Deep Learning With Convolutional Neural Networks for Sleep Arousal Detection

Dongya Jia[1], Shengfeng Yu[2], Cong Yan[1], Wei Zhao[1], Jing Hu[1], Hongmei Wang[1], Tianyuan You[1]

[1]Guangzhou Shiyuan Electronics Co, Ltd, Guangzhou, China
[2]South China University of Technology, Guangzhou, China

Abstract

Sleep arousal influences the quality of sleep and causes health problems. Polysomnography (PSG), a group of biological signals, is often used to diagnose sleep arousal. But it is costly for sleep experts to identify sleep arousal via PSG. Thus, we designed an automatic algorithm to analyze PSG for sleep arousal identification.

According to the nature of sleep arousal, we selected electroencephalogram (EEG) from PSG for further analysis. To extract frequency domain information, Welch algorithm was applied to EEG to obtain power spectral density (PSD). And then PSD was fed into a 34-layer convolutional neural network (CNN) for further feature extraction and classification. Shortcut connections were employed across every two convolutional layer to speed up the training process and realize identity mapping.

Our model was trained on 900 subjects' PSGs and validated on 94 subjects' PSGs. And 989 subjects' PSGs were used as test set. Our method achieved an AUPRC of 0.10 on the full test set. There is some room for improvement comparing with other methods.

1. Introduction

During the process of sleep, multiple potential factors will influence the quality of sleep and consequently lead to many other health problems. Among these factors, multiples kinds of sleep arousal account for a large proportion. Sleep arousal does not mean completely awakening from sleep, but partial "arousal" from slow wave sleep, which can occur at any sleep stage. Abrupt change of the pattern brain wave activity is an important characteristic of sleep arousal.

In Physionet Challenge 2018, competitors are asked to design a sleep arousal detection algorithm. The sleep arousals to be identified mainly consist of respiratory effort related arousal (RERA), spontaneous arousal, hypoventilations, vocalization, snores, periodic leg movements, partial airway obstructions. 13-channel PSG data are provided for analysis. According to the nature of sleep arousal, these multiple kinds of arousals share an identical characteristic: variation of brain wave pattern. Therefore, in order to identify these arousals in the same manner, we only analyzed 6-channel EEG for arousal detection.

Due to strong capability of feature extraction and convenient usage, deep learning has been one of the most popular techniques to solve various image or signal classification problems. There have been a number of proposed methods using deep learning technique to analyze physiological signals, especially EEG. Akara's group proposed DeepSleepNet to automatically score sleep stage based on single-channel EEG[1]. R. Schirrmeister's group designed a CNN to decode EEG pathology[2]. But there is few work applying deep learning on PSG data to detect sleep arousals. H. Espiritu's group used adaptive segmentation strategy and hand-crafted features to identify sleep arousals[3], but the evaluated data set was small and the variety of arousal was poor.

In this study, we proposed a data-driven, EEG-oriented method to identify sleep arousals. Our method was mainly composed of two steps, data preprocessing and classification. Our method achieved an AUPRC of 0.114 and an AUROC of 0.646 on official test set of Physionet Challenge 2018.

2. Dataset Overview

The PSG data are provided by the Physionet Challenge 2018. The dataset was officially partitioned into two parts base on subjects, training set (n=994) and test set (989). Additionally, a group of records of 94 subjects were randomly selected from training set, as validation set. Each record lasts from 7 to 9 hours, and each sample is labeled as arousal (1), normal (0) and unscored (-1). The dataset contains six kinds of physiological signals, including: EEG, electrooculography (EOG), electromyography (EMG), electrocardiolography (EKG), oxygen saturation and respiratory airflow. In this challenge, regions containing Obstructive Sleep Apnea Hypopnea Syndrome are excluded and not scored. More information can be found in the official website[4].

