IEICE TRANSACTIONS

on Communications

VOL. E103-B NO. 12
DECEMBER 2020

A PUBLICATION OF THE COMMUNICATIONS SOCIETY
The Institute of Electronics, Information and Communication Engineers
Kikai-Shinko-Kaikan Bldg., 5-8, Shibakoen 3chome, Minato-ku, TOKYO, 105-0011 JAPAN

PAPER Special Section on IoT Sensor Networks and Mobile Intelligence

ECG Classification with Multi-Scale Deep Features Based on Adaptive Beat-Segmentation

Huan SUN[†], *Nonmember*, Yuchun GUO[†a)], *Member*, Yishuai CHEN[†], *and* Bin CHEN[†], *Nonmembers*

SUMMARY Recently, the ECG-based diagnosis system based on wearable devices has attracted more and more attention of researchers. Existing studies have achieved high classification accuracy by using deep neural networks (DNNs), but there are still some problems, such as: imprecise heart beat segmentation, inadequate use of medical knowledge, the same treatment of features with different importance. To address these problems, this paper: 1) proposes an adaptive segmenting-reshaping method to acquire abundant useful samples; 2) builds a set of hand-crafted features and deep features on the inner-beat, beat and inter-beat scale by integrating enough medical knowledge. 3) introduced a modified channel attention module (CAM) to augment the significant channels in deep features. Following the Association for Advancement of Medical Instrumentation (AAMI) recommendation, we classified the dataset into four classes and validated our algorithm on the MIT-BIH database. Experiments show that the accuracy of our model reaches 96.94%, a 3.71% increase over that of a state-of-the-art alternative.

key words: *ECG classification, adaptive beat segmentation, multi-scale deep features, channel attention module*

1. Introduction

With the development of IoT sensor networks and the popularization of electronic health applications, ECG-based diagnosis system has been paid more and more attention. According to the 2018 World Health Statistics Report, in 2016, global cardiovascular diseases (CVD) caused a total of 17.9 million deaths, accounting for 44% of all noncommunicable diseases (NCD). At the same time, in China, about 290 million people were suffering from CVD in 2016. CVD deaths accounted for more than 40% of the deaths of residents, ranking first, higher than cancer and other diseases. As the number of cardiovascular patients is increasing rapidly, the need for dynamic real-time monitoring of heart activity is also increasing. Wu et al. [1] presented a wearable sensor network system for Internet of Things (IoT) connected safety and health applications. The wearable sensors on different subjects can communicate with each other and transmit the data to a gateway via a LoRa network which forms a heterogeneous IoT platform with Bluetooth-based medical signal sensing network. This provides abundant biometrical data for electronic health applications. Therefore, in the ECG-based diagnosis system, we need to improve the accuracy of the algorithm.

However, existing methods cannot attain the diagnosis accuracy needed, because of the following problems: 1) infeasibility of obtaining the large sample sets needed as deep neural network input deep neural network; 2) how to complete and appropriate feature expression for a single sample; 3) how to build a classifier with clear boundary. An ECG aided diagnosis system mainly consists of three parts, i.e. beat segmentation, feature extraction, and classification [2], [3]. To simplify the beat segmentation and meet the requirement that input data is of fixed length, existing approaches emphasize fixed length beat segmentation [4]–[6]. Xiang et al. [7] segment PQRST waves with a fixed length. In fact, since the length of a beat or a wave varies, the existing fixed-length segmentation method may result in that the beat or wave is incomplete or redundant. Actually, this incompleteness or redundancy affects the precision of feature extraction and confuses the classifier. As for feature extraction, hand-crafted features and deep features have been proven to be useful, but a very few studies use both of them for beat representation. Since the hand-crafted feature is often based on medical knowledge, the combination of these two features can compensate for the lack of medical knowledge of the neural network. Meanwhile, beat, local waves and multiple beats are important for a medical expert to classify the beat. However, the beat feature is widely used, and inner-beat feature, which is extracted from PQRST waves, is merely used in the work of Xiang et al. [7]. The inter-beat feature is ignored for classification, which we define as the correlation among multiple beats. For classification, neural network tends to be used as classifier recently, so the learned feature is of high dimension. However, since current work widely uses a typical deep neural network, each dimension is of equal importance. But obviously, different dimensions have different importance. Therefore, the learning ability of the classifier in the feature dimension is urgently needed to be improved.

To improve the performance of ECG-based diagnosis systems, we classify ECG signals with multi-scale features based on adaptive beat segmentation. Specifically, our study has three parts. Firstly, for beat segmentation, we propose to segment beat, wave or multiple beats with unfixed-length adaptively, and reshape the beat, wave or multiple beats using adaptive window lengths to fit the feature extraction and classifier and keep the waveform vital for feature extraction and classification unchanged. Then, for feature extraction, we introduce both hand-crafted feature and deep feature on multi-scales including inner-beat, beat, inter-beat scale

Manuscript received December 30, 2019.
Manuscript revised April 16, 2020.
Manuscript publicized July 1, 2020.

