Computer Methods and Programs in Biomedicine 171 (2019) 1–10

Contents lists available at ScienceDirect

Computer Methods and Programs in Biomedicine

journal homepage: www.elsevier.com/locate/cmpb

A hierarchical method based on weighted extreme gradient boosting in ECG heartbeat classification

Haotian Shi[a], Haoren Wang[a], Yixiang Huang[a], Liqun Zhao[b], Chengjin Qin[a], Chengliang Liu[a,*]

[a] School of Mechanical Engineering, Shanghai Jiao Tong University, 800 Dongchuan Road, Shanghai 200240, PR China
[b] Department of Cardiology, Shanghai First People's Hospital Affiliated to Shanghai Jiao Tong University, 100, Haining Road, Shanghai 200080, PR China

ARTICLE INFO

Article history:
Received 29 December 2018
Revised 3 February 2019
Accepted 9 February 2019

Keywords:
Electrocardiogram (ECG)
Heartbeat classification
Extreme gradient boosting (XGBoost)
Hierarchical classifier

ABSTRACT

Background and objective: Electrocardiogram (ECG) is a useful tool for detecting heart disease. Automated ECG diagnosis allows for heart monitoring on small devices, especially on wearable devices. In order to recognize arrhythmias automatically, accurate classification method for electrocardiogram (ECG) heartbeats was studied in this paper.
Methods: Based on weighted extreme gradient boosting (XGBoost), a hierarchical classification method is proposed. A large number of features from 6 categories are extracted from the preprocessed heartbeats. Then recursive feature elimination is used for selecting features. Afterwards, a hierarchical classifier is constructed in classification stage. The hierarchical classifier is composed of threshold and XGBoost classifiers. And the XGBoost classifiers are improved with weights.
Results: The method was applied to an inter-patient experiment conforming AAMI standard. The obtained sensitivities for normal (N), supraventricular (S), ventricular (V), fusion (F), and Unknown beats (Q) were 92.1%, 91.7%, 95.1%, and 61.6%. Positive predictive values of 99.5%, 46.2%, 88.1%, and 15.2% were also provided for the four classes.
Conclusions: XGBoost was improved and firstly introduced in single heartbeat classification. A comparison showed the effectiveness of the novel method. The method was more suitable for clinical application as both high positive predictive value for N class and high sensitivities for abnormal classes were provided.

1. Introduction

As the aggravation of social aging, more and more attentions have been paid to health monitoring. Heart monitoring is one of the most important aspects in health monitoring. In a cycle of normal heartbeat, electrical impulse emanates from sinus node. Then the electrical impulse gets to atria, atrioventricular node, and ventricle sequentially. Abnormal site of origin, or abnormal conduction of the electrical impulse can lead to abnormal heart rhythms [1]. These conditions are called arrhythmias. Some arrhythmias may cause serious complications and even lead to sudden death [2]. The electrocardiogram (ECG) is a noninvasive and effective tool for detecting arrhythmia. To solve the limitations of visual inspection and prepare for wearable devices, automated classification of heartbeats appears and grows rapidly. Some monitoring frameworks or automated monitoring systems were invented for arrhythmias diagnosis [3,4].

This paper focuses on the classification of ECG heartbeat. It is difficult for accurate classification as the ECG patterns of different individuals were quite different, even for the same heartbeat type. For performance comparison, the Association for the Advancement of Medical Instrumentation (AAMI) [5] proposed a standard. Five classes of heartbeats were recommended: normal (N), supraventricular (S), ventricular (V), fusion (F), and Unknown beats (Q). Although many feature extraction and classification methods have been developed, there were still some problems (see Section 2 for details).

In this study, a recent and efficient algorithm, named extreme gradient boosting (XGBoost), is introduced and improved. XGBoost is derived from the gradient boosting decision tree [6] and proposed by Chen et al. [7]. As an ensemble classifier, XGBoost has excellent performance on generalization. In addition, XGBoost introduces a regularization term to control the complexity of the model, which prevents overfitting of the model. XGboost has provided superior results in many machine learning fields. Hence, the performance of XGBoost on single heartbeat classification is studied.

* Corresponding author.
E-mail address: chlliu@sjtu.edu.cn (C. Liu).

https://doi.org/10.1016/j.cmpb.2019.02.005

The aim of this study is to develop an accurate heartbeat classifier for clinical use. To this end, this paper proposes a hierarchical method based on weighted XGBoost. N, S, V, and F classes are selected and discriminated. Primarily, usable heartbeats are obtained via preprocessing. Afterwards, various kinds of features are extracted. Subsequently, recursive feature elimination is applied for feature selection. Eventually, the feature vectors are inputted into a hierarchical classifier and the predicted labels are obtained. In the hierarchical classifier, heartbeats are first classified into N–S class or V–F class based on XGBoost. Then heartbeats in N–S class are distinguished via a threshold and heartbeats in V–F class are classified also using XGBoost classifier. Considering the unbalance of the ECG dataset, the XGBoost classifier is improved with weights. To verify the effectiveness of the proposed method, an inter-patient experiment conforming AAMI standard is set up. The results are compared with several previous works.

The main novelty of the proposed method is in its application of XGBoost in heartbeat classification. Although XGBoost is well-known, it is first introduced in single heartbeat classification. In addition, a weighted XGBoost classifier is presented for unbalanced heartbeat dataset with multiple classes. The novelty of the proposed method is also explained by the hierarchical classifier. Moreover, recursive feature elimination is employed for feature selection from 6 categories of features.

The remainder of this paper is organized as follows. The next section introduces the background on previous works and analyzes the limitations. Section 3 presents the ECG data used in this work and the data division. In Section 4, preprocessing, features, classifier, and the hierarchical classification method are described in detail. The proposed methods are evaluated via experiment using performance metrics in Section 5. The results are also provided in this section. Section 6 of this paper discusses the results and makes some comparisons with previous works. Finally, Section 7 summarizes all the works and draws conclusions.

2. Background

Various methods have been presented in the literature for automated classification of heartbeats. Support vector machine (SVM) was the most popular method and was used in many researches [8–10]. Other conventional methods employed include linear discriminant (LD) [11,12], K-nearest neighbor (KNN) [13], Gaussian mixture model (GMM) [14], and decision tree (DT) [15]. Classifiers based on artificial neural network (ANN), such as Multilayer perceptrons (MLPs) [16], fuzzy neural networks [17], radial basis networks [18], and convolutional neural network (CNN) [19–21] were also utilized. Other researches have made use of some novel classifiers, such as random forest (RF) [22], optimum-path forest [23], and conditional random fields [24]. The parameters of the classifiers were usually optimized by grid search or genetic algorithm [25,26].

The features extracted from the ECG signal have great influences on the classification result. Besides RR interval, which was used in almost every research, a variety of features or combinations of features have also been reported to describe the heartbeats. Conventional features include interval features, morphological features [27], statistic features [16], higher order statistic features [28], and vectorcardiogram [29]. Features based on Hermite functions [30] and wavelet transform [11] have also been proposed. In addition, some novel features include wavelet packet entropy [31], linear Prediction features [8], abstract features [32], and Auto-Encoder [33]. Meanwhile, there have been several feature selection or dimensionality reduction methods, such as *F*-score [34], mutual information (MI) criterion [24], sequential floating feature selection [11], principal component analysis [8], and genetic algorithm [35].

Table 1
Inter-patient division scheme of MIT-BIH arrhythmia database.

Datasets	Number of heartbeats				
	N	S	V	F	Total
DS1[a]	45,824	943	3788	414	50,969
DS2[b]	44,218	1836	3219	388	49,661
Total	90,042	2779	7007	802	100,630

[a] Recordings in DS1: 101, 106, 108, 109, 112, 114, 115, 116, 118, 119, 122, 124, 201, 203, 205, 207, 108, 109, 215, 220, 223, and 230.
[b] Recordings in DS2: 100, 103, 105, 111, 113, 117, 121, 123, 200, 202, 210, 212, 213, 214, 219, 221, 222, 228, 231, 232, 233, and 234.

