

Multi-label Classification of Electrocardiogram With Modified Residual Networks

Shan Yang¹, Heng Xiang¹, Qingda Kong¹, Chunli Wang¹

¹Chengdu Spaceon Electronics Co, Ltd, Chengdu, China

Abstract

In this study, an end-to-end deep residual neural network with one dimensional convolution is presented to identify the 12-lead ECGs for the PhysioNet/Computing in Cardiology Challenge 2020 with different durations and sampling frequency. Firstly, a split attention block that enables attention across feature-map groups is introduced to the modified residual networks based on one dimensional convolutional neural network. Secondly, referring to the improved residual network (iResNet), the modified network architecture for residual learning based on stages is present to provide a better path for information to propagate through network's layers. Moreover, an improved projection shortcut that reduces the information loss is proposed in our proposed modified residual network. Finally, fine-tune is adopted through pre-trained model, which avoids training the network from scratch and virtually expands the dataset and enhances the generalization performance.

Our ensemble of five models trained on different data folds is further validated on the PhysioNet challenge hidden test, achieving a score of 0.208 on the full test data, but is not ranked due to omissions in the submission (team name: SpaceOn Flattop). Results suggest that the proposed method achieves competitive performance in the multi-label ECGs classification.

1. Introduction

Cardiovascular disease is one of the most serious diseases that harm human life [1]. The standard 12-lead electrocardiogram (ECG) reflecting the physiology activities of heart has been widely used to diagnose a variety of cardiac abnormalities, and predicts cardiovascular morbidity and mortality for its simplicity and non-invasive property [2]. The PhysioNet/Computing in Cardiology Challenge 2020 focused on automated, open-source approaches for classifying cardiac abnormalities from 12-lead ECGs [3, 4].

Recently, artificial intelligence and machine learning techniques have shown great potential in automatic

detection and classification of cardiac abnormalities which can assist physicians in the diagnosis of the growing number of ECGs recorded. Especially, one dimensional (1D) convolutional neural networks (CNN) have gained a lot of interest in physiological signal processing due to their strong capabilities in learning complex features by being directly applied on raw data without extracting any hand-crafted features [5,6].

Residual networks (ResNets) represent a powerful type of convolutional neural network architecture, widely adopted and used in various tasks. The ResNets are easy to optimize and can easily enjoy accuracy gains from greatly increased depth, producing results substantially better than previous networks [7]. The arrhythmia detection algorithm based on 1D CNN with residual blocks has achieved outstanding performance [6, 8, 9]. Rajpurkar et al. [6, 8] proposed a 1D CNN classifier that used a 34-layer convolutional neural network with residual blocks which maps a sequence of ECG samples to a sequence of rhythm classes. By testing their model against board-certified cardiologists, they concluded that the CNN model exceeds the individual expert performance on both recall, and precision on the test set. S. Yang et al. [9] proposed 1D densely connected CNN following the basic structure of densely connected convolutional networks [10] and optimizing CNN model architecture to show optimal performance for ECG arrhythmia classification with different durations.

Inspired by the studies mentioned above, we present a modified deep residual network with one dimensional convolution to identify the rhythm/morphology abnormalities from 12-lead ECGs records. The proposed method contains modified residual network based on stages and split attention blocks, which not only avoids lacking cross-channel interaction in original ResNets model, but also makes the substantially convolutional networks stronger ability of feature representation. The main contributions in this study are: (1) adopting noise addition, y-axis shift, band-pass and wavelet-based filter methods for data enhancement, splitting and resampling the ECG records of 10 second length with a sampling frequency of 500 Hz for training (Sect. 2.1). (2) a network architecture for residual learning based on stages is introduced, which facilitates the learning process by

providing a better path for information to propagate through network’s layers (Sect. 2.2). (3) a modular split attention block that enables attention across feature-map groups is introduced. By stacking these split attention blocks, a ResNets-style model is constructed based on 1D CNN, which reserves the overall ResNets structure without introducing additional computational costs (Sect. 2.2). (4) a simple method of transfer learning: fine-tune is adopted though pre-trained model, which avoids training the network from scratch and virtually expands the dataset and improves the generalization performance (Sect. 2.3).

2. Methods

2.1. Data Preprocessing

Data Resampling: The data set from 4 different sources is sampled at different frequencies of 257 Hz, 500 Hz and 1000Hz. To ensure the consistency of the data as much as possible, the ECG records with frequency of 257 Hz and 1000 Hz are resampled to 500Hz.

Data Augmentation: To increase the diversity of dataset and improve the generalization performance, this data set can be further extended by data augmentation. And some common techniques of data augmentation used in this study consist of noise addition, corresponding channel switch, y-axis shift, band-pass and wavelet-based filter, and so on.