[†]The authors are with School of Electronic and Information Engineering, Beijing Jiaotong University, Beijing, China.
a) E-mail: ychguo@bjtu.edu.cn

DOI: 10.1587/transcom.2020SEP0002

to integrate more medical information with deep learning method. Finally, for classification, we introduce a modified channel attention module into 1-D CNN to optimize classifier ability. Our algorithm is validated on the MIT-BIH dataset, and the results show that the performance of our algorithm is superior to other related work.

The rest of this paper is organized as follows. Section 2 reviews the previous work. Section 3 explains our method in detail. Section 4 shows the results of the experiments. Finally, Sect. 5 concludes this paper.

2. Related Work

Our work involves three parts: beat segmentation, feature extraction and classification. The related work of each topic is as follows.

Beat segmentation. With R peak as a reference position, it is a routine to segment beat with a number of samples before and after R peak. Cimen et al. [8] firstly downsampled the raw data, then they segmented beat with 61 samples before R peak and 38 samples after R peak. Kiranyaz et al. [9] also downsampled the raw data, after that they extracted 64 samples from both side of R peak. Zubair et al. [5] extracted 128 samples from both side of R peak without downsampling the raw data. Ye et al. [4] segmented beat with 100 samples before R peak and 200 samples after R peak. Furthermore, Xiang et al. [7] downsampled the raw data at first, then they segmented beat with 26 samples before R peak and 37 samples after R peak. Based on the obtained beat, Xiang et al. [7] segmented PQRST waves. The first 20 samples represent P wave, and the second 20 samples represent QRS waves, while the last 24 samples represent T wave. All of them segmented beat or wave with fixed length, without considering the length of beat or wave varies due to different individuals and even for the same person, the length of beat or wave will change in different conditions. In fact, these works of beat or wave segmentation are simplified in an unreasonable way.

Feature extraction. For beat representation, the existing work can be divided into two categories. The first one is hand-crafted feature and it can be obtained from frequency domain [10], wavelet transform [11], hermite function [12] and morphology [13]. The second one is deep feature. Many neural networks are used, such as artificial neural network [14], convolutional neural network [6], recurrent neural network [15] and block-based neural network [16]. The above two feature extraction methods are based on single beat. Xiang et al. [7] introduced PQRST waves to improve the accuracy of classification. To sum up, the current studies mostly extract beat feature from a single beat, while few considers the inner-beat feature, not to mention the inter-beat feature. Besides, hand-crafted feature and deep feature are proved very useful for classification, but a very few studies classify ECG signals combining hand-crafted feature and deep feature.

Classification. The choice of classifier is variable, but generally, it can be divided into two categories. The first one is traditional machine learning classifier, i.e. random forest [17], SVM [18] and KNN [19]. Clustering algorithm, k-means [8], is also used for classification. The second one is neural network method. The most widely used classifier is MLP [9]. Though the neural network has better performance than traditional machine learning classifiers, the number of classifiers in different works is different. Kiranyaz et al. [9] and Xiang et al. [7] trained a classifier for each specific subject, while Zubair et al. [5] trained a classifier for all subjects. A classifier trained for a specific subject with his/her ECG records and corresponding diagnose labels can achieve better accuracy for the same person than for others. Such a classifier is not feasible in practice as it cannot work well for a new patient with no records being collected. Therefore, we are going to train a general classifier for all patients. Moreover, Hu et al. [20] proved that the channel information is very helpful for many image tasks, such as image segmentation. Woo et al. [21] extended the work and verify its performance. The classifier used for ECG classification is just a simple application of a commonly used classifier. The limitation of commonly used classifier is that the learned feature map of each channel has the same importance. In fact, different feature channel has different importance. The classifier can be improved by modifying channel attention module for ECG classification.

3. Method

As shown in Fig. 3, the entire multi-input 1-D CNN classification system consists of three parts: segmentation, feature extraction and deep classifier network. In the segmentation part, we explain adaptive beat segmentation (ABS), overlapping wave segmentation (OWS), and multiple beats segmentation (MBS), which are important for multi-scale deep feature extraction. Then, we detail multi-scale hand-crafted feature extraction (HFE). After that, we introduce a modified channel attention module to enhance the ability of our classifier on extracting channel information.

3.1 Adaptive Beat Segmentation

To segment beat, the start and the length of beat are necessary. As is shown in Fig. 1, the waveform between T wave and P wave is flat, so the boundary of current beat and posterior beat is vague. But the R peak can be identified quickly and accurately with the method by Pan et al. [22]. It is widely accepted to use the length of RR interval, i.e. the length between current R peak and posterior R peak, to estimate the length of beat. So we define the real length of beat i as the length of RR interval.

$$l_s^i = R_{i+1} - R_i \tag{1}$$

where R_i means R peak of beat i, and l_s^i means the real length of beat i. The real length of beat is varying from Eq. (1).

