SE-ECGNet: Multi-scale SE-Net for Multi-lead ECG Data

Jiabo Chen¹, Tianlong Chen⁴, Bin Xiao¹, Xiuli Bi¹*, Yongchao Wang¹, Han Duan¹, Weisheng Li¹, Junhui Zhang²*, Xu Ma³

¹ Chongqing Key Laboratory of Image Cognition, Chongqing University of Posts and Telecommunications, Chongqing, China
² The First Affiliated Hospital of Chongqing Medical University, Chongqing, China
³ Human Genetics Resource Center, National Research Institute for Family Planning, Beijing, China
⁴ Southwest Jiaotong University, Chengdu, China

Abstract

Cardiovascular disease is a life-threatening condition, and more than 20 million people die from heart disease. Therefore, developing an objective and efficient computer-aided tool for diagnosis of heart disease has become a promising research topic. In this paper, we design a multi-scale shared convolution kernel model. In this model, two paths are designed to extract the features of electrocardiogram (ECG). Paths have different convolution kernel sizes, which are 3 × 1 and 5 × 1, respectively. Such multi-scale design enables the network to obtain different receptive fields and capture information at different scales, which significantly improves the classification effect. And SE-Net is added to every path of the model. The attention mechanism of SE-Net learns feature weight according to loss, which makes the effective feature maps have significant weights and the ineffective or low-effect feature maps have small weights. Our team name is CQUPT_ECG.

1. Introduction

Cardiovascular diseases are very common in the world and threaten people’s health. Electrocardiogram (ECG) is a non-invasive method of diagnosing cardiovascular diseases. It’s fast, effective, and harmless, but ECG diagnosis relies on doctors’ experience. Therefore, it is very important to develop a computer-aided ECG diagnostic technique.

In this paper, a multi-scale shared convolution kernel model is designed to predict cardiovascular diseases. It extracts the common characteristics of different leads by sharing convolution kernels. It combines multi-scale depth characteristics extracted by the convolution modules of different paths and age and gender characteristics.

2. Challenge Data analysis

The open training data published in challenge comes from multiple sources [6]. There are currently four datasets from China, the United States, Russia and Germany. Statistically speaking, there are a total of 43101 patient records in the training set with 111 symptoms. There are 27 classes in which scores are calculated. As shown in Figure 1, the imbalance of samples in the datasets is very serious.

![Figure 1. Sample distribution of 27 classes that calculate scores.](image)

The data distribution in the dataset is extremely imbalanced. At the same time, it can be seen from figure 2 that there is a certain correlation between diseases and age and gender. For example, the disease codenamed 427084000 is more common among people between 50 and 70 years of age, and the prevalence rate is higher in men than in women. The correlations between age and sex are also different for different diseases.
Figure 2. The incidence of disease codenamed 427084000 is related to gender and age.

There is similarity between different leads in electrocardiogram. The beats of different leads are consistent, and the peak positions are almost the same.

Figure 3. The similarity between different leads in 12-lead electrocardiogram.

In summary, the three characteristics are the most important characteristics of electrocardiogram data. How to make the best of these characteristics is the key to build our model.

3. Method

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 model proposed in this paper is then used to extract the features of ECG fragments. The depth characteristics obtained are then combined with the age and gender characteristics of the patients. Finally, fully connected layers are used to classify and obtain the record level results based on a one-vote-in-favour voting decision strategy. Figure 4 is the network diagram of the model. In the figure, 50x1, 15x1, 5x1 and 3x1 are the sizes of the convolution kernels. In fact, there are 9 residual blocks used in every branch in the network, but only 3 of them are drawn in the figure for simplification.

3.1. Data processing

3.1.1 Patch segment

In this paper, complicated preprocessing such as signal denoising is not adopted. Instead, the ECG records are divided into 10 seconds long segments through a sliding window. Note that 10 seconds is far longer than a complete cardiac cycle. And if there are records less than 10 seconds, using 0 padding to 10 seconds. At the same time, the method of dividing ECG records is helpful to expand the dataset.

3.1.2 Data enhancement

The amount of data for some labels are too small and imbalanced, which is not conducive to the training of the model, so the data enhancement operation is carried out.

This paper uses two methods of data enhancement: random shift signals and random addition of Gaussian noise to the signals. It should be noted that it is not suitable to use vertical flip for data expansion here, because there are two types of left deviation of the electric axis and right deviation of the electric axis in the labels. If using vertical flip, it will cause the problem of label error.

