ECG Heartbeat Classification Based on ResNet and Bi-LSTM

Yang Zhou, Haoxi Zhang, Yuan Li and Guangjian Ning
School of Cyberspace Security, Chengdu University of Information Technology, Chengdu 610225, China
Email: haoxi@cuit.edu.cn

Abstract. As a reliable cardiovascular system, ECG has been widely used in the detection of heart rhythms. This paper presents an attention-based ResNet processing of ECG data. Respectively are MIT and PTB Diagnostics datasets. The model is prominent in both data sets. And the accuracy in MIT datasets is 96.2%. And in PTB datasets the accuracy is 99.6%.

1. Introduction

An electrocardiogram (ECG, or EKG) is a technique that uses an electrocardiogram machine to record changes in the electrical activity of the heart during each cardiac cycle from the body surface.

According to the world health organization, several people die each year from arrhythmias, which have become the leading cause of death in humans. At present, heart rate can be divided into five categories, namely non-ectopic (N), supraventricular ectopic (S), ventricular ectopic (V), fused heartbeat (F), and unknown heartbeat (Q). So heart rate detection has become urgent. Early diagnosis and treatment can be achieved by detecting a patient's heart rate in advance.

At present, heart rate detection is mainly based on traditional machine learning methods such as SVM and deep learning. However, deep learning methods often only use one model, and some only use CNN to extract features from heart rate data and filter redundant features. We believe that ECG data are interdependent in time series, and LSTM is very good at processing time series data. Therefore, this paper proposes a CNN structure ResNet combined with LSTM model.

2. Related Work

For ECG heart classification algorithm has many, mainly is the traditional machine learning algorithm and the current epidemic of neural network, and by virtue of the size of the MIT data sets each classified data difference is very big, so they resolved MIT AAAI standard data for cutting unbalanced data sets, enhanced F class and SVB class objective function of the power of expression. Make it easier to distinguish from other classes. Due to the strong support vector machine (SVM), SVM combined with other algorithms, such as support vector machine (SVM) combined with fuzzy theory, support vector machine (SVM) combined with genetic algorithm, were also proposed for ECG classification.

Neural networks are also widely used in ECG heartbeat classification, before deep learning had no fire. A hybrid neural network method was proposed to reduce the training time of MLP and increase its generalization. CNN is now widely used in ECG classification.[1] et al. A CNN model-based auxiliary radar for ECG monitoring.[2] et al. proposed a deep convolution model to classify ECG signals of myocardial infarction, which could successfully distinguish 10 different types.[3] a convolution model is proposed to identify ECG of different animals, which is the first method of biometric recognition, and a method to expand ECG data is proposed to train CNN.
3. Datasets
In this paper, ECG heartbeat is classified using the following two data sets:

3.1. MIT-bih
Data Source: Physionet's MIT-BIH Arrhythmia Dataset
There are five categories in this dataset, namely non-ectopic (N), supraventricular ectopic (s), ventricular ectopic (V), fused heartbeat (F), and unknown heartbeat (Q). The number of each category is shown in table 1:

| Type | N  | S   | V   | F   | Q    |
|------|----|-----|-----|-----|------|
| count| 18118 | 556 | 1448 | 162 | 1608 |

3.2. The PTB Diagnostic ECG Database
Data Source: Physionet's PTB Diagnostic Database
This data is of a dichotomous type, abnormal and normal.

| Type   | Normal | Abnormal |
|--------|--------|----------|
| count  | 10506  | 4046     |

4. Methods
We proposed the ResNet+ Bi-LSTM architecture, which combines the advantages of Resnet and Bi-LSTM to extract local relevant features. It also captures the back and forth dependencies between the data.

4.1. ResNet
The emergence of Alexnet led to the development of deep learning, and since then, people have found that deeper networks are more capable of expressing features, and have achieved quite good results in many fields. Such as the classic network VGG, and GoogLeNet in the image processing competition shine.

However, if the network is too deep, it will lead to the problem of gradient disappearance or gradient explosion. As the depth increases, the accuracy will also decrease, because the back propagation may eventually fail to change the parameters.

ResNet, in order to solve this problem, allows direct access to be added to the network. The output of the layer above the input layer and the output layer can be connected to the output layer, so that the output layer can not only obtain the output of each layer, but also continue to receive input. See figure 1 for details.

