ECG-Based Heartbeat Classification Using Two-Level Convolutional Neural Network and RR Interval Difference

Yande XIANG†, Jiahui LUO††, Taotao ZHU††, Sheng WANG††, Nonmembers, Xiaoyan XIANG†††, Member, and Jianyi MENG††††, Nonmember

SUMMARY Arrhythmia classification based on electrocardiogram (ECG) is crucial in automatic cardiovascular disease diagnosis. The classification methods used in the current practice largely depend on hand-crafted manual features. However, extracting hand-crafted manual features may introduce significant computational complexity, especially in the transform domains. In this study, an accurate method for patient-specific ECG beat classification is proposed, which adopts morphological features and timing information. As to the morphological features of heartbeat, an attention-based two-level 1-D CNN is incorporated in the proposed method to extract different grained features automatically by focusing on various parts of a heartbeat. As to the timing information, the difference between previous and post RR intervals is computed as a dynamic feature. Both the extracted morphological features and the interval difference are used by multi-layer perceptron (MLP) for classifying ECG signals. In addition, to reduce memory storage of ECG data and denoise to some extent, an adaptive heartbeat normalization technique is adopted which includes amplitude unification, resolution modification, and signal difference. Based on the MIT-BIH arrhythmia database, the proposed classification method achieved sensitivity \( Sen = 93.4\% \) and positive predictivity \( Ppr = 94.9\% \) in ventricular ectopic beat (VEB) detection, sensitivity \( Sen = 86.3\% \) and positive predictivity \( Ppr = 80.0\% \) in supraventricular ectopic beat (SVEB) detection, and overall accuracy \( OA = 97.8\% \) under 6-bit ECG signal resolution. Compared with the state-of-the-art automatic ECG classification methods, these results show that the proposed method acquires comparable accuracy of heartbeat classification though ECG signals are represented by lower resolution.

key words: electrocardiogram (ECG), beat classification, convolutional neural network (CNN), biomedical signal processing, multi-layer perceptron (MLP)

1. Introduction

There is a large number of people suffering from cardiovascular diseases among which is arrhythmia. There are various types of arrhythmias and each type has its own characteristics in the morphology or wave frequency. Even for the Electrocardiogram (ECG) signals of a healthy subject, the shape of P waves, QRS complexes, T waves, and heart rates are various under different times and circumstances. Generally speaking, there are two kinds of arrhythmias. One is life-threatening which contains ventricular fibrillation and tachycardia. The other is not life-threatening but needs to be prevented from further problems. The ECG provides an analytical tool for detecting disorder of rhythm and change in the morphological pattern. The heartbeat classification is an important step to identify an arrhythmia. In addition, real-time and long-term health monitoring and analysis about ECG are vital to diagnose cardiac disease, improve patient safety, and enhance nursing efficiency. Hence, the light-weight wearable devices and the wireless sensor networks (WSNs) have been invested. As to the light-weight wearable devices and the WSNs, battery-powered wireless sensors are usually used to capture the ECG signals and transmit essential signals to remote telecardiology center for further analysis. Generally, the sensors should be implemented with ECG data analysis, compression, storage, and transmission [1]. Therefore, it is critical to improve the efficiency of automatic heartbeat classification under low-memory-storage implementation.

In recent years, many approaches on automatic ECG classification have been proposed. The types of heartbeats could be distinguished from time domain [2], [3], frequency domain [4], projected-domain [5], wavelet transform [6], [7], stockwell transform (ST) [8], Hermite basis function [9], and hidden Markov modeling (HMM) [10]. In addition, many machine learning based algorithms have been proposed, such as artificial neural networks (ANNs) [11], block-based neural networks (BBNs) [12], mixture-of-experts approach [13], support vector machine (SVM) [14], [15], and genetic algorithms [16]. Although, the above classification methods present high accuracy in their experimental datasets, their performance largely depends on extracted fixed and hand-crafted manual features. In addition, the ECG patterns are physiological variations due to temporal, personal, or different circumstances. Moreover, extracting hand-crafted features manually may introduce significant computational complexity of overall process, especially in the transform domains.

To avoid extracting fixed and hand-crafted manual features, deep learning frameworks have been developed for heartbeat classification. The learning process is based on representation learning which allows a machine to discover the necessary representations for classification automatically. Kiranyaz et al. [17] presented a 1-D convolutional neural network (CNN) for ECG feature extraction and classification. It used 5 minutes patient-specific ECG signals as well as 245 common heartbeats for training, and classi-
fied beats into five types of heartbeat according to the Association for the Advancement of Medical Instrumentation (AAMI) standard. The method achieved 99.0% ventricular ectopic beats (VEBs) and 97.6% supraventricular ectopic beats (SVEBs) classification accuracy. Rahhal et al. [18] adopted unsupervised learning way by using stacked denoising autoencoders. Then the expert-assist method was adopted by adding uncertain heartbeats from testing data to the training data. Softmax regression layer was added on the top of the hidden layer for active arrhythmia classification. Teijeiro et al. [19] proposed a knowledge-based automatic beat classification method that demanded domain-dependent knowledge base. In Yang et al. [20] research work, a stacked sparse auto-encoders framework was proposed to extract features of ECG arrhythmia data hierarchically. This method prevented the features lost caused by input signal corruption in the denoising auto-encoders. Despite these great efforts, it is still challengeable to improve accuracy of heartbeat classification on the basis of low-memory-storage implementation.