However, in some studies [13–28], heartbeats from the same individuals were used in both training and testing process. These works are known as intra-patient classification [36], which can lead to bias of the results [27]. What the model actually learned were the particularities of the patient's heartbeat, rather than the particularities of the diseases. Corresponding to clinical environment, training and test set should contain heartbeats from different individuals, namely, inter-patient classification. In terms of the classes of heartbeats, only part of these methods [11–31] conformed the standards recommended by AAMI and obtained comparable results. In addition, most of the previews studies focused on the classification of N, S, and V classes or the classification of N and V classes. Despite a good overall accuracy, the sensitivities of S class obtained in many studies were not satisfying. All these problems have been considered in this study.

XGBoost has been applied to various machine learning problems. In fault detection of wind turbines [37], radar emitter classification [38], and short-term load forecasting [39], XGBoost gave state-of-the-art results. In biomedical fields, XGBoost has been used in the prediction of essential protein [40] and the classification of patients with epilepsy [41]. Especially, Sodmann et al. [42] trained a convolutional neural network for ECG annotation and employed XGboost to classify atrial fibrillation (AF). Goodfellow et al. [43] also using XGboost for the classification of AF. But these researches were not aimed at single heartbeat classification.

3. Materials

In this study, the proposed method was applied to the Massachusetts Institute of Technology-Beth Israel Hospital (MIT-BIH) arrhythmia database [44]. The database contains over 10,900 heartbeats obtained by the BIH Arrhythmia Laboratory between 1975 and 1979. There are 48 recordings collected from 47 individuals. Each recording, around 30 min are sampled with a frequency of 360 Hz and 11-bit resolution over a 10 mV range. In addition, almost all the heartbeats are annotated by experts. The annotations are generally located at the R-wave peak with sufficient accuracy for study.

Each heartbeat belongs to one of 15 heartbeat types. According to their physiological origin, AAMI proposed a standard, which recommends all the heartbeats to be grouped in 5 classes. In MIT-BIH arrhythmia database, four recordings, which contain paced beats, are removed in this study. For the remaining recordings, an inter-patient division scheme proposed in [27] is used. 44 recordings are divided into two datasets, as summarized in Table 1. The first dataset (DS1), containing 22 recordings, is used for training and determining parameters. The second dataset (DS2) is only used as test set for final performance evaluation. The 44 recordings all contain modified limb lead II (MLII) and only the data of MLII are used in this study.

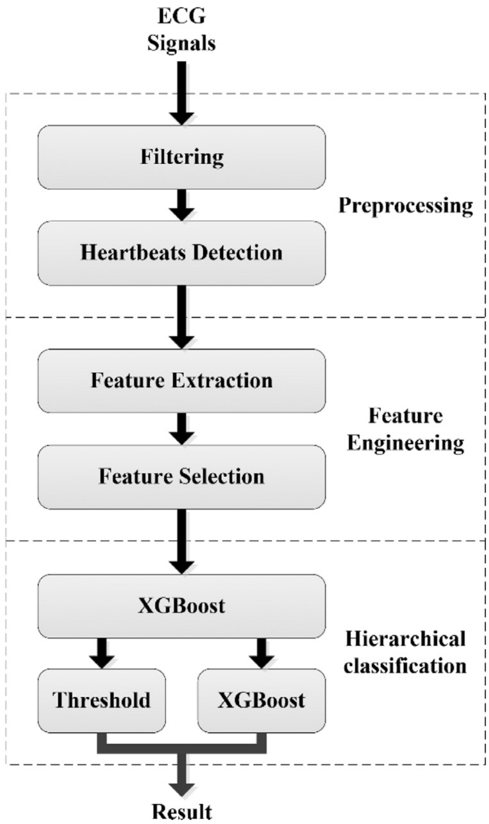

Fig. 1. Flowchart of the proposed method.

4. Methods

As illustrated in Fig. 1, the classification method proposed in this paper is composed of three main stages: preprocessing, feature engineering, and hierarchical classification. In preprocessing stage, the ECG signal is filtered via the wavelet method. Afterwards, all heartbeats are located with fiducial points. In feature engineering, various kinds of features from time domain and time-frequency domain are extracted. Then the features are selected using recursive feature elimination. In classification stage, a hierarchical classifier based on XGBoost classifier and threshold is employed to obtain the final results. The detailed descriptions of each stage are given in this section.

4.1. Preprocessing

The final classification results can be influenced by noises in ECG signal. These noisy signals include power line interference, baseline drift, contact noise, electrosurgical noise, and muscle artifacts. In this study, wavelet denoising method is applied to the raw ECG signal for removing noises and unwanted components. To remove baseline drift, which usually exists in frequency components less than 3 Hz, the raw signal is decomposed into six levels by Daubechies 6 (db6) wavelet. As the sampling frequency is 360 Hz, the frequency range of the sixth level approximation subband is 0 Hz–2.8125 Hz. Therefore, wavelet coefficients of the 6th approximation are replaced with zeros. Meanwhile, the frequency range of the second level detail subband is 45 Hz–90 Hz, which is considered to contain useless information and high-frequency noise. Eventually, only the coefficients between the 3rd and 6th detail subbands are reserved and used for reconstructing denoised ECG signal.

Before feature extraction, the locations of all heartbeats are required. As many heartbeats detection methods have been proposed and proved to be effective, heartbeats detection method is not discussed in this study. The locations of the annotations in MIT-BIH arrhythmia database are taken as the fiducial points. Thus, the locations of heartbeats are obtained.

4.2. Feature extraction

Many configurations of features have been applied in the literature. In order to fully utilize the information hidden in the ECG signal, a variety of features are extracted in this stage. Some of these features were proposed in [13,16,17]. All the features are summarized in Table 2 and described below.

1. *RR interval features:* RR interval is defined as the interval between the fiducial points of two heartbeats. In this study, anterior RR interval, posterior RR interval, average RR interval of a recording, and local average RR interval of the surrounding ten heartbeats are used as features. Moreover, anterior RR interval, posterior RR interval, local RR interval are normalized and taken as additional features, namely, the ratio of them to the average RR interval of the entire recording.
2. *Morphological features:* Morphology features are obtained by uniformly sampling the ECG signal. The sampling windows are located by fiducial points (FP). The signal between FP − 50 ms and FP + 100 ms is considered as QRS wave and sampled by 10 points. The second window between FP + 150 ms and FP + 500 ms is sampled by 8 points.
3. *Statistical features:* This study takes 90 points before and after the fiducial point as a heartbeat. The max value, min value, Max–Min ratio, variance, skewness, and kurtosis are calculated for QRS wave and heartbeat respectively. The peak width of QRS wave at 70% max value and mean of heartbeats are also extracted.
4. *Higher order statistic (HOS):* The cumulants of the 2nd, 3rd, and 4th order are calculated and sampled by 10 equally spaced points. Then, variance, sum of absolute value, number of zero crossing points, and symmetry of the 2nd, 3rd, and 4th order cumulants are extracted. For a signal $x(i)$ of length L, the symmetry is calculated by

$$\text{symmetry} = \sum_{i=1}^{L} |x(i) - x(L+1-i)| / \sum_{i=1}^{L} |x(i)| \tag{1}$$

5. *Wavelet transform features:* The heartbeat is decomposed into five levels using Daubechies 4 (db4) wavelet. Then the wavelet coefficients of the 3nd, 4rd and 5th detail subbands are used respectively to reconstruct the signal components of the three frequency bands. For these three signal components, max value, min value, difference and distance between max value and min value, mean, standard deviation, skewness and relative energy are taken as features.
6. *Wavelet packet entropy:* Also using Daubechies 4 (db4) wavelet, Wavelet packet decomposition (WPD) is applied to the signal with level 6. The wavelet packet entropy is calculated as follows:

$$E_{i,j,k} = \left\| c_{i,j,k} \right\|^2 \tag{2}$$

$$E_{i,j} = \sum_{k=1}^{N} E_{i,j,k} \tag{3}$$

$$P_{i,j,k} = E_{i,j,k} / E_{i,j} \tag{4}$$

$$SE_{i,j} = -\sum_{k=1}^{N} p_{i,j,k} log p_{i,j,k} \tag{5}$$

Table 2
Features extracted in this method.

