Multi-Label Classification of Multi-lead ECG Based on Deep 1D Convolutional Neural Networks With Residual and Attention Mechanism

Yamin Liu[1], Hanshuang Xie[1], Qineng Cao[1], Jiayi Yan[1], Fan Wu[1], Huaiyu Zhu[2], Yun Pan[2]

[1]Research and Development Department, Hangzhou Proton Technology, Co., Ltd., Hangzhou, China
[2]College of Information Science and Electronic Engineering, Zhejiang University, Hangzhou, China

Abstract

Arrhythmia automatic analysis techniques provide convenience for the prevention and diagnosis of cardiac disease. Aiming at classify cardiac abnormalities into 27 classes with either 12-lead, 6-lead, 4-lead, 3-lead or 2-lead multi-label ECG recordings, we develop a deep 1-Dimensional Convolutional Neural Network (1D CNN) with residual block and squeeze-and-excitation (SE) attention mechanism (namely 1D RANet). First, we introduce residual block and SE attention block into 1D CNN to extract deep features adaptively and avoid vanishing gradient or network degradation. Second, to improve the robustness, we use data augmentation techniques such as band-pass and wavelet-based filter, noise addition, time-frequency switch and so on. For the convenience of training, all recordings are cropped or padded to 60 seconds with resampled rates of 300 Hz. Our classifiers received challenge scores of 0.73, 0.53, 0.59, 0.57 and 0.47 for the 12-lead, 6-lead, 4-lead, 3-lead, and 2-lead versions of the public training set, 0.56, 0.54, 0.56, 0.54 and 0.55 for the hidden validation set and 0.52, 0.49, 0.51, 0.50 and 0.49 for the hidden test sets (Team name: Proton). However, our entries did not rank since we missed the preprint submitting deadline of the Computing in Cardiology Conference.

1. Introduction

Cardiovascular disease is one of the most serious diseases that endanger human health [1]. Therefore, a rapid and accurate diagnosis of arrhythmia is of great importance for the prevention and treatment of heart disease. The electrocardiogram (ECG) have been widely used to cardiovascular disease diagnosis as a physiology signals generated by the heart's excitement. The PhysioNet/Computing in Cardiology Challenge 2021 encourages challengers to design effective multi-label ECG arrhythmia classification algorithm of multi-lead ECG signals including 12, 6, 4, 3, and 2 leads [2, 3].

In recent years, with the rapid development of artificial intelligence theory, researches on arrhythmia detection techniques based on machine learning have emerged one by one. The arrhythmia detection algorithm based on machine learning mainly consist of pre-process of ECG signals, feature extraction and arrhythmia classification. However, the feature extraction methods severely depend on the artificial experience and expert knowledge, which results in the limitation of arrhythmia classification. The ability of deep learning to adaptively extract features provides new solution to such problems. Compared with the traditional shallow models, deep neural networks like convolutional neural network (CNN) [4] and recurrent neural network (RNN) [5] are able to adaptively extract deep features with strong representation ability from raw data, which avoids the problem such as insufficient or redundant of extracted hand-crafted features and achieves good performance in arrhythmia detection as well. For example, Pourbabaee et al. [4] proposed an end-to-end atrial fibrillation detection method based on CNN, and verified the strong representation ability of deep features compared with the traditional hard-crafted features. Hunnun et al. [6] and Rajpurkar et al. [7] proposed a 34-layer CNN with residual blocks (ResNet34) algorithm for arrhythmia detection of single-lead ECG signals and obtained good performance. Wang et al. [8] came out with a 12-lead ECG arrhythmia detection algorithm based on 1-dimensional (1D) CNN.

Inspired by the researches above, we proposed an end-to-end multi-lead ECG multi-label arrhythmia detection method based on 1D CNN with residual blocks and attention mechanism (1D RANet). Detailed information of the proposed method is described in Section 2.

The database used in Cardiology Challenge 2021 including the CPSC database [9], the INCART database [10], the PTB database [11], the PTB-XL database [12], the Chapman-Shaoxing Database [13], and the Ningbo Database [14]. Each ECG recording of the above dataset has one or more labels that describe cardiac abnormalities.

 ISSN: 2325-887X DOI: 10.22489/CinC.2021.035

2. Methods

2.1. Description of dataset

The dataset used in Cardiology Challenge 2021 have been presented in Section 1. And for more details, the challengers are required to train five models corresponding to the ECG recording of 12-lead, 6-lead, 4-lead, 3-lead and 2-lead. The reduced-lead ECG recordings are obtained according to the references [2,3,15].

