well-established diagnostic tool. It is widely used to evaluate heart function. However, for the analysis of long-term ECG recording, beat-by-beat manual examination is tedious and time-consuming, especially for junior doctors. Therefore, in the past decades, computer-aided ECG analysis has been a research hotspot.

A typical practice scenario is to automatically identify ectopic heartbeats from a long-term ECG recording. There have been numerous works [1-5], where supervised machine learning methods were generally employed, i.e., learning a classifier on training dataset and evaluation the corresponding classifier on test dataset. Those works can be categorized into "intra-patient" and "inter-patient" classification paradigms. Intra-patient paradigm [1, 2] partitions the whole data set into training and testing subsets based only on the beat label, where an ECG recording may partly appear in both the data subsets. By this scheme, the classifiers usually produce over optimistic results. In clinical practice, the classification performance declines due to the inter-individual variation. With inter-patient paradigm, the training and testing subsets were constructed from different ECG recordings so that inter-individual variation was taken into account and the classifier would exhibit better generalization ability. This paradigm has been adopted [3-6] to evaluate the classification performance on MIT-BIH arrhythmia (MIT-BIH-AR) data set [7].

Most of the studies followed the Advancement of Medical Instrumentation (AAMI) recommendation [8] which specifies heartbeats using five labels, namely, normal or bundle branch block beat (N), supraventricular ectopic beat (S), ventricular ectopic beat (V), fusion of ventricular and normal beat (F), and heartbeats that cannot be classified (Q). This recommendation makes possible a fair comparison among various heartbeat classifiers.

Regarding the ECG features commonly employed for classification, features surrounding RR intervals are most widely used, such as pre-RR, post-RR, local-RR, average-RR [6] and other RR-based features [3]. Other time domain features including PP interval, P duration, QRS duration, PR interval, T duration and QT interval are also considered. Moreover, "morphology" features of ECG samples in P wave, QRS complex and T wave as well as the morphological distances [9] between the beats and the median beat have also been used. These features have been clinically studied and related diagnostic standards have also been stipulated [10]. Although vectorcardiogram (VCG) based features [3, 11] can provide comprehensive information about heart conditions, reconstruction of VCG requires more leads and thus the applicability of these features is rather limited. Besides, frequency domain feature analysis can also provide deep insight into ECG signals. Signal processing methods include wavelet decomposition (WT) [1, 2], principal component analysis (PCA) and independent component analysis (ICA) [4]. Although these features are associated with clear mathematical interpretations, they do not have physiological meaning that allows doctors to comprehend in an intuitively way.

An ECG recording consists of multi-lead signals, and those signals can be considered as observations on the same cardiac activity from different position. The classifiers fusion mechanisms with Bayesian framework [12, 13] and no-Bayesian framework [14] have been well studied. The fusion method with Bayesian framework imposes a class prior probability and conditional probabilities of member classifiers to estimate a joint posterior probability. The method with no-Bayesian framework uses linear combination of member classifiers to improve the performance of classification. In this study, we employ SVM classifier as member classifier and the prior probabilities can be estimated from the output of SVM, therefore we adapt Bayesian based fusion method to fuse multi-lead decisions.

ECG DATA

In this study, the MIT-BIH-AR database is used for training and testing. The database consists of 48 two-lead recordings obtained from 47 subjects (recordings 201 and 202 were obtained from the same subject), each recording was measured for approximately 30 minutes and sampled at 360 Hz. Of the 48 recordings, 23 of them (the "100 series"') were collected from routine ambulatory practice and the remaining 25 (the "200 series") were selected to include examples of uncommon but clinically important arrhythmias cases that were not well represented in the 23 100-series recordings. The ECG leads varied among the subjects due to physical limitation of electrode placement [15]. For most of the recordings, the first channel was a modified limb lead II (MLII) (only 114 recordings used V5 as the first lead and MLII as the second lead; the leads were then swapped in the study). The second channel was usually V1 (sometimes V2, V4 or V5, depending on subjects). The database contains annotations for both QRS position and beat class information, verified by at least two independent experts. The annotated beat types include Left bundle branch block (L), Right bundle branch block (R),