We made some modifications to the ResNet network because we were processing one-dimensional data, while the ResNet model was based on two-dimensional data processing. We set ResNet network as three modules, and each module has the Block mentioned above. However, different from the original text, we modified the layer number of neural network and the size of convolution kernel to make it more suitable for one-dimensional data, as shown in figure 2:
Figure 1. This is one Block of the Residual Network. It is the main module of the residual network. It is composed of two convolutional layers.

Figure 2. This is a complete ResNet, which is made up of many blocks stacked and divided into four modules. The number of corresponding blocks is 3, 5, 7 and 3, respectively. Besides, the convolution kernel of the convolutional neural network in the Block is also corresponding to 16, 32, 64 and 128.

4.2. Bi-LSTM

LSTM when it is put forward based on RNN, on the basis of RNN able to capture the dependence of series data, however, because of RNN as the sequence is too long and too deep, the same problems so LSTM in order to solve this problem, the introduction of the three doors, respectively enter door, forget the door.

Before introducing the three gates, we need to introduce the concept of a unit state, which is calculated based on the output of last time and the input of this time. The intuitive understanding is that it is based on the output of neural network layer by layer.

\[
\tilde{c}_t = \tanh(W_c \cdot [h_{t-1}, x_t] + b_c)
\]

Input gate, Where \( w_i \) is the weight matrix of the forgetting gate and \( b_i \) is the bias term.

\[
i_t = \sigma(W_i \cdot [h_{t-1}, x_t] + b_i)
\]

Forget gate, Where \( w_f \) is the weight matrix of the forgetting gate and \( b_f \) is the bias term.

\[
f_t = \sigma(W_f \cdot [h_{t-1}, x_t] + b_f)
\]

New unit state, intuitively understand, forget state before a layer of the unit will be passed on to the door, get forgotten to retain information, as well as the state of this unit to enter the door, can extract the information they need to form this unit, this new unit state had previously information and information now:

\[
c_t = f_t \cdot c_{t-1} + i_t \cdot \tilde{c}_t
\]

Output gate, Where \( w_o \) is the weight matrix of the forgetting gate and \( b_o \) is the bias term.

\[
o_t = \sigma(W_o \cdot [h_{t-1}, x_t] + b_o)
\]
Final output:

\[ h_t = o_t \cdot \tanh(c_t) \]

This paper adopts Bi-LSTM, which is a bidirectional LSTM, that is, the output not only contains the previous information, but also the information behind. The reason for this bidirectional approach is that we consider the ECG data to be a continuous data, and that the future data is related to the previous one, which also affects the later one.

4.3. ResNet+Bi-LSTM

We put forward a kind of based on Resnet + Bi - LSTM serial structure of neural network. First use of Bi - LSTM capture data dependencies between, then get the output of the LSTM, to connect the output of a LSTM, then through the ResNet convolution feature extraction, finally through the softmax classification, get the final output. The network structure as shown in figure 3.

![Figure 3](image)

**Figure 3.** Our network structure, the first one is the data, which is processed by ResNet and Bi-LSTM respectively and then sent to a fully connected one. Finally, classification is made.

5. Experiment

5.1. Experimental Setup

In this experiment, the number of training sessions was 32 on the PTB dataset and 16 on the MIT dataset the learning rate was set to 0.001, the batch-size was set to 128, the size of the convolution kernel was 3, and the step size was 2. The optimizer selected SGD and Adam respectively.

5.2. Evaluation Indicators

Before introducing evaluation indicators, four concepts are introduced: That is to say, several common terms of model evaluation are introduced here at first. Now, it is assumed that there are only two categories of classification objectives, namely positive examples and negative examples:

1) True positive (TP): the number of positive examples that are correctly divided into positive examples, i.e. the number of instances (samples) that are actually positive examples and are divided into positive examples by the classifier;

2) False positive (FP): the number of instances that are incorrectly classified as positive, i.e. the number of instances that are actually negative but are classified as positive by the classifier;

3) False negatives (FN): is divided into the number of negative example, wrongly, that the real positive cases but were classifier divided into negative cases on the number of instances

4) True negatives (TN): is the number of correctly classified as negative cases, the actual negative cases and is the instance number classifier divided into negative cases

1) accuracy
Accuracy is our most common evaluation index, accuracy = (TP+TN)/(P+N), which is easy to understand, that is, the number of samples divided by all samples. Generally speaking, the higher the accuracy, the better the classifier.