Data Segmentation: The length of dataset is different, and the records of 10 seconds length are in the majority. To make the length of data fed into network is equal, zero padding and data truncation are introduced, which have a certain destructive effect on the ECG information. Specifically, the records less than 10 s are padded by zeros to a size of 10 s and the records more than 10 s are truncated to take the first 10 s data as the current data.

2.2. Model Architecture

In this work, the proposed model is mainly composed of multiple basic blocks and four modified residual convolutional network stages, as is shown in Figure 1. Each consists of three different types of residual block, start Res-Block, middle Res-Block and end Res-Block, which are shown in Figure 2.

Firstly, the modified residual convolutional network can be split into four main stages [11]. Each can contain a number of Res-Blocks. There are three Res-Blocks for stage 1, four for stage 2, six for stage 3 and three for stage 4. Each main stage is divided into three parts: one Start Res-Block, a number of Middle Res-Blocks ((1, 2, 4, 1) Middle Res-Blocks for the corresponding stages) and one End Res-Block. Each Res-Block has a different design depending on the position in the stage.

In start Res-Block, there is a batch normalization (BN) layer after the last convolution, which provides a normalized signal, preparing it for the element-wise addition with the projection shortcut. And the End Res-block is end up with a BN layer and rectified linear units (ReLU) activation functions, which can be seen as preparation for the next stage. It offers a better path for information to propagate through the network.

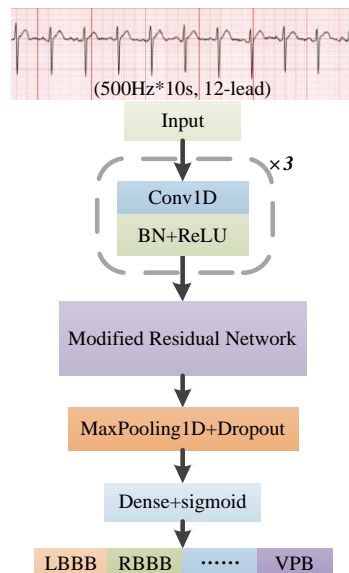


Figure 1. Proposed model based on modified residual convolutional network

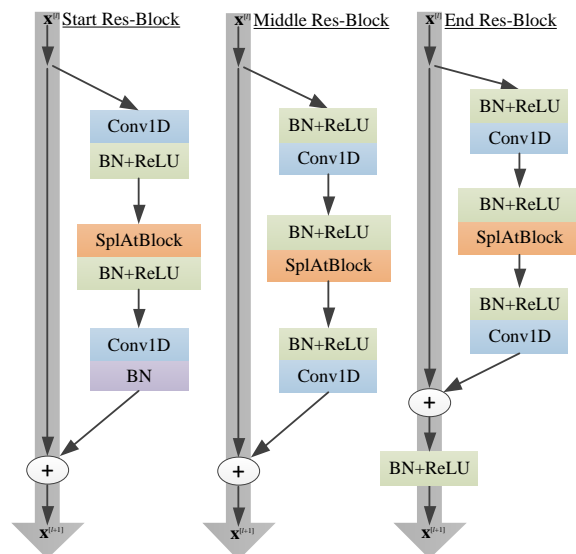


Figure 2. Three different types of residual blocks

As is shown in Figure 3 and Figure 4, the split attention block (SplAtBlock) in Res-Block divides the feature into several feature-map groups, just like ResNeXt blocks [12], the number of which is given by a *cardinality* hyper parameter k . And the *radix* parameter r

that indicates the number of splits within each cardinal group.

The combined representation of each cardinal group can be obtained by fusing via an element-wise summation across multiple splits. While the cardinal group representations are concatenated along the channel dimension. The SplAtBlock in Res-Block generalizes the channel-wise attention into feature-map group representation, which can be modularized and accelerated using unified CNN operators [13]. In addition, it greatly improves the ability of feature representation.

