Multi-label Classification of Abnormalities in 12-Lead ECG Using Deep Learning

Ao Ran[1], Dongsheng Ruan[2], Yuan Zheng[3], Huafeng Liu[1,*]

[1]College of Optical Engineering, Zhejiang University, China
[2] College of Computer Science and Technology, Zhejiang University, China
[3]College of Aeronautics and Astronautics, Zhejiang University, China

Abstract

Identifying arrhythmias from electrocardiogram(ECG) signals remains an intractable challenge. This study aims to develop an effective and non-invasive approach to realize the recognition of arrhythmias based on 12-lead ECG for the PhysioNet/Computing in Cardiology Challenge2020. To this end, we propose a deep learning-based diagnosis approach, called EASTNet which captures the characteristics of cardiac abnormalities and correlation between heartbeats sampled randomly from 12-lead ECG records by a 34-layer 1D-deep squeeze-and-excitation network. Experimenting in the multi-label arrhythmia classification task, our team, EASTBLUE, was unable to rank and score in the hidden validation and test sets, but achieved diagnostic performance with 0.7030 ± 0.0090 metric score using 5-fold cross-validation on the training set. We also investigate the effect of beat sampling on diagnostic performance, and find that the beat sampling plays a role in data augmentation that effectively alleviates network overfitting. These results demonstrate that our approach has good potential application prospects in clinical practice, especially in the auxiliary diagnosis of abnormalities.

1. Introduction

According to the World Health Organization, cardiac abnormalities are the leading cause of death globally: they kill more people each year than any other cause of death. Cardiac abnormality is a kind of common circulatory system disease, which includes atrial fibrillation, premature beat and other arrhythmias. These symptoms are usually examined and classified clinically by means of non-invasive electrocardiogram (ECG). Among them, 12-lead ECG is the gold standard in diagnosing cardiovascular abnormalities, which can more accurately capture abnormalities and timely identify various cardiovascular abnormalities. In view of the large amount of data collected and analyzed by 12-lead dynamic ECG

examination, automatic ECG classification algorithm is required to be both accurate and fast, which is also the focus of our study. The PhysioNet/Computing in Cardiology Challenge2020 focus on intelligent diagnosis of 12-lead ECG[1].

Before this, a variety of machine learning and data processing methods have been used to improve the accuracy of ECG detection. In[2], the author proposed a method of ECG signal classification combining wavelet transform and decision tree classification, and used MIT-BIH arrhythmia database as ECG signals for classification. Finally, the accuracy of the system reached 96.31%. In[3], the author proposed an automatic arrhythmia classification system based on Artificial Neural Network (ANN). The experimental results showed that the classification accuracy is 86.67%. In[4],a 34-layer deep neural network was trained to classify 12 rhythm classes with 91,232 single-lead ECGs and achieved a better result than cardiologists.
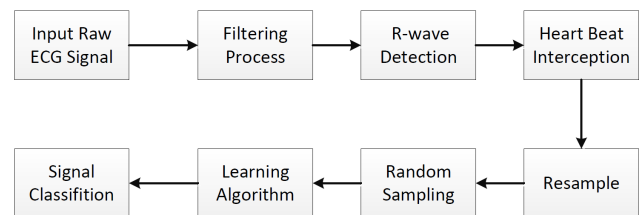

Figure 1. Flow chart of the proposed methodology

In our paper, we develop a novel deep residual network (ResNet) [5] with multi-layer perceptron [6] and 1D-Squeeze-and-Excitation (SE) [7] architectures. The experiments show our method is able to achieve high challenge score on training set. Figure 1 is the framework for the proposed system.

2. Methods

Currently, the popular method is to directly learn the entire ECG signal or to learn with a single heartbeat after the heartbeat is intercepted. However, such a method has the problems of low learning efficiency, low data

utilization and data imbalance.

In response to the above problems, we propose a data enhancement method for heartbeat sampling. Specifically, we first intercept a single heartbeat, and then combine the heartbeats in a random manner. After experimentation, we select four heartbeats as input, and finally achieve outstanding results. The advantage of this is that we effectively solve the problem of data imbalance, allowing the network to fully learn the characteristics of each cardiac abnormalities. At the same time, our data volume has also been significantly enhanced so that we can well solve the problem of model overfitting. In addition, our method combines the characteristics of multiple heartbeats to make the judgment of the model more comprehensive and accurate.