In this paper, we trained five models based on the proposed basic model 1D RANet using the data of different leads, respectively. Therefore, we only take the 12-lead as an example to display our work on the data preprocessing, model construction and experiment results.

2.2. Data preprocessing

The sampling rate of these different datasets is different. For the convenience of the next data processing, we resampled the data at 300 Hz.

2.2.1. Data augmentation

To enhance the diversity of samples and improve the robustness of model, we adopted data augmentation techniques including band-pass filter (0.05 Hz - 100 Hz), noise addition, time-frequency transform and data selection. Take one of the samples (JS26892) as an example, Figure 1 provides an intuitive display about the raw ECG sample and the corresponding samples after applied data augmentation. Figure 1 (a) means the raw ECG signals in time domain, and Figure 1 (b)-(d) represent the filtered signal, the signal added random noise and the spectrum after fast Fourier transform (FFT) respectively. When the signal length is more than 60s, we cut it into multiple pieces of data with a sliding window of 10s, and then randomly select one of them for training. It is obvious that the data augmentation is able to enlarge the diversity of samples and provide more useful information which contributes to improve the robustness of model.

2.2.2. Data segmentation

For the convenience of model training, we treat all ECG data to the same length. According to out statistics, the length and number of ECG signals are shown in Table 1. Although the most ECG signals length is less than 20s, we have verified that our model can obtain the best performance when the length of data is set 60s. This may be due to the data of 20s discarding too much useful information when training the model. Thus, we finally process the data into 60s. Specifically, the signals less than 60s are padded by zero into 60s, the signals more than 60s

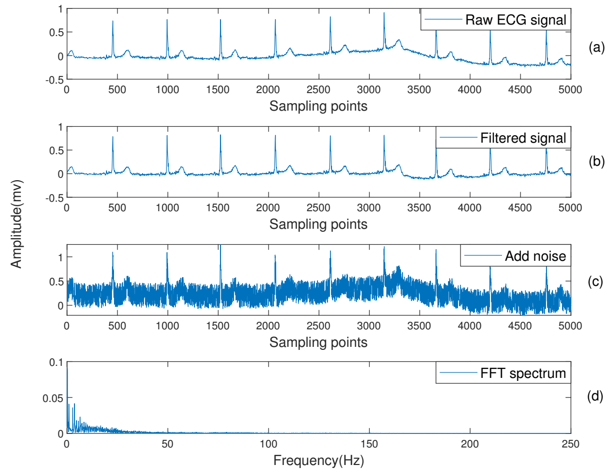


Figure 1: Illustration of data and the corresponding data after data augmentation

Table 1: The length and number of ECG signals.

Length	<20s	<30s	<60s	<180s	30min
Number	85549	1194	878	558	74

are processed by the data selection method mentioned in Section 2.2.1.

2.3. Model construction

This paper introduces residual blocks and SE attention mechanism [16] into the 1D CNN, the network structure is shown in the left of Figure 2. The SE attention block and residual block are shown in the right of Figure 2. In order to simplify model training and avoid over-fitting, we add a Dropout layer after each Basicblock, which is set to 0.2.

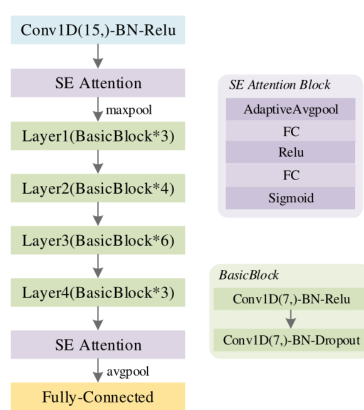


Figure 2: Structure of proposed model

3. Results

3.1. Experiment setting

This paper firstly conducts data resampling and data segmentation, and then uses multi-lead ECG signals with a sampling rate of 300 Hz and a length of 60s (18000 sampling points) as the input of the proposed model. Finally construct an end-to-end 1D CNN with attention mechanism and residual block which classifies the ECG signals into 27 categories, including 26 arrhythmia types and sinus rhythm.

In the experiment, we divided the dataset provided into training set, validation set and test set in proportion of 8:1:1. During the model training, the batch size is set to 64, learning rate is set to 0.01. The model is optimized by stochastic gradient descent algorithm with sigmoid active function and cross-entropy loss. The model proposed in this article is built under the Pytorch framework and trained on NVIDIA RTX 2080Ti GPU.

