Motor imagery 디코딩을 위한 Residual Dense Network

Permana Deny, 최계원
성균관대학교

denypermana@g.skku.edu, kaewonchoi@skku.edu

Residual Dense Network for motor imagery decoding

Permana Deny, Kae Won Choi
Sungkyunkwan University

요 약

이 논문은 residual dense network 을 사용한 Motor Imagery 의 디코딩 프로세스에 대한 연구이다. 본 연구에서는 뇌파(EEG)를 사용하여 Motor Imagery 중 뇌파의 신호를 수집하고 처리한 다음 신호를 세 가지 다른 Motor Imagery 신호로 디코딩하는 모델을 학습시켰다. 학습된 모델에서 피험자의 의도를 예측하기 위한 코헨(Cohen)의 카파 상관계수(Kappa)는 0.82 로 계산되었다.

Abstract

The paper studied a decoding process of motor imagery using adopted residual dense network. We acquired and processed the brain signal during motor imagery using electroencephalogram (EEG), then train the model to decode the signal into three different imagery intention. The proposed method achieved Cohen's value in predicting the subject the intention by 0.82.

Ⅰ. Background

Hand movement phases during motor imagery provide insight into the underlying neural mechanism in controlling human limbs. Using an EEG to record brain activity, we investigated the feasibility of brain signals during motor imagery (MI) from different imagery actions to establish a control path between human and computer.

This paper studied a brain-computer interface (BCI) to decode brain activity during motor imagery. The challenge of the decoding method lay on the brain signal that has non-stationary characteristics along the time. Signal representation method through short-time Fourier transform (STFT) and decoding method such as deep learning are proposed in the previous studies to address these challenges.

The decoding method aims to identify the signal distinction of motor imagery performed by the subject through transforming the signal to a smaller matrix representing the subject intention.

We propose a residual dense network architecture [1] consisting deep layer to decode the subtle characteristic from the different intention of movement phases during motor imagery. The adopted model consists of the convolutional network that initially extracts local features of image by accessing the previous input from concatenation and estimating the error through residual connection. These paths lift vanishing gradient in designing deeper layer towards the training process. Therefore, the model is able to generalize the feature to an acceptable degree.

In all, we evaluate the proposed method on the EEG signal during motor imagery to verify the effectiveness of decoding the motor imagery in predicting subject intention.

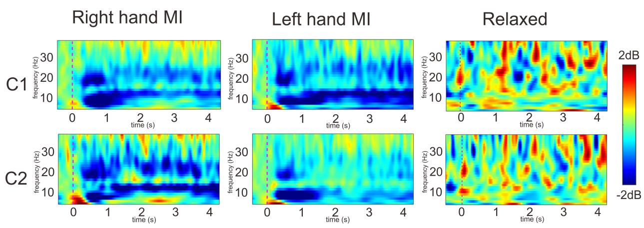

Figure 1. The average event related changes in spectral power of hand movement phases during motor imagery for channel C1 and C2. The STFT heighten the spectral change throughout the time, signified by deep blue color that indicates event related desynchronization (ERD).

Ⅱ. Discussion

We utilized an openBCI device which is consisted of 16 EEG channels, with 10-10 standard channel mapping. In order to acquire the motor imagery data, we asked a subject to perform imagery movement of the right hand, left hand, and relaxed.

This experiment has a total 3000 trials (1000 trials for each type of imagery intention) in the dataset. The EEG was then, bandpass-filtered between 1 and 50Hz followed by a notch-filter at 60Hz. Finally, we extracted the signal from time period one-second pre-instruction and 4 seconds during motor imagery and did an STFT.

We completed the signal analysis to deliver a dataset, to then train a residual dense model. In the model framework, deep residual connection x aims to find underlying map h(x) from F(x) which is defined by:

$$h(x) := F(x) + x \quad (1)$$

The connection makes it possible to estimate the error of x, by estimating the error from residual function. Thus, when x is optimal, then $F(x) \rightarrow 0$ is achieved during training. Therefore, the approach becomes:

$$F(x) := h(x) - x \quad (2)$$

Furthermore, the adopted DenseNet [2] in the weight layer is to fit the residual value from x. The presence of DenseNet architecture is to design deep layer without severe vanishing gradient, by employing skip connection of previous layer to the end of the output.

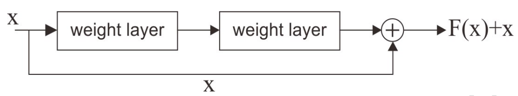

Figure 2. A building block of residual learning [3]. The tensor x represents layer input from previous connection. In this research, the residual layer consists of layer connection from DenseNet architecture.

The model consists of certain depth of residual and concatenation layers that is preserved on figure 3. It consisted of $\theta \approx 36M$ trainable parameters accounting each layer. As the result of feature representation, we impose an input size $x := [16,256,2]$ to the model, while the output $y := [1,3]$. In the training phase, we occupy a callback function to update the learning rate by monitoring loss value. The model is trained for 200 epochs with batch size 32, and implemented Adam optimizer with initial learning rate 0.0001, together with categorical cross entropy as the loss function.

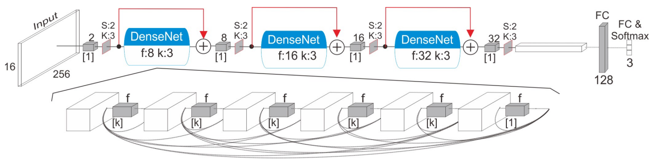

Figure 3. Abstract architecture of the adopted Residual Dense Network. The grey box indicates the Conv2D coupled with BatchNormalization layers. The above diagram is network with residual connection. The below diagram is DenseNet architecture with concatenation. The f represents filter number, while k represents [k,k] kernel