2.1. Dataset

Since our team EASTBLUE failed to score and rank on the hidden validation and test sets, we had to use the 5-fold cross-validation method on the training set to test our model. First of all, we take the removal of erroneous and unnecessary data as the first step in data processing. In addition, we collect statistics on the patient data of each cardiac abnormalities and then selectively discard and copy, so as to balance the data distribution as much as possible. Finally, we get 6720 valid records with 201600 heartbeats.
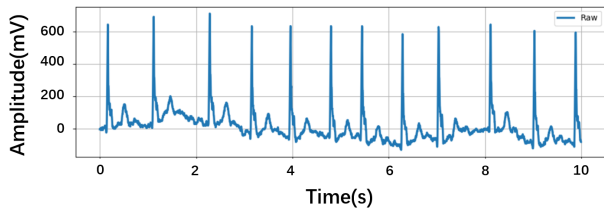
2.2. Data Processing
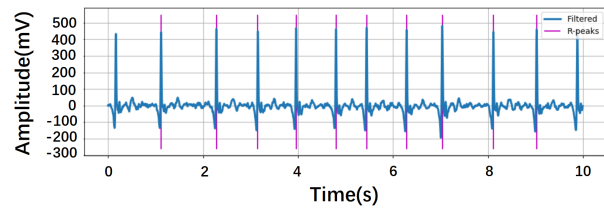

Figure 2. unprocessed waveforms of lead II

Figure 3. processed waveforms of lead II

Firstly, after eliminating baseline drift, we filter the original ECG signals (Figure 2) by 3~45 Hz finite impulse response band-pass filter. Then we obtain the position of the R peak in the lead record by R-peak[8] detection (Figure 3).Finally according to the specific position of the R peak in a given signal, extending a

certain length in the front and back, the heartbeat can be intercepted for use in the following algorithm. After the heartbeat is intercepted, we can resample processed signal by resample function. Finally, we get some records with a sampling point of 300 as the Figure 4 shows.
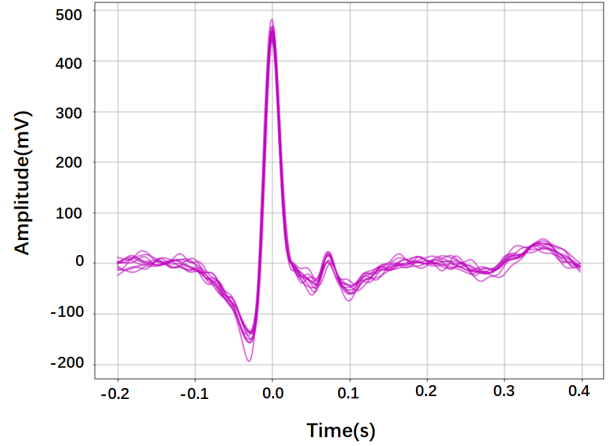

Figure 4. An example of ECG segments after heartbeat segmentation.
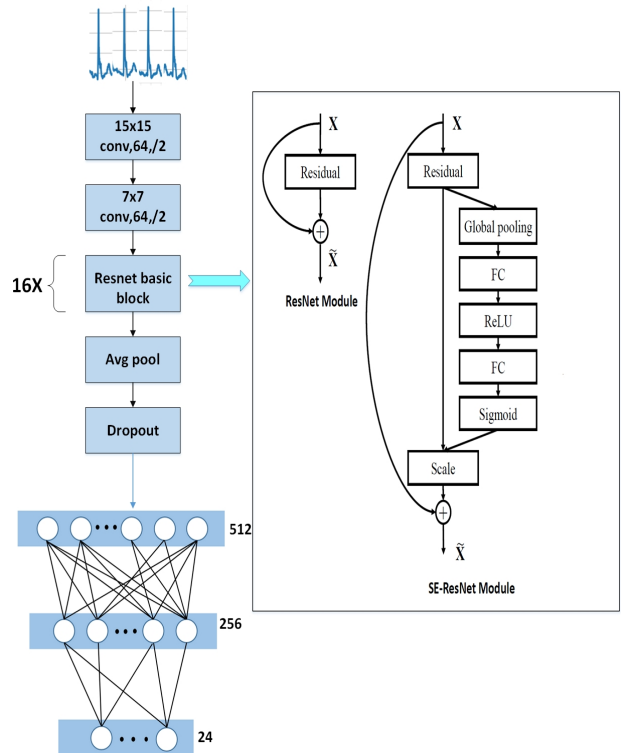
2.3. Network Structure

