Detection of First-Degree Atrioventricular Block on Variable-Length Electrocardiogram via a Multimodal Deep Learning Method

Dongya Jia[1], Wei Zhao[1], Jing Hu[1], Hongmei Wang[1], Cong Yan[1], Zhenqi Li[1]

[1] Central Research, Guangzhou Shiyuan Electronic Technology Company Limited, Guangzhou, China

Abstract

Automatic detection of first-degree atrioventricular block (I-AVB) from electrocardiogram (ECG) is of great importance in prevention of more severe cardiac diseases. I-AVB is characterized by a prolonged PR interval. However, due to various artifacts and diversity of ECG morphology, existing ECG delineation algorithms is unable to provide robust measurement of the PR interval. Deep neural network is good at extracting high-level feature from ECG waveform, but merely using waveform as input of neural network may aggravate overfitting when lack of I-AVB records. In this paper, we propose a multimodal-input deep learning method to effectively detect I-AVB from 12-lead ECG records. We utilize ECG waveform and delineation result as the multimodal input of neural network. Our neural network, mainly composed of convolutional neural network and Long Short-term Memory, is well designed to adapt to variable-length ECG. Our method is evaluated on dataset of CPSC2018, and outperforms the baseline methods in F1 score.

1. Introduction

The electrocardiogram (ECG), which measures electrical activity of heart, is an important tool for clinical diagnosis of multiple cardiac diseases. From the end of the P wave to the beginning of the QRS complex is called the PR interval, indicating the pulse conduction time between the atria and ventricles. AV block (AVB) is a type of heart block occurring in the transmission between the atria and ventricles[1]. In terms of severity, AVB is divided into three degrees. The first-degree AVB (I-AVB) is characterized by a prolonged PR interval, usually greater than 0.2 seconds[1]. Although most patients with I-AVB are asymptomatic, but I-AVB is associated with more severe degree of AVB.

The development of I-AVB automatic detection method can reduce workloads and increase work efficiency for cardiologists. Automatic measurement of the PR interval is an explicit way to distinguish I-AVB. Up to now, various ECG delineation methods are developed. In [2], Martinez uses discrete wavelet transform (DWT) to enhance ECG waveform and apply an adaptive threshold to detect ECG fiducial points. In [3] and [4], some frequency features are extracted and fed into a Hidden Markov Model (HMM) to model the transition process between different waves. However, due to effect of artifacts and diversity of ECG morphology associated with cardiac diseases, it is not robust enough to detect I-AVB merely based on the delineation result.

In recent years, utilizing deep learning method to automatically detect abnormalities from ECG in a end-to-end manner, has been widely investigated. These methods directly analyze ECG waveform and output the classification result. In [5], a Convolutional Neural Network (CNN) is proposed to classify the single-lead ECG into multiple kinds of rhythm, and achieved performance comparable to the average level of cardiologist. In [6], Yao proposed an attention-based time-incremental CNN to address multi-class arrhythmia detection from 12-lead ECG. In [7], features from CNN are integrated with some expert features, and then fed into a XGBoost[8] classifier to identify multiple abnormalities. Detection of I-AVB needs accurate measurement of the PR interval, merely using ECG waveform as input of deep neural network may introduce redundancy information and aggravate overfitting.

In this paper, we propose a multimodal deep learning method to detect I-AVB from variable-length 12-lead ECG records. We use [9]' s algorithm to delineate ECG records, the delineation result is concatenated with 12-lead ECG waveform as the input of out neural network. The structure of our network is composed of CNN and Long Short-term Memory[10]. Our method can deal with variable-length ECG without any segmentation or padding process. The method is evaluated on the dataset of the China Physiological Signal Challenge 2018[11], which contains multiple abnormalities such as Atrial fibrillation and Premature ventricular contraction. Our method achieves an average F1-score of 0.889 over 5-fold cross validation, exceeding state-of-the-art methods.