3.2. Experiment results

To evaluate the method proposed, we use accuracy, F-measure and the challenge metric as evaluation indictors. The comparative results on the public dataset of 12-lead model are shown in Table 2. The challenge metric results of different leads obtained by the proposed method (1D RANet) are shown in Table 3.

Table 2: The comparative results of 12-lead model on validation set. The Challenge metric score is obtained from the testing set by splitting the public dataset provided by the committee.

	Accuracy	F-measure	Challenge metric
ResNet34	0.42	0.58	0.71
ResNet34 with focal loss [17]	0.42	0.59	0.70
Resnet34 with data augmentation	0.46	0.61	0.72
1D RANet with 20s segments	**0.54**	0.63	0.72
1D RANet (proposed method)	0.50	**0.67**	**0.73**

Table 3: Challenge scores for our final selected entry (team Proton) on the testing set divided from the public dataset, repeated scoring on the hidden validation set, and one-time scoring on the hidden test set.

Leads	Our testing set	Hidden validation	Hidden test	Ranking[1]
12	0.73	0.562	0.52	N/A
6	0.53	0.54	0.49	N/A
4	0.59	0.56	0.51	N/A
3	0.57	0.54	0.50	N/A
2	0.47	0.55	0.49	N/A

[1] Our entries did not rank since we missed the preprint submitting deadline of the Computing in Cardiology Conference.

4. Discussion and Conclusions

4.1. Discussion

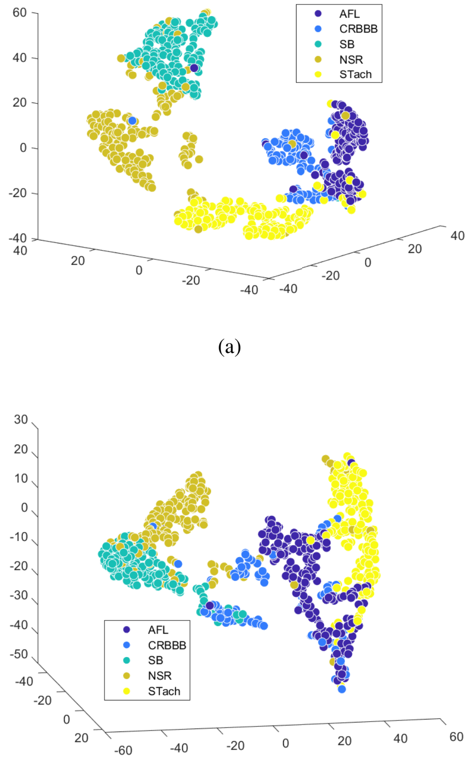
It is easy to see from Table 2 and Table 3 that the proposed method 1D RANet outperforms the other comparative methods in terms of accuracy, F-measure and challenge metric, which indicates that residual blocks and attention mechanism can improve the effectiveness of 1D CNN. Meanwhile, it also shows that the data augmentation techniques can promote the generalization performance of the models.

To further identify the effectiveness of the SE attention mechanism in the proposed method, we visualize the extracted 512-dimentional deep features from our method and the model without SE module by using t-SNE algorithm [18]. For easy observation, we only choose atrial flutter (AFL), complete right bundle branch block (CRBBB), sinus bradycardia (SB), sinus tachycardia (Stach) four kinds of arrhythmias and sinus rhythm (NSR) from the given 27 classes to visualize as shown in Figure 3. Figure 3 (a) visualizes the features extracted from the proposed method and Figure 3 (b) shows the features extracted from 1D CNN without SE attention mechanism. The axis in Figure 3 represents the 3-dimentional features obtained by t-SNE dimension reduction. From Figure 3 we can see that SE attention mechanism contributes to extract deep features with stronger representation by learning the importance of different channels.

4.2. Conclusions

In conclusion, this paper proposed a multi-label arrhythmia classification algorithm of multi-lead ECG signals based on 1D CNN with residual and SE attention mechanism. The method 1D RANet firstly improves the performance of deep neural network by introducing residual block to tackle with the problems of vanishing/exploding gradient. Secondly, it learns the importance of different channels through the SE attention mechanism to help extract deep features with strong representation which is contributes to arrhythmia classification of ECG signals. Finally, the method improves the generalization ability of the model by increasing the diversity of samples through band-pass filtering, adding noise and other data augmentation techniques.