Figure 5. The architecture of the proposed method

The structure of the proposed network is shown in Figure 5. The proposed approach can be divided into two main modules from the top to down: feature encoder

module which we introduce 1D-Squeeze-and-Excitation block on the basis of 1D-ResNet and Multi-layer Perceptron (MLP) module. The first module is used to extract the features of ECG signals. The last MLP module, maps the learned feature representation in ECG domain to predict the cardiac abnormalities.

2.4. Feature Encoder

During the clinical diagnosis of cardiac problems, the interpretations of 12-lead ECG signals by medical workers mostly depends on its local characteristics, such as PR Interval or T-Wave duration, and P-Wave flag and so on. Inspired by this, we choose ResNet as the base model due to its good performance in feature extraction. ResNet has proven to be very successful in the image field. We modify the original two-dimensional network structure to one-dimensional so that it can be well adapted to our task of ECG classification. In order to better fit the characteristics of the ECG signals, we increase the size of the first convolution kernel to 15 to greatly increase its receptive field, then we add a convolutional layer with a convolution kernel size of 7x7 to obtain information at different scales. As we all know, the maximum pooling layer will cause the loss of spatial information, and it cannot take the relationship between multiple channels into account, so we cancel the maximum pooling layer and replace it with a convolutional layer. Meanwhile, we also add a series of dropout layers to enhance the robustness of the network.

2.5. Squeeze and Excitation block

In medicine, there is a certain relationship between 12 leads. Doctors can roughly make judgments about certain abnormalities based on the images of specific leads. So we want to pay attention to the relationship between each lead, the model can automatically learn the importance of different channels. We want to capture the importance of each feature channel by enabling the machine to learn autonomously. Therefore we embed the 1D-Squeeze-and-Excitation block in our basic block to learn hidden features between different leads. As we all know, convolution will get the feature map. The SE module first squeezes this feature map to obtain channel-level global features. After the excitation of the global features, the relationship between each channel will be determined by the SE module, so that the weight of different channels will be determined. The final feature can be obtained by multiplying the weight by the original feature map. In essence, the SE module performs attention or gating operations in the channel dimension, which is conducive to channel features with the largest amount of information to get the most attention and is helpful to ignore the unimportant channel features.

3. Experimental Results

3.1. Challenge Score

We compare the proposed model with several alternatives, such as Support Vector Machine (SVM)[9] and Long Short Term Memory (LSTM)[10]. For fairness, we only replace the model part in the complete code without changing anything else. The experimental results are presented in Table 1, which involves the challenge scores.

Table 1. Challenge score on different models.

Algorithms	Challenge score
EASTNet	0.7030 ± 0.0090
SVM	0.5611 ± 0.0156
LSTM	0.5937 ± 0.0119
1D-ResNet34	0.6760 ± 0.0044
1D-ResNet34+SEblock	0.6856 ± 0.0132
1D-ResNet34+MLP	0.6875 ± 0.0121

As shown in Table 1, the proposed model get a challenge score of 0.7030 ± 0.0090. Although both SVM and LSTM are state-of-the-art methods, the experimental results indicate that they fail to perform as well as our model under the same conditions.

3.2. Comparison Experiment

In this part, we focus on the impact of our data processing methods. After intercepting a single heartbeat, we think that the number of heartbeats entered is a very worthy question. Intuitively speaking, a large number of heartbeats will introduce more noise and more invalid information, which will lead to difficulties in network learning, but a small number of heartbeats cannot fully reflect the characteristics of all abnormalities. There is a whole versus part relationship here. Therefore, we need to explore the most suitable number of heartbeats. At the same time, continuous heartbeats and random heartbeats are also part of our comparative experiment. The final experimental results are shown in the Table 2 and Table 3 below.

Table 2. Score from continual different heartbeats.