 ISSN: 2325-887X DOI: 10.22489/CinC.2019.194

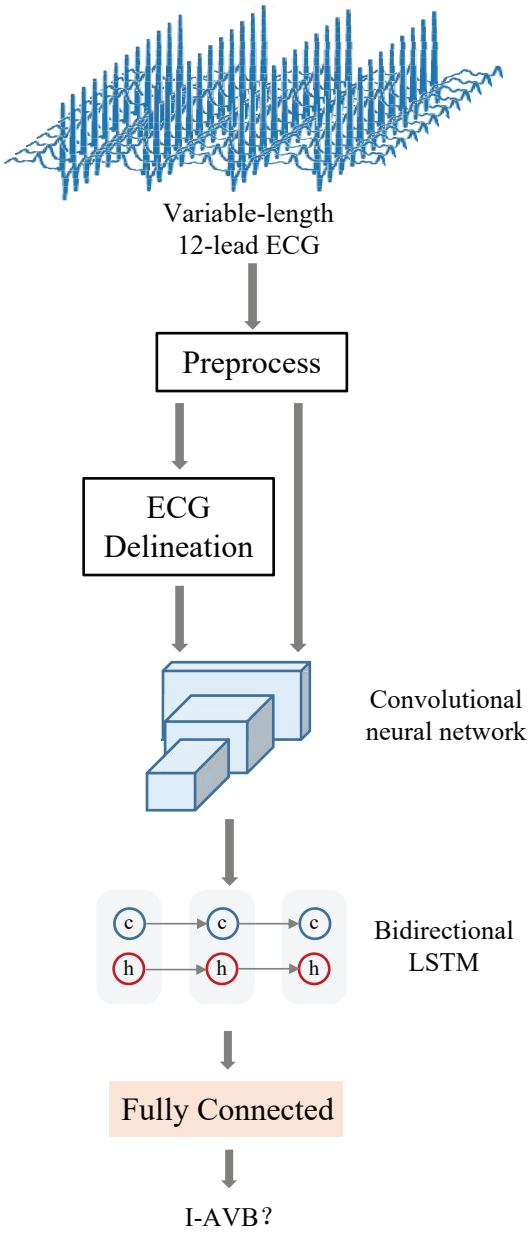

Figure 1. The framework of our scheme.

2. Methods

The framework of our scheme is shown in Figure 1. Our scheme is composed of preprocessing, ECG delineation and classification by neural network. Each variable-length 12-lead ECG record is preprocessed first. Then ECG delineation algorithm is applied to ECG waveform yielding locations of P wave, QRS complex and T wave. The delineation result and ECG waveform are combined and fed into our well-designed neural network for classification.

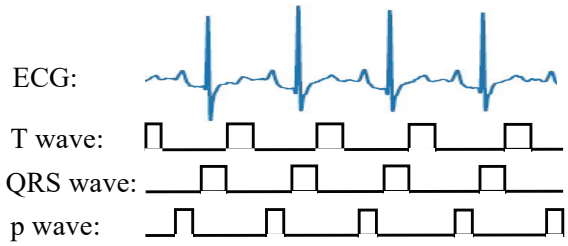

Figure 2. An example of ECG delineation result of three characteristic waves.

2.1. ECG preprocessing and delineation

To remove the high-frequency noise and baseline wandering noise, each lead of ECG is processed by a bandpass Butterworth filter, with a range of 0.5Hz–48Hz. Then, each ECG recording is delineated by algorithm[9], so as to obtain the location of ECG fiducial points, including the onset and offset of the P wave, the onset and offset of the QRS complex and the offset of the T wave.

The delineation result is transformed into three channels, each of which represents a kind of characteristic wave. An example is shown in Figure 2. The three delineation channels are concatenated with 12-lead ECG signal together, as the multimodal input of our neural network.

2.2. Network architecture

An overview of the proposed network architecture is shown in Figure 3. Our neural network is composed of a CNN, a bidirectional LSTM and a fully connected layer.

For each convolutional layer, the kernel size is 15 and the stride is 1. Additionally, each convolutional layer is padded with 0 to keep the length of feature map unchanged. Batch normalization layer is utilized after each convolutional layer to lessen internal covariate shift and accelerate training process[12]. ReLU activation function is adopted to provide nonlinearity for the model. The block of Conv-BN-ReLU-Conv-BN is repeated for times, and shortcut connection is adopted to handle degradable problem in deep neural network[13].

After the CNN, a bidirectional LSTM with hidden size of 128 is adopted to extract long-term correlation between features far apart. In typical routine, the inputs are resized to the same size, and the feature map is averaged globally along the time dimension before transmitting to the fully connected layer. In our case, to address variable-length inputs, each proportion of feature map along the time dimension is processed individually by one identical fully connected layer, and their results are averaged to yields a final prediction.

