A Branched Deep Neural Network for End-to-end Classification from ECGs with Varying Dimensions

Han Duan[1], Junchao Fan[1], Junhui Zhang[2], Bin Xiao[1], Xiuli Bi[1], Xu Ma[3]

[1] Chongqing Key Laboratory of Image Cognition, Chongqing University of Posts and Telecommunications, Chongqing, China
[2] Health Management Center, the First Affiliated Hospital of Chongqing Medical University, Chongqing, China
[3] Human Genetics Resource Center, National Health Commission, Beijing, China

Abstract

Objective: Computerized electrocardiogram (ECG) interpretation plays a critical role in the clinical ECG workflow. For 12-lead ECGs, deep neural networks (DNNs) could match state-of-the-art algorithms when trained in openly available datasets (e.g., 2020 PhysioNet/CinC Challenge data). However, there is limited evidence to demonstrate the utility of reduced-lead ECGs for capturing a wide range of diagnostic information. Thus, in this work, we aimed to develop a deep learning-based algorithm to identify clinical diagnoses from twelve-lead, six-lead, four-lead, three-lead, and two-lead ECG recordings.

Methods: We have employed a branched DNN that utilizes two CNNs (similar to the residual network, but adapted to signals with varying dimensions) with different filter sizes. This architecture can effectively capture abnormal patterns of diseases and suppress noise interference by introducing double branch and skip connections. Moreover, the Squeeze-and-Excitation block is added to every branch of the architecture, which allows the DNNs to learn more effective features from input signals.

Results: Our (CQUPT_Dontwant) method is evaluated using the twelve-lead, six-lead, four-lead, three-lead, and two-lead of the Challenge validation datasets, and we received a 5-fold cross-validation Challenge metric score of 0.639, 0.610, 0.638, 0.619, and 0.625, respectively.

Conclusion: The results demonstrate the effectiveness of the proposed method in achieving robustness towards the classification of ECGs with varying dimensions.

1. Introduction

Elctrocardiogram (ECG) is the most basic, convenient and economical routine examination approach. It is very commonly performed for clinical medical screening of many cardiac diseases, and provides important reference information for clinicians [1]. With the development of Artificial Intelligence (AI), the importance of automated analysis of ECG has vecoe invreasingly prominent.

Over the past decade, a large number of automatic analysis algorithms have been introduced [2, 3]. Theses methods through reasonably combining feature extraction and classifier, they still have some common defects: 1) They require manual feature engineering based on a considerable amount of expert knowledge. 2) The manual feature of different diseases characteristics may be sightly different, therefore the generalization ability of the model is restricted [4]. Deep neural networks (DNNs) have recently achieved striking success in ECG prcessing. For single-lead ECGs, a convincing preliminary study of the use of DNNs was presented in [5]. For twelve-lead ECGs, Min *et al.* reported a DNNs-based network to classify clinical cardiac abnormalities [6].

The 2021 PhysioNet/CinC challenge focused on automated, open-source approaches for identify clinical diagnoses from twelve-lead, six-lead, four-lead, three-lead and two-lead ECG recordings [7]. In this paper, as part of the PhysioNet/Computing in Cardiology Challenge 2021, we developd a branched DNN that utilizes two residual convolutional neural network to classify clinical cardiac abnormalities from this ECG recordings. We will describe our approaches to tackle the challenge.

2. Methods

In this section, we first briefly describe the dataset, and then introduce the implementation details of our methods. The overall system design is shown in Figure 1. The ECG records are divided into 10 seconds fragments by sliding window. If the length of original ECG record does not reach 10 seconds, it is filled with zero. The deep neural network proposed in this paper is then used to extract the features of ECG fragments. Finally, fully connected lay-

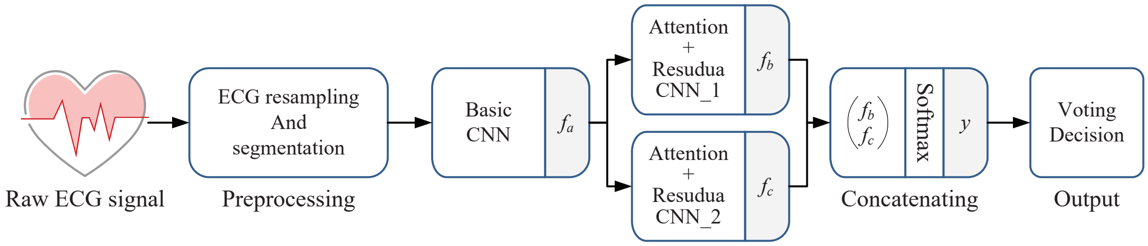

Figure 1. The overall system of our proposed method.

ers are used to classify and obtain the record level results based on a onevote-in-favour voting decision strategy.