To simplify signal processing as well as combine the advantages of using hand-crafted manual features and non-manual features, this paper adopts attention-based two-level 1-D CNN for extracting morphological features automatically and computing RR interval difference manually as a dynamic feature. It is worth noting that visual attention models have been applied in computer vision problems for fine-grained object detection [21] and fine-grained categorization [22]. The attention model is able to process candidate regions for classification with different resolution and reduce processing cost by focusing on a restricted set of regions. With the help of the attention models, discriminatory power could be focused on the specific parts of the input data, which helps to classify input data [23].

In order to improve accuracy of heartbeat classification in the condition of low memory storage implementation, an adaptive beat normalization technique as well as two-level 1-D CNN are proposed. The beat normalization technique reduces ECG data memory storage as well as denoise to some extent. The two-level 1-D CNN extracts different grained features for final heartbeat classification. The contribution of our work is in the following aspects: (1) the resolution of the ECG signals is normalized and reduced at pre-processing stage; (2) a novel two-level 1-D CNN structure is proposed, which adopts object-level CNN and part-level CNN for different grained feature extraction; (3) the RR interval difference is computed as dynamic feature of ECG arrhythmia instead of using RR interval directly. The performance of the proposed beat classification system was evaluated by using the Massachusetts Institute of Technology and the Boston Hospital (MIT-BIH) arrhythmia database and the St. Petersburg Institute of Cardiological Technics 12-lead Arrhythmia (INCART) database. In addition, to estimate the effect of ECG signal resolution on the capability of classification, we also evaluated the classification accuracy with different resolution of heartbeats.

The rest of the paper is organized as follows. The processing of ECG signals and overview of the proposed approach are presented in Sect. 2. Section 3 describes the details of proposed heartbeat classification method. The experiment results and corresponding analyses are presented in Sect. 4. Section 5 makes a conclusion of the paper.

2. Materials and Overview

2.1 ECG Samples

The ECG signals from the MIT-BIH arrhythmia database and the INCART database were used in this study [24]. We divided the heartbeats of the MIT-BIH arrhythmia database into training and testing parts, and adopted the beats from the INCART database to estimate the robustness of the proposed classification system.

The MIT-BIH arrhythmia database contains 48 ECG recordings from 47 subjects, and each recording was sampled at 360 Hz for 30 minutes with 11-bit resolution over a 10 mV range. The database consists of annotations for both heartbeat class information and R-peak position information verified by two or more expert cardiologists. The 17 beat types can be classified into five beat classes defined by the AAMI which follows the American National Standard for Ambulatory ECGs (ANSI/AAMI EC38:2007) recommendations. The five standard classes include N (beats with sinus mode), S (supraventricular ectopic beats), V (ventricular ectopic beats), F (fusion beats), and Q (unclassifiable beats). The INCART database contains 75 annotated recordings of 12-lead ECG signals. Each of them is 30 minutes long, sampled at 257 Hz, and gained varying from 8-bit to 10-bit resolution over a 1 mV range. To match with the MIT-BIH arrhythmia database, each recording of the INCART database was resampled at 360 Hz. This database contains over 175,000 beats annotations which were detected automatically and then corrected manually. There is no paced beat in the database, and most of beats are VEBs.

2.2 Overview of the Proposed Method

The overview diagram of the proposed method consists of four steps, including heartbeat normalization, RR interval difference computation, morphological features extraction, and heartbeat classification. The overview diagram is shown in Fig. 1. First, an adaptive heartbeat normalization technique is used to normalize each heartbeat. The process of normalization consists of amplitude unification, resolution modification, and signal difference. The adaptive normalization method normalizes various amplitude of beats into same resolution and amplitude. Then, RR interval difference is computed as dynamic feature for indicating timing information of a given heartbeat. In addition, two-level CNN is applied for focusing on different parts of the heartbeat and extracting different grained morphological features. The CNN adopts hierarchical architecture, and each layer consists of 1-D convolution and subsampling processes. The coarse-grained features are extracted by object-level CNN.
and the fine-grained features are extracted by part-level CNN. Both the morphological features and the RR interval difference are fed into MLP for heartbeat classification.

3. Proposed Methods

3.1 Heartbeat Normalization

As mentioned in Sect. 2.1, each raw ECG recording was quantified into 11-bit resolution over a 10 mV range in the MIT-BIH arrhythmia database and varying from 8-bit to 10-bit resolution over a 1 mV range in the INCART database. To eliminate the impact of various amplitudes of heartbeats on the classification performance and reduce the redundance of resolution, each subject is normalized to same amplitude and re-quantified with low resolution. The processes of the amplitude unification and the resolution modification can be formulated as follows:

\[ q_s = \frac{b_{\text{max}} - b_{\text{min}}}{Q_g} \]  
\[ b_q(n) = \begin{cases} \frac{b_{\text{raw}}(n) - b_{\text{min}}}{Q_g}, & b_q(n) \neq b_{\text{max}} \\ 1, & b_q(n) = b_{\text{max}} \end{cases} \]

where \( b_{\text{max}} \) and \( b_{\text{min}} \) are the maximum and minimum ECG data values of each heartbeat respectively. \( Q_g \) is a constant that represents the grain of quantitation. \( b_q \) is quantified step value. \( b_q(n) \) is the raw ECG heartbeat data at time \( n \), and \( b_q(n) \) is the normalized ECG heartbeat data at time \( n \). \( \lfloor \cdot \rfloor \) represents flooring process.