 ISSN: 2325-887X DOI: 10.22489/CinC.2018.371

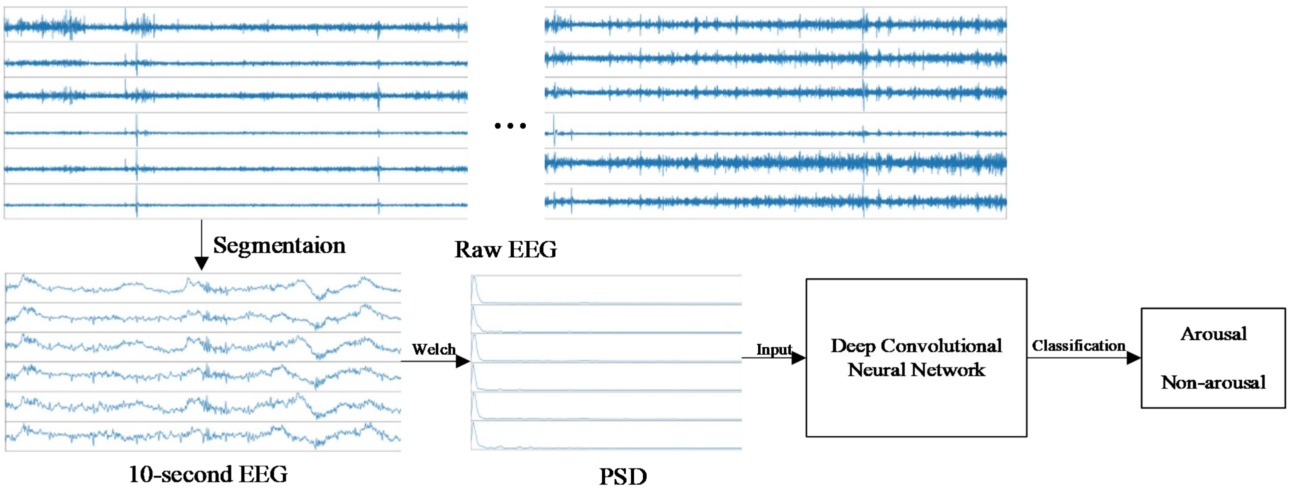

Figure 1. Flow chart of the proposed method

3. Methods

In this chapter, we first describe our EEG preprocessing method. And then the proposed convolutional neural network is explained. The flow chart of the proposed method is shown in figure 1.

3.1. Data Preprocessing

Due to the high complexity of the sample-wise classification, we implemented a simplified segment-wise classification method. Firstly, each record is divided into 10-second segments using a 10-second sliding window. Segment-wise label is defined by the most sample-wise labels in a segment. In training process, segments labeled as "-1" were excluded.

To address the imbalance between negative and positive class, the step size of sliding window was adaptively adjusted according to corresponding segment. We applied a step size of 10 seconds or 2 seconds, when the label of current segment is judged as 0 or 1, respectively.

Each record was downsampled from 200Hz to 100Hz to reduce complexity. Then the power spectrum density of each 10-second EEG record was computed using Welch's algorithm[5]. Welch's algorithm first divides the record into several overlapping sub-segments, and then computes PSG for each sub-segment and finally averages them. The length of sub-segment was set to 256 and the overlapping length was 128. Then we obtained a 129-point PSD segment from a 10-second EEG record. Finally, the PSD segments were normalized by maximum normalization.

3.2. Classification

After the preprocessing, each 10-second EEG was transformed to a PSD segment. 6 channels of PSD segment were together fed into the proposed convolutional neural network, and was predicted as arousal or non-arousal. The architecture of the proposed CNN is shown in Figure 2.

The proposed CNN is mainly composed of multiple basic blocks with shortcut identity connections[6], which can make deep neural network trainable and speed up the training process. Each block consists of two 1-d convolutional layers, two activation functions and two batch normalization layer. The filter size of each convolutional layer is 1x3, except the first convolutional layer whose filter size is 1x7. The filter number is doubled when feature map size is halved by convolutional layers that have a stride of 2. A batch normalization layer is applied after each convolutional layer to ensure the outputs satisfy normal distribution, which can speed up the convergence process and reduce impact from poor weight initialization. And then rectified linear activation unit (ReLU) is applied after the BN layer to improve the model' s capability of describing nonlinearity. After all the basic blocks, an average pooling layer with kernel size of 1x5 is followed to smooth the feature map and reduce the redundancy features. The final unit is a fully-connected layer with a softmax activation function, which is used to execute classification task according to features extracted by preceding convolutional blocks.

The proposed deep learning model was developed using Pytorch 0.4.1 framework.