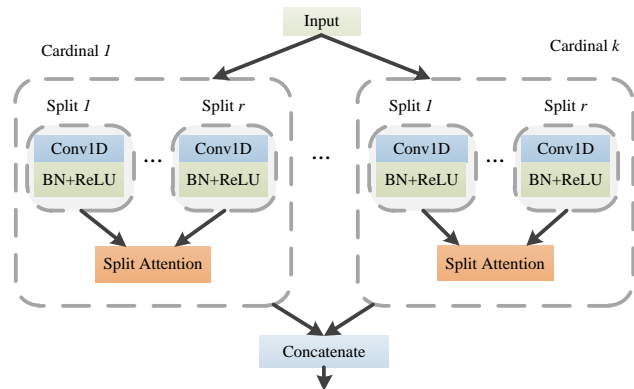


Figure 3. The structure of SplAtBlock in Res-Block

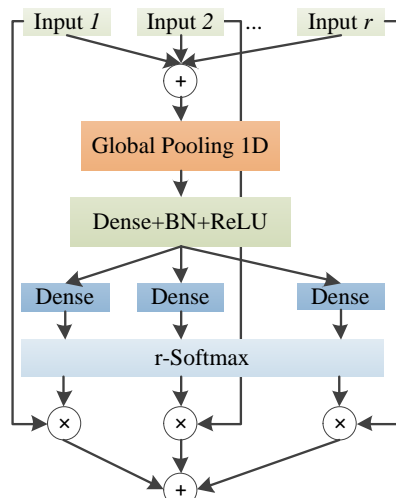


Figure 4. The detail of split attention unit in SplAtBlock

2.3. Transfer Learning

In this study, inspired by the application of transfer learning in biological signals [15], transfer learning is introduced to avoid training the network from scratch and improve the generalization performance. Fine-tune as the simplest method of deep transfer learning is adopted by using pre-trained model which has been previously trained on another datasets, we can directly use the

obtained weights and architecture and apply the learning on identifying abnormalities from 12-lead ECGs records.

The dataset for pre-trained composes of 40,000 medical ECG records and 55 abnormalities in total provided by the Engineering Research Center of the Ministry of Education for mobile Health Management System of Hangzhou Normal University, China. We get access to the worthy dataset through the algorithm competition of tianchi.aliyun.com. Each record has 8 leads, namely I, II, V1, V2, V3, V4, V5, and V6. We can also calculate the data of the remaining 4 leads by using the following formula: $III=II-I$, $aVR=-(I+II)/2$, $aVL=I-II/2$, $aVF=II-I/2$. Also, the ECG record is sampled at a frequency of 500 HZ, of 10 seconds length, and of 4.88 microvolts unit voltage.

Since the size of the data is large as well as there is high data similarity, we just customize and modify the output layers according to our problem statement. And we retrain the proposed model using the weights of pre-trained model as initial weights with a smaller learning rate. The global fine-tune based on pre-trained model not only avoids training the network from scratch and accelerates the model convergence, but also virtually expands the dataset and enhances the robustness and generalization.

2.4. Learning mechanism

In this study, the model takes 12-lead ECG records of 10 seconds length and of 500 Hz frequency as input, the batch size parameter of which is set to 16, and produced a multi-label classification for each ECG record. The training set only considering containing 27 classes that are scored with the challenge metric, is randomly split into 80% training and 20% validation for 5-fold cross validation. The rectified adaptive momentum estimation (RADam) optimizer [14], with a constantly decreasing learning rate according to the assessment criteria (which is the score used in the challenge), is used to optimize the network parameters.

Aim at the multi-label classification, we use binary cross entropy (BCE) as loss function and sigmoid as the activation function. Since the identification of the 12-lead ECGs for the PhysioNet/Computing in Cardiology Challenge 2020 is a class unbalanced problem, class weight is introduced in BCE loss function.

The proposed model is developed using Pytorch framework¹. Training is done on an NVIDIA RTX 2080Ti GPU, the time of which is about 24 hours with five folds cross-validation. Training is performed for a maximum of 80 epochs, and early-stopping is used if the challenge score on the validation set does not improve for 10 consecutive epochs.

¹An implementation can be found at <https://github.com/yshanyes/Pytorch-ECG-Classifer-Cinc2020-Official>.

3. Results and Discussion

To evaluate the performance of our method, we perform a five-fold cross-validation on the training dataset, which consisted of a set of 37749 ECG records only considering the scored classes. Table 1 shows the challenge score of final ensemble model, measured on the four hidden datasets of Physionet challenge.

The proposed method based on modified residual convolutional network can achieve a score of 0.208 on the full test data, but is not ranked due to omissions in the submission (team name: SpaceOn Flattop).

Table 1. The score of the ensemble model on hidden set

| Hidden datasets | Score of ensemble model |
|-----------------|-------------------------|
| Validation Set | 0.681 |
| Test Database 1 | 0.871 |
| Test Database 2 | 0.219 |
| Test Database 3 | 0.126 |
| Full Test Set | 0.208 |

4. Conclusions

In this paper, we have proposed a method based on modified residual convolutional network for identifying the cardiac abnormalities from 12-lead ECG records which are provided in the 2020 Physionet challenge database. The split attention block is introduced in modified residual network to make the substantially convolutional networks stronger ability of feature representation. And the network architecture for residual learning based on stages is adopted to provide a better path for information to propagate through network's layers. Finally, transfer learning method namely fine-tune, and ensemble model obtained through computing the average prediction of the five-fold cross-validation models are utilized to improve the generalization performance and robustness of the model.