Then a normalized ECG signal \( b_q \) is obtained by making subtraction between adjacent normalized signal data, which is shown as follows:

\[ b_q(n) = b_q(n) - b_q(n - 1) \]

where \( b_q(n) \) is the normalized difference ECG heartbeat data at time \( n \).

3.2 RR Interval Difference

In addition to the morphological features of ECG signals, RR interval is computed as dynamic feature. It is worth noting that the RR interval is not used directly as [5], [25]. Instead, the difference between previous RR interval and post RR interval is adopted. The previous RR interval is the distance between R-peak position of a given heartbeat and its previous one. The post RR interval is the distance between R-peak position of a given heartbeat and its next one. The difference between the the two RR intervals is adopted, which is computed as Eqs. (4)–(7). Only the difference is adopted as a dynamic feature, instead of taking both previous RR interval and post RR interval as two separate features into consideration.

\[ r_{rs} = \frac{r_{\text{post}} - r_{\text{pre}}}{Q_g} \]  
\[ r_{\text{pre}} = r_{\text{cur}} - r_{\text{pre}} \]  
\[ r_{\text{post}} = r_{\text{post}} - r_{\text{cur}} \]  
\[ r_{r_{\text{diff}}} = \frac{r_{r_{\text{pre}}} - r_{r_{\text{post}}}}{r_{rs}} \]

where \( r_{\text{pre}}, r_{\text{cur}}, \) and \( r_{\text{post}} \) are R-peak position of previous beat, current beat, and following beat respectively. \( Q_g \) is a constant that represents the grain of quantitation. \( r_{rs} \) is the quantified distance between R-peak position of previous beat and following beat. \( r_{r_{\text{pre}}} \) is the previous RR interval and \( r_{r_{\text{post}}} \) is the post RR interval. \( r_{r_{\text{diff}}} \) is the difference between previous RR interval and post RR interval.

3.3 Attention-Based Two-Level Feature Extraction and Classification

To extract different grained morphological features from normalized difference ECG signals, an automatic feature extraction system comprised of object-level and part-level 1-D CNNs is proposed, which is illustrated in Fig. 2.
The two-level CNN adopts hierarchical structure, in which different abstract features are extracted from different layers. In the low-level layer, low-level features are extracted through convolution and subsampling computation. Next, the extracted features are propagated to the following hidden layers for extracting higher-level features. In order to extract different level features corresponding to different parts of a heartbeat ECG signal, the number of layers among object-level and part-level CNNs can be implementation defined. All of the features extracted by the two-level CNN are concatenated and sent to MLP for heartbeat classification.

3.3.1 Object-Level CNN

The given R-peak labelled position information was utilized to locate heartbeats. Each beat is represented by 64 samples or 128 samples. For 64 samples, 26 samples at left of R-peak and 37 samples at right of R-peak. For 128 samples, 51 samples at left of R-peak and 76 samples at right of R-peak. The heartbeat composed of the selected 64 samples or 128 samples contains a whole heart cycle including the P wave, QRS complex, and T wave. The comparison between the classification accuracy of heartbeats represented by 64 samples and 128 samples was also taken in experimental section. The object-level 1-D CNN is used to extract coarse-grained features corresponding to a whole heartbeat.

3.3.2 Part-Level CNN

The normalized difference ECG signal is firstly divided into three parts, including P wave, QRS complex, and T wave. In order to get the QRS complex segmentation, a specific number of samples from each side of the R-peak position of the heartbeat are selected. Moreover, P wave and T wave segmentations are represented by samples from the first third part and last third part of whole heartbeat samples respectively. For 64 samples, 20 samples are used to represent P wave and QRS complex respectively, and 24 samples are used to represent T wave. For 128 samples, 40 samples are used to represent P wave and QRS complex respectively, and 48 samples are used to represent T wave. Then, the three parts of a given beat are paid attention and fine-grained features corresponding to each part are extracted.

3.3.3 Training and Classification

The heartbeat classification process is composed of two steps that include training and classification. These two levels of the system are combined for training with backpropagation (BP) scheme. The training process is used to optimize weights and biases, and then the neural network configured with the optimal weights and biases are used to classify heartbeat types. Figure 3 presents the relationships among CNN and MLP layers, and each CNN layer contains convolution and subsampling computation processes. The intermediate value of the \(k\)th neuron at CNN layer \(l\) is computed as Eq. (8), and the output of the \(k\)th neuron at CNN layer \(l\) is computed as Eq. (9). The output of the \(j\)th neuron at MLP layer \(m\) is computed as Eq. (10).

\[
d_k^l = f\left(\sum_{i=1}^{N_l} 1Dconv(w_{ki}^{l-1}, s_i^{l-1}) + b_k^l\right) \\
s_k^l = 1Dsubs(d_k^l) \\
s_j^m = f\left(\sum_{k=1}^{N_l} w_{kj}^l s_k^l + b_j^m\right)
\]
neuron $k$ at layer $l$. $b^l_k$ is the bias value of neuron $k$ at layer $l$. 1Dconv and 1Dsubs represent 1-D convolution and sub-sampling processes respectively. $f(.)$ represents activation function. In this study, the rectified linear unit (Relu) activation function is used in CNN layer as well as MLP hidden layer, and the Softmax activation function is used in MLP output layer for classifying five heartbeat types.

