Predicting Neurological Recovery Following Coma After Cardiac Arrest Using the R(2+1)D Network Based on EEG Signals

Meng Gao, Rui Yu, Zhuhuang Zhou, Shuicai Wu, Guangyu Bin

Faculty of Environment and Life, Beijing University of Technology, Beijing, China

Abstract

Electroencephalogram(EEG) signals can reflect the electronic activities of the brain and are commonly used for disease diagnosis. For this reason, as part of the George B. Moody PhysioNet Challenge 2023, we BJUT-bme team developed a deep learning model for automated analysis of EEG signals to predict patients ' clinical outcomes after 12, 24, 48, and 72 hours following cardiac arrest. This model used a convolutional neural network with R(2+1)D module and extracted the power spectrum features of EEG and brain functional connectivity features as inputs to the model, then output prediction outcomes. We got scores of 0.602, 0.522, and 0.406 separately on the training set, validation set, and test set. But our team was unable to be ranked.

1. Introduction

Our team participated in the 2023 George B. Moody PhysioNet Challenge, this competition asked participants to develop an automatic analysis algorithm for analysing longitudinal electroencephalogram (EEG) signals and other patient records, and predicting the neurological recovery of patients after cardiac arrest [1,2]. EEG signals can reflect the electrical activities of the brain and are commonly used to analyse diseases. In clinical practice, this technology can help clinicians predict the neurological recovery status of patients after cardiac arrest, improving the accuracy of prognosis diagnosis [3]. Thus, our challenge entry uses a convolutional neural network with R(2+1)D module [4], extracting power spectrum features of EEG and brain functional connectivity features as input of the model.

2. Methods

2.1. Database analysis

The database comes from 1,020 adults who experienced cardiac arrest and recovered cardiac function but remained in a coma [5]. Starting from a few hours after cardiac arrest, all patients received continuous EEG monitoring with 19 channels, recording for hours or days, with low-quality signals, non-physiological artifacts, and transient interruptions. In addition, patients also recorded signals of electrocardiogram and/or other clinical data. The database also provides the cerebral performance category (CPC) scale score. Patients are classified into two categories: good outcome and poor outcome based on CPC scores and clinical results. The records contain EEG signals, electrocardiogram (ECG) signals, reference (REF) electrodes, and other signals. Since all patients had EEG signals, only EEG signals were used for analysis in this paper.

2.2. Data preprocessing

We excluded EEG signals with a recording duration of less than five minutes. Next, we applied bandpass filtering to all lead data within 0.5 to 40 Hz. The EEG data was subsequently segmented into 2-second epochs with a 50% overlap. Any epochs containing noise were discarded based on the noise judgment criteria: an epoch is classified as noise if its maximum-minimum range exceeds 300 uV or if a differential value surpasses 30 uV.

2.3. Feature extraction

After segmenting the EEG data and removing noisy segments, we employed the smoothed periodogram method for spectral analysis.

We smoothed EEG segments using the Hamming window and performed the Fourier transform to obtain the power spectrum:

$$\mathrm{X(k)} = \sum_{\mathrm{n=0}}^{\mathrm{N-1}} \mathrm{x(n)} \times e^{-j2\pi k\frac{n}{N}}, \quad (1)$$

$$\mathrm{PSD} = \frac{|\mathrm{X(k)}^2|}{\mathrm{U \times fs}}. \quad (2)$$

We obtained power spectrum density for one lead,

Fp1, of EEG data shown in Figure 1(a), after extracting all leads of EEG, we got a total of 19×60 power spectrum features for the EEG data of 19 leads, with a range of 0-60Hz (Figure 1(c)).

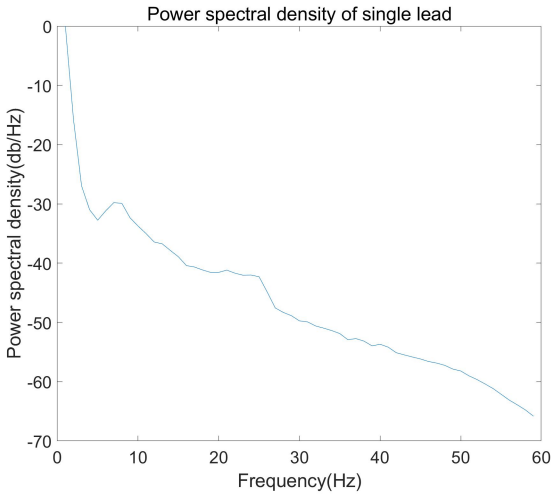
In addition, research has shown that brain damage may cause functional connectivity between different regions of the brain [6]. Therefore, we considered using functional connectivity features of the brain, and analyzed the spectral coherence features between each pair of electrodes using the Spectral coherence method [7]. The formula for calculating Spectral coherence is shown below, which $|S_{xy}(k)|$ represents the amplitude of the cross-spectrum density, and $C_{xy}(k)$ represents the coherence matrix that indicates the coherence between signals x and y at frequency k:

$$C_{xy}(k) = \frac{|S_{xy}(k)|^2}{S_{xx}(k) \cdot S_{yy}(k)}. \quad (3)$$

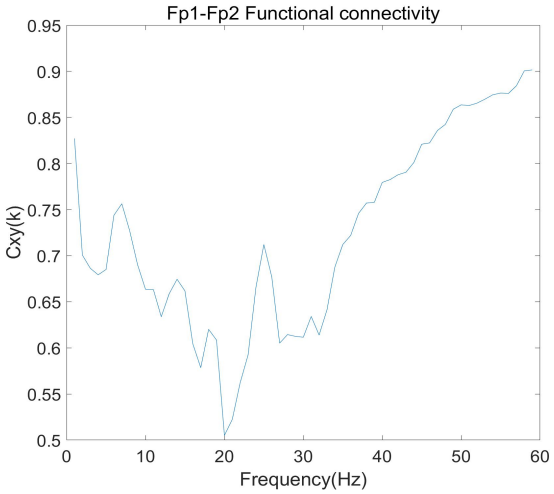
For each pair of electrodes, Spectral coherence features were calculated, as shown in Figure 1(b), resulting in a total of 172×60 functional connectivity features (Figure 1(d)). The power spectrum features and coherence features were concatenated to form a 191 × 60 feature matrix (Figure 1(e)).

By performing the above feature extraction operation on all hours of data, a series of discontinuous feature maps would be formed within 72 hours, as shown in Figure 2(a). Since not all datasets have complete 72 hours data, for the 12h, 24h, 48h records and missing features during records, they were completed by linear interpolation based on each pixel, and finally a 3D (72 × 60 × 191) feature map was formed, as shown in Figure 2(b).

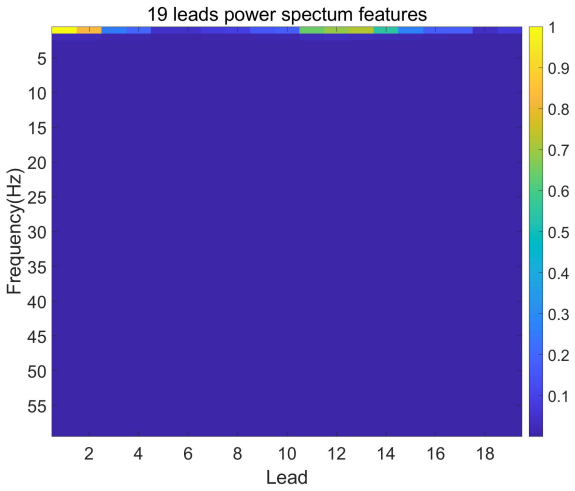
The feature maps within 72 hours were then deleted by 20%, and then filled by linear interpolation again to achieve data enhancement, as shown in Figure 2(c).



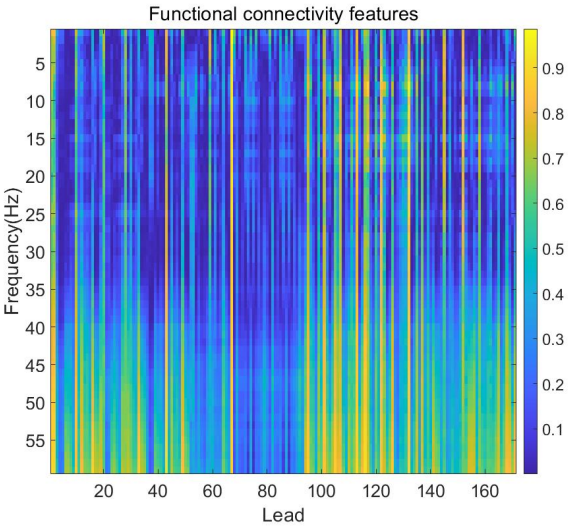
(a) Power spectral density of single lead