To ensure a fair comparison with the results in related literatures [4, 6] and to comply with the AAMI recommendation, the four recordings with paced beats, namely, 102, 104, 107 and 217, are discarded in the study. Also, all original heartbeat labels are mapped to the AAMI labels with the mapping rules listed in Table 1. We also follow the training set (DS1) and testing set (DS2) division scheme adopted in [6]. Note that the AAMI Q class (unclassified and paced heartbeats) is discarded since this class is marginally represented in MIT-BIH-AR. The numbers of heartbeats in DS1 and DS2 data sets are listed in Table 1 by class.

Table 1. Size of data sets DS1 and DS2 and the mapping between AAMI and MIT-BIH-AR labels.

Data set[1]	N			S					V		F	Q	
	N	L	R	e	j	A	a	J	S	V	E	F	f	Q
DS1[2]	38052	3946	3779	16	16	807	100	32	2	3664	105	414	0	8
		45777				973				3769		414	8	
DS2[3]	36413	4123	3475	0	213	1735	50	51	0	3215	1	388	0	7
		44011				2049				3216		388	7	

[1]Detailed acronyms locate at http://www.physionet.org/physiobank/database/html/mitdbdir/intro.htm#symbols
[2]Recordings in DS1: 101, 106, 108, 109, 112, 114, 115, 116, 118, 119, 122, 124, 201, 203, 205, 207, 208, 209, 215, 220, 223, 230.
[3]Recordings in DS2: 100, 103, 105, 111, 113, 117, 121, 123, 200, 202, 210, 212, 213, 214, 219, 221, 222, 228, 231, 232, 233, 234.

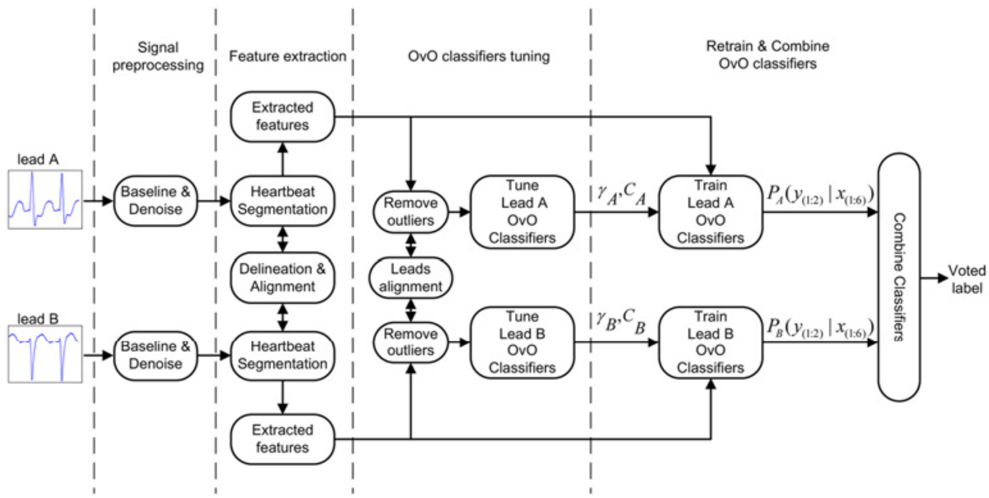

Fig. 1. The fully automatic heartbeat classication procedure.

METHODOLOGIES

As shown in Fig. 1, the whole classification strategy consists of four stages: signal preprocessing, feature extraction, classifier tuning and test.

ECG preprocessing

Since the ECG signals in the MIT-BIH-AR database were collected by Holter devices, the signals were contaminated by baseline wandering and noises from power-line as well as high-frequency electromyography disturbance. Following the approaches adopted in previous work [6, 16], all the ECG signals are first preprocessed using a 200-ms width median filter to remove P wave and QRS complex, then a 600-ms width median filter to remove T wave. The resulted signals are then regarded as the baseline which is subsequently subtracted from the original signals to yield the baseline-corrected ECG signals. A 12-order FIR low-pass filter with a 35Hz cut-off frequency is then used to remove power-line and high-frequency noise. Finally, the filtered two-lead ECG signals are fed into the next stage for further processing.