Beats	Challenge score
Single beat	0.6870 ± 0.0109
Double beats	0.6875 ± 0.0062
Triple beats	0.6931 ± 0.0140
Four beats	0.6849 ± 0.0195
Five beats	0.6847 ± 0.0055

Table 3. Score from random different heartbeats.

Beats	Challenge score
Double beats	0.6985 ± 0.0126
Triple beats	0.7027 ± 0.0119
Four beats	0.7030 ± 0.0090
Five beats	0.7090 ± 0.0097
Six beats	0.7036 ± 0.0163

In addition, in order to test the robust performance of our model, we also design a comparison experiment that use different k-fold cross-validation which k = 3,4,5. The result is shown in the Table 4. We can see that our model performs well on robustness.

Table 4. Comparison with other k-fold cross-validation.

k-fold	Challenge score
3-fold	0.6703 ± 0.0250
4-fold	0.6968 ± 0.0106
5-fold	0.7030 ± 0.0090

4.　　Conclusion

In this experiment, EASTNet is proposed to diagnose cardiac abnormalities with the aid of 12-lead ECG only. The proposed method combines local-encoder and the relationship between 12-leads to obtain a better feature representation of ECG records. It is a end-to-end model that doesn't need other additional information like Magnetic Resonance Imaging (MRI) or Computed Tomography (CT).

From the Table 2 and Table 3, the score of the model is the highest when the number of random heartbeats is five, but we have a trade-off between score and computational complexity, so we choose four random heartbeats as our input. Too few heartbeats will lead to insufficient effective features and too many heartbeats will introduce some noise and invalid information, which will make model learning difficult. So finding the best number of random heartbeats is the key to improve the challenge score.

In the future, there are two potential directions worth exploring. Firstly, we will explore in more depth the performance of different abnormalities under the same heartbeat and the impact of different heartbeat capture methods on the model. Another potential direction that we want to study is to combine our model with the model suitable for extracting the signal characteristics of time series, which is beneficial to obtain more useful information from ECG signals.

Acknowledgments

References

[1] Erick A. Perez Alday, Annie Gu, et al, "Classification of 12-lead ECGs: the PhysioNet/Computing in Cardiology Challenge 2020," *Physiol. Meas*, 2020.

[2] Zhang L, Peng H, Yu C, "An approach for ECG classification based on wavelet feature extraction and decision tree," *2010 International Conference on Wireless Communications & Signal Processing (WCSP),* pp. 1–4, Oct, 2010.

[3] Jadhav S M, Nalbalwar S L, Ghatol A, "Artificial neural network based cardiac arrhythmia classification using ECG signal data," *2010 International Conference on Electronics and Information Engineering*, vol. 1, pp. V1-228-V1-231, Aug, 2010.

[4] Hannun A Y, Rajpurkar P, Haghpanahi M, et al, "Cardiologist-level arrhythmia detection and classification in ambulatory electrocardiograms using a deep neural network," *Nature Medicine*, vol. 25, no. 1, pp. 65–69, Jan, 2019.

[5] He K, Zhang X, Ren S, et al, "Deep residual learning for image recognition," *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 770–778, Jun, 2016.

[6] Zhang Z, Lyons M, Schuster M, et al, "Comparison between geometry-based and gabor-wavelets-based facial expression recognition using multi-layer perceptron," *Proceedings Third IEEE International Conference on Automatic Face and Gesture Recognition*, pp. 454-459, Apr, 1998.

[7] Hu J , Shen L , Albanie S , et al, "Squeeze-and-excitation networks," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, pp. 99, Sep, 2017.

[8] Hamilton P, "Open source ECG analysis," *Computers in Cardiology*, pp. 101-104, Sep, 2002.

[9] Cortes C, Vapnik V, "Support-vector networks," *Machine Learning*, vol. 20, no. 3, pp. 273-297, Sep, 1995.

[10] Sundermeyer M, Ney H, Schlüter R, "From feedforward to recurrent LSTM neural networks for language modeling," *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, vol. 23, no. 3, pp. 517-529, Mar, 2015.

Address for correspondence:

Name: Huafeng Liu
Address: Teaching Building 3, Yuquan Campus, Zhejiang University, no. 38, Zhejiang University Road, Hangzhou ,China
Email: liuhf@zju.edu.cn