Detection of sleep apnea syndrome by CNN based on ECG

Yunxiang Bai, Luqiao Zhang*, Dechao Wan, Yu Xie, Hanghang Deng
Chengdu University of Information Technology, No.24, Xuefu Road, Chengdu 610225, China

*9357094@qq.com

Abstract: Sleep apnea syndrome is a sleep disease that may lead to sudden death. Long term apnea syndrome can cause chronic cerebral hypoxia, hypertension, cardiovascular and cerebrovascular complications. At present, PSG is the most reliable method for diagnosis. But the diagnosis of PSG is complex and expensive. Electrocardiograph (ECG) and portable medical equipment have been widely used nowadays, which makes the acquisition of ECG signal more and more popular and convenient. In this paper, a convolution neural network based on ECG signal is proposed to predict apnea syndrome, the accuracy and sensitivity of this CNN model for apnea syndrome classification are 94% and 88% respectively. The results show that this method has the advantages of low cost and low complexity.

Key words: apnea syndrome, ECG, CNN

1. Introduction
Sleep apnea syndrome (SAS) is a kind of sleep disease which seriously affects people's sleep quality and health. The common manifestations of sleep are snoring, short-term asthma, respiratory arrest, body twitch and even shock. It can lead to chronic cerebral hypoxia, hypertension, arrhythmia, myocardial infarction and other diseases in the long term. Polysomnogram (PSG) monitoring is the "gold standard" [1] in the diagnosis of SAS. The monitoring signals of PSG are various and complex, and the equipment is expensive. The monitoring process and results need professional and technical personnel to monitor, analyze and process, which makes it difficult to popularize and apply in clinical practice because of the high professional requirements for medical personnel. At the same time, the physiological signal acquisition of PSG needs to wear a variety of sensors on patients. The complicated wearable monitoring methods increase the physical and mental burden of patients, and it is difficult to monitor successfully at one time. Due to the inconvenience of PSG, electrocardiograph and portable medical equipment have been widely used, which makes the acquisition of ECG signal more and more popular and convenient.

Based on RR interval series and ECG morphology, Chazal [2] and others calculated 88 characteristic parameters, including respiratory signal (EDR) parameters derived from R wave amplitude, and compared the classification accuracy before and after using EDR parameters. The results show that the detection accuracy is high after adding EDR parameters. The automatic detection algorithm designed by kesper et al. [3] achieves 80.5% correct detection rate by analyzing HRV and EDR parameters. Acharya et al. [4] proposed to train BP neural network using five nonlinear parameters such as ECG correlation dimension and fractal dimension to automatically classify sleep apnea syndrome. Van steenkiste et al. [5] proposed a new method for training LSTM network on
respiratory signal itself, which can detect OSA, CSA and hypoventilation. Ajit et al. [6] used CNN based method to classify apnea of ECG signals per minute segment with an accuracy of 86.22%. In this paper, a self coding convolution neural network based on ECG signal is proposed to predict apnea syndrome, the accuracy and sensitivity of this CNN model for apnea syndrome classification are 88% and 94% respectively.

2. Method

2.1. Data set
The data set used in this paper is based on www.phsionet.org Apnea ECG database [7] [8] was completed by Al Goldberger et al. There are 70 records in the data set, 35 in the training set and 35 in the test set. The records range from 7 to 10 hours, and the sampling frequency is 100 Hz. Each ECG record in the training set is labeled with the length of 1 minute, apnea or normal. The original test set does not contain the label. In the label file, each label represents whether there is apnea within 1 minute from the beginning of the sampling point. That is, the sampling point 0 is marked as N, indicating that the sampling points 0-5999 is normal. The training set is divided into training set and validation set according to the ratio of 8:2. When the performance of the model on the validation set reaches the expectation, it is tested on the test set.

2.2. Data processing
ECG signal is divided into 1 minute length, and invalid segments are removed according to sampling points in the annotation file, and 17023 minutes marked ECG segments are obtained. The ECG segment signal is filtered by the encoder to obtain new ECG segments, which are used as input of neural network, as shown in Figure 1.