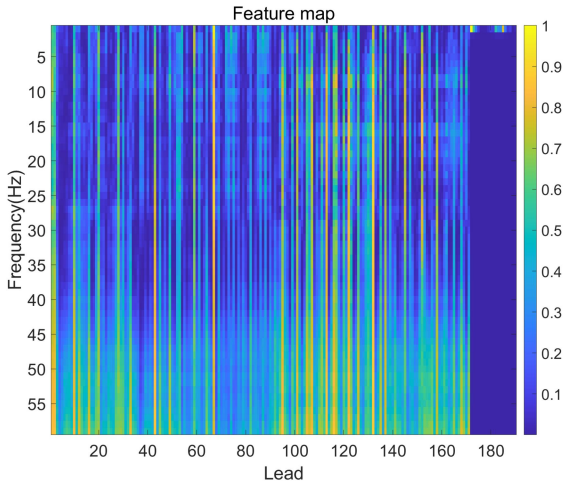
(b) Fp1-Fp2 Functional connectivity



(c) 19 leads Power spectrum features



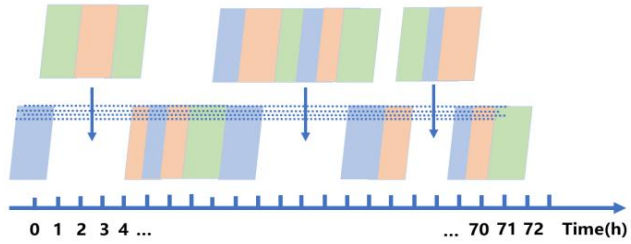
(d) Functional connectivity features



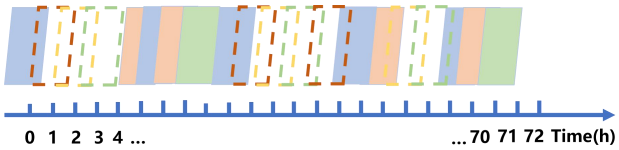
(e) Combined feature map

Figure 1. Feature extraction:
(a) Power spectrum feature of Fp1 lead
(b) Functional connectivity of one pair of electrodes
(c) Power spectrum features of all 19 leads of EEG
(d) Functional connectivity of all pair of electrodes
(e) Combined feature map

(a)



(b)



(c)

Figure 2. Time domain feature extraction process:
(a) Split the signal and extract discontinuous feature maps;
(b) Interpolate the features to pad the missing part;
(c) Remove some feature maps randomly and fill with linear interpolation;

2.4. Model selection

Convolutional neural networks (CNNs) are a type of neural network model that is widely used in computer vision and image processing tasks [8]. Its core idea is to automatically learn the features of images or other data through convolutional layers and pooling layers. 3D convolutional neural networks are an extension of CNNs and are used to process 3D data , such as videos, medical images, or time series data .

R(2+1)D module is a modular structure proposed to improve the effect of 3D convolutional neural networks. It introduces the separation of time and space dimensions in the 3D convolution layer to reduce network parameters and computational complexity. The R(2+1)D module includes two steps: first, apply 1D convolution in the time dimension, and then apply 2D convolution in the spatial dimension [4]. By separating time and space operations, the R(2+1)D module can better model the spatiotemporal relationship in video sequences and provide better feature representation ability. The convolution structure is shown in Table 1.

Layer name	In channels	Out channels	Kernel size
conv1	1	64	3×7×7
conv2	64	64	3×3×3
conv3	64	128	3×3×3
conv4	128	256	3×3×3
conv5	256	512	3×3×3

Table 1. R(2+1)D convolutional neural network structure

The model used Adam optimizer [9]. Weighted cross-entropy was used as the loss function. By assigning different weights to samples of different categories, the model can pay more attention to those categories that are not very common but very important during the training process. To balance between computational efficiency and gradient stability, we set the upper limit of training to 200 epochs to ensure that the model has enough training iterations to master the features of the data. We initialized the learning rate to 0.001. To promote a more stable learning process in the early stages of training, we used a 10-step warming algorithm. In the first 10 steps, the learning rate started from 0.0001 and increased by 0.0001 per step. Finally, we used Kaiming initialization [10] to initialize the model parameters.

3. Results

The scores of our team are shown in Table 2. Unfortunately, our team was unable to be ranked because it could not be retrained. We got scores of 0.602, 0.522, and 0.406 separately on the training set, validation set, and test set. And cross-validation results on the training set are shown in Table 3. The Challenge winner is the

team with the officially ranked entry with the highest Challenge score 0.792 (TPR at a FPR of 0.05) on the test set at 72 hours after ROSC.

Training	Validation	Test	Ranking
0.602	0.522	0.406	None

Table 2. Challenge score (TPR at a FPR of 0.05) on the training, validation, and test sets of our final entry (team BJUT-bme). Our team was unable to be scored on the hidden data.