2.1. Datasets

We would carry out the experiments on the 2021 PhysioNet/CinC Challenge Dataset [7]. For clarity, we summarized the basic information of this datasets in Table 1.

The 2021 PhysioNet/CinC Challenge dataset is a cross-corpus achieve of ECG recordings from six sources in four countries. These databases include over 100,000 twelve-lead ECG recordings with 88,253 ECGs shared publicly as training data, 6,630 ECGs retained privately as validation data, and 16,630 ECGs retained privately as test data. In our experiments, we only employed the publicly available training set of this dataset.

Table 1. Detailed profiles for the available ECG datasets.

Data source	Number of Recordings	leagth(s)	Sampling Frequency(Hz)
CPSP	10330	6-144	500
INCART	74	1800	257
PTB	22353	10-120	500 or 1000
Georgia	10334	5-10	500
Shaoxing	45152	10	500

2.2. Data Pre-processing

The ECG recordings should naturally be diverse due to the differences in individuals and data acquisition environments. To better prepare the data for model training, we adopted the following data preprocessing techniques. First, all ECGs were resampled to the frequency of 500 Hz. Then, to allow a fixed input size in the deep learning model, each ECG are divided into 10 seconds long segments through a sliding window. Third, we uses two methods of data enhancement: random shift signals and random addition of Gaussian noise to the signals. In addition we dropped records that only consist of classes not intersecting with the evaluation metric.

2.3. Model overview

Our branched DNN is composed of three main components: 1) Basic CNNs for learning shared low-level features, 2) multiple branch sub-networks to learn the high-level scale-specific signal features using different scales convolution kernels collaboratively, 3) multi-scale features fusion for integrating features from sub-networks. Figure 2 shows the configuration of our proposed network.

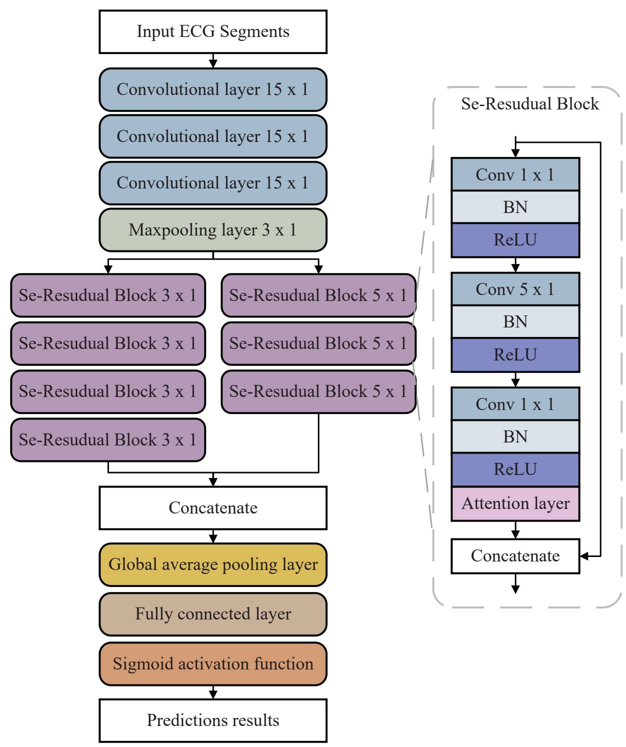

Figure 2. Illustration of our branched deep neural network.

We use the first three convolution layers for shared learning and obtain the feature maps f_a. Then the features f_a are fed into two sub-branches with different numbers of stacked Se-Resudual blocks in eachbranch to extract scale-specific feature maps f_b and f_c. Concatenate different feature maps f_b, f_c to obtain the fusion cross-scale features y. At last, the fusion features are employed for loss optimization.