The goal of training is to minimize cost function $E$ by adjusting the kernel weights and biases:

$$E = E(y_1, y_2, \ldots, y_N) = \sum_{j=1}^{N} (t_j - y_j)^2$$

where $N$ is the total neuron number of output layer. For a given input vector $u$, [$y_1, y_2, \ldots, y_N$] and [$t_1, t_2, \ldots, t_N$] are the corresponding predicted output vector and target output vector respectively. The weights and the biases are updated with the learning rate $\eta$ as shown in Eqs. (12) and (13). We set the initial learning rate as $\eta = 0.09$, and slightly decrease it by 0.01% during each training iteration.

$$w^l_k(t) = w^l_k(t-1) - \eta \frac{\partial E}{\partial w^l_k(t-1)} \tag{12}$$

$$b^l_k(t) = b^l_k(t-1) + \eta \frac{\partial E}{\partial b^l_k(t-1)} \tag{13}$$

After the two-level CNN and the MLP are trained, a given normalized difference heartbeat signal and several parts of the signal are sent to the neural network system for heartbeat classification.

4. Experimental Results

4.1 System Configuration

As mentioned earlier, each heartbeat is represented by 64 or 128 samples. For 64 samples, 26 samples at left of R-peak and 37 samples at right of R-peak. For 128 samples, 51 samples at left of R-peak and 76 samples at right of R-peak. As to the two-level CNN, it is feasible to adopt different depths focusing on different parts of the heartbeat. Two CNN layers are used for object-level, and one CNN layer is used for focusing on each part of the heartbeat. The detail configuration of the two-level CNN is presented in Table 1. The outputs of the two levels are concentrated and fed into two MLP layers for classification. The first MLP layer contains 10 neurons which are fully connected with neurons of the following layer. The second MLP layer contains 5 neurons corresponding to five types of heartbeats.

4.2 Evaluation Method

For training as well as testing the proposed two-level CNN, except for four ECG recordings containing paced heartbeats, the ECG records of the MIT-BIH arrhythmia database are divided into two parts. One part comprises 20 records range from 100 to 124, in which some representative heartbeats are selected as common training beats. The other part consists of remaining 24 records range from 200 to 234, which contains uncommon but clinically significant heartbeats, such as supraventricular beats, ventricular beats, and nodal beats. The training dataset contains common and patient-specific ECG arrhythmias. Similarly to [17], the common part of the training dataset contains 245 representative heartbeats selected from each heartbeat class randomly, including 75 from heartbeat class N, S, V respectively, 13 from class F as well as 7 from class Q. First 5 minutes ECG signals of each recording from the second part of the dataset are selected as patient-specific training beats. According to the recommendations of the AAMI, the testing dataset is built by considering three different scenarios. Scenario 1 adopts 11 testing records (200, 202, 210, 213, 214, 219, 221, 228, 231, 233, 234) for VEB detection and 14 testing records (200, 202, 210, 212, 213, 214, 219, 221, 222, 228, 231, 232, 233, 234) for SVEB detection. Scenario 2 adopts 24 testing records from 200 to 234, and Scenario 3 adopts all 44 testing records for both VEB and SVEB detection.

The measured metrics used for evaluating classification performance are as follows: accuracy $Acc = (TP + TN)/(TP + FP + TN + FN)$, sensitivity $Sen = TP/(TP + FN)$, specificity $Spe = TN/(FP + TN)$, and positive predictivity $Ppr = TP/(FP + TP)$. The above four metrics are computed by the quantity of true positive $(TP)$, false positive $(FP)$, true negative $(TN)$, and false negative $(FN)$.

4.3 Evaluation on the MIT-BIH Arrhythmia Database

4.3.1 VEB and SVEB Classification Accuracy Evaluation

To verify the effectiveness, the performance of the proposed system is compared with existing methods which also comply with the AAMI norm [6], [8], [11]–[13], [15], [17], [19]. Although some other researches also meet the AAMI standard [5], [14], since they use different test datasets, it is not suitable to compare the proposed method with their results directly. According to the recommendations of the AAMI,

| Table 1 | Detail description of the proposed two-level 1-D CNN configuration |
|---------|---------------------------------------------------------------|
|         | Object-level | Part-level |
|         | P   | QRS | T   | P   | QRS | T   |
| CNN Layer 1 | 1Dconv$^a$ | 3   | 3   | 3   | 3   | 3   |
|           | 1Dsubs$^b$ | 2   | 2   | 2   | None | None |
| Neurons$^c$ | 32  | 4   | 4   | 4   | None | None |
| CNN Layer 2 | 1Dconv$^a$ | 3   | 3   | 3   | 3   | 3   |
|           | 1Dsubs$^b$ | 2   | None | None | None |
| Neurons$^c$ | 16  | None | None | None |