As shown in Fig. 1, with R peak as a reference position, the start of beat i can be determined by shifting leftward with a length l_f. It is known that P wave occupies a relatively

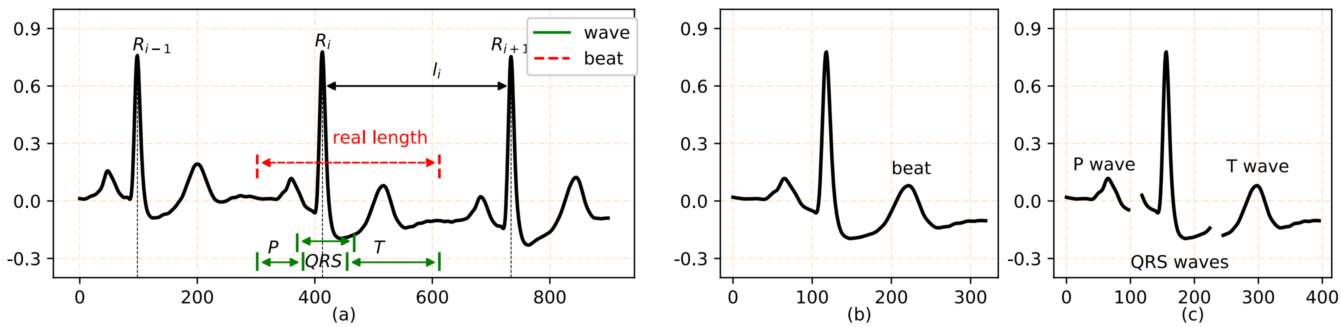

Fig. 1 (a) The black straight solid line shows the length of RR interval. The red straight dashed line shows the real length of beat. Based on the obtained beat, the three green solid lines shows the proposed overlapping wave segmentation method. (b) The result of proposed adaptive beat segmentation method. (c) The result of proposed overlapping wave segmentation method.

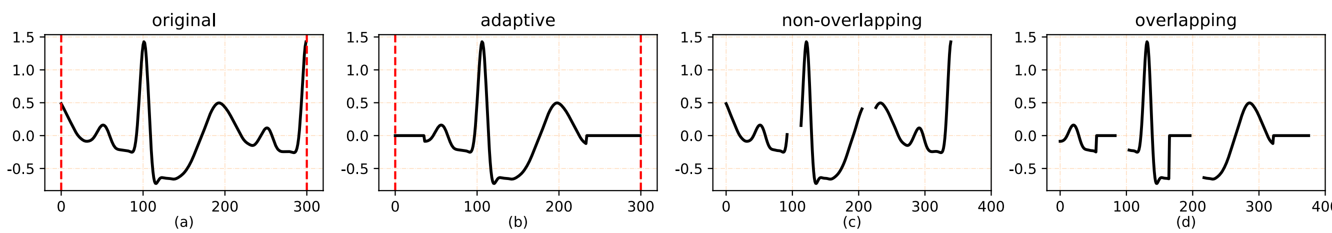

Fig. 2 Examples of different segmentation methods. (a) Beat segmentation method [4]. (b) Proposed beat segmentation method. (c) Non-overlapping wave segmentation method [7] based on (a). (d) Proposed overlapping wave segmentation method based on (b).

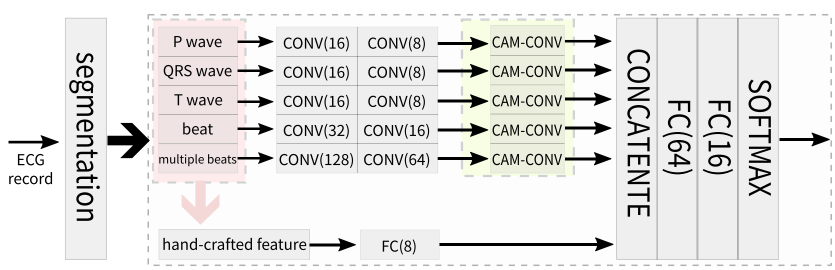

Fig. 3 The overview of proposed method. ECG record is cut into a set of beats. Then, wave and multiple beats are obtained. Based on this, hand-crafted feature is extracted. If the CAM is used, there are three convolutional module in neural network. Otherwise, there are two convolutional module.

fixed proportion of a beat. Then, l_f^i is calculated as follows,

$$l_f^i = n_f \times l_s^i \tag{2}$$

where n_f means the proportion of P wave in a beat and it is the same for all beats. The choice of parameters affects the precision of segmented beat and will be discussed in detail later.

Through the above two steps, such beats with varying length need to be reshaped to beats with equal length. An intuitive solution is scaling, but beats will be changed by scaling differently, which hurts feature extraction and classification. Thence we use padding & truncating method. As shown in subgraph (b) of Fig. 2, we set a parameter L_s, which is defined as valid length. When the l_s^i is bigger than L_s, since the tail of T wave is useless for classification, we truncate the beat at the end. Otherwise, we adopt the padding strategy. Firstly, we set each R peak of beat in the same index, and then pad the beat with zeros at the head and the end. Moreover, the choice of the L_s should be made carefully and reasonably, because the beat is incomplete if L_s is too small or the beat is complete but has too many zeros at the end when L_s is too big. In Sect. 4, we will choose L_s based on the distribution of l_s in MIT-BIH database.