Categories	Features	Number
RR interval features	Anterior, posterior, local average, average, anterior (normalized), posterior (normalized), local average (normalized)	7
Morphological features	Sampled QRS (10 samples), sampled T (8 samples)	18
Statistical features	Max (QRS, beat), Min (QRS, beat), Max-Min ratio (QRS, beat), variance (QRS, beat), skewness (QRS, beat), kurtosis (QRS, beat), peak width at 70% max (QRS), mean (beat)	14
Higher order statistic	Sampled (10 samples) (2, 3, 4), variance (2, 3, 4), sum of absolute value (2, 3, 4), number of zero crossing points (2, 3, 4), symmetry (3, 4)	41
Wavelet transform features	Max, Min, Max-Min, distance between Max and Min, mean, standard deviation, skewness, relative energy (3,4,5)	24
Wavelet packet entropy	All the nodes in level 6	64

where $c_{i,j,k}$, $E_{i,j,k}$, and $p_{i,j,k}$ are the value, energy, and probability of the kth coefficient of the jth node at ith level. $E_{i,j}$ and $SE_{i,j}$ are the energy and entropy of the jth node at ith level.

4.3. Recursive feature elimination (RFE)

The number of features is too large if all of the 168 features are employed. RFE, a feature selection approach based on wrapper, is used in this study. In the process of RFE, features are ranked by repeatedly training model and removing features with the smallest score. The flow of RFE for a feature set of M features can be described as:

1. Initialize original feature set $S=[f_1, f_2,..., f_M]$ and ranked feature set $R=[]$.
2. If $S \neq []$, Do:
   - Train model using samples with features in S.
   - From the model, calculate the score of each feature: c_k, $k=1, 2, \ldots, M$
   - Find the feature with the smallest score, $p=\text{argmin}(c_k)$.
   - Remove p from S, and update: $R=[p, R]$.

c_k is the score of the kth feature. As support vector machine is usually taken as classifier in RFE, the score is calculated as $c_k = w_k^2$.w_k is the weight of the feature in SVM. Eventually, a ranking of features is obtained. Features with high rankings are selected for classification.

4.4. XGBoost

Similar with gradient boosting, XGBoost [7] combining weak base classifier into a stronger classifier. At each iteration of the training process, the residual of a base classifier is used in the next classifier for optimizing the objective function, as shown in Fig. 2. Suppose the base classifiers are trees with a number of K. For an input sample x_i, the output is calculated by

$$\hat{y}_i = \sum_{i=1}^{K} f_k(x_i),\ f_k \in F \tag{6}$$

where $f_k(x_i)$ is the output of the kth trees and F is the space of all regression trees. Based on gradient boosting, XGBoost makes some improvements by regularizing the objective function:

$$L(\Theta) = \sum_i l\left(y_i, \hat{y}_i\right) + \sum_k \Omega(f_k) \tag{7}$$

where the former term is a loss function that measures the difference between the prediction $\hat{y}_i$ and label y_i. The latter term is a regularization term that measures the complexity of the trees.

The whole objective function cannot be optimized directly. Instead, additive manner is considered. Let $\hat{y}_i^{(t)}$ be the prediction of the ith sample at the tth iteration, the objective function is written

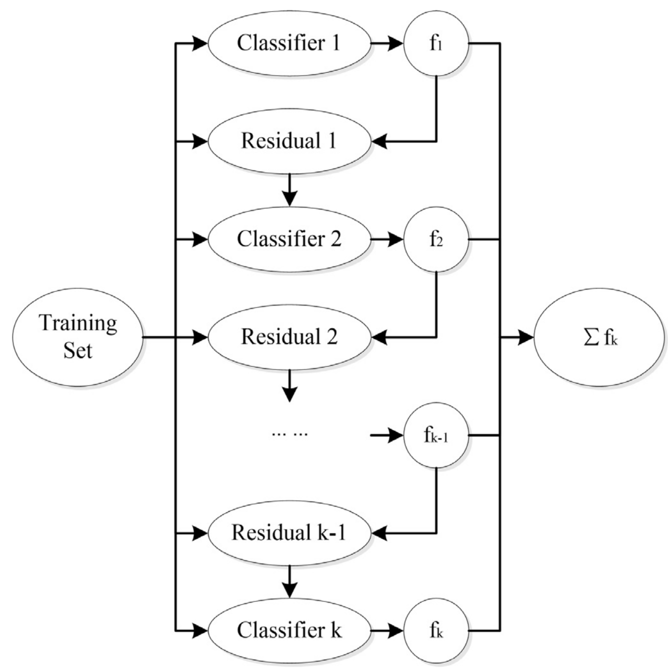

Fig. 2. Structure of extreme gradient boosting.

as:

$$\begin{aligned} L^{(t)} &= \sum_{i=1}^{n} l\left(y_i, \hat{y}_i^{(t-1)} + f_t(x_i)\right) + \Omega(f_t) \\ &\simeq \sum_{i=1}^{n} \left[l\left(y_i, \hat{y}_i^{(t-1)}\right) + g_i f_t(x_i) + \frac{1}{2} h_i f_t^2(x_i) \right] + \Omega(f_t) \end{aligned} \tag{8}$$

where g_i is the first order partial derivative of the loss function and h_i is the second order partial derivative of the loss function. Hence, the loss function must be twice differentiable. The constant terms of Eq. (8) is removed and the objective function is simplified as follows:

$$\tilde{L}^{(t)} = \sum_{i=1}^{n} \left[g_i f_t(x_i) + \frac{1}{2} h_i f_t^2(x_i) \right] + \Omega(f_t) \tag{9}$$

The regularized term is defined by

$$\Omega(f) = \gamma T + \frac{1}{2}\lambda \|\omega\|^2 \tag{10}$$

where T is the number of leaves in the tree. ω is the T-dimension vector of scores on leaves. γ and λ are constant coefficients representing the complexity of leaves and scale of penalty. The space of trees is defined as $F=\{f(x)=\omega_{q(x)}\}$. $q(x)$ is a map assigning the sample to the corresponding leaf. Let the instance set of leaf j is I_j.

Thus Eq. (9) can be expanding as follows:

$$\begin{aligned}\tilde{L}^{(t)} &= \sum_{i=1}^{n}\left[g_i f_t(x_i) + \frac{1}{2}h_i f_t^2(x_i)\right] + \gamma T + \frac{1}{2}\lambda\sum_{j=1}^{T}\omega_j^2 \\ &= \sum_{j=1}^{T}\left[\left(\sum_{i\in I_j} g_i\right)\omega_j + \frac{1}{2}\left(\sum_{i\in I_j} h_i + \lambda\right)\omega_j^2\right] + \gamma T \end{aligned} \tag{11}$$

Eq. (11) can be further compressed by defining $G_j = \sum_{i\in I_j} g_i$ and $H_j = \sum_{i\in I_j} h_i$. Assuming the structure of the tree is fixed, the optimal value of all leaved can be calculated by Eq. (12). And the corresponding value of objective function can be obtained using Eq. (13).

$$\omega_j^* = -\frac{G_j}{H_j + \lambda} \tag{12}$$

$$\tilde{L}^{(t)}(q) = -\frac{1}{2}\sum_{j=1}^{T}\frac{G_j^2}{H_j + \lambda} + \gamma T \tag{13}$$

As the structures of trees can be evaluated, a measurement for the splitting nodes is defined in Eq. (14). Define I_L and I_R as the instance sets after the split.

$$\text{Gain} = \frac{1}{2}\left[\frac{G_L^2}{H_L + \lambda} + \frac{G_R^2}{H_R + \lambda} - \frac{(G_L + G_R)^2}{H_L + H_R + \lambda}\right] - \gamma \tag{14}$$

4.5. Weighted XGBoost classifier for heartbeats

In MIT-BIH arrhythmia database, the number of each class is imbalanced. About 89% of the heartbeats in database belong to N class. The final performance of the model can be distorted by the imbalance. The model after training tends to have a good performance on the classification of N class. In order to reduce the impact of imbalance, weights are introduced in the model.