The following aspects could be considered in the future study to further improve the performance of ECG abnormal identification: (1) optimising the threshold of corresponding to classes for multi-label classification, the threshold can achieve the highest challenge metric on the training set (2) introducing some classical hand-crafted ECG features to improve performance, (3) combining recurrent neural network with CNN to further represent features of ECG.

References

[1] Benjamin EJ, Muntner P, Alonso A, et al., "Heart Disease and Stroke Statistics—2019 Update: A Report From the American Heart Association," *Circulation*, vol. 139, no. 10, pp. e56–e528, Jan. 2019.
[2] Kligfield P., "The Centennial of the Einthoven

Electrocardiogram," *Journal of Electrocardiology*, vol. 35, no. 4, pp. 123–129, 2002.
[3] Goldberger AL, Amaral LA, Glass L, et al., "PhysioBank, PhysioToolkit, and PhysioNet: Components of a New Research Resource for Complex Physiologic Signals," *Circulation*, vol. 101, no. 23, pp. e215–e220, Jun, 2000.
[4] Erick A. Perez Alday, Annie Gu, Amit Shah, Chad Robichaux, An-Kwok Ian Wong, Chengyu Liu, Feifei Liu, Ali Bahrami Rad, Andoni Elola, Salman Seyedi, Qiao Li, Ashish Sharma, Gari D. Clifford, Matthew A. Reyna., "Classification of 12-lead ECGs: the PhysioNet/Computing in Cardiology Challenge 2020," *Physiol. Meas.*, Aug, 2020.
[5] Pourbabae B, Javan Roshtkhari M, Khorasani K, "Deep Convolutional Neural Networks and Learning ECG Features for Screening Paroxysmal Atrial Fibrillation Patients," *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, pp. 1–10, Jun. 2017.
[6] Awni Y. Hannun, Pranav Rajpurkar, Masoumeh Haghpanahi, et al., "Cardiologist-level Arrhythmia Detection and Classification in Ambulatory Electrocardiograms Using a Deep Neural Network," *Nat. Med.* 25, pp. 65–69, Jan. 2019.
[7] K. He, X. Zhang, S. Ren and J. Sun., "Deep Residual Learning for Image Recognition," *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 770–778, Jun. 2016.
[8] Pranav Rajpurkar, Awni Y. Hannun, Masoumeh Haghpanahi, et al., "Cardiologist-Level Arrhythmia Detection with Convolutional Neural Networks," *arXiv:1707.01836 [cs.CV]*, Jul. 2017.
[9] Wang C., Yang S., Tang X., Li B., "A 12-Lead ECG Arrhythmia Classification Method Based on 1D Densely Connected CNN," *Machine Learning and Medical Engineering for Cardiovascular Health and Intravascular Imaging and Computer Assisted Stenting*, vol. 11794, pp. 72–79, Oct. 2019.
[10] G. Huang, Z. Liu, L. Van Der Maaten and K. Q. Weinberger, "Densely Connected Convolutional Networks," *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 2261–2269, Jul. 2017.
[11] Ionut Cosmin Duta, Li Liu, Fan Zhu, Ling Shao, "Improved Residual Networks for Image and Video Recognition," *arXiv:2004.04989 [cs.CV]*, Apr. 2020.
[12] S. Xie, R. Girshick, P. Dollár, Z. Tu and K. He, "Aggregated Residual Transformations for Deep Neural Networks," *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 5987–5995, Jul, 2017.
[13] Hang Zhang, Chongruo Wu, Zhongyue Zhang, et al., "ResNeSt: Split-Attention Networks," *arXiv:2004.08955 [cs.CV]*, Apr. 2020.
[14] Liyuan Liu, Haoming Jiang, Pengcheng He, et al., "On the Variance of the Adaptive Learning Rate and Beyond," *the Eighth International Conference on Learning Representations [ICLR 2020]*, 2020.
[15] Jordan J. Bird, Jhonatan Kobylarz, Diego R. Faria, et al., "Cross-Domain MLP and CNN Transfer Learning for Biological Signal Processing: EEG and EMG," *IEEE Access*, vol. 8, pp. 54789–54801, Mar. 2020.

Address for correspondence:

Shan Yang.
50 Jinke East Road, Jinniu District, Chengdu, Sichuan, China.
yangshanbuaa@163.com