	Score
CPC MSE	2.247
CPC MAE	1.239

Table 3. Cross-validation results on the training set

4. Discussion and Conclusions

In our first submission, we generated 2D feature maps using only the most recent 5 minutes of EEG data and used them to train and test the ResNet18 model. We obtained a challenge score of 0.25. When we used 3D models and 3D data, we achieved a challenge score of 0.52, indicating that 3D models can effectively utilize temporal information. However, our data augmentation method had some drawbacks. Since we first filled the data to 72 hours using linear interpolation, this reduced the reliability of the information to some extent. In addition, the results for subjects with short sampling times were more unreliable. During model training, we found that the loss curve of the validation set has been fluctuating. This may be due to unstable gradients. As memory limitations of the device used, we did not explore further. Using gradient clipping and Kaiming initialization [10] could slightly reduce this fluctuation. There was also severe overfitting during training; if a pre-trained model or training from scratch is used, the training set will converge quickly, but as training progresses, the validation set fluctuates. We used early stopping algorithms to simply avoid this overfitting problem. We also calculated the challenge score on the validation set during training and found that the value varies considerably and is inconsistent with the trend of loss and accuracy. Later, we used focal loss as the loss function, but this phenomenon did not disappear, which may be related to class imbalance. Loss and accuracy both tried to make the model's judgments as accurate as possible for all categories, so improving the accuracy of judgments for a large number of categories was more likely to improve the overall accuracy of the model. On the other hand, challenge scores were designed to improve the accuracy of positive samples, which lead to inconsistent optimization directions for models. Trying to balance data or incorporating challenge scores into loss functions would be part of our future work.

Acknowledgments

This work was supported in part by the Beijing Natural Science Foundation (Grant No. 4222001), the National Natural Science Foundation of China (Grant No. 11804013).

References

[1] Goldberger AL, Amaral LA, Glass L, Hausdorff JM, IvanovPC, Mark RG, et al. PhysioBank, PhysioToolkit, and PhysioNet: Components of a new research resource for complex physiologic signals. *Circulation 2000*;101(23):e215–e220.

[2] Reyna MA, Amorim E, Sameni R, Weigle J, Elola A, Bahrami Rad A, et al. Predicting neurological recovery from coma after cardiac arrest: The George B. Moody PhysioNet Challenge 2023. *Computing in Cardiology 2023*;50:1–4.

[3] E. A. Samaniego, S. Persoon and C. A. C. Wijman, "Prognosis After Cardiac Arrest and Hypothermia: A New Paradigm," *Curr Neurol Neurosci Rep*, vol. 11, no. 1, pp 111–119, Feb. 2011.

[4] Tran, D., Wang, H., Torresani, L., Ray, J., LeCun, Y., & Paluri, M. "A closer look at spatiotemporal convolutions for action recognition," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit*., 2018, pp. 6450-6459.

[5] Amorim E, Zheng WL, Ghassemi MM, Aghaeeaval M, Kandhare P, Karukonda V, Lee JW, Herman ST, Adithya S, Gaspard N, Hofmeijer J, van Putten MJAM, Sameni R, Reyna MA, Clifford GD, Westover MB. The International Cardiac Arrest Research (I-CARE) Consortium Electroencephalography Database. *Critical Care Medicine*; Oct. 2023; doi:10.1097/CCM.0000000000006074.

[6] D. Cui, J. Wang, Z. Li and X. Li, "A new coherence estimating method: The magnitude squared coherence of smoothing minimum variance distortionless response," in *CISP-BMEI*, 2016, pp. 1440-1445.

[7] F. H. Duffy and H. Als, "A stable pattern of EEG spectral coherence distinguishes children with autism from neuro-typical controls - a large case control study," in *BMC Med.*, vol. 10, no. 1, Dec. 2012, pp. 64.

[8] Tran, D, Bourdev, L, Fergus, R, Torresani, L and Paluri, M, "Learning spatiotemporal features with 3d convolutional networks," in *Proc. IEEE Int. Conf. Comput. Vis.*, 2015, pp. 4489-4497.

[9] D. P. Kingma and J. Ba, "Adam: A Method for Stochastic Optimization," in *Phys. Org.*, Aug. 2017, pp. 1-2.

[10] K. He, X. Zhang, S. Ren, and J. Sun, "Delving deep into rectifiers: Surpassing human-level performance on imagenet classification," in *Proc. IEEE Int. Conf. Comput. Vis.*, 2015, pp. 1026-1034.

Address for correspondence:
Guangyu Bin
Department of Biomedical Engineering, Faculty of Environment and Life, Beijing University of Technology, Beijing 100124, China.
guangyubin@bjut.edu.cn