![Fig. 1 Pretreatment of ECG signal.](image)

The self encoder can be used as a feature detector in the data pre training of deep neural network. Firstly, the data learn the features through the encoder, and then reconstruct the data through the decoder. Its basic form is as in Formula 1:

\[ x \xrightarrow{h} [g \, g] \xrightarrow{f} x'. \]  

(1)

The simplest loss function is MSE as in Formula 2:

\[ L(x, x') = \frac{1}{n} \sum_{i=1}^{n} (x_i - x_i')^2. \]  

(2)

After the reconstruction of the self decoder, the noise can be filtered out and the main components can be extracted. The comparison between the original signal and the processed signal is shown in Figure 2.
2.3. 1D convolution neural network model

CNN is a supervised learning network with automatic feature extraction, which has been widely used in the field of image processing and its effectiveness has been proved. CNN network can extract the intrinsic characteristics of signals by convolution operation. Convolution operation can be expressed as in Formula 3:

\[ h = \text{ReLU}(X*W + b). \]  \hspace{1cm} (3)

\( X \) is the input signal, \( W \) is the convolution kernel, and \( b \) is the bias. ReLU is a linear correction unit, as in Formula 4:

\[ \text{ReLU}(x) = \max(0,x). \]  \hspace{1cm} (4)

In this paper, we use CNN network to do 1D convolution instead of 2D convolution in image processing. The network consists of two convolution blocks (1 volume layer, 1 batch normalization (BN) layer, 1 dropout layer, 1 ReLU layer, 2 max pooling layers and 3 full connection layers. BN layer and dropout layer are between the volume layer and pool layer, and BN layer can make the model converge faster, and dropout layer avoids model over fitting. As shown in Figure 3:

\[ \text{ReLU}(x) = \max(0,x). \]  \hspace{1cm} (4)

In the first convolution block of the proposed CNN model, we choose the expansion filling. The kernel size of 5 and step size of 1 are used to convolute the data to extract 100 features. We initialize the weight matrix with a normal initialization program. In order to speed up the training process, after convolution, batch normalization is carried out. In order to prevent overload, a dropout layer with a probability of 0.5 is introduced. Finally, ReLU function is used to activate. The kernel size of max pooling layer is 10. In the second convolution block, the parameters are consistent and 10 features are extracted. After max pooling layer, the data is flattened and sent to the full connection layer. Batch normalization and ReLU activation are also used to improve the nonlinear expression ability of each
layer. Finally, the output is sent to the classifier. The cross entropy loss function is used to update the network.

2.4. Hyper-parameters setting
In machine learning or deep learning, researchers need to set some parameters according to their own experience and professional knowledge before learning, and select a group of relatively optimal parameters to train the model. These parameters are not updated iteratively according to the optimization algorithm, which are called Hyper-parameters. After many experiments to adjust the parameters, the following parameters are selected to train proposed convolution neural network.

a) Learning rate: it determines whether the loss function (loss) can converge to the minimum value in the appropriate time. If the learning rate is too small, the convergence speed will be very slow. If the learning rate is too large, the gradient may vibrate back and forth near the minimum value, so that the convergence can not be achieved. After many small-scale adjustments, the network efficiency is better when the learning rate is set to 0.01.

b) Mini batch: in model training, take out one mini batch each time, calculate the loss, and update the network weight and bias according to the optimization algorithm. If the value of mini batch is too small, it can not be applied to the fast operation of matrix library, so learning becomes slow. If the value selection is too large, the weight cannot be updated well. Therefore, we need to choose a compromise value to optimize the operation rate and weight update. In this experiment, the mini batch is set to 32.

c) Epoch: each time the network is trained, all the training sets are called an epoch cycle. If the epoch setting is too small, the network will not have enough data and time to train to get the optimal parameters. If the epoch setting is too large, the network training will be over fitted, resulting in very low test accuracy. In this experiment, epoch is set to 15, and the training effect is better. Generally speaking, in order to get better accuracy, researchers usually design the network model deeper and get more advanced features. But the deeper the layer, the more parameters. Theoretically speaking, the more parameters, the better the fitting degree of the model, but the more prone to over fitting. Table 1 shows the details of the proposed architecture and the network hyper-parameters.