Using the proposed ABS, we obtain beat with a fixed length of L_s, which is called the valid length, and the real length of beat varies. The comparison of different beat segmentation methods is shown in Fig. 2.

3.2 Overlapping Wave Segmentation

To segment PQRST waves respectively and precisely, the start and length of each wave are necessary as well. However, as shown in Fig. 1, similar to the problem of beat detection, the boundaries of PQRST waves are also not clear enough to detect automatically, but each wave occupies a relatively fixed proportion. To solve the problem that the start of each wave is hard to detect, we extract wave by overlapping sampling. Then the length of each wave can be obtained as follows,

$$\begin{cases} l_p^i = (n_p + n_o) \times l_s^i & (3a) \\ l_r^i = (n_r + 2 \times n_o) \times l_s^i & (3b) \\ l_t^i = (n_t + n_o) \times l_s^i & (3c) \\ n_p + n_r + n_t = 1 & (3d) \end{cases}$$

where l_p^i, l_r^i, l_t^i mean the real length of P wave, QRS waves and T wave for beat i respectively. n_p, n_r, n_t are parameters that represent the proportion that each wave occupies in a beat. n_o is an overlapping parameter. All these four parameters are the same to all beats. The obtained waves also need to be reshaped into waves with equal length. As shown in subgraph (d) of Fig. 2, different from the beat reshaping method, zeros are padding to the end of each wave. Subgraph (c) of Fig. 2 shows the results of wave segmentation method by Xiang et al. [7]. They extract PQRST waves by dividing the beat into three non-overlapping parts based on beat with fixed length. Their method may result in that wave is incomplete or redundant, or even worse, the P wave of posterior beat is mistakenly considered as the T wave of the current beat, which is a serious disturbance to classifier. Subgraph (d) of Fig. 2 is the result of proposed wave segmentation method, and the waves is still precise.

3.3 Multiple Beats Segmentation

Heart rate variability (HRV) means the variation in time between consecutive heartbeats [23]. A variety of physiological phenomena affect HRV, so HRV can be a significant feature to classify different types of beats. To simplify the procedure of feature extraction, we choose the consecutive three beats to express HRV.

To segment multiple beats, the start and the length of multiple beats are still necessary. But firstly, we should define multiple beats. For medical experts, multiple beats are used to extract the correlation among beats, which in fact is to find the difference between current beat and previous or posterior beat. To enhance this difference, we define the real length of multiple beats as three times of the real length of the current beat. For beat i, there are both l_s^i samples contained before and after the current beat. And the real length of multiple beats is as follows,

$$l_{tr}^i = 3 \times l_s^i \tag{4}$$

where l_{tr}^i means the real length of multiple beats i and varies based on l_s^i. We also do the zero-padding or tail-truncating at the end of multiple beats.

3.4 Hand-Crafted Feature Extraction

Based on beat, wave and multiple beats segmentation, multi-scale hand-crafted feature can be extracted. To reflect the average voltage strength and the changing degree in voltage strength of wave or beat, we calculate five parameters, i.e. the mean, the variance, the absolute mean, the square mean, and the square root mean of absolute value. Therefore, 20 hand-crafted feature is obtained. And for beat, the real length is another hand-crafted feature used. To reflect the statistical feature among beats, the ratio of adjacent RR intervals can is used. In this paper, five kinds of parameters which are related to RR intervals are used as follows,

$$\begin{cases} r_1^i = \dfrac{RR_{i-1}}{RR_i} - 1 & (5a) \\ r_2^i = \dfrac{RR_i}{RR_{i+1}} - 1 & (5b) \\ r_3^i = \dfrac{RR_{i-1}}{RR_{i+1}} - 1 & (5c) \\ r_4^i = \dfrac{|RR_{i-1} - RR_{i+1}|}{RR_{i-1} + RR_{i+1}} & (5d) \\ r_5^i = \dfrac{RR_i}{RR_{i-1} + RR_{i+1}}. & (5e) \end{cases}$$

where RR_i represents the RR interval of beat i, and $r_1^i, r_2^i, r_3^i, r_4^i$, and r_5^i are all ratios of different RR intervals for beat i.

3.5 Modified Channel Attention Module

To solve the problem that CNN has limited capabilities learning channel information, we introduce channel attention module, which helps CNN learn channel information before classification and be proved useful in image task. As is shown in Fig. 4, the size of input raw feature map is $C \times L$. Similar to Woo et al. [21], for each channel, max-pooling and average-pooling are used. After that, a shared full-connected (FC) module with one hidden layer is followed. The number of neurons for the hidden layer is set to $\frac{C}{r}$, where r is a hyperparameter called reduction ratio. The number of neurons for output layer is set to C. Then, we add the two tensors and multiply the new tensor with the raw feature map. Finally, we obtain a new feature map. We can compute the CAM as follows,

$$\begin{aligned} F_n &= F \times \sigma(MLP(AP(F))) + \sigma(MLP(MP(F)))] \\ &= F \times [\sigma(W_1\sigma(W_0(F_a))) + \sigma(W_1\sigma(W_0(F_m)))] \end{aligned} \tag{6}$$

where σ denotes ReLU function, W_0 and W_1 are the FC module weights, shared by tensors generated by max pooling and average pooling. As is shown in Fig. 4, there are several differences between our CAM and module by Woo et al. [21]. Firstly, in the work by Woo et al. [21], the activation of the output layer in the shared FC module is sigmoid function. But we find the accuracy will be about 1% higher if the ReLU function is used. Secondly, we drop the activation operation after the adding operation. The reason why we change the module is that we find the accuracy will be higher based on the proposed method.