Weights are taken as coefficients of the first term of objective function. The weight of each class is inversely proportional to the number of samples contained in each class, as calculated by

$$w_i = N/N_{C(i)} \tag{15}$$

where N is the total number of samples in all the class. N_C is the number of samples in class C. $C(i)$ is a map assigning the ith sample to the corresponding label. Hence, the objective can be rewritten as follows:

$$L(\Theta) = \sum_i w_i l(y_i, \hat{y}_i) + \sum_k \Omega(f_k) \tag{16}$$

The classification of heartbeats is a multi-classification problem. Although the problem is converted into binary-class classification in the hierarchical classification below, expanding to multi-class problem is considered. Hence, for the loss function of XGboost classifier in this study, softmax loss is used.

A one-to-many strategy is employed. For a training set with k classes, k trees are trained in each iteration. The output of a tree for the ith sample is defined as $\hat{y}_i = (\hat{y}_{i,1}, \ldots, \hat{y}_{i,k})$, namely, the score of each class. Then $p_{i,j}$, the probability that the ith sample belongs to class j is calculated using softmax function Eq. (17). The real target is written as a k-dimension unit vector $y_i = (y_{i,1}, \ldots, y_{i,k})$. If the real label of the ith sample is j, $y_{i,j}$ is equal to 1. The real target is represented as $1\{y_i = j\}$. Then the loss function is defined in Eq. (18).

$$p_{i,j} = \exp\left(\hat{y}_{i,j}\right) / \sum_{l=1}^{k}\exp\left(\hat{y}_{i,l}\right) \tag{17}$$

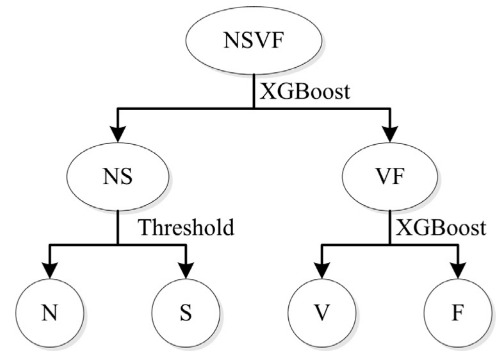

Fig. 3. Construction of the hierarchical classifier.

$$l\left(y_i, \hat{y}_i\right) = -\sum_{j=1}^{k} 1\{y_i = j\} log p_{i,j} \tag{18}$$

As described above, the loss is only depended on the first and second order partial derivatives of the loss function. Using softmax loss, g_i and h_i are calculated as follows:

$$l\left(y_i, \hat{y}_i\right) = -\left[1\{j = y_i\} log p_{i,j} + \sum_{j\neq c} 1\{c = y_i\} log p_{i,c}\right] \tag{19}$$

$$g_i = \frac{\partial l}{\partial \hat{y}_{i,j}} = -\left[1\{j = y_i\}\frac{1}{p_{i,j}}\frac{\partial p_{i,j}}{\partial \hat{y}_{i,j}} + \sum\nolimits_{j\neq c} 1\{c = y_i\}\frac{1}{p_{i,c}}\frac{\partial p_{i,c}}{\partial \hat{y}_{i,j}}\right] \tag{20}$$

$$\begin{aligned} h_i = \frac{\partial^2 l}{\partial \hat{y}_{i,j}^2} &= -\Bigg[1\{j = y_i\}\left[-\frac{1}{{p_{i,j}}^2}\left(\frac{\partial p_{i,j}}{\partial \hat{y}_{i,j}}\right)^2 + \frac{1}{p_{i,j}}\frac{\partial^2 p_{i,j}}{\partial \hat{y}_{i,j}^2}\right] \\ &+ \sum_{j\neq c} 1\{c = y_i\}\left[-\frac{1}{{p_{i,c}}^2}\left(\frac{\partial p_{i,c}}{\partial \hat{y}_{i,j}}\right)^2 + \frac{1}{p_{ic}}\frac{\partial^2 p_{i,c}}{\partial \hat{y}_{i,j}^2}\right]\Bigg] \end{aligned} \tag{21}$$

where $\frac{\partial p_{i,j}}{\partial \hat{y}_{i,j}}$ is calculated by Eq. (22). Other partial derivatives are computed in a similar way.

$$\begin{aligned}\frac{\partial p_{i,j}}{\partial \hat{y}_{i,j}} &= \frac{\exp\left(\hat{y}_{i,j}\right)}{\sum_{l=1}^{k}\exp\left(\hat{y}_{i,l}\right)} - \frac{\exp\left(\hat{y}_{i,j}\right) * \exp\left(\hat{y}_{i,j}\right)}{\left(\sum_{i=1}^{k}\left(\hat{y}_{i,l}\right)\right)^2} \\ &= \frac{\exp\left(\hat{y}_{i,j}\right)}{\sum_{l=1}^{k}\exp\left(\hat{y}_{i,l}\right)}\frac{\sum_{l=1}^{k}\exp\left(\hat{y}_{i,l}\right) - \exp\left(\hat{y}_{i,j}\right)}{\sum_{l=1}^{k}\exp\left(\hat{y}_{i,l}\right)} \\ &= p_{i,j}\left(1 - p_{i,j}\right)\end{aligned} \tag{22}$$

The final result is obtained in Eqs. (23) and (24).

$$g_i = p_{i,j} - 1\{j = y_i\} \tag{23}$$

$$h_i = p_{i,j}\left(1 - p_{i,j}\right) \tag{24}$$

4.6. Hierarchical classification

Among the four classes of heartbeats in this study, S class is difficult to be distinguished, as waveform of S class is similar with that of N class. Considering similarities between N class and S class, or the similarity between V class and F class, different features should be employed in the classification of different types of heartbeats. Hence, a hierarchical classifier is constructed for final classification, as illustrated in Fig. 3. Firstly, all the heartbeats are divided into N–S class and V–F class using an XGBoost classifier.

Table 3
Optimized parameters of XGBoost classifier.

Parameter	Range of grid	Final value	
		XGBoost1[a]	XGBoost2[b]
n_estimators (number of trees)	[500:1000], interval: 50	850	1000
max_depth (maximum depth of a tree)	[3:9], interval: 1	6	6
min_child_weight(minimum sum of the sample weights in a leaf node)	[1:5] interval: 1	2	5
gamma (the minimum loss function degradation required for node splitting)	[0:0.4], interval: 0.1	0.1	0.1
Subsample (the proportion of random sampling)	[0.5:0.9], interval: 0.1	0.9	0.5
colsample_bytree (the proportion of column in random sampling)	[05:0.9], interval: 0.1	0.6	0.8
reg_alpha (L1 regularization term)	[0.001,0.01,0.1,1,2,3,10]	2	0.001
reg_lambda (L2 regularization term)	[0.001,0.01,0.1,1,2,3,10]	3	0.1
eta (learning rate)	[0.01,0.05,0.1,0.2]	0.01	0.1

[a] XGBoost 1 represent the XGBoost for 'N–S' and 'V–F' class.
[b] XGBoost 2 represent the XGBoost for V class and F class.

For the classification of N class and S class, all the features are analyzed separately. The 5th feature, normalized anterior RR interval (the ratio of anterior RR interval to the average RR interval of the entire recording) is found to have good differentiation in classification of N and S. Hence, a threshold of normalized anterior RR interval is used directly. 70% of the heartbeats in S class have normalized anterior RR interval less than 0.75 and 90% of the heartbeats in S class have it less than 0.86. The best threshold is searched between 0.75 and 0.86 on training set. Heartbeats with normalized anterior RR interval which were smaller than the threshold are considered as S class. Then heartbeats in V–F class are also classified using an XGBoost classifier. The two XGBoost classifiers are trained by corresponding data respectively. And features are processed by RFE respectively.