$^a$ 1-D convolution kernel length
$^b$ 1-D subsampling factor
$^c$ Number of neurons
Table 2  VEB and SVEB classification performance of the proposed system and comparison with relevant state-of-the-art methods

| Methods                  | Resolution | VEB Acc | VEB Sen | VEB Spe | VEB Ppr | SVEB Acc | SVEB Sen | SVEB Spe | SVEB Ppr |
|--------------------------|------------|---------|---------|---------|---------|---------|---------|---------|---------|
| Ince et al. (2009) [6]a  | 11-bit     | 97.9    | 90.3    | 98.8    | 92.2    | 96.1    | 81.8    | 98.5    | 63.4    |
| Manab et al. (2014) [8]a | 11-bit     | 99.0    | 95.7    | 99.6    | 96.3    | 98.2    | 84.9    | 98.9    | 82.6    |
| Jiang et al. (2007) [11]a| 11-bit     | 98.9    | 94.3    | 99.4    | 95.8    | 97.5    | 74.9    | 98.8    | 78.8    |
| Hu et al. (1997) [13]a   | 11-bit     | 94.8    | 78.9    | 96.8    | 75.8    | N/A     | N/A     | N/A     | N/A     |
| Kiranyaz et al. (2016) [17]a | 11-bit     | 98.9    | **95.9**| 99.4    | 96.2    | 96.4    | 68.8    | **99.5**| 79.2    |
| Proposeda                | 6-bit      | 98.8    | 93.9    | 99.5    | 95.8    | 97.9    | **86.4**| 98.5    | 77.5    |
| Ince et al. (2009) [6]b  | 11-bit     | 97.6    | 83.4    | 98.1    | 87.4    | 96.1    | 62.1    | 98.5    | 56.7    |
| Manab et al. (2014) [8]b | 11-bit     | 98.5    | 91.4    | 99.1    | 91.8    | 97.5    | 74.0    | 98.7    | 73.6    |
| Jiang et al. (2007) [11]b| 11-bit     | 98.1    | 86.6    | 99.3    | 93.3    | 96.6    | 50.6    | 98.8    | 67.9    |
| Shadmand et al. (2016) [12]b | 11-bit     | 98     | 87.4    | 98.8    | 88.6    | 97.4    | 58.6    | 99     | 71.3    |
| Kiranyaz et al. (2016) [17]b | 11-bit     | 98.6    | **95.0**| 98.1    | 89.5    | 96.4    | 64.6    | 98.6    | 62.1    |
| Proposedb                | 6-bit      | **98.9**| 93.4    | **99.5**| **94.9**| **98.3**| **86.3**| **98.9**| **79.9**|

Best results are highlighted

- a The results are based on 11 recordings for VEB detection and 14 recordings for SVEB detection
- b The results are based on 24 recordings for VEB and SVEB detection
- c The results are based on all recordings for VEB and SVEB detection

the VEB and SVEB detection problems are considered respectively and the results are presented in Table 2. As shown in the table, the sensitivity and positive predictivity scores of SVEB classification are lower than VEB classification. There are two main reasons for explaining the worse performance. One reason is that the 245 representative heartbeats were randomly selected from the first part of the dataset, and the beats from the first 5 minutes of each patient may not cover all representative types of SVEBs. The class S beats are underrepresented in the training step for SVEB detection due to this reason. The other reason is that the P wave of each heartbeat is difficult to be classified. The P wave is a critical characteristic for classification of type S beats. Therefore, the classification of many type S beats are confused with normal heartbeats. As to the resolution of ECG heartbeat signal, the proposed classification system only adopts 64 samples to represent each beat, and each sample is 6-bit instead of 11-bit in other compared methods.

4.3.2 Different Resolution Evaluation

Figure 4 shows the influence of different resolution of the ECG signals on the heartbeat classification performance in three scenarios. As we can see from the figure, higher resolution of ECG signal is represented, almost higher accuracy and specificity are achieved. For 11-bit resolution, the accuracy and specificity are 99.2% and 99.8% for VEB detection, as well as 99.0% and 99.7% for SVEB detection in scenario 1. In scenario 2, the accuracy and specificity are 99.3% and 99.5% for VEB detection, as well as 99.0% and 99.6% for SVEB detection. In scenario 3, the accuracy and specificity are 99.5% and 99.8% for VEB detection, as well as 99.3% and 99.7% for SVEB detection. With regard to the positive predictivity value, the high resolution improves the value to some extent, especially in classification of SVEB. Compared with the positive predictivity in 6-bit resolution, the values are increased by 2.3%, 1.2%, and 0.1% for VEB detection as well as 16.1%, 10.3%, and 12.5% for SVEB detection from scenario 1 to 3 respectively with 11-bit resolution. However, the increase of resolution has little effect on the sensitivity improvement. As presented in Table 2 and Fig. 4, compared with other state-of-the-art detection methods with 11-bit resolution, the proposed classification approach can achieve well overall performance with lower resolution.

4.3.3 Different Numbers of Samples Evaluation

The classification performance comparison between different numbers of samples representing a heartbeat is also con-
ducted. As shown in Fig. 5, except for the sensitivity of SVEB classification, other values are improved with the increase of the sample number in the three scenarios. For VEB detection, all the maximum increment of accuracy, sensitivity, specificity, and positive predictivity are achieved in scenario 2, which are 0.3%, 1.8%, 0.1%, and 1.4% respectively. For SVEB detection, all the maximum increment of accuracy, specificity, and the positive predictively are achieved in scenario 1, which are 0.3%, 0.4%, and 6.6% respectively. However, the maximum decrement of sensitivity in aspect of SVEB classification is 1.5% obtained in scenario 3. Compared with number of type S beats represented by 64 samples, larger number of type S heartbeats represented by 128 samples are misclassified as normal as well as type V beats. In the light of the similar classification performance with the two different number of samples, 64 samples are preferred to be adopted in the proposed heartbeat classification system, which also reduce computation cost compared with 128 samples.