Heartbeat delineation and segmentation

Before extracting the heartbeat features, the ECG time sequence must be segmented into individual heartbeats based on the positions of the R-peak in QRS complexes. Based on these positions, Other fiducial points including onsets and offsets of P wave, QRS complex and T wave can be delineated. This procedure is called ECG measurement or delineation, and is a key step in automatical heartbeat classification. In the present study, we focus on developing a multi-lead fused heartbeat classifier, thus the existing delineation method are considered. For convenience and to operate well with the WaveForm DataBase (WFDB), we employ the QRS annotation included in MIT-BIH-AR and use the tool "ecgpuwave" (a QRS detector and waveform limit locator available from the *physionet* at http://www.physionet.org/physiotools/ecgpuwave/) to detect the above fiducial points. When P-wave is absent from a heartbeat, the onset and offset of the heartbeat are assigned to the next QRS-onset point. As a result, the heartbeat would have zero P-duration, a truncated PR interval and a truncated post-PP interval. Similarly, when T-wave is absent from a heartbeat, the onset and offset are assigned to the previous QRS-offset, resulting in zero T-duration and a truncated QT-interval.

Feature extraction

Based on the detected fiducial points, as illustrated in Fig. 2, we first extract interval based features and area morphology based features. Then other morphological features including

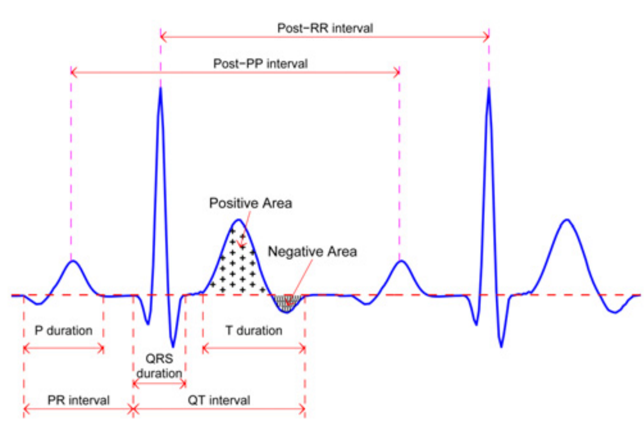

Fig. 2. Illustration of area-based and interval-based feature in ECG: regions lled with + and / sign to represent positive and negative areas.

Table 2. Features used in this study.

Feature group	Indices	Feature Description
Inter-beat intervals	1	Pre-RR interval
	2	Post-RR interval
	3	Local-10 Average RR-interval
	4	Average RR-interval
	5	Post-PP interval
Intra-beat intervals	6	P duration, P offset – P onset
	7	QRS duration, QRS offset QRS onset
	8	T duration, T offset – T onset
	9	PR interval, QRS onset – P onset
	10	QT interval, T offset – QRS onset
Morphological amplitudes	11-20	P morphology, 10 samples between P onset and P offset
	21-30	QRS morphology, 10 samples between QRS onset and QRS offset
	31-39	ST morphology, 9 samples between QRS offset and T offset
Morphological areas	40,41	Positive and negative areas of P wave
	42,43	Positive and negative areas of QRS complex
	44,45	Positive and negative areas of T wave
Morphological distance	46	DTW distance between a beat and the median beat of a recording
	47-143	97 wavelet coefficients

morphological amplitudes, distance and wavelet coefficients are extracted to describe a heartbeat. As a result, a total of 143 features per-beat and per-lead are considered in the study, which are listed in Table 2.