Table 1. F1 score of I-AVB on CPSC2018 dataset.

Methods	F1 score
Ours	0.889
VGG-60[6]	0.776
TI-CNN[6]	0.809
ATI-CNN[6]	0.850
Expert+Deep[7]	0.870

2.3. Implementation details

The weights of convolutional layers, LSTM and fully connected layer are initialized using Xavier initialization method[14]. Initial hidden state and cell state of LSTM are initialized with zero. Cross entropy loss is used and our model is optimized using Adam optimizer[15]. The total epoch of training is 40. The learning rate is initially set to 0.001 and multiplied by 0.1 at the 25th epoch.

The L2 regularization with a factor of 0.001 is adopted to alleviate overfitting. At training phase, each ECG record is split into 6-second length segments so as to form mini-batch and accelerate the training process. At test phase, no segmentation is applied to ECG records.

3. Experiments and Discussions

3.1. Dataset Description

The dataset is from the China Physiological Signal Challenge 2018[11]. The ECG recordings were collected from 11 hospitals. The dataset is split into training set and test set, but the labels of test set are unreleased. So our experiment is conducted on the training set. The official training set contains 6877 12-lead ECG recordings lasting from 6s to 60s. There are totally 1098 recordings labeled I-AVB. Apart from I-AVB, some other abnormalities also exist in the dataset, including atrial fibrillation, left bundle branch block, right bundle branch block, premature atrial contraction, premature ventricular contraction, ST-segment depression and ST segment elevated. In our experiment, all recordings without I-AVB label are regarded as non I-AVB.

3.2. Results

We adopt F1 score as our evaluation metric, which considers both the recall and the precision. It is calculated by harmonically averaging the recall and precision, as shown in Equation(1).

$$F_1 \text{ score} = \frac{2 \cdot \text{precision} \cdot \text{recall}}{\text{precision} + \text{recall}} \quad (1)$$

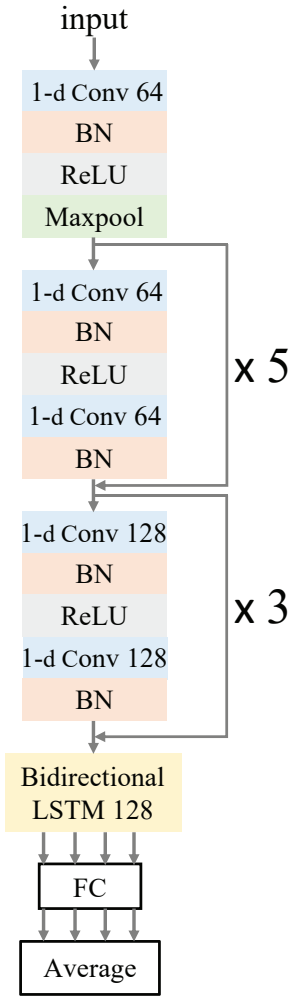

Figure 3. The architecture of our neural network. The input has 15 channels, including 12-lead ECG waveform and 3 channels of delineation result. The numbers (64 and 128) represent output channels of convolutional layer and bidirectional LSTM.

We perform 5-fold cross-validation on the official training set to evaluate our proposed method, the experimental results are shown in Table 1. The reference methods are [7] and [6]. Their experiments are conducted in a multi-class classification manner, the reported F1 score of I-AVB is used for comparison. As can be seen from Table 1, on F1 score of I-AVB detection, our method exceeds the baseline methods with a large margin.

3.3. The Influence of Multimodal Input

I-AVB is clinically recognized by the prolonged PR interval, which is the distance between the onset of the P wave and the onset of the QRS complex. ECG delineation

Table 2. Comparison between different inputs.

Methods	F1 score
Waveform input	0.881
Multimodal input	0.889

algorithms is able to detect location of multiple kinds of fiducial points, such as the onset of the P wave. But due to effect of various artifacts and noise, it is not robust enough to detect I-AVB only based on the delineation result. Deep learning method can extract high-level features from signal. Since I-AVB is only related to the PR interval, end-to-end learning only using ECG waveform may make model focus more on the other portion of ECG morphology and lead to overfitting.

In our method, apart from 12-lead ECG waveform, some auxiliary channels, which are transformed from the delineation result, are concatenated with the original channels as the input of network. The experimental results show that multimodal input yields better performance on 5-fold cross-validation, as shown in Table 2.