| Layer Type | Kernel Size | Output Shape | Number of parameters |
|------------|-------------|--------------|----------------------|
| Input Layer | [1,6000] | 0 |
| Conv1d1 | [100,1,5] | [100,6000] | 600 |
| BN1 | | [100,6000] | 0 |
| Dropout1 | | [100,6000] | 0 |
| Activation1(relu) | | [100,6000] | 0 |
| Max1 | [1,10] | [100,600] | 0 |
| Conv1d2 | [10,1,5] | [10,600] | 5010 |
| BN2 | | [10,600] | 0 |
| Dropout2 | | [10,600] | 0 |
| Activation2(relu) | | [10,600] | 0 |
| Max2 | [1,10] | [10,60] | 0 |
| FC1 | | [128] | 76801 |
| FC2 | | [16] | 2049 |
| FC3 | | [2] | 33 |

Tab. 1 Architecture of CNN model.
2.5. Performance parameter

We used four indicators to evaluate the performance of the model, true positive (TP), false positive (FP), true negative (TN), false negative (FN). TP means that the model predicts normal and the actuality is normal; FP means that the model predicts normal and the actuality is abnormal. TN is the model prediction is abnormal, the actualy is abnormal, FN is the model prediction is abnormal, the actualty is normal.

\[
SN = \frac{TP}{TP+FN} \times 100\% \quad (5)
\]

\[
SP = \frac{TN}{TN+FP} \times 100\% \quad (6)
\]

\[
AC = \frac{TP+TN}{TP+FP+TN+FN} \times 100\%, \quad (7)
\]

SN in Formula 5, SP in Formula 6, and AC in Formula 7 represent sensitivity, specificity, and accuracy.

3. Results

In this paper, the ECG signal is preprocessed to get the filtered signal, and then the processed signal is used as the input of CNN model. Through the experiment, it is found that when epoch exceeds 15, the accuracy can not be effectively improved and slightly decreased. Therefore, epoch is set to 15. At the same time, the results of general neural network and LSTM are compared, and it is found that the model in this paper is obviously better than the other two. The accuracy and sensitivity of the proposed model are 94% and 88% on the validation set and 92% and 90% on the test set, respectively. Figure 4 shows the accuracy of the model on training set, validation set and test set under different epochs.

![Fig. 4](image)

(a) Accuracy of model on three datasets (b) Comparison of the proposed model with MLP and LSTM

4. Conclusion

In this study, we propose a CNN based classification method for apnea syndrome. We use the self encoder to preprocess the ECG signal, and filter the noise of the signal through the self encoder. Our method based on CNN can be used as a reliable method to detect sleep apnea. This technology can use portable ECG signal collector to detect apnea, which can solve the problem of complex, expensive and inconvenient diagnosis of apnea syndrome. Our next step is to use a portable ECG signal collector to collect signals and apply this method, which will be improved in the future.

References

[1] Epstein L J and et al 2009 Clinical guideline for the evaluation, management and long-term care of obstructive sleep apnea in adults Journal of Clinical Sleep Medicine Jcsm Official Publication of the American Academy of Sleep Medicine 3 263
[2] De Chazal P and et al 2003 Automated processing of the single-lead electrocardiogram for the detection of obstructive sleep apnoea IEEE transactions on bio-medical engineering 6 686
[3] Kesper K and et al. 2012 "ECG signal analysis for the assessment of sleep-disordered breathing and sleep pattern Medical & Biological Engineering & Computing pp 135-144
[4] Acharya, U Rajendra and et al 2011 Automated detection of sleep apnea from electrocardiogram signals using nonlinear parameters Physiological Measurement pp 287-303
[5] Van S, Tom and et al 2018 Automated Sleep Apnea Detection in Raw Respiratory Signals Using Long Short-Term Memory Neural Networks IEEE Journal of Biomedical and Health Informatics p 1
[6] Sinam, Ajit and S. Majumder 2019 A NOVEL APPROACH OSA DETECTION USING SINGLE-LEAD ECG SCALOGRAM BASED ON DEEP NEURAL NETWORK Journal of Mechanics in Medicine and Biology 4 1950026
[7] Penzel T and et al 2000 The apnea-ECG database." Computers in Cardiology IEEE, 2000
[8] Goldberger A and et al 2000 PhysioBank, PhysioToolkit, and PhysioNet: Components of a new research resource for complex physiologic signals. Circulation (Online) pp 215–220