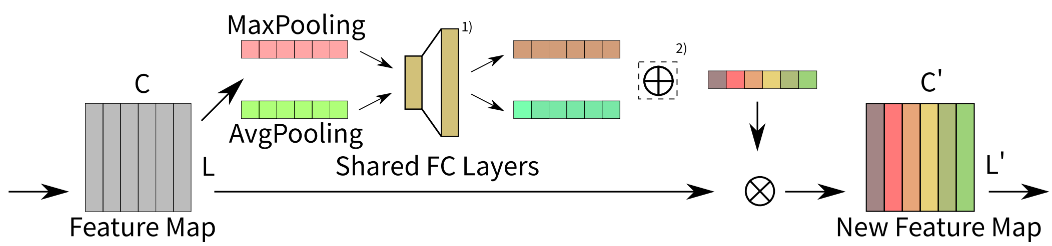

Fig. 4 Modified channel attention module. There are two improvements: 1) the activation function of the output layer is changed to ReLU function; 2) the activation operation is omitted.

4. Experiments

4.1 Dataset

The ECG dataset from MIT-BIH arrhythmia database is used in this paper, which is considered as one of most famous standard databases. The dataset contains 48 half-hour ECG records, obtained from 47 subjects studied by the BIH Arrhythmia Laboratory between 1975 and 1979. All records are passed through a band pass filter at 0.1–100 Hz and are sampled at 360 Hz. For classification, AAMI recommendations are adopted in this study. AAMI recommends that each beat be classified into the following five beat types: N (normal beats), S (supraventricular ectopic beats), V (ventricular ectopic beats), and F (fusion beats), and Q (unclassifiable beats). The following 4 records, including 102, 104, 107 and 217, are excluded in this study because these beats do not preserve sufficient signal quality for reliable processing. For the other 44 records, modified-lead II signals are used.

There are more than 100000 beats in the database. Similar to Zubair et al. [5], for the first 20 records (100–124) from the MIT-BIH database, 75 beats are randomly selected from type-N, type-S and type-V and all beats of type-F and type-Q are selected. According to the AAMI recommended procedure [24], a set of these 245 beats and the beats from the first 5 minutes of the second 24 records (200–232) are used for training. Actually, the number of beats used for training is no more than 10%. As shown in Fig. 5, the distribution of samples for the first 5 minutes is consistent with that for the entire dataset. All the other beats are used as test dataset. 5000 beats are randomly selected from the test dataset as validation dataset. The aforementioned dataset partitioning method is adopted in all experiments. In order to reduce data redundancy, a downsampling method has also been adopted.

4.2 Setup

The choice of the parameters defined in Sect. 3 will be discussed detailedly in this section. For the choice of L_s, firstly, we calculate the real length of beat for the entire MIT-BIH database, and then count the number of beat in each interval with a length of 50. The result shows that the beats whose

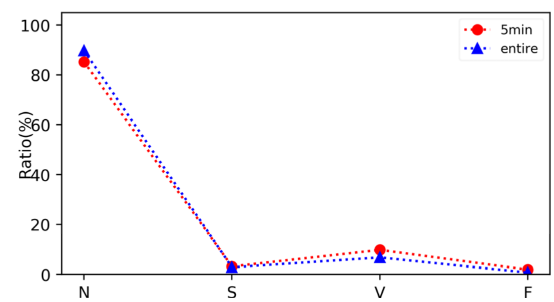

Fig. 5 The distribution of samples for the first 5 minutes and the entire dataset.

real length range from 300 to 350 occupies the largest number on the dataset. Moreover, the most real lengths in the database are less 300, which means most beats with fixed length will be complete after the reshaping operation. So we set L_s to 300. As the downsampling method is used in our work, we finally set L_s to 100. For the choice of other parameters, such as n_f, n_p, n_r, n_t, n_o, first of all, we delineate a range based on medical knowledge and select values with a fixed step. Then, for different values, we draw each beat from the first 5 minutes of all 44 records and select the value corresponding to the best performance. For example, we set n_f with a list of [0.20, 0.25, 0.30, 0.35, 0.40, 0.45], and we segment beat with ABS. Then, we draw the curve and judge whether the beat is complete. We find the ABS method performs best when n_f is 0.35. We choose the value of n_p, n_r, n_t, n_o in the same way. After a series of experiments, we set n_p and n_r to 0.25, n_t to 0.5 and n_o to 0.03. The reduction ratio, r, is set to 4 in this study.