Inter-beat intervals

Five inter-beat intervals are defined as the interval between successive heartbeat fiducial points. Based on the R-peak points, there are four RR related intervals. The *pre-RR interval* is the RR-interval between a given heartbeat and the previous heartbeat. The *post-RR interval* is the RR-interval between a given heartbeat and the following heartbeat. The *local-10 average RR-interval* is determined by averaging the ten RR-intervals surrounding a heartbeat. The *average RR-interval* is the mean of the RR-intervals for a recording and has the same value for all heartbeats in a recording. Finally, based on P-peak points, the *post-RR interval* is defined as the interval between a given heartbeat's R-peak and the following heartbeat's R-peak. The post-RR intervals characterize ventricular periods and the post-PP intervals characterize atrial period. As illustrated in Fig 2, post-RR interval, post-PP interval and intra-beat intervals are interdependent.

Intra-beat intervals

An intra-beat interval is defined as the interval between a posterior fiducial point and an anterior fiducial point in a heartbeat. The five intra-beat intervals used in this study are depicted in Fig. 2 and their related fiducial points are also listed in the corresponding rows and the final column of Table 2. We use these intervals as candidate features because they are well-known and widely used in clinical practice.

Morphological amplitudes

The morphological amplitude features presented in [6] are also adopted in this study. To depict ECG morphology, a group of values are derived by down-sampling the signal amplitude in a specific window. The *P morphology* is defined as physical amplitudes of 10 samples between P onset and P offset. Similarly, the *QRS morphology* and *ST morphology* are defined as physical amplitudes of 10 samples between QRS onset and QRS offset and physical amplitudes of 9 samples between QRS offset and T offset respectively.

Morphological areas

Given the importance of morphological and interval features in automatic heartbeat classification [6, 16], we introduce wave-area-based features in this study to take heartbeat morphology into account for evaluating the effect of feature selection and improving the classification performance. Specifically, as illustrated in Fig. 2, it is clear that some regions are enclosed by the wave intervals, characteristic ECG waves and the baseline. The sum of the area of the regions above baseline is thus defined as a positive area (AreaPos), whereas the area below the baseline is defined as a negative area (AreaNeg). With these enclosed areas as measures, 6 area-based features are obtained, i.e., the AreaPos and AreaNeg of P wave, QRS complex and T wave respectively, which collectively represent the combined characteristics of wave interval and sample amplitude.

Morphological distance

Despite individual variations, normal ECG usually possesses similar morphology between heartbeats of a lead in a recording. To characterize this similarity between heartbeats, Wiens *et al.* [9] present a novel feature based on the dynamic time warping (DTW) distance between a given beat and the median beat of a recording. This feature has excellent discrimination power for ventricular ectopic beats and supraventricular ectopic beats. In this study, ECG signal started at 250 ms left of R-peak and finished at 380 ms right of R-peak within each beat are segmented and down-sampled with 50 Hz. The DTW distance from each segment to the median segment is then considered as the morphological distance.

Morphological wavelet coefficients

ECG is non-stationary signal, traditional Fourier transform (FT) only provides a global characterization of the heartbeats frequency content. In contrast, wavelet transform (WT) provides a characterization in both temporal and frequency domains. In this study, Daubechies wavelets with 8 orders (db8) are adopted to decompose a heartbeat for their similarity with most characteristic QRS waveform. The same heartbeat segments used in the aforementioned morphological distance are decomposed by the four-level db8 wavelet. We use the detail coefficients at level 3 (D3) and 4 (D4) as well as the approximation coefficients at level 4 (A4), these 97 coefficients (28 from A4, 28 from D4, and 41 from D3) are considered as morphological wavelet coefficients.

It is worth noting that the WT-based features used in this paper are different from that used in [4] in which the wavelet coefficients are projected into a PCA space, and the dimension reduced coefficients are used as heartbeat features. The PCA projecting matrix is learned with the unsupervised method, which ignores the annotated label of heartbeat and may lose some useful information. Although the dimension of the original wavelet coefficients is high (94), SVM has the capability to process high-dimensional data by kernel trick.