4.3.4 Overall Accuracy Evaluation

The overall classification accuracy of all ECG recordings represented by 64 samples is shown in Table 3. All compared methods classify heartbeats from the 44 recordings into five types according to the AAMI standard. It is obvious that the better performance is achieved by adopting the proposed method compared with existing methods. As described in the table, the proposed system implemented with only 6-bit ECG signal resolution outperforms the others in the aspect of the overall accuracy (OA). As the resolution increases, the improvement of overall accuracy is also obtained by using the proposed classification approach.

4.3.5 Robustness to Noise Evaluation

To assess the robustness of the proposed classification system to the noise, we added Gaussian noise with four different signal noise ratio (SNR) to the raw ECG signals with
Fig. 5  Beat classification performance with different samples in each scenario: (a) VEB in scenario 1, (b) SVEB in scenario 1, (c) VEB in scenario 2, (d) SVEB in scenario 2, (e) VEB in scenario 3, and (f) SVEB in scenario 3

Table 3  Comparison of overall accuracy with different resolution

| Methods          | Resolution | OA  |
|------------------|------------|-----|
| Ince et al. (2009) [6] | 11-bit     | 95.6 |
| Raj et al. (2016) [15] | 11-bit     | 89.1 |
| Kiranyaz et al. (2016) [17] | 11-bit     | 96.6 |
| Teijeiro et al. (2016) [19] | 11-bit     | 97.7 |
|                 | 6-bit      | 97.8 |
|                 | 7-bit      | 97.9 |
|                 | 8-bit      | 98.1 |
|                 | 9-bit      | 98.2 |
|                 | 10-bit     | 98.4 |
|                 | 11-bit     | 98.5 |

Table 4  Heartbeat classification performance with noise added and without noise

| Methods | Without noise | SNR in noise added (dB) |
|---------|---------------|-------------------------|
|         | Acc | Sen | Spe | Ppr | 40  | 30  | 20  | 10  |
| VEB     | 99.2 | 99.1 | 99.2 | 99.0 | 98.4 |
|         | 93.7 | 94.2 | 94.2 | 91.5 | 85.2 |
|         | 99.6 | 99.5 | 99.6 | 99.6 | 99.4 |
|         | 94.8 | 93.0 | 94.1 | 94.2 | 91.3 |
| SVEB    | 98.9 | 98.8 | 98.7 | 98.4 | 98.3 |
|         | 85.0 | 80.5 | 80.4 | 73.5 | 65.8 |
|         | 99.3 | 99.4 | 99.3 | 99.2 | 99.3 |
|         | 78.2 | 80.4 | 77.7 | 73.9 | 73.6 |
|         | 97.8 | 97.8 | 97.7 | 97.2 | 96.4 |

6-bit resolution, and the corresponding classification results are represented in Table 4. Compared with classification performance using noise-free ECG signal, the overall accu-
The training database selection and measurement methods mate the VEB classification capability and overall accuracy. In this study, the INCART database was also used to estimate combined with noise, so that the proposed two-level CNN normalization approach can eliminate noise to some extent. The other is that the ECG signals in training dataset are also normalized of a given heartbeat. In order to facilitate feature extraction mentioned above, an adaptive ECG signal normalization technique is adopted for heartbeat preprocessing, which can normalize resolution of ECG signal, reduce data memory storage, as well as denoise to some extent. The experimental results obtained on the MIT-BIH arrhythmia database and the INCART database show that the proposed heartbeat classification system provides robustness to noise and achieves high accuracy of heartbeat classification with low signal resolution.

Table 5

| Methods       | Resolution | OA  | Sen | Ppr  | OA  | Sen | Ppr  |
|---------------|------------|-----|-----|------|-----|-----|------|
| Li et al. (2014) [2] | 8-10 bits/mV | 94.0 | 94.0 | 99.1 | 93.4 | 66.5 |
| Allami et al. (2017) [3] | 8-10 bits/mV | 94.2 | N/A  | 87.5 | 92.7 |
| Llamedo et al. (2011) [7] | 8-10 bits/mV | N/A  | N/A  | N/A  | 82.0 | 88.0 |
| Manab et al. (2014) [8] | 8-10 bits/mV | N/A  | N/A  | N/A  | 94.3 | 89.1 |
| Proposed      | 6 bits/beat | 97.8 | 99.3 | 98.2 | 86.0 | 94.2 |

Best results are highlighted.

racy is only decreased by 1.4% with 10 dB SNR. We also observe that when the SNR is larger than 20 dB, the overall accuracy is close to the value without noise added. Moreover, some evaluating indicators are higher than corresponding noise-free indicators, such as sensitivity in VEB detection, specificity and positive predictivity in SVEB detection. The reasons behind the robustness to noise mainly focus on two aspects. One is that the proposed adaptive heartbeat normalization approach can eliminate noise to some extent. The other is that the ECG signals in training dataset are also combined with noise, so that the proposed two-level CNN can extract features of ECG signals with noise added.