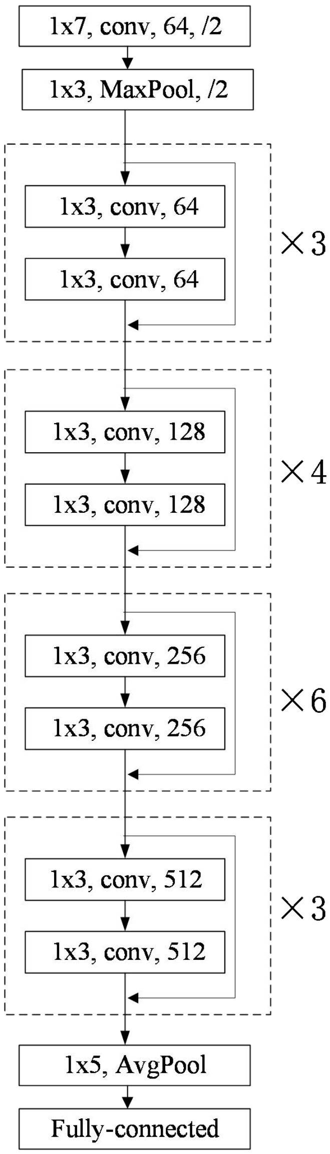

Figure 2. The architecture of our neural network. In this figure, each “conv” block contains a convolutional layer, a batch normalization layer and a ReLU activation function, sequentially. Feature map is halved at the first convolutional layer of each dotted block.

4. Results

4.1. Experiment Design

Weights of Convolutional layers and fully-connected layer were initialized using He’s method[7]. And weights of batch normalization layer were initialized by filling 1. The learning rate was initially set to 0.001 and was halved every 30 epochs. All segments from different subjects in training set were shuffled. The mini-batch size was set to 1024. Binary cross entropy loss function was adopted, and a L2 regularizer with coefficient of 0.001 was added to loss function for the purpose of addressing over-fitting. Adam optimizer[8] with backpropagation was adopted to update model parameters.

4.2. Experimental Results

The model yielded best result on validation set was assessed on official test set provided by Physionet Challenge 2018, consisting of 989 subjects’ records. The evaluation metrics are AUPRC and AUROC. The experimental results are shown in Table 1.

In addition, Figure 3 shows the curves of AUROC and AUPRC on training set and validation set, with the training epoch increasing. We can see that although the evaluation result of training set become better with epoch increasing, however, the result of validation set is always very bad.

Table 1. Final evaluation results on different sets.

	AUPRC	AUROC
Training set	0.125	0.674
Validation set	0.110	0.636
Test set	0.100	\

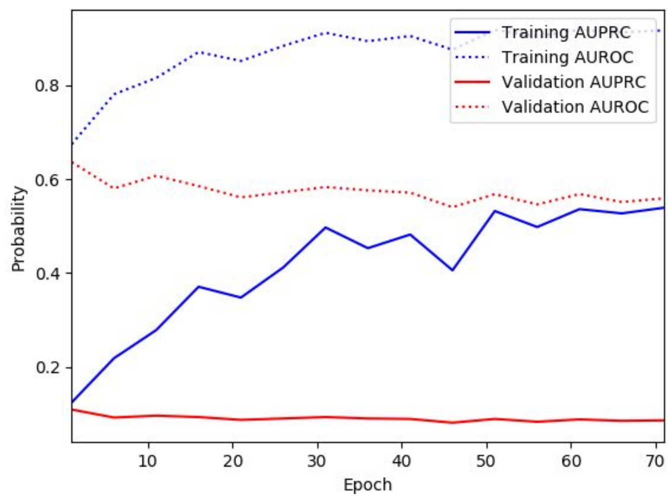

Figure 3. The curves of AUPRC and AUROC on training set and validation set.

5. Discussions

5.1. Model

Official training data (135GB) are stored in format of int16, and after being transformed to format of float64 for computing, the volume of data will be much larger. Large database processing has high demand of memory capacity. In consideration of limited hardware resources as well as the physiological nature of sleep arousal, we selected 6-channel EEG signals for sleep arousal detection.

In official training set, sleep arousal region is very sparse. According to the statistics, the ratio of the number of sample labeled as “1” to labeled as “0” is approximately 1:22. Therefore, to address data imbalance, we implement an adaptive sliding window technique to increase the proportion of positive class.