2.3.1. Multi-Perspective DNN

As shown in Figure 2, our Multi-Perspective DNN has varieties of branches that have different numbers of stacked modules. The shallow brach tend to exploit the information directly stored in the signal amplitude. In contrast, the deeper branch are able to extract high-level features, which are relative to the local morphology structure and global shape information. These multi-perspective representations are then concatenated and fed into a softmax layer to predict the clinical diagnoses of ECG recordings.

The receptive field in convolutional neural network is important. In order to obtain the appropriate receptive field and increase the diversity of model features, a multi-scale feature network is designed, which is mainly carried out by parallel multi-branches. The convolutional kernel sizes of each branch are different, which are 3 and 5, respectively. After that, different scales of features extracted by different paths are concatenated.

Inspired by ResNet [8], each branch has multiple serial residual modules. In the residual modules, this paper uses several convolutional kernels. Pre-activation is used for each convolution operation. In short, this paper put batch normalization (BN) layer and Relu activation layer before weight layers. The design of pre-activation has two benefits. First, the optimization is futher eased. Second, using BN as pre-activation improves regularization of the models.

2.3.2. Attention layer

In order to realize the characteristics weighted between channels and to ensure that more useful information will be sent to the subsequent layer for futher feature extraction, we introduce a lightweight channel attention mechanism (named SENet [9]) behind each Resudual block.

As shown in Figure , the attention mechanism can be divided into two steps: compression and activation. In the process of compression, global average pooling is used to compress the spatial dimension information of the feature maps into channel description vectors. This can overcome the problem that the receptive field of convolution is too small to make use of useful contextual information outside the receptive field. The cth element of the channel descriptors can be calculated using formula:

$$Z_c = F_{sq}(x_c) = \frac{1}{L}\sum_{i=1}^{L} x_c(i) \tag{1}$$

where L represents the size of the input feature maps $X = [x_1, x_2, ..., x_c]$. During activation, the Sigmoid function is used to extract dependencies between channels, as shown below:

$$s = F_{ex}(z, w) = \sigma(g(z, w)) = \sigma(w_2 \delta(w_1 z)) \tag{2}$$

where σ stands for Sigmoid function and δ stands for ReLu function, $w_1 \in R^{\frac{c}{r} \times c}$ and $w_2 \in R^{c \times \frac{c}{r}}$ stand for two fully connected layer parameters. Parameter r is used to control the number of parameters in the network and increase model generalization performance. Finally, use the channel description vectors and the input feature maps to multiply the corresponding between the channels to weight the features. The output feature maps can be defined by formula:

$$v_c = F_{scale}(x_c, s_c) = x_c \times s_c \tag{3}$$

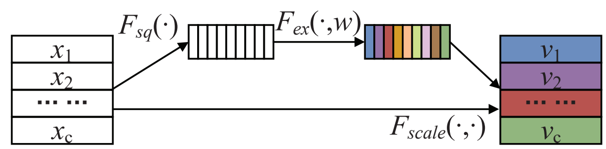

Figure 3. The SENet layer.

2.3.3. Voting decision strategy

In the data processing stage, this paper cut ECG records into several segments. So how to convert the diagnostic results of the segmented fragments into the result of original record is critical. Given that cardiac anomalies are more localized in the local areas of electrocardiogram, a simple majority voting strategy should not be used, which would miss important local information. Therefore, a voting decision strategy is designed. In short, a union of the diagnostic results of all the fragments is the diagnosis of the original record.

2.4. Training Setup

Our branched DNN model was trained for 13 epoches with a batch size of 64 on a machine with 64 GB RAM, 8-core CPU and one NVIDIA GeForce RTX 2080 Ti GPU. The model parameters were optimized with the Adam optimizer [10]. The learning rate during training was set as 0.001, and rescheduled to 0.0001 at the 10^{th} epoch.