4.4 Evaluation on the INCART Database

In this study, the INCART database was also used to estimate the VEB classification capability and overall accuracy. The training database selection and measurement methods are complied with [2]. The performance results are compared with other published studies adopting same validating datasets [2], [3], [7], [8], which are presented in Table 5. As we can see from the table, although adopting 6-bit resolution, the proposed classification approach outperforms other studies in the aspects of overall accuracy, sensitivity of Non-VEB detection, and positive predictivity of VEB detection.

5. Conclusion

In this paper, a patient-specific automatic heartbeat classification system is proposed, which adopts morphological features and timing information. The coarse-grained and fine-grained morphological features are extracted using attention-based two-level 1-D CNN, negating the necessity to extract features manually. The RR interval difference is computed as dynamic feature for indicating timing information of a given heartbeat. In order to facilitate feature extraction mentioned above, an adaptive ECG signal normalization technique is adopted for heartbeat preprocessing, which can normalize resolution of ECG signal, reduce data memory storage, as well as denoise to some extent. The experimental results obtained on the MIT-BIH arrhythmia database and the INCART database show that the proposed heartbeat classification system provides robustness to noise and achieves high accuracy of heartbeat classification with low signal resolution.

References

[1] E.B. Mazomenos, D. Biswas, A. Acharyya, T. Chen, K. Maharatna, J. Rosengarten, J. Morgan, and N. Curzen, “A low-complexity ECG feature extraction algorithm for mobile healthcare applications,” IEEE J. Biomed. Health Inform., vol.17, no.2, pp.459–469, March 2013. DOI: 10.1109/TITB.2012.2231312.
[2] P. Li, C. Liu, X. Wang, D. Zheng, Y. Li, and C. Liu, “A low-complexity data-adaptive approach for premature ventricular contractions recognition,” Signal, Image Video Process., vol.8, no.1, pp.111–120, Jan. 2014. DOI: 10.1007/s11760-013-0478-6.
[3] R. Allami, A. Stranieri, V. Balasubramanian, and H.F. Jelinek, “A count data model for heart rate variability forecasting and premature ventricular contraction detection,” Signal, Image, Video Process., vol.11, no.8, pp.1427–1435, Nov. 2017. DOI: 10.1007/s11760-017-1103-x.
[4] K. Minami, H. Nakajima, and T. Toyoshima, “Real-time discrimination of ventricular tachycardia from Fourier-transform neural network,” IEEE Trans. Biomed. Eng., vol.46, no.2, pp.179–185, Feb. 1999. DOI: 10.1109/10.740880.
[5] S. Chen, W. Hua, Z. Li, J. Li, and X. Gao, “Heartbeat classification using projected and dynamic features of ECG signal,” Biomed. Signal Process. Control, vol.31, pp.165–173, Jan. 2017. DOI: 10.1016/j.bspc.2016.07.010.
[6] T. Ince, S. Kiranyaz, and M. Gabbouj, “A generic and robust system for automated patient-specific classification of ECG signals,” IEEE Trans. Biomed. Eng., vol.56, no.5, pp.1415–1426, May 2009. DOI: 10.1109/TBME.2009.2183934.
[7] M. Llamedo and J.P. Martinez, “Heartbeat classification using feature selection driven by database generalization criteria,” IEEE Trans. Biomed. Eng., vol.58, no.3, pp.616–625, March 2011. DOI: 10.1109/TBME.2010.2068048.
[8] K.D. Manab and A. Samit, “Patient-specific ECG beat classification technique,” Healthcare Technol. Lett., vol.1, no.3, pp.98–103, Sept. 2014. DOI: 10.1049/hlt.2014.0072.
[9] H. Sharma and K.K. Sharma, “An algorithm for sleep apnea detection from single-lead ECG using Hermite basis functions,” Comput. Biol. Med., vol.77, no.1, pp.116–124, Oct. 2016. DOI: 10.1016/j.compbiomed.2016.08.012.
[10] C. Song, K. Liu, X. Zhang, L. Chen, and X. Xian, “An obstructive sleep apnea detection approach using a discriminative hidden Markov model from ECG signals,” IEEE Trans. Biomed. Eng., vol.63, no.7, pp.1532–1542, July 2016. DOI: 10.1109/TBME.2015.2498199.
[11] W. Jiang and S.G. Kong, “Block-based neural networks for personalized ECG signal classification,” IEEE Trans. Neural Netw., vol.18, no.6, pp.1750–1761, Nov. 2007. DOI: 10.1109/TNN.2007.900239.
[12] S. Shadmand and B. Mashoufi, “A new personalized ECG signal classification algorithm using block-based neural network and particle swarm optimization,” Biomed. Signal Process. Control, vol.25, pp.12–23, March 2016. DOI: 10.1016/j.bspc.2015.10.008.
[13] Y.H. Hu, S. Palreddy, and W.J. Tompkins, “A patient-adaptable ECG beat classifier using a mixture of experts approach,” IEEE Trans. Biomed. Eng., vol.44, no.9, pp.891–900, Sept. 1997. DOI: 10.1109/10.623058.
[14] Z. Zhang, J. Dong, X. Luo, K.-S. Choi, and X. Wu, “Heartbeat classification using disease-specific feature selection,” Comput. Biol. Med., vol.46, pp.79–89, March 2014. DOI: 10.1016/j.compbiomed.2013.11.019.
[15] S. Raj, K.C. Ray, and O. Shankar, “Cardiac arrhythmia beat classification using DOST and PSO tuned SVM,” Comput. Methods Programs Biomed., vol.136, pp.163–177, Nov. 2016. DOI:
10.1016/j.cmpb.2016.08.016.