As shown in Fig. 3, after a large number of experiments, we find that for the 1-D CNN without CAM, two convolutional modules can reach a higher accuracy, with fewer parameters and less training time. And for the 1-D CNN with CAM, there are three convolutional modules. As Hu et al. [20] proved that the SE module in the later convolutional module is more helpful, the CAM is only used in the third convolutional module. The number of neurons in each layer is indicated in the parentheses, which is determined by a lot of experiments. For the first convolutional module, a convolutional layer is followed by a max pooling layer and a batch normalization layer, while for the second one, there is a dropout layer after the convolutional layer. Another thing worth mentioning is that we set the training times to 30. What's more, to avoid overfitting, L2 regular-

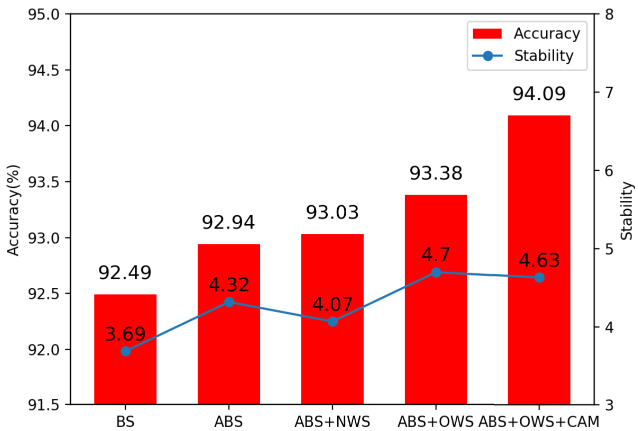

Fig. 6 The accuracy and variance of different methods. BS notes original beat segmentation; ABS notes adaptive beat segmentation; NWS notes non-overlapping wave segmentation; OWS notes overlapping wave segmentation; CAM notes 1-D CNN with channel attention module.

ization and early stop are also used in our network. In order to reduce the randomness of the neural network, we repeat all experiments for ten times, and the final experimental results are the mean of the ten experiments. Besides, we define the model stability as follows,

$$stability = -\log_{10}[\frac{1}{n}\sum_{i=1}^{n}(a_i - \overline{a})] \tag{7}$$

where n is the times of repeating experiments and is 10 in this paper. a_i is the accuracy of the experiment i. $\overline{a}$ is the mean of ten experments. The larger stability is, the more stable the model is.

4.3 Results of Different Segmentation Methods

To prove that beat or wave segmentation with adaptive length is better than that with fixed length, a set of experiments using different beat or wave segmentation methods are done in this part. In the following part, to simplify expression, we call the beat segmentation with fixed length BS, the non-overlapping wave segmentation with fixed-length NWS, which is based on fixed-length segmentation. And CAM represents the 1-D CNN with CAM.

In Fig. 6, the proposed ABS method outperforms BS method in accuracy obviously, the result indicates that BS leads into many noises, which is because that fixed-length segmentation method may result in beat is incomplete or redundant. For feature extraction and classification, this incompleteness or redundancy actually is a kind of noise. The accuracy of combination of ABS and OWS is not higher than 1% of that of combination of ABS and NWS, but stability has improved significantly. Moreover, we also validate the segmentation method on CAM. The combination of ABS and OWS on CAM outperforms the combination of ABS and OWS, that proves the validity of CAM. In a word, the proposed methods are more accurate and stable than related works.

Table 1 The comparison of related work. Proposed without HF means the network without hand-crafted feature. The accuracy in the parentheses is the result of 1-D CNN with CAM.

Methods	Acc
Tang et al. [25]	91.7%
Zubair et al. [5]	92.7%
Acharya et al. [26]	93.47%
Zhai et al. [27]	92.06%
Proposed without HF	95.70%
Proposed	**96.90%(96.94%)**

4.4 Results of Multi-Scale Feature

In order to prove that the multi-scale feature, i.e. inner-beat, beat, inter-beat and hand-crafted feature, is helpful for classification, a set of experiments are done. From Table 1, the accuracy of network without head-crafted feature (HF) is more than 1% lower than the accuracy of the network with it, which proves that the combination of hand-crafted feature and deep repesentation is valid.

Table 1 shown the comparison of related work. It is clear that the proposed method has a better performance and achieves a great improvement in classification. As other related work only repeats experiments for one times, so we cannot show the comparison of model stability. And the stability of the proposed model without CAM is 5.73. Though the accuracy of the 1-D CNN with CAM (shown in parentheses of Table 1) is still higher than the accuracy of the 1-D CNN without the module. But the difference between two models is much smaller, which means that multi-scale feature can make up for the shortcomings of the 1-D CNN without CAM. And the model stability of neural network with CAM is 4.81. This means that on the one hand, the models are more stable as the multi-scale feature is introduced. On the other hand, the model with CAM is more complicated and has more parameters. But to speed up the training process and make different models have the same hyperparameters, we use a few neurons, which may result in that the 1-D CNN with CAM is underfitting. So the results of the network with CAM have a bigger fluctuation. In fact, as the variance is very small (small than 10^{-5}), it is normal to have some fluctuation.