3. Results

We train and evaluate our model using 5-fold corss-validation, whereby four folds are used for model training and the remaining acts as a held-out test set. This is repeated five times to produce five trained models and the challenge score of each is averaged to get an estimate about how well the model is performing. The averaged results are shown in Table 2. We report Accuracy and F-measure scores as well as Challenge metric. From the results, our DNN approach demonstrated its ability to classify the cardiac abnormalities from ECGs with varying dimensions.

Table 2. 5-fold cross-validation results showing challenge scoring for different models.

Model	Accuracy	F-measure	Challenge metric
12-lead	0.312	0.500	0.639
6-lead	0.306	0.462	0.610
4-lead	0.313	0.490	0.638
3-lead	0.309	0.483	0.619
2-lead	0.308	0.462	0.625

4. Conclusion

This paper introduced a end-to-end branched deep learning method for ECG signal classification by utilizing the multi-scale ECG signal features. At the same time, a lightweight channel attention mechanism from SENet [9] is introduced to realize the characteristics weighted between channels. Compared with the existing deep learning methods for ECG analysis using single scale, our approach can effectively achieve multi-scale feature extraction and cross-scale information complementarity of ECG signals. The experimental results convince the effectiveness of the proposed method. In the future, we will apply the branched model to other physiological signal analysis and processing requirements.

Acknowledgments

This work was partly supported by the National Science &. Technology Major Project (2016YFC1000307-3), the National Natural Science Foundation of China (61572092), the Natural Science Foundation of Chongqing (cstc2018jcyjAX0117, cstc2016jcyjA0407), the Scientific & Technological Key Research Program of Chongqing Municipal Education Commission (KJZD-K201800601), and the Chongqing research and innovation project of graduate students (CYS18245).

References

[1] Schläpfer J, Wellens HJ. Computer-interpreted electrocardiograms: benefits and limitations. Journal of the American College of Cardiology 2017;70(9):1183–1192.

[2] Monasterio V, Laguna P, Martinez JP. Multilead analysis of t-wave alternans in the ecg using principal component analysis. IEEE Transactions on Biomedical Engineering 2009; 56(7):1880–1890.

[3] Jun TJ, Park HJ, Minh NH, Kim D, Kim YH. Premature ventricular contraction beat detection with deep neural networks. In 2016 15th IEEE International Conference on Machine Learning and Applications (ICMLA). IEEE, 2016; 859–864.

[4] Wang R, Fan J, Li Y. Deep multi-scale fusion neural network for multi-class arrhythmia detection. IEEE journal of biomedical and health informatics 2020;24(9):2461–2472.

[5] Hannun AY, Rajpurkar P, Haghpanahi M, Tison GH, Bourn C, Turakhia MP, Ng AY. Cardiologist-level arrhythmia detection and classification in ambulatory electrocardiograms using a deep neural network. Nature medicine 2019; 25(1):65–69.

[6] Min S, Choi HS, Han H, Seo M, Kim JK, Park J, Jung S, Oh IY, Lee B, Yoon S. Bag of tricks for electrocardiogram classification with deep neural networks. In 2020 Computing in Cardiology. IEEE, 2020; 1–4.

[7] Reyna MA, Sadr N, Alday EAP, Gu A, Shah AJ, Robichaux C, Rad AB, Elola A, Seyedi S, Ansari S, et al. Will Two Do? Varying Dimensions in Electrocardiography: The PhysioNet - Computing in Cardiology Challenge 2021 (version 1.02). PhysioNet (2021);URL `https://doi.org/10.13026/jz9p-0m02`.

[8] He K, Zhang X, Ren S, Sun J. Deep residual learning for image recognition. In Proceedings of the IEEE conference on computer vision and pattern recognition. 2016; 770–778.

[9] Hu J, Shen L, Sun G. Squeeze-and-excitation networks. In Proceedings of the IEEE conference on computer vision and pattern recognition. 2018; 7132–7141.

[10] Kingma DP, Ba J. Adam: A method for stochastic optimization. arXiv preprint arXiv14126980 2014;.

Address for correspondence:

Junchao Fan
School of Computer Science
No.2, Chongwen Road, Nan'an district, Chongqing, China
fanjc@cqupt.edu.cn

Junhui Zhang
The First Affiliated Hospital of Chongqing Medical University
No. 1 Yixueyuan Road, Yuzhong District, Chongqing, China
2275610878@qq.com