Outliers removing from training dataset

Although ecgpuwave is a robust tool for detecting QRS and locating waveform limit, it may still fail to detect fiducial points in extremely noisy segments and lead to faulty measures and abnormal features. A beats with any abnormal feature can be considered as an outlier. Removing these outliers will improve the performance of classifier.

In this study, we employe a rule-based procedure to identify and filter the outliers. Here, the interval features are restricted to a reasonable range. Any beats out of the range are considered as outliers. With reference to the physiological function of heart [10], the lower and upper bounds of the related features listed in Table 3 are used to qualify a beat. Note that when an outlier beat of a lead is removed, in order to align with the beats of different leads, the corresponding beats in the other leads are removed manually. Note that the outlier removal procedure is only applied to training dataset in the preprocessing stage. In the final testing procedures to be presented in Section 4, the original beats are used and no outlier is removed.

Table 3. Lower and upper bounds of features for identifying a heartbeat.

Feature name	Lower bound	Upper bound
P duration (ms)	0	200
QRS duration (ms)	0	500
T duration (ms)	0	800
PR interval (ms)	0	400
QT interval (ms)	0	1200
Inter-beat intervals (ms)	20	4000
Morphology amplitude (mV)	-3.5	4

Tuning SVM classifiers

Because of excellent generalization ability, SVM is employed in this study as the heartbeat classifier. This classifier attempts to identify an optimal hyperplane that maximizes the separation margin between two different classes. In this task, suppose a training set consists of N samples $\{(y_i, \mathbf{x}_i), i = 1, ..., N\}$, where $\mathbf{x}_i \in \mathbb{R}^d$ denotes the d-dimensional feature vector of the ith example and $y_i \in \mathbb{R}$ denotes the corresponding class label and $y_i \in \{\pm 1\}$. The optimal hyperplane is represented by a decision function $f(\mathbf{x})$ learned from the training set to predict the class label in the subsequent tests. By using kernel trick, the decision function can be formulated as

$$f(\mathbf{x}) = \mathrm{sign}\left(\sum_{i \in SV_s} \alpha_i y_i K(\mathbf{x}_i, \mathbf{x}) + b\right), \quad (1)$$

where $K(.,.)$ is the kernel function which maps the data into a higher dimensional space, and α_i is the Lagrange multiplier for each training data example. The radial basis function (RBF) kernel $K(\mathbf{x}_i, \mathbf{x}_j) = exp(-\gamma \cdot \|\mathbf{x}_i - \mathbf{x}_j\|^2)$ is usually employed to extend SVM to a nonlinear classifier. A few of α_i's are usually nonzero and the corresponding training data examples are termed as support vectors (SVs). While a number of software packages can be used to learn the decision function, the *libsvm* software package [17] is selected as the bin-classifier solver in the study due to its high efficiency and MATLAB-friendly interface.

As SVM is intrinsically a binary classifier, to extend it for multi-class classification task, specifically for a 4-class heartbeat classification, the one-versus-rest (OvR) and the one-versus-one (OvO) methods are commonly utilized. OvR method worsens the imbalance of class distributions, in contrast, OvO method still maintain the original class distributions and this is convenient for tuning SVM parameters. In this study, we use OvO scheme to decompose a multi-class classification into several binary classification.

We also note from Table 1 that the number of training instances is imbalanced for the four classes. To balance the sensitivities in a binary classifier, the geometric mean (g-means) of the two predicted sensitivities are used to assess the predictor of SVM, which is given by

$$g = \sqrt{Se^{+} \cdot Se^{-}}\,, \tag{2}$$

where Se^{+} and Se^{-} are the predicted sensitivity of the positive and negative classes respectively. This metric has been widely used to deal with imbalanced data sets [18] to take into consideration the classification results on both positive and negative classes.

In each binary classifier, it is necessary to adjust the penalty parameter C and the width parameter γ of the RBF kernel. The optimal parameters are tuned by a 22-fold (leave-one) cross-validation strategy on DS1, where C is searched over the grid $\{2^0, 2^2, ..., 2^{20}\}$ and γ over $\{2^{-10}, 2^{-7}, ..., 2^4\}$. With these parameters, the OvO training procedures are re-executed on the training set to obtain the corresponding SVM binary classifiers. These procedures are individually executed on each lead. Thus, there are 6 binary classifiers on each lead and a total of 12 binary classifiers for the two-lead MIT-BIH-AR. In the following subsection, we will show how to fuse all these binary classifiers and vote for the final prediction class.