2) sensitivity
Sensitivity = TP/P, which represents the proportion of pairs in all positive examples, and measures the classifier's ability to recognize positive examples.

3) precision
Precision is a measure of accuracy, indicating the proportion of positive examples in the example divided into positive examples. Precision = TP/(TP+FP).

In this experiment, ACC was used to evaluate the MIT data set as a whole, and sensitivity and precision were used to evaluate each type of MIT data set. Only ACC evaluation index is used for PTB data set, because there are only two types of data set.

6. Result
The performance of our model in the MIT data set PTB data set is shown in figure 4 and figure 5. In figure 5 and figure 6, we can also see the comparison results of ResNet+ Bi-LSTM and ResNet and Bi-LSTM models. As a result, on the surface of the two data sets, the performance of my model is better than that of the two models. The overall performance is ok, and the accuracy reaches 96.2% and 99.6%, respectively. We also made a number of comparisons between our model and others, and the results are shown in table 6. According to this table, we can see that our model performs better than these models, and also shows good precision and sensitivity in SVE, F and Q heart beats that some models are not good at.

![Figure 4. Representation of the Three Models on the MIT Dataset](image1)

![Figure 5. Representation of the Three Models on the PTB Dataset](image2)
7. Summary
In this paper, we proposed a ResNet combined with LSTM network structure to classify ECG heartbeat, and finally achieved good results, achieving 96.2% accuracy on MIT dataset and 99.6% accuracy on PP dataset. However, there are still some shortcomings, because our neural network is too deep, so our training time is particularly long, and there is still a big gap in the efficiency, which can be improved in the future.

Table 3. Performance Comparisons of Different Models on MIT Datasets

| Method            | Acc (%) | N Se/+/P (%) | S Se/+/P (%) | V Se/+/P (%) | F Se/+/P (%) | O Se/+/P (%) |
|-------------------|---------|--------------|--------------|--------------|--------------|--------------|
| Our model         | 96.2    | 98.2         | 78.3         | 26.5         | 55.7         | 78.6         |
| Yu and Chou [9]   | 75.2    | 78.3         | 79.2         | 1.8          | 5.9          | 83.9         |
| Song et al. [10]  | 76.3    | 78.0         | 83.9         | 27.0         | 48.3         | 80.8         |
| Güler and Übeyli  | 66.7    | 69.2         | 72.1         | 0.0          | 0.0          | 78.8         |

8. Acknowledgments
This work was supported by the Sichuan Science and Technology Program under Grant 2019YFH0185. (Corresponding author: Haoxi Zhang.)

9. References
[1] Yin W, Yang X, Zhang L, et al. ECG Monitoring System Integrated with IR-UWB Radar Based on CNN [J]. IEEE Access, 2016, PP (99):1-1.
[2] Chang, Pei-Chann, Lin, Jyun-Jie, Hsieh, Jui-Chien, & Weng, Julia. Myocardial infarction classification with multi-lead eeg using hidden markov models and gaussian mixture models. Applied Soft Computing, 12(10), 3165-3175.
[3] Simonyan K, Zisserman A. Very Deep Convolutional Networks for Large-Scale Image Recognition [J]. Computer Science, 2014.
[4] Krizhevsky A, Sutskever I, Hinton G. ImageNet Classification with Deep Convolutional Neural Networks [C] NIPS. Curran Associates Inc. 2012.
[5] Szegedy C, Liu W, Jia Y, et al. Going Deeper with Convolutions [J]. 2014.
[6] Hochreiter S, Schmidhuber, Jürgen. Long Short-Term Memory [J]. Neural Computation, 1997, 9(8):1735-1780.
[7] Yu, S. N., & Chou, K. T.. (2008). Integration of independent component analysis and neural networks for eeg beat classification. Expert Systems with Applications, 34(4), 2841-2846.
[8] Asl, B. M., Setarehdan, S. K., & Mohebbi, M. . (2008). Support vector machine-based arrhythmia classification using reduced features of heart rate variability signal. , 44(1), 51-64.
[9] nan Güler, & Elif Derya übeylı (2005). Ecg beat classifier designed by combined neural network model. Pattern Recognition, 38(2), 199-208.
[10] Kachuee, M., Fazeli, S., & Sarrafzadeh, M. . (2018). Ecg heartbeat classification: a deep transferable representation.