3.1.3 Age and sex characteristics

As for the relationship between diseases and age and gender mentioned in section 2, this paper add age and gender characteristics, normalize age information and one-hot code of gender information, which is helpful for the
model to better learn age and gender characteristics. Age and gender characteristics are combined with the depth characteristics extracted from the model, and then processed by the fully connected layer.

3.2. Shared convolution kernel

In general, a one-dimensional convolution kernel is often used to process the one-dimensional ECG signals. Different leads use different convolution kernels.

But considering that the similarities between the different leads mentioned in the second section, this paper extracts the common characteristics of different leads by sharing convolution kernels.

As shown in the figure 5, the length, width and channel of the original ECG data are 5000, 12 and 1 respectively. In this paper, the input ECG data is transposed to obtain a two-dimensional image with channel of 1, constant length, and height of 12. After that, the operation of two-dimensional convolution whose kernel size is $50 \times 1 \times C$ is used to extract the features of the image. So the different leads of ECG share the same convolution kernel. Such a design can help the network make the best use of similarity of leads.

![Figure 5. Different leads of ECG share the same convolution kernel.](image)

3.3. Multi-scale residual module

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.

![Figure 6. Multi path residual module.](image)

Inspired by ResNet[2], 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 layer(BN) and Relu activation layer before weight layers. The design of pre-activation has two benefits. First, the optimization is further eased. Second, using BN as pre-activation improves regularization of the models.

The dual pooling design is adopted at the end of each branch. The features obtained from each branch are processed through maxpooling layer and avgpooling layer respectively. In general, only avgpooling layer is used to retain significant features, reduce feature dimensions, and increase kernel receptive field. The purpose of adding maxpooling layer here is to enable the network to better capture local abnormalities, because heart disease is often reflected in a local area of ECG. The characteristics obtained from different branches and the age and gender characteristics extracted are fused together and then sent to the fully connection layer for processing.

3.4. Attention mechanism

This paper introduces a lightweight channel attention mechanism from SENet[5] in order to realize the characteristics weighted between channels and to ensure that more useful information will be sent to the subsequent layer for further feature extraction.

Specifically, 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 $C$th element of the channel descriptors can be calculated using formula 1:
\[ Z_c = F_{eq}(x_c) = \frac{1}{L} \sum_{i=1}^{L} x_c(i) \]  

(1)

Where L represents the size of the input feature maps \( X = [x_1, x_2, \ldots, x_c] \). During the convolution, the Sigmoid function is used to extract dependencies between channels, as shown below:

\[ s = F_{es}(z, w) = \sigma(g(z, w)) = \sigma(w_2\delta(w_1z)) \]  

(2)

Where \( \sigma \) stands for Sigmoid function and \( \delta \) stands for ReLU function, \( w_1 \in \mathbb{R}^{c \times c} \) and \( w_2 \in \mathbb{R}^{c \times c} \) 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 3:

\[ x_c = F_{scale}(x_c, s_c) = x_c \cdot s_c \]  

(3)

3.5. A one-vote-in-favour 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 one-vote-in-favour voting decision strategy is designed. In short, a union of the diagnostic results of all the fragments is the diagnosis of the original record.

4. Result

Our team name is CQUPT_ECG, in official phase, the method based on the proposed model achieves a Challenge metric score of 0.640.

And five-fold cross-validation is carried out on the training set. Meanwhile, It is compared with ResNet[3] and ResNest[7].

Table 1. 5-fold cross validation experiment results on the training set.

<table>
<thead>
<tr>
<th>Model</th>
<th>Fbeta</th>
<th>Gbeta</th>
<th>Challenge metric</th>
</tr>
</thead>
<tbody>
<tr>
<td>ResNext</td>
<td>0.497</td>
<td>0.256</td>
<td>0.5546</td>
</tr>
<tr>
<td>ResNest</td>
<td>0.4896</td>
<td>0.2554</td>
<td>0.5236</td>
</tr>
<tr>
<td>Proposed</td>
<td>0.5742</td>
<td>0.3054</td>
<td>0.6134</td>
</tr>
</tbody>
</table>

5. Conclusion

In this paper, a model with residual modules is proposed, which uses the features of different scales. At the same time, the similarity between different leads of electrocardiogram is utilized by using shared convolution kernels. A lightweight channel attention mechanism from SENet[5] is introduced to realize the characteristics weighted between channels. According to the characteristics of electrocardiogram, a one-vote-in-favour voting decision strategy is designed in this paper.

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


Address for correspondence:

Xiuli Bi  
School of Computer Science  
No.2, Chongwen Road, Nan’an district, Chongqing, China  
bixl@cqupt.edu.cn

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