Combining classifiers

ECG data usually involve signals from multiple leads, which can be regarded as making observations of the same cardiac activity from different positions (perspectives). Integrating the signals from different leads can enhance the classification accuracy. In this study, we employ the product rule to fuse the same pair of binary classifiers on the two-lead of MIT-BIH-AR. It is based on the Bayesian theory and robust performance has been reported [12]. As discussed previously, since two independent training procedures are executed for the data acquired from each lead, we simply assume the equal prior probability.

Given M classes and L classifiers, the final predicted class y_m can be decided by the average rule and is given by

$$\arg\max_{m} \frac{1}{L}\sum_{l=1}^{L} P(y_m|\kappa_l)\,, \tag{3}$$

where $P(y_m|\kappa_l)$ is the posteriori probability predicted on the class y_m by the classifier κ_l. The fusion procedure is executed over two individual binary classifiers trained on the two leads, i.e., $M = 2$ and $L = 2$.

In a binary classifier, $y_1 = +1$ and $y_2 = -1$ denotes positive and negative class respectively. Given a future observation $\mathbf{x}$, the pairwise posteriori probability $P(y_m, x)$ of $\mathbf{x}$ being positive or negative class can be estimated by the sigmoid function [19]:

$$P(y_m|\kappa_1) = \frac{1}{1+\exp(-y_m f(\mathbf{x}))}\,, \tag{4}$$

where $f(\mathbf{x})$ is the decision value given by (1) which is the distance from $\mathbf{x}$ to the hyperplane.

After a series of fused decisions is made, the bin-label $\{+1, -1\}$ can be obtained to make a prediction on the multi-class problem. Then the majority vote rule below is applied,

$$\arg\max_{c} \sum_{r} \delta_{cr}\,, \tag{5}$$

where δ_{cr} is assigned as 1 when the class c is predicted by the rule r, elsewise it is assigned as 0. Thus $\Sigma_{r=1}^{6} \delta_{cr}$ simply counts the number of votes received for class c from each binary classifier, $c = 1, ..., 4$ (i.e, class N, S, V and F), and $r = 1, ..., 6$.

It is worth noting that the multi-lead fusion scheme is different from that used in [4], in which the OvO binary classifiers are first vote to achieve a multi-class classifier per-lead and the classifiers of the two leads are then fused to get the final decision. In such a way, a biased binary classifier has no chance to be compensated by the other leads, and the voted label may be uncorrect. In this paper, we first fuse the two binary classifiers from the two leads, then the final decision are voted by the fused binary classifiers, this scheme is most resilient to binary classification errors.

RESULTS

We evaluate the performance on the MIT-BIH-AR. The penalty parameters and the width parameters of the RBF kernel in each OvO rule and each lead are independently searched over the grid $\{2^0, 2^2, ..., 2^{20}\}$ and $\{2^{-10}, 2^{-7}, ..., 2^4\}$ respectively. In the searching procedure, the average g-mean of 22-fold (leave-one) cross-validations over DS1 is used an indicator of the optimal parameters.

With the searched parameters, a series of optimal OvO binary classifiers is re-trained on DS1, then the prediction and combination are executed on DS2. Following the AAMI recommendation, the multi-class classification result is evaluated with a confusion matrix to illustrate the performance of the classifiers and the detailed distribution of the misclassified samples. Based on the confusion matrix, the

Table 4. Performance on DS2 of MIT-BIH-AR with reference classifiers and the proposed method.