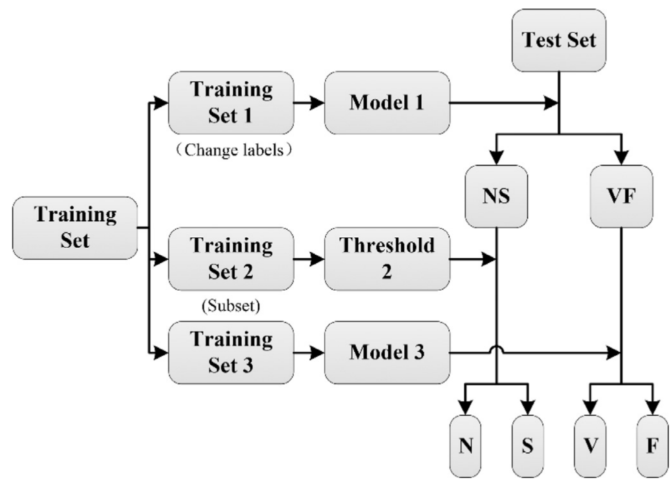

Fig. 4. Training and testing process of the hierarchical classifier.

5. Experiment and results

5.1. Experimental setup

As mentioned earlier, the proposed classification method was applied to MIT-BIH arrhythmia database. The ECG signals in all the recordings were preprocessed by filtering and detection of heartbeats. Afterwards, 168 features of 6 categories were extracted from each recording. Subsequently, the features were selected by RFE according to the corresponding training set, as different training sets were used in different stages of the hierarchical classification. The top 15 features in the ranking were retained. For the XGBoost classifiers in the hierarchical classifier, parameters were selected by grid search using 22-fold cross-validation. In cross-validation, 21 recordings of the training set were used for training. The remaining 1 recording was served as validation data. The optimized parameters of XGBoost, range of grid search, and final values are presented in Table 3. Eventually, classification results were obtained by employing the classifier on the test set.

Specifically, the hierarchical classifier was trained and tested as Fig. 4. In the first stage, the labels of all the training heartbeats were changed to 'N–S' or 'V–F'. They were named training set 1 and an XGBoost classifier was trained. In the second stage, all the training heartbeats of N class and S class were selected to training set 2. The threshold of normalized anterior RR interval was determined according to them and was set to 0.8. Then heartbeats of V class and F class were selected to training set 3 and another XGBoost classifier was trained. Model 1, threshold 2, and model 3 were used to classify the heartbeats in test set into the four classes.

5.2. Evaluating method

In this study, the overall accuracy (*Acc*), sensitivity (*Sen*), and positive predictive value ($+P$) were calculated for evaluating the performance of the classifier. As *Acc* can be distorted by the major class, *Sen* and $+P$ were mainly used for comparison.

$$Sen = TP/(TP + FN) \times 100\% \tag{25}$$

$$+P = TP/(TP + FP) \times 100\% \tag{26}$$

$$Acc = TP + TN/(TP + TN + FP + FN) \times 100\% \tag{27}$$

where TP (true positive) represents the number of correctly detected heartbeats. FP (false positive) represents the number of falsely detected heartbeats. TN (true negative) represents the number of correctly undetected heartbeats. FN (false negative) represents the number of falsely undetected heartbeats. The sensitivity and positive predictive value of each class were calculated respectively.

5.3. Results

Fig. 5 holds the final features for heartbeat classification. For 'N–S' and 'V–F' class, the top 15 features were selected. Fig. 5(a) presents the features and the ranks. For the classification of V and F class, 19 features obtained the ranking of 1. Therefore, these 19 features were all retained, as illustrated in Fig. 5(b). To study the statistical significance of the features, one-way analysis of variance was conducted. The *p*-value is less than 0.001 for all the features used. As an example, the mean values, standard deviation values, and *p*-values of the top five features are shown in Table 4. Fig. 6 illustrates the probability density function (PDF) plot of the top five features. It is evident that the PDF characteristics of the features

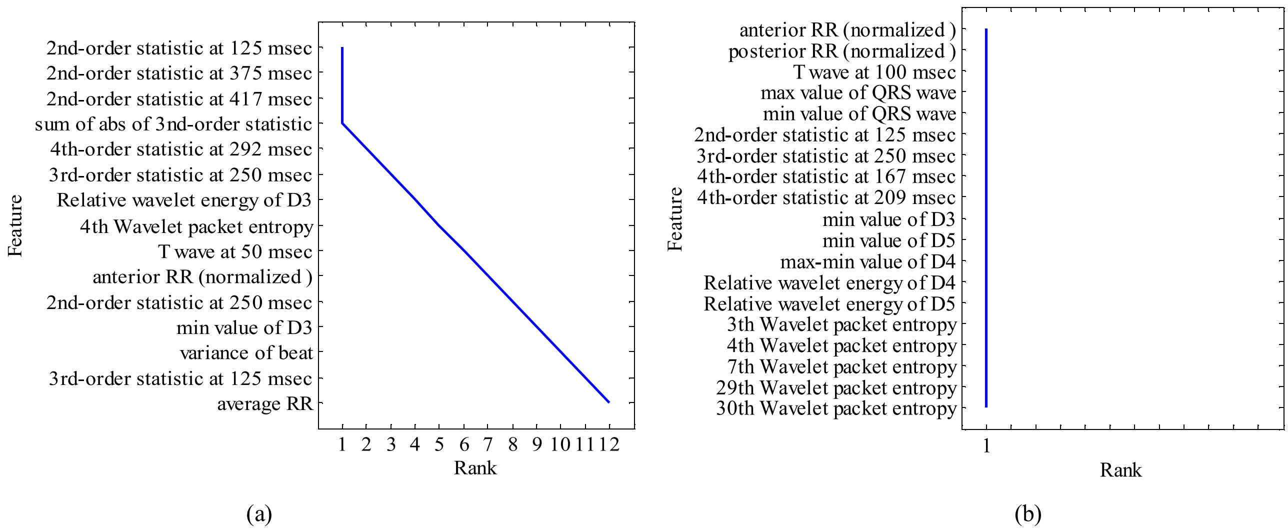

Fig. 5. The final features and ranks – (a) final features for 'N–S' and 'V–F' class, (b) final features for V and F class.

Table 4
Top five features and their statistic values.

Feature rank	Classification of 'N–S' and 'V–F'			Classification of V and F		
	N–S	V–F	*p*-value	V	F	*p*-value
1	−0.008 ± 0.0149	−0.070 ± 0.0849	0	0.7526 ± 0.1928	0.9583 ± 0.0979	<0.001
2	−0.011 ± 0.0158	−0.076 ± 0.0880	0	1.2041 ± 0.3129	0.8117 ± 0.1259	<0.001
3	0.0025 ± 0.0142	−0.032 ± 0.0686	0	0.0677 ± 0.2317	0.22217 ± 0.1482	<0.001
4	0.0028 ± 0.0237	−0.030 ± 0.0801	0	1.0934 ± 0.5649	1.7164 ± 0.3508	<0.001
5	−0.027 ± 0.0445	−0.030 ± 0.1257	<0.001	−0.769 ± 0.3855	−0.511 ± 0.1904	<0.001

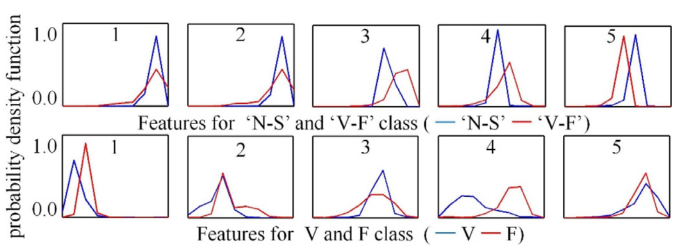

Fig. 6. The probability density function of the top five features.