In our proposed method, power spectra density of EEG, instead of waveform of EEG, is used as input of neural network. The reason is that EEG is a kind of stochastic signal and its waveform information is of little significance. Comparing with time domain, the pattern of frequency domain information is much easier to learned by neural network. Additionally, PSD segment is shorter than waveform segment, which means lower computational complexity.

Deep neural network is well capable of extracting features of high-level abstraction. Therefore, we adopted deep learning model, instead of using traditional combination of hand-crafted features and a standalone classifier, as feature extractor and classifier. But deep learning method has some drawbacks. Firstly, the training process requires large dataset, otherwise serious overfitting would occur. Secondly, training or evaluating a deep neural network is computationally costly, and high-performance GPU is necessary to save time. Our deep learning model was trained using the Titan XP.

5.2. Distribution difference between subjects

Before training, the official training set (994 subjects) are randomly divided into two parts: training and validation. During training, although the evaluation result of training set become better with epoch increasing, however, the evaluation result of validation set is always very bad. The results are demonstrated in Figure 3.

From the first look, it seems to be an overfitting problem. We have tried some techniques to address overfitting, such as L2 regularization, dropout layer, data augmentation, etc. But they all failed to work. In a common overfitting problem, the performance curve of validation set normally first goes up and then fall down. But in our case, the performance curve of validation set never goes up regardless of training for how many epochs.

So, we think that the problem is not about overfitting but distribution difference between subjects. Although to some extent, we have a sufficiently large training set, however, the validation result never improves. To address this problem, we tried normalizing each PSD segment individually. Although this trick speeded up the convergence while training, it still failed to solve the problem.

6. Conclusions

The purpose of Physionet Challenge 2018 is accurately detecting sleep arousal region via a group of physiological signals. We proposed a data-driven and EEG-oriented method to address this problem. At preprossing stage, raw EEG signals were divided into short segments and their power spectra densities were computed. At classification stage, we proposed a deep neural network to classify each PSD segment as sleep arousal or not. Our method achieved a AUPRC of 0.114 and a AUROC of 0.646 on official test set. In the future, we plan to analyze more kinds of physiological signals, except for EEG, and improve our method to solve the problem of distribution difference.

References

[1] Akara Supratak, Hao Dong, Chao Wu and Yike Guo. DeepSleepNet: A Model for Automatic Sleep Stage Scoring Based on Raw Single-Channel EEG. IEEE Transactions on Neural Systems and Rehabilitation Engineering, Vol 25, No. 11, Nov 2017

[2] R. Schirrmeister, L. Gemein, K. Eggensperger, F. Hutter and T. Ball. Deep Learning with Convolutional Neural Networks for Decoding Visualization of EEG Pathology. arXiv: 1708.08012v3, 11 Jan 2018

[3] Hugo Espiritu, Vangelis Metsis. Automated Detection of Sleep Disorder-Related Events from Polysomnographic Data. International Conference on Healthcare Informatics. Oct 2015

[4] Mohammad M Ghassemi, Benjamin E Moody, Li-wei H Lehman, Christopher Song, Qiao Li, Haoqi Sun, Roger G. Mark, M Brandon Westover, Gari D Clifford, You Snooze, You Win: the PhysioNet/Computing in Cardiology Challenge 2018, Computing in Cardiology Volume 45. Maastricht, Netherlands, 2018. pp 1-4

[5] P. Welch. The Use of the Fast Fourier Transform for the Estimation of Power Spectra: A Method Based on Time Averaging over Short, Modified Periodograms. IEEE Trans. Audio Electroacoust. Vol 15. pp. 70-73. Jun 1967

[6] Kaiming He, Xiangyu Zhang, Shaoqing Ren, Jian Sun. Deep Residual Learning for Image Recognition. arXiv: 1512.033855v1, 10 Dec 2015

[7] Kaiming He, Xiangyu Zhang, Shaoqing Ren, Jian Sun. Delving Deep into Rectifiers: Surpassing Human-Level Performance on ImageNet Classification. arXiv: 1502.01852, 6 Feb 2015

[8] Diederik P. Kingma, Jimmy Ba. Adam: A Method for Stochastic Optimization. The 3rd ICLR. 2015

Address for correspondence.
Dongya Jia

Central Research Institute, Guangzhou Shiyuan Electronics Co., Ltd, No. 6, 4th Yunpu Road, Huangpu District, Guangzhou, Guangdong, China
jiadongya@cvte.com