Ref. [4]		Predicted						Our method		Predicted					
		N	S	V	F	Total	Se (%)			N	S	V	F	Total	Se (%)
True	N	39157	931	1284	2816	44188	88.61	True	N	39009	830	518	3654	44011	88.63
	S	502	1199	252	12	1965	61.02		S	431	1521	78	19	2049	74.23
	V	284	160	2624	139	3207	81.82		V	55	193	2832	136	3216	88.06
	F	199	1	110	76	386	19.69		F	90	1	12	285	388	73.45
Total		40142	2291	4270	3043	49746	62.79	Total		39585	2545	3440	4094	49664	81.09
P^+ (%)		97.55	52.34	61.45	2.50	53.46	86.55	P^+ (%)		98.54	59.76	82.33	6.96	61.90	87.88

Classifier[1]	N		S		V		F		Average			g-mean	
	Se	P^+	Se	P^+	Se	P^+	Se	P^+	Se	P^+	Acc	Se	P^+
Ref. [4]	88.61	97.55	61.02	52.34	81.82	61.45	19.69	2.50	62.79	53.46	84.79	54.33	29.75
Our method	**88.63**	**98.54**	**74.23**	**59.76**	**88.06**	**82.33**	**73.45**	**6.96**	**81.09**	**61.90**	**87.88**	**80.77**	**42.86**

[1]In each metric column, the bold metric denotes the best one among the four methods.

classification performance is measured in term of class sensitivity Se_c and class positive predictive value P_c^+ for class c, which has been formulated in detail in [3] and used to measure inter-patient classification performance across various databases and classifiers. Besides, the averages of accuracy, sensitivity, and positive predictivity, and the g-means of sensitivity and positive predictivity are used to evaluate the performance of classification. Using the same training and testing set division scheme, the reference classifier [4] and our proposed method are compared and the detailed confusion matrices and the derived metrics are reported in Table 4. The results shows that our method outperformed the Ref. [4]. Especially on the positive predictivity of N, in clinical practice, the positive predictive value of class N is a more important metric than the sensitivity of class N, since misdiagnosing a patient as a healthy person may delay therapy and aggravate the illness. For class S, V and F, the high sensitivity the better.

CONCLUSION

We have presented a multi-lead fused method for heartbeat classification. This method decomposes the AAMI four-class classification task in to a series of binary classifier by OvO scheme then these member classifiers are combined over lead and the final labels predicted by voting the fused binary classifiers. The classification diagram fulfils the inter-patient schema and complies with the AAMI recommendation. We have evaluate the method using the MIT-BIH-AR database with the inter-patient schema, based on the inter-beat intervals, intra-beat intervals, morphological amplitudes, morphological areas morphological distances and morphological wavelet coefficients features. With typical DS1/DS2 division, the sensitivities of the classification system are found to be 88.63%, 74.23%, 88.06% and 73.45% respectively for the class N, S, V and F, and the corresponding values of positive predictivitis are 98.54%, 59.76%, 82.33% and 6.96%. In summary, the average sensitivity, positive predictivity and accuracy are 81.09%, 61.90% and 87.88% respectively for the four classes, the geometric means of sensitivities and positive predictivities are 80.77% and 42.86%.

Regarding future work, one important research direction is to develop patient specific classifier by making use of heartbeats in the first five minutes as auxiliary information to improve the classification performance of the long-term recording that follows. Recently, manual expert-assisted approach [16, 20] and unsupervised machine learning methods [21] have been employed for research in this direction. Further investigations will be carried by studying other full-automatic clustering or semi-supervised methods.

ACKNOWLEDGEMENTS

We would like to thank the anonymous reviewers for their diligent work and efficient efforts.

This work was supported by the National Natural Science Foundation of China (61103128), the China Postdoctoral Science Foundation (2013M541601) and the Jiangsu Planned Projects for Postdoctoral Research Funds (1301079C).

CONFLICT OF INTEREST STATEMENTS

Zhang Z declares that he has no conflict of interest in relation to the work in this article. Luo X declares that she has no conflict of interest in relation to the work in this article.

REFERENCES

[1] Martis RJ, Chakraborty C, Ray AK. A two-stage mechanism for registration and classification of ECG using gaussian mixture model. Pattern Recogn. 2009; 42(11):2979-88.