4. Conclusion

In this paper, we propose a novel multimodal deep learning method to detect I-AVB from 12-lead variable-length ECG records. Apart from ECG waveform, the delineation results of ECG are also adopted as part of input of our neural network. The architecture of our neural network consists of CNN and LSTM. To cope with variable-length records, output of all time steps of LSTM are utilized to compute final predictions. From the experimental results, the proposed method outperforms other baseline methods with a large margin. In the future, we plan to expand our method to multi-class classification problem and apply to other abnormalities on ECG.

References

[1] Surawics B, Childers R, Deal BJ, Gettes LS. Aha/accf/hrs recommendations for the standardization and interpretation of the electrocardiogram. Circulation 2009;119(10):235–240.

[2] Martínez JP, Almeida R, Olmos S, Rocha AP, Laguna P. A wavelet-based ECG delineator: evaluation on standard databases. IEEE Trans Biomed Engineering 2004; 51(4):570–581.

[3] Andreão RV, Dorizzi B, Boudy J. ECG signal analysis through hidden markov models. IEEE Trans Biomed Engineering 2006;53(8):1541–1549.

[4] Hughes NP, Tarassenko L, Roberts SJ. Markov models for automated ECG interval analysis. In Advances in Neural Information Processing Systems 16 [Neural Information Processing Systems, NIPS 2003, December 8-13, 2003, Vancouver and Whistler, British Columbia, Canada]. 2003; 611–618.

[5] Rajpurkar P, Hannun AY, Haghpanahi M, Bourn C, Ng AY. Cardiologist-level arrhythmia detection with convolutional neural networks. CoRR 2017;abs/1707.01836.

[6] Yao Q, Wang R, Fan X, Liu J, Li Y. Multi-class arrhythmia detection from 12-lead varied-length ECG using attention-based time-incremental convolutional neural network. Information Fusion 2020;53:174–182.

[7] Liu Z, Meng X, Cui J, Huang Z. Automatic identification of abnormalities in 12-lead ecgs using expert features and convolutional neural networks. In International Conference on Sensor Networks and Signal Processing, October 28-31, 2018, Xi'an, China. 2018; 163–167.

[8] Chen T, Guestrin C. Xgboost: A scalable tree boosting system. In Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, San Francisco, CA, USA, August 13-17, 2016. 2016; 785–794.

[9] Jia D, Zhao W, Li Z, Hu J, Yan C, Wang H, You T. An electrocardiogram delineator via deep segmentation network. In Engineering in Medicine and Biology Conference, July 23-27, 2019, Berlin, Germany. 2019, accepted; .

[10] Hochreiter S, Schmidhuber J. Long short-term memory. Neural Computation 1997;9(8):1735–1780.

[11] Liu F, Liu C, Zhao L, Zhang X, Wu X, Xu X, Liu Y, Ma C, Wei S, He Z, Li J, Kwee Y, Ng E. An open access database for evaluating the algorithms of electrocardiogram rhythm and morphology abnormality detection. Journal of Medical Imaging and Health Informatics 2018;8(7):1368–1373.

[12] Ioffe S, Szegedy C. Batch normalization: Accelerating deep network training by reducing internal covariate shift. In Proceedings of the 32nd International Conference on Machine Learning, ICML 2015, Lille, France, 6-11 July 2015. 2015; 448–456.

[13] He K, Zhang X, Ren S, Sun J. Deep residual learning for image recognition. In 2016 IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2016, Las Vegas, NV, USA, June 27-30, 2016. 2016; 770–778.

[14] Glorot X, Bengio Y. Understanding the difficulty of training deep feedforward neural networks. In Proceedings of the Thirteenth International Conference on Artificial Intelligence and Statistics, AISTATS 2010, Chia Laguna Resort, Sardinia, Italy, May 13-15, 2010. 2010; 249–256.

[15] Kingma DP, Ba J. Adam: A method for stochastic optimization. In 3rd International Conference on Learning Representations, ICLR 2015, San Diego, CA, USA, May 7-9, 2015, Conference Track Proceedings. 2015; .

Address for correspondence:

Dongya Jia

No.6 Yunpu 4th Road, Huangpu District, Guangzhou, China

jiadongya@cvte.com