Table 5
Confusion matrix for the proposed method.

Real labels	Predictions			
	n	s	v	f
N	40,740	1999	295	1295
S	87	1683	65	1
V	46	69	3062	42
F	94	1	54	239

are different for two classes. These indicate the features are statistically significant and discriminative for different classes.

The confusion matrix for the proposed hierarchical classifier is presented in Table 5 The numbers of correctly classified samples of each class were recorded in the diagonal cells. Most of the heartbeats were classified correctly, except for two main mistakes. One mistake was that some heartbeats of N class were misclassified as S class. This was because of the similar waves of these two categories. And the normalized anterior RR intervals of these heartbeats are less than the threshold. Another mistake was that some heartbeats of N class were misclassified as F class. The morphology of N class has more similarities with F class than V class. Hence, for the heartbeats of N class which were classified as 'V–F' in the first layer, most of them were classified as F. Moreover, V class obtained the best performance because of its distinct wide QRS wave.

The trained model was also evaluated via each recording. The accuracies, sensitivities and positive predictive values of all the recordings in DS2 were shown in Table 6. The accuracy represented the correctly classified samples in a recording. The higher sensitivity a category obtained, the more samples of that category were successfully detected. Similarly, lower positive predictive value of a category indicates that more samples were falsely classified to that category. Among the 22 test recordings, 16 recordings obtained accuracies exceeding 95%. An overall accuracy of 92.1% was provided by the proposed method. For N class, the sensitivity was 92.1%, and the positive predictive value was 99.5%. Sensitivity of 91.7% and positive predictive value of 46.2% were provided for S class. The sensitivity was 95.1%, and positive predictive value was 88.1% for V class. F class obtained a sensitivity of 61.6% and a positive predictive value of 15.2%.

6. Discussion

6.1. Performance of the proposed method

The proposed method was applied to an inter-patient experiment conforming AAMI standard. Hence, the experimental conditions were similar with clinical application and the results was convictive. As shown in Table 6, the sensitivities provided by the proposed method was high and balanced. The sensitivities of N, S, and F all exceeded 90%. The positive predictive values were also acceptable, especially 99.5% for N class. Actually, for each class, the improvement of sensitivity means a lower positive predictive value. Considering clinical application, more abnormal heartbeats

Table 6
Classification performance of the proposed method on each recording of test set (DS2).

Record	Number of heartbeats				*Acc*(%)	N (%)		S (%)		V (%)		F (%)	
	N	S	V	F		*Sen*	*+P*	*Sen*	*+P*	*Sen*	*+P*	*Sen*	*+P*
100	2237	33	1	–	99.74	100.00	99.73	81.81	100.00	100.00	100.00	–	–
103	2080	2	–	–	99.38	99.47	99.90	0.00	0.00	–	–	–	–
105	2524	–	41	–	93.06	93.01	100.00	–	0.00	92.68	23.03	–	0.00
111	2121	–	1	–	97.22	97.27	99.95	–	0.00	0.00	0.00	–	–
113	1787	6	–	–	97.16	97.15	100.00	100.00	10.53	–	–	–	–
117	1532	1	–	–	99.74	99.74	100.00	100.00	20.00	–	–	–	–
121	1859	1	1	–	95.33	95.37	100.00	100.00	1.14	0.00	–	–	–
123	1513	–	3	–	98.48	98.48	100.00	–	0.00	100.00	100.00	–	–
200	1742	30	825	2	96.84	99.08	97.68	20.00	13.04	95.15	100.00	0.00	0.00
202	2059	55	19	1	65.28	65.42	99.70	54.54	4.17	84.21	25.40	0.00	–
210	2421	22	195	10	94.64	95.55	99.18	68.18	25.42	92.31	71.43	0.00	0.00
212	2746	–	–	–	100.00	100.00	100.00	–	–	–	–	–	–
213	2639	28	220	362	55.96	52.82	94.70	7.14	100.00	83.63	70.23	65.75	15.73
214	2001	–	256	1	98.76	99.70	99.90	–	0.00	91.80	97.91	0.00	–
219	2080	7	64	1	83.64	83.85	99.43	0.00	0.00	87.50	98.25	0.00	–
221	2029	–	396	–	97.32	97.34	99.80	–	0.00	97.22	100.00	–	–
222	2272	209	–	–	75.82	74.38	99.00	91.39	24.90	–	0.00	–	–
228	1686	3	362	–	99.37	99.64	99.76	33.33	25.00	98.62	98.35	–	–
231	1566	1	2	–	99.94	100.00	99.94	0.00	–	100.00	100.00	–	–
232	397	1381	–	–	99.72	99.75	99.25	99.71	99.93	–	0.00	–	–
233	2229	7	830	11	98.96	99.91	99.51	0.00	0.00	98.43	98.43	9.09	12.50
234	2698	50	3	–	99.13	100.00	99.12	52.00	100.00	100.00	100.00	–	–
Total	44,218	1836	3219	388	92.07	92.13	99.45	91.67	46.22	95.12	88.09	61.60	15.16

are expected to be detected even the cost of misdiagnose increases. A heartbeat classified as normal heartbeat is supposed to really belong to normal class, namely, normal class require high positive predictive value. An abnormal heartbeat is supposed to be classified as abnormal heartbeat, namely, abnormal class require high sensitivity. The result of this study followed this principle. Hence, a high sensitivity of S class was obtained at the cost of a low positive predictive value. For F class, the result was worse than other classes. It was because heartbeats of F class were too less. Although the loss function was weighted, diversity is lacked between these heartbeats. More efforts can be made to improve the performance of F class.

The misclassified samples in the test set were mainly contained in several recordings. In recording 213, 1239 heartbeats of N class were classified as F class, which cause the low positive predictive value of F. In recording 202, 219, and 222, the numbers of normal heartbeats classified as S class were 689, 336, and 576. Similarly, the positive predictive value of S was reduced. Meanwhile, results of these recording reduced the whole sensitivity of N class. Additional attention needs to be paid to these recording in the future.

The proposed method is effective, as XGBoost is an excellent classifier. Moreover, a large number of features are extracted from various categories. Enough diversity exists in these features. A more important reason was that the hierarchical classifier fully utilized these features according to the characteristic of different heartbeat class. In addition, the weighted loss function made the training more effective. In recent years, deep learning (DL) models were introduced in ECG analysis and developed rapidly, as DL can learn implicit features automatically. Compared with DL, although feature extraction is necessary, this study focuses on the performance of XGBoost with extracted features. On the other hand, the proposed method does not require as much data as deep learning.

6.2. Contrast experiments

Several contrast experiments were added in this part, as presented in Table 7. The results of models with and without weights were compared. The evaluating indexes of method without weights were all lower than weight-based results, which demonstrated the effectiveness of weights. Furthermore, three novel entropies, e.g. bubble entropy [45], dispersion entropy [46], and state space correlation entropy [47] were used for comparing with wavelet package entropy. The new entropy, also calculated from each signal produced by WPD, only replaced 64 wavelet package entropy features. Higher accuracy and sensitivity of N class were provided by these entropy measures. But the sensitivity of V class and the performance on F class were poor. The positive predictive value of N was also decreased. To sum up, wavelet package entropy was more suitable in ECG heartbeat classification compared with the three novel entropy measures.

6.3. Comparison with previous works

The performance of the proposed method was compared with several previous works, as presented in Table 8. The performance of the proposed method was notably better than [16] and [27]. Although [9] and [31] achieved high overall accuracies, they can hardly distinguish S class. And high overall accuracy was caused by the high accuracy of the major class. For [29], the sensitivities of N class and F class were superior to this study. However, the sensitivity of S class was much lower than the proposed method. Moreover, compared with these results, the proposed method provided both high positive predictive value for N class and high sensitivities for abnormal classes.