4.5 Performance on Each Type Beat

In order to analyze the improvement of the model in each type, we calculate the precision, recall and F1 score. Since the network with CAM is underfitting, we show the results of network without CAM in Table 2. Moreover, there are a large number of beats constructed on their own in [26] work and [27] only show the confusion matrix of the second 24 records from MIT-BIH database in their paper, we cannot compare with them work. We compare our work with [5]. What's more, though the number of beats for the whole database is more than 100000, there are only no more than 15 beats of type Q (unclassifiable beats), which is meaningless. So we only show the results of the following types: N

Table 2 The precision, recall and F1 score of different types based on different methods. The results of [5] method are shown in parentheses.

Type	N	S	V	F
Precision	**98.03%** (97.24%)	**72.33%** (44.66%)	**92.08%** (64.26%)	**74.59%** (64.50%)
Recall	**98.98%** (96.54%)	**57.14%** (35.08%)	**89.47%** (79.20%)	**64.06%** (61.02%)
F1 score	**0.99** (0.97)	**0.64** (0.39)	**0.91** (0.71)	**0.69** (0.63)

(normal beats), S (supraventricular ectopic beats), V (ventricular ectopic beats), and F (fusion beats). The type N has more than 77000 beats. The type S has about 2500 beats. The type S has about 6000 beats and the type F has no more than 800 beats. It is clear that the results of proposed method are superior to [5] work for all parameters, especially for the types with less samples, which also are abnormal types. For example, the precision, reall and F1 score of type S based on our methods are improved about 0.25. The other types also have a significant improvement.

5. Conclusion

In this paper, we propose a novel ECG classification method and verify its performance with extensive experiments. Our contributions are as follows: 1) we propose a novel method of sample acquisition, that is, to segment and reshape variable-length beats, waves or multi beats, which has higher accuracy than other related works and meets the requirements of fixed length without changing the extracted features; 2) to enrich the feature expression, we extract multi-scale hand-crafted and deep features i.e. inner-beat, beat and inter-beat scale, based on the proposed segmentation method; 3) in order to improve the learning ability of 1-D CNN, we introduce a modified channel attention module and prove its effectiveness. We evaluate the proposed method on MIT-BIH database, and the results show that this method is better than existing related works in effectiveness and stability. In the future, our work may be improved by more accurate beat segmentation and more adequate utilization of heart rate variability.

Acknowledgments

This work was supported in part by the National Science Foundation of China under Grant No. 61572071 and 61872031.

References

[1] F. Wu, T. Wu, and M.R. Yuce, "Design and implementation of a wearable sensor network system for iot-connected safety and health applications," 2019 IEEE 5th World Forum on Internet of Things (WF-IoT), pp.87–90, 2019.

[2] T.J. Jun, H.M. Nguyen, D. Kang, D. Kim, D. Kim, and Y.H. Kim, "ECG arrhythmia classification using a 2-D convolutional neural network," arXiv preprint arXiv:1804.06812, 2018.

[3] S. Kaplan Berkaya, A.K. Uysal, E. Sora Gunal, S. Ergin, S. Gunal, and M.B. Gulmezoglu, "A survey on ECG analysis," Biomedical Signal Processing and Control, no.43, pp.216–235, 2018.

[4] C. Ye, B.V. Kumar, and M.T. Coimbra, "Combining general multi-class and specific two-class classifiers for improved customized ECG heartbeat classification," Proceedings of the 21st International Conference on Pattern Recognition (ICPR2012), pp.2428–2431, IEEE, 2012.

[5] M. Zubair, J. Kim, and C. Yoon, "An automated ECG beat classification system using convolutional neural networks," 2016 6th international conference on IT convergence and security (ICITCS), pp.1–5, IEEE, 2016.

[6] S. Kiranyaz, T. Ince, R. Hamila, and M. Gabbouj, "Convolutional neural networks for patient-specific ECG classification," Engineering in Medicine and Biology Society (EMBC), 2015 37th Annual International Conference of the IEEE, pp.2608–2611, IEEE, 2015.

[7] Y. Xiang, J. Luo, T. Zhu, S. Wang, X. Xiang, and J. Meng, "ECG-based heartbeat classification using two-level convolutional neural network and RR interval difference," IEICE Trans. Inf. & Syst., vol.E101-D, no.4, pp.1189–1198, April 2018.

[8] E. Cimen and G. Ozturk, "Arrhythmia classification via k-means based polyhedral conic functions algorithm," Computational Science and Computational Intelligence (CSCI), 2016 International Conference on, pp.798–802, IEEE, 2016.

[9] 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, 2016.

[10] M. Kropf, D. Hayn, and G. Schreier, "ECG classification based on time and frequency domain features using random forests," Computing, vol.44, p.1, 2017.