At present, the automatic detection of multi-label arrhythmia of multi-lead ECG has attracted much attention. In the future research, we will consider how to combine the artificial features and deep features to design a more efficient arrhythmia detection model.



(a)

(b)

Figure 3: Feature visualization: (a) proposed method, (b) 1D CNN without SE module

References

[1] Heart Disease and Stroke Statistics-2019 Update: A Report From the American Heart Association[J]. Circulation, 2019, 139(10):E56-E528.

[2] Perez Alday EA, Gu A, Shah A, Robichaux C, Wong AKI, Liu C, et al. Classification of 12-lead ECGs: the PhysioNet/Computing in Cardiology Challenge 2020. Physiological Measurement 2020, 41.

[3] Reyna MA, Sadr N, Perez Alday EA, Gu A, Shah A, Robichaux C, et al. Will Two Do? Varying Dimensions in Electrocardiography: the PhysioNet/Computing in Cardiology Challenge 2021. Computing in Cardiology 2021, 48:1–4.

[4] Pourbabaee B, Roshtkhari M J, Khorasani K. Deep Convolutional Neural Networks and Learning ECG Features for Screening Paroxysmal Atrial Fibrillation Patients[J]. IEEE Transactions on Systems, Man, and Cybernetics: Systems, 2018, 48(12):2095-2104.

[5] Huesken M, Stagge P. Recurrent Neural Networks for Time Series Classification[J]. Neurocomputing, 2003, 50(Jan):223-235.

[6] Hannun A Y, Rajpurkar P, Haghpanahi M, et al. Cardiologist-level Arrhythmia Detection and Classification in Ambulatory Electrocardiograms Using a Deep Neural Network[J]. Nature medicine, 2019, 25(1):65-69.

[7] Rajpurkar, Pranav, et al. Cardiologist-level Arrhythmia Detection with Convolutional Neural Networks[J]. 2017, arXiv preprint arXiv:1707.01836.

[8] Wang C, Yang S, Tang X, et al. A 12-lead ECG Arrhythmia Classification Method Based on 1D Densely Connected CNN[M]//Machine Learning and Medical Engineering for Cardiovascular Health and Intravascular Imaging and Computer Assisted Stenting. Springer, Cham, 2019, 11749:72-79.

[9] Liu F, Liu C, Zhao L, Zhang X, Wu X, Xu X, et al. 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.

[10] Tihonenko V, Khaustov A, Ivanov S, et al. St Petersburg INCART 12-lead arrhythmia database[J]. PhysioBank, PhysioToolkit, and PhysioNet, 2008.

[11] Bousseljot R, Kreiseler D, Schnabel A. Nutzung der EKG Signaldatenbank CARDIODAT der PTB uber das Internet. Biomedizinische Technik 1995, 40(S1):317–318.

[12] Wagner P, Strodthoff N, Bousseljot RD, Kreiseler D, LunzeFI, Samek W, et al. PTB-XL, a Large Publicly Available Electrocardiography Dataset. Scientific Data 2020, 7(1):1–15.

[13] Zheng J, Zhang J, Danioko S, Yao H, Guo H, Rakovski C. A 12-lead Electrocardiogram Database for Arrhythmia Research Covering More Than 10,000 Patients. Scientific Data 2020, 7(48):1–8.

[14] Zheng J, Cui H, Struppa D, Zhang J, Yacoub SM, et al. Optimal Multi-Stage Arrhythmia Classification Approach. Scientific Data 2020, 10(2898):1–17.

[15] Goldberger AL, Amaral LA, Glass L, Hausdorff JM, Ivanov PC, Mark RG, et al. PhysioBank, PhysioToolkit, and PhysioNet: Components of a New Research Resource for Complex Physiologic Signals. Circulation 2000, 101(23):e215–e220.

[16] Jie H, Li S, Gang S, et al. Squeeze-and-Excitation Networks[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2017, 42(8):2011-2023.

[17] Lin T Y, Goyal P, Girshick R, et al. Focal Loss for Dense Object Detection[C]// IEEE Transactions on Pattern Analysis & Machine Intelligence. IEEE, 2017:2999-3007.

[18] Pezzotti N, Lelieveldt B, Laurens V, et al. Approximated and User Steerable tSNE for Progressive Visual Analytics[J]. 2015, 23(7):1739-1752.

Address for correspondence:

Yamin Liu
Room A-405, Innovation Building, East Software Park, No.90
Wensan Road, Hangzhou, Zhejiang, China
yamin.liu@protontek.com