[16] A. Khazaee and A. Ebrahimzadeh, “Classification of electrocardiogram signals with support vector machines and genetic algorithms using power spectral features,” Biomed. Signal Process. Control, vol.5, no.4, pp.252–263, Oct. 2010. DOI: 10.1016/j.bspc.2010.07.006.

[17] S. Kiranayaz, T. Iace, and M. Gabbouji, “Real-time patient-specific ECG classification by 1-D convolutional neural networks,” IEEE Trans. Biomed. Eng., vol.63, no.3, pp.664–675, March 2016. DOI: 10.1109/TBME.2015.2468589.

[18] M.M.A. Rahhal, Y. Bazi, H. Alhichri, N. Alajlan, F. Melgani, and S. Kiranyaz, T. Ince, and M. Gabbouj, “Real-time patient-specific ECG classification by 1-D convolutional neural networks,” IEEE Trans. Biomed. Eng., vol.63, no.3, pp.664–675, March 2016. DOI: 10.1109/TBME.2015.2468589.

[19] T. Teijeiro, P. Félix, J. Presedo, and D. Castro, “Heartbeat classification using abstract features from the abductive interpretation of the ECG,” IEEE J. Biomed. Health Inform., vol.19, no.6, pp.1645–1652, June 2016. DOI: 10.1109/JBHI.2016.2531247.

[20] H. Yang, C. Kan, G. Liu, and Y. Chen, “Spatiotemporal differentiation of myocardial infarctions,” IEEE Trans. Autom. Sci. Eng., vol.10, no.4, pp.938–947, Oct. 2013. DOI: 10.1109/TASE.2013.2263497.

[21] A. Gonzalez-Garcia, A. Vezhnevets, and V. Ferrari, “An active search strategy for efficient object class detection,” IEEE Conf. on Comput. Vis. Pattern Recognit. (CVPR), pp.3022–3031, 2015.

[22] N. Zhang, J. Donahue, R. Girshick, and T. Darrell, “Part-based R-CNNs for fine-grained category detection,” European Conf. on Comput. Vis. (ECCV), vol.8689, pp.834–849, 2014. DOI: 10.1007/978-3-319-10590-1_54.

[23] T. Xiao, Y. Xu, K. Yang, J. Zhang, Y. Peng, and Z. Zhang, “The application of two-level attention models in deep convolutional neural network for fine-grained image classification,” IEEE Conf. on Comput. Vis. Pattern Recognit. (CVPR), pp.842–850, 2015.

[24] 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, and H.E. Stanley, “PhysioBank, PhysioToolkit, and PhysioNet: components of a new research resource for complex physiologic signals,” Circulation, vol.101, no.23, pp.E215–E220, June 2000. DOI: 10.1161/01.CIR.101.23.E215.

[25] A. Kennedy, D.D. Finlay, D. Gudenring, R.R. Bond, K. Moran, and J. McLaughlin, “Automated detection of atrial fibrillation using R-R intervals and multivariate-based classification,” J. Electrocardiol., vol.49, no.6, pp.385–387, Nov.-Dec. 2016. DOI: 10.1016/j.jelectrocard.2016.07.033.

[26] Jiahui Luo received the Bachelor of Engineering degree in electronic and information engineering from Zhejiang University, in 2012. She is currently a Ph.D student at the Institute of VLSI Design, Zhejiang University. Her research interests include biomedical signal processing and ultra-low-power IC design.

[27] Taotao Zhu received the B.S. degree in information engineering from Zhejiang University, Hangzhou, China, in 2012, where he is currently pursuing the Ph.D. degree from the Institute of VLSI Design. His current research interests include the system of extremely low power architecture, and near-threshold and sub-threshold circuits designing.

[28] Sheng Wang received the B.S. degree in Information Engineering from Zhejiang University, Hangzhou, China, in 2012, where he is currently pursuing the Ph.D. degree from the Institute of VLSI Design. His current research interests include the system of extremely low power architecture, EDAC circuit design, and near-threshold and sub-threshold circuits designing.

[29] Xiaoyan Xiang received the B.S. and Ph.D degrees from Zhejiang University, Hangzhou, China, in 2008 and 2013, respectively, all in Institute of VLSI Design. She has been a lecturer in the State Key Lab of ASIC and System, Fudan University from 2013. Her research interests include low-voltage low-power circuits for resilient processor, and the CPU architecture optimization in IOT field.

[30] Jianyi Meng received his B.S. degree and Ph.D degree in Department of Information Science & Electronic Engineering, Zhejiang University, China, 2004 and 2009. He joined Institute of VLSI Design, College of Electrical Engineering, Zhejiang University in 2009 as a post-doctor. He visited IMEC, Belgium, as short term scholar visitor in 2010. He is now a professor of State Key Laboratory of ASIC and System, Department of Microelectronics, Fudan University. He has been in charge of or has participated as a main contributor in Special Grand National Science-technology Project, Pre-research Program, 863 Program, and etc. He has published over 30 academic papers index by SCI/ EI, and applied over 20 patents. His research mainly focuses on computer architecture, ultra low power digital circuit design and application specific SoC design.