6.4. Experimental environment and computation time

This work utilized Python together with the scikit-learn library and the xgboost library. The hardware used for training the model is a GPU server equipped with an Intel Xeon E5-2620 CPU and two NVIDIA GTX 1080 GPU. The dynamic memory of the computer is 128GB. For XGBoost, the computational complexity is $O(Kd||x||_0+||x||_0\log n)$. Here K is the total number of trees, d is the maximum depth of the tree, and $||x||_0$ is number of non-missing entries in the training data. Actually, the training cost 18.7761 s and the test cost 0.0350 s. The average training and test time for each heartbeat were 0.3684 ms and 0.7043 µs. Hence, real-time diagnosis can be accomplished in clinical application.

Table 7
Results of the contrast experiments.

Methods	Acc(%)	N(%)		S(%)		V(%)		F(%)	
		Sen	+P	Sen	+P	Sen	+P	Sen	+P
Original method	92.1	92.1	99.5	91.7	46.2	95.1	88.1	61.6	15.2
Method without weights	91.6	91.7	99.4	91.6	46.4	94.7	86.5	57.2	12.9
Using bubble entropy	93.5	95.4	98.2	93.1	43.2	78.4	94.9	4.6	30.5
Using dispersion entropy	93.8	95.4	98.4	93.3	42.7	81.8	95.9	8.3	27.6
Using state space correlation entropy	93.8	95.4	98.4	93.3	43.7	82.0	95.8	8.0	29.5

Table 8
Comparisons of results between the proposed method and previous works.

Methods	Acc(%)	N(%)		S(%)		V(%)		F(%)	
		Sen	+P	Sen	+P	Sen	+P	Sen	+P
Chazal [27]	85.8	86.8	99.1	75.9	38.5	77.7	81.9	89.4	8.6
Mar [16]	89.0	89.6	99.1	83.2	33.5	86.8	75.9	61.1	16.6
Li [31]	94.6	94.7	99.7	20.0	0.16	94.2	89.8	50.0	0.5
Kiranyaz [19]	95.1	97.1	98.0	64.6	62.1	95.0	89.5	76.1	80.6
Chen [9]	93.1	98.4	95.4	29.5	38.4	70.8	85.1	0.0	0.0
Proposed	92.1	92.1	99.5	91.7	46.2	95.1	88.1	61.6	15.2

7. Conclusion

In this study, a novel hierarchical method based on weighted extreme gradient boosting was proposed for ECG heartbeat classification. 168 features of 6 categories were extracted and RFE was conducted for feature selection. The hierarchical classifier consisted of two layers. In the first layer, heartbeats were divided into 'N–S' class or 'V–F' class using an XGBoost classifier. In the second layer, threshold classifier and another XGBoost classifier were employed respectively for further classification. Applied to an inter-patient experiment conforming AAMI standard, the proposed method provided sensitivities of 92.1%, 91.7%, 95.1%, and 61.6% for N, S, V, and F classes. The positive predictive value for these four classes were 99.5%, 46.2%, 88.1%, and 15.2%.

To the best of our knowledge, it is the first time that the XGBoost classifier is applied to the classification of single heartbeats. And the hierarchical classifier fully utilizes the extracted features. Moreover, the XGBoost classifier was weighted to reduce the impact of imbalance in the dataset. Compared with previous works, higher positive predictive value for N class and higher sensitivities for abnormal classes were obtained, which were more suitable for clinical applications. The comparison demonstrated that the results of ECG heartbeats classification were improved by the proposed method. However, the positive predictive value for S class and the performance for V class need further improvement. Future work will focus on solving the limitations: (1) study the improvement of positive predictive values for abnormal classes. (2) collect and use more heartbeats for the rare classes.

Conflict of interest

None.

Acknowledgements

This study is supported by the National Key Technology R&D Program of China (No. 2017YFB1302004).

References

[1] G.J. Taylor, 150 Practice ECGs: Interpretation and Review, Blackwell, Oxford U.K., 2002.
[2] H.V. Huikuri, A. Castellanos, R.J. Myerburg, Sudden death due to cardiac arrhythmias, N. Engl. J. Med. 345 (2001) 1473–1482, doi:10.1056/NEJM200203213461218.
[3] H. Mora, D. Gil, R.M. Terol, J. Azorín, J. Szymanski, An IoT-based computational framework for healthcare monitoring in mobile environments, Sensor 17 (2017) 2302, doi:10.3390/s17102302.
[4] H.T. Shi, H.R. Wang, Y.X. Huang, Y.F. Zhang, C.L. Liu, A mobile intelligent ECG monitoring system based on IOS, in: International Conference on Sensing, Diagnostics, Prognostics, and Control, 2017, pp. 149–153, doi:10.1109/SDPC.2017.37.
[5] ANSI/AAMI, Testing and reporting performance results of cardiac rhythm and ST segment measurement algorithms, Association for the Advancement of Medical Instrumentation (AAMI), 2008, American National Standards Institute, Inc. (ANSI), 2008 ANSI/AAMI/ISO EC57, 1998-(R).
[6] J.H. Friedman, Greedy function approximation: a gradient boosting machine, Ann. Statist. 29 (2001) 1189–1232, doi:10.2307/2699986.
[7] T. Chen, C. Guestrin, XGBoost: a scalable tree boosting system, in: The International Conference on Knowledge Discovery and Data Mining, 2016, pp. 785–794, doi:10.1145/2939672.2939785.
[8] R.J. Martis, U.R. Acharya, K.M. Mandana, A.K. Ray, C. Chakraborty, Application of principal component analysis to ECG signals for automated diagnosis of cardiac health, Expert Syst. Appl. 39 (2012) 11792–11800, doi:10.1016/j.eswa.2012.04.072.
[9] S. Chen, W. Hua, Z. Li, J. Li, X. Gao, Heartbeat classification using projected and dynamic features of ECG signal, Biomed. Signal Process. Control 31 (2017) 165–173, doi:10.1016/j.bspc.2016.07.010.
[10] H. Huang, J. Liu, Q. Zhu, R. Wang, G. Hu, A new hierarchical method for inter-patient heartbeat classification using random projections and RR intervals, Biomed. Eng. Online 13 (2014) 1–26, doi:10.1186/1475-925X-13-90.
[11] M. Llamedo, J.P. Martínez, Heartbeat classification using feature selection driven by database generalization criteria, IEEE Trans. Biomed. Eng. 58 (2011) 616–625, doi:10.1109/TBME.2010.2068048.
[12] Y.C. Yeh, W.J. Wang, C.W. Chiou, Cardiac arrhythmia diagnosis method using linear discriminant analysis on ECG signals, Meas. J. Int. Meas. Confed. 42 (2009) 778–789, doi:10.1016/j.measurement.2009.01.004.
[13] S. Faziludeen, P. Sankaran, ECG beat classification using evidential K -nearest neighbours, Procedia Comput. Sci. 89 (2016) 499–505, doi:10.1016/j.procs.2016.06.106.
[14] R.J. Martis, U.R. Acharya, H. Prasad, C.K. Chua, C.M. Lim, Automated detection of atrial fibrillation using Bayesian paradigm, Knowledge-Based Syst. 54 (2013) 269–275, doi:10.1016/j.knosys.2013.09.016.
[15] R.J. Martis, U.R. Acharya, H. Adeli, H. Prasad, J.H. Tan, K.C. Chua, Computer aided diagnosis of atrial arrhythmia using dimensionality reduction methods on transform domain representation, Biomed. Signal Process. Control 13 (2014) 295–305, doi:10.1016/j.bspc.2014.04.001.
[16] T. Mar, S. Zaunseder, J.P. Martnez, M. Llamedo, R. Poll, Optimization of ECG classification by means of feature selection, IEEE Trans. Biomed. Eng. 58 (2011) 2168–2177, doi:10.1109/TBME.2011.2113395.
[17] L.Y. Shyu, Y.H. Wu, W. Hu, Using wavelet transform and fuzzy neural network for VPC detection from the Holter ECG, IEEE Trans. Biomed. Eng. 51 (2004) 1269–1273, doi:10.1109/tbme.2004.824131.
[18] A.S. Al-Fahoum, I. Howitt, Combined wavelet transformation and radial basis neural networks for classifying life-threatening cardiac arrhythmias, Med. Biol. Eng. Comput. 37 (1999) 566–573, doi:10.1007/bf02513350.
[19] S. Kiranyaz, T. Ince, M. Gabbouj, Real-time patient-specific ECG classification by 1-d convolutional neural networks, IEEE Trans. Biomed. Eng. 63 (2016) 664–675, doi:10.1109/TBME.2015.2468589.
[20] O. Faust, Y. Hagiwara, T.J. Hong, O.S. Lih, U.R. Acharya, Deep learning for healthcare applications based on physiological signals: a review, Comput. Methods Programs Biomed 161 (2018) 1–13, doi:10.1016/j.cmpb.2018.04.005.