[2] Minhas FU, Arif M. Robust electrocardiogram (ECG) beat classification using discrete wavelet transform. Physiol Meas. 2008; 29(5):555-70.

[3] Llamedo M, Martínez JP. Heartbeat classification using feature selection driven by database generalization criteria. IEEE T Biomed Eng. 2011; 58(3):616-25.

[4] Ye C, Kumar BV, Coimbra MT. Heartbeat classification using morphological and dynamic features of ecg signals. IEEE T Biomed Eng. 2012; 59(10):2930-41.

[5] de Lannoy G, Francois D, Delbeke J, Verleysen M. Weighted conditional random fields for supervised interpatient heartbeat classification. IEEE T Biomed Eng. 2012; 59(1):241-7.

[6] de Chazal P, O'Dwyer M, Reilly RB. Automatic classification of heartbeats using ecg morphology and heartbeat interval features. IEEE T Biomed Eng. 2004; 51(7):1196-206.

[7] Goldberger AL, Amaral LAN, Glass L, Hausdorff JM, Ivanov PC, Mark RG, Mietus JE, Moody GB, Peng C-K, Stanley HE. Physiobank, physiotoolkit, and physionet: components of a new research resource for complex physiologic signals. Circulation. 2000; 101(23):e215-20.

[8] Testing and reporting performance results of cardiac rhythm and ST segment measurement algorithms. AAMI. 2013. https://standards.aami.org/kws/public/projects/project/details?project_id=30. Accessed 18 Nov 2014.

[9] Wiens J, Guttag JV. Active learning applied to patient-adaptive heartbeat classification. Conf Proc Neural Inf Process Syst. 2010; 1:2442-50.

[10] ECGpedia, a free electrocardiography (ECG) tutorial and textbook. http://en.ecgpedia.org/wiki/Main_Page. Accessed 16 Nov 2014.

[11] Christov I, Gómez-Herrero G, Krasteva V, Jekova I, Gotchev A, Egiazarian K. Comparative study of morphological and time-frequency ECG descriptors for heartbeat classification. Med Eng Phys. 2006; 28(9):876-87.

[12] Kittler J, Hatef M, Duin RPW, Matas J. On combining classifiers. IEEE T Pattern Anal Mach Intell. 1998; 20(3):226-39.

[13] Poh N, Kittler J. A unified framework for biometric expert fusion incorporating quality measures. IEEE T Pattern Anal Mach Intell. 2012; 34(1):3-18.

[14] Terrades OR, Valveny E, Tabbone S. Optimal classifier fusion in a non-bayesian probabilistic framework. IEEE T Pattern Anal Mach Intell. 2009; 31(9):1630-44.

[15] Moody GB, Mark RG. The impact of the mit-bih arrhythmia database. IEEE Eng Med Biol Mag. 2001; 20(3):45-50.

[16] de Chazal P, Reilly RB. A patient-adapting heartbeat classifier using ecg morphology and heartbeat interval features. IEEE T Biomed Eng. 2006; 53(12):2535-43.

[17] Chang C-C, Lin C-J. LIBSVM: a library for support vector machines, software. http://www.csie.ntu.edu.tw/-cjlin/libsvm. Accessed 31 Dec 2010.

[18] Wu M, Ye J. A small sphere and large margin approach for novelty detection using training data with outliers. IEEE T Pattern Anal Mach Intell. 2009; 31(11):2088-92.

[19] Wu T-F, Lin C-J, Weng RC. Probability estimates for multi-class classification by pairwise coupling. J Mach Learn Res. 2004; 5:975-1005.

[20] Ince T, Kiranyaz S, Gabbouj M. Ageneric and robust system for automated patient-specific classification of ecg signals. IEEE T Biomed Eng. 2009; 56(5):1415-26.

[21] Llamedo M, Martínez JP. An automatic patient-adapted ecg heartbeat classifier allowing expert assistance. IEEE T Biomed Eng. 2012; 59(8):2312-20.