[11] C.I.M. Cruz, J.P. Marasigan, A.P.G. Perez, J.E.V. Pillejera, N.P. Veron, and A.R. dela Cruz, "A comparative study between DWT-ANFIS and DWT-SVM in ECG classification," Region 10 Conference (TENCON), 2016 IEEE, pp.2052–2055, IEEE, 2016.

[12] A. Ebrahimzadeh, M. Ahmadi, and M. Safarnejad, "Classification of ECG signals using hermite functions and MLP neural networks," J. AI and Data Mining, vol.4, no.1, pp.55–65, 2016.

[13] C. Ye, B.V. Kumar, and M.T. Coimbra, "Heartbeat classification using morphological and dynamic features of ECG signals," IEEE Trans. Biomed. Eng., vol.59, no.10, pp.2930–2941, 2012.

[14] A. Vishwa, M. Lal, S. Dixit, and P. Vardwaj, "Clasification of arrhythmic ECG data using machine learning techniques," International J. Interactive Multimedia and Artificial Intelligence, vol.1, no.4, pp.67–70, 2011.

[15] M. Limam and F. Precioso, "Atrial fibrillation detection and ECG classification based on convolutional recurrent neural network," Computing, vol.44, p.1, 2017.

[16] 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, 2007.

[17] T. Li and M. Zhou, "ECG classification using wavelet packet entropy and random forests," Entropy, vol.18, no.8, p.285, 2016.

[18] R. Smíšek, J. Hejč, M. Ronzhina, A. Němcová, L. Maršánová, J. Chmelík, J. Kolářová, I. Provazník, L. Smital, and M. Vítek, "SVM based ECG classification using rhythm and morphology features, cluster analysis and multilevel noise estimation," Computing, vol.44, p.1, 2017.

[19] K. Kadir, A. Safie, M.I. Yusof, H. Nasir, and L. Petropoulakis, "Multiple pulse k-nearest neighbors authentication for malay ECG based class attendance system," International Conference on Engineering Technology & Technopreneuship, 2014.

[20] J. Hu, L. Shen, and G. Sun, "Squeeze-and-excitation networks," Proc. IEEE Conference on Computer Vision and Pattern Recognition, pp.7132–7141, 2018.

[21] S. Woo, J. Park, J.Y. Lee, and I. So Kweon, "CBAM: Convolutional block attention module," Proc. European Conference on Computer Vision (ECCV), pp.3–19, 2018.

[22] J. Pan and W.J. Tompkins, "A real-time QRS detection algorithm," IEEE Trans. Biomed. Eng, vol.32, no.3, pp.230–236, 1985.

[23] U.R. Acharya, K.P. Joseph, N. Kannathal, L.C. Min, and J.S. Suri, “Heart rate variability,” Advances in Cardiac Signal Processing, pp.121–165, 2018.
[24] A. ECAR, “Recommended practice for testing and reporting performance results of ventricular arrhythmia detection algorithms,” Association for the Advancement of Medical Instrumentation, p.69, 1987.
[25] X. Tang and L. Shu, “Classification of electrocardiogram signals with RS and quantum neural networks,” International J. Multimedia and Ubiquitous Engineering, vol.9, no.2, pp.363–372, 2014.
[26] U.R. Acharya, S.L. Oh, Y. Hagiwara, J.H. Tan, M. Adam, A. Gertych, and R. San Tan, “A deep convolutional neural network model to classify heartbeats,” Comput. Biol. Med., vol.89, pp.389–396, 2017.
[27] X. Zhai and C. Tin, “Automated ECG classification using dual heartbeat coupling based on convolutional neural network,” IEEE Access, vol.6, pp.27465–27472, 2018.

Huan Sun received her B.S. degree in School of Electronics and Information Engineering at the Beijing Jiaotong University, in 2017.

Yuchun Guo received her B.S. and M.S. degrees from the Department of Electronics, Beijing Normal University, Beijing, China in 1990 and 1993. After that she has been with the School of Electrical and Information Engineering, Northern Jiaotong University, Beijing China. She spent an academic year at Delft University of Technology, the Netherlands, as a visiting scientist, sponsored by NUFFIC (Netherlands organization for cooperation in higher education) and CSC (Chinese scholarship council). She started her PhD study in October 2002 in Northern Jiaotong University with a focus on network measurement and modelling.

Yishuai Chen received his B.S., M.S., and Ph.D. degrees from School of Electronics and Information Engineering at the Beijing Jiaotong University, in 1998, 2001, and 2010, respectively. He is currently an Associate Professor at the School of Electrical and Information Engineering of the Beijing Jiaotong University. From 2010 to 2012, he was a Postdoctoral Fellow with the Research Center of Ubiquitous Sensor Networks at the University of Chinese Academy of Sciences, Beijing, China. From 2001 to 2007, he worked at Lucent Bell-labs (China) on intelligent network systems as a Technical Staff. His research interests include peer to streaming media, web and internet services, and intelligent learning.

Bin Chen received his B.S. and M.S. degrees in School of Electronics and Information Engineering at the Beijing Jiaotong University, in 2016, 2019, respectively.