[21] O. Yıldırım, P. Pławiak, R. Tan, U.R. Acharya, Arrhythmia detection using deep convolutional neural network with long duration ECG signals, Comput. Biol. Med. 102 (2018) 411–420, doi:10.1016/j.compbiomed.2018.09.009.

[22] R. Tan, M. Perkowski, Toward improving electrocardiogram (ECG) biometric verification using mobile sensors: a two-stage classifier approach, Sensor 17 (2017) 410, doi:10.3390/s17020410.

[23] E. Luz, T.M. Nunes, V.H.C. De Albuquerque, J.P. Papa, D. Menotti, ECG arrhythmia classification based on optimum-path forest, Expert Syst. Appl 40 (2013) 3561–3573, doi:10.1016/j.eswa.2012.12.063.

[24] G. Lannoy, D. Francois, J. Delbeke, M. Verleysen, Weighted conditional random fields for supervised interpatient heartbeat classification, IEEE Trans. Biomed. Eng. 59 (2012) 241–247, doi:10.1109/TBME.2011.2171037.

[25] P. Pławiak, Novel genetic ensembles of classifiers applied to myocardium dysfunction recognition based on ECG signals, SWARM EVOL. COMPUT 39 (2018) 192–208, doi:10.1016/j.swevo.2017.10.002.

[26] P. Pławiak, U.R. Acharya, Novel Deep Genetic Ensemble of Classifiers for Arrhythmia Detection Using ECG Signals, Neural Computing and Applications, 2018, doi:10.1007/s00521-018-03980-2.

[27] P. deChazal, M. O'Dwyer, R.B. Reilly, Automatic classification of heartbeats using ECG morphology and heartbeat interval features, IEEE Trans. Biomed. Eng. 51 (2004) 1196–1206, doi:10.1109/tbme.2004.827359.

[28] F.A. Elhaj, N. Salim, A.R. Harris, T.T. Swee, T. Ahmed, Arrhythmia recognition and classification using combined linear and nonlinear features of ECG signals, Comput. Methods Programs Biomed 127 (2016) 52–63, doi:10.1016/j.cmpb.2015.12.024.

[29] M. Soria, J. Martínez, Analysis of multidomain features for ECG classification, The Computers in Cardiology Conference 2009 561–564.

[30] G.D. Lannoy, D. François, J. Delbeke, M. Verleysen, Weighted SVMs and feature relevance assessment in supervised heart beat classification, Commun. Comput. Inf. Sci. 127 (2011) 212–223, doi:10.1007/978-3-642-18472-7_17.

[31] T. Li, M. Zhou, ECG classification using wavelet packet entropy and random forests, Entropy 18 (2016) 1–16, doi:10.3390/e18080285.

[32] T. Teijeiro, P. Félix, J. Presedo, D. Castro, Heartbeat classification using abstract features from the abductive interpretation of the ECG, IEEE J. Biomed. Health Inform. 22 (2018) 409–420, doi:10.1109/JBHI.2016.2631247.

[33] L. Zaoralek, J. Platos, V. Snasel, Patient-adapted and inter-patient ECG classification using neural network and gradient boosting, Neural Network World 28 (2018) 241–254, doi:10.14311/NNW.2018.28.015.

[34] Z. Zhang, J. Dong, X. Luo, K.S. Choi, X. Wu, Heartbeat classification using disease-specific feature selection, Comput. Biol. Med. 46 (2014) 79–89, doi:10.1016/j.compbiomed.2013.11.019.

[35] P. Pławiak, Novel methodology of cardiac health recognition based on ECG signals and evolutionary-neural system, Expert Syst. Appl. (2017), doi:10.1016/j.eswa.2017.09.022.

[36] E. Luz, W.R Schwartz, G. Cámara-Chávez, D. Menotti, ECG-based heartbeat classification for arrhythmia detection: a survey, Comput. Methods Programs Biomed. 127 (2016) 144–164, doi:10.1016/j.cmpb.2015.12.008.

[37] D. Zhang, L. Qian, B. Mao, C. Huang, Y. Si, A data-driven design for fault detection of wind turbines using random forests and XGboost, IEEE Access 6 (2018) 1–1, doi:10.1109/ACCESS.2018.2818678.

[38] W. Chen, K. Fu, J. Zuo, X. Zheng, T. Huang, W. Ren, Radar emitter classification for large data set based on weighted-xgboost, IET Radar Sonar. Navig. 11 (2017) 1203–1207, doi:10.1049/iet-rsn.2016.0632.

[39] H. Zheng, J. Yuan, L. Chen, Short-term load forecasting using EMD-LSTM neural networks with a xgboost algorithm for feature importance evaluation, Energies 10 (2017) 1168, doi:10.3390/en10081168.

[40] J. Zhong, Y. Sun, W. Peng, M. Xie, J. Yang, X. Tang, XGBFEMF: an XGBoost-based framework for essential protein prediction, IEEE Trans. Nanobiosci. 17 (2018) 243–250, doi:10.1109/TNB.2018.2842219.

[41] L. Torlay, M. Perrone-Bertolotti, E. Thomas, M. Baciu, Machine learning–XGBoost analysis of language networks to classify patients with epilepsy, Brain Inform. 4 (2017) 159–169, doi:10.1007/s40708-017-0065-7.

[42] P. Sodmann, M. Vollmer, N. Nath, Lars Kaderali1, A convolutional neural network for ECG annotation as the basis for classification of cardiac rhythms, Physiol. Meas 39 (2018) 14, doi:10.1088/1361-6579/aae304.

[43] S.D. Goodfellow, A. Goodwin, R. Greer, P.C. Laussen, M. Mazwi1, D. Eytan, Classification of atrial fibrillation using multidisciplinary features and gradient boosting, in: Computing in Cardiology Conference, 2017, pp. 1–4, doi:10.22489/CinC.2017.361-352.

[44] A.L. Goldberger, L.A.N. Amaral, L. Glass, J.M. Hausdorff, P.C. Ivanov, R.G. Mark, J.E. Mietus, G.B. Moody, C.-K. Peng, H.E. Stanley, Physiobank, physiotoolkit, and physionet: components of a new research resource for complex physiologic signals, Circulation 101 (2000) 215–220. database and tools available at:http://www.physionet.org/.

[45] G. Manis, M. Aktaruzzaman, R. Sassi, Bubble entropy: an entropy almost free of parameters, IEEE Trans. Biomed. Eng. 64 (2017) 2711–2718, doi:10.1109/TBME.2017.2664105.

[46] M. Rostaghi, H. Azami, Dispersion entropy: a measure for time-series analysis, IEEE Signal Process. Lett. 23 (2016) 610–614, doi:10.1109/LSP.2016.2542881.

[47] R.K. Tripathy, S. Deb, S. Fandapat, Analysis of physiological signals using state space correlation entropy, IET Healthc. Technol. Lett. 4 (2017) 30–33, doi:10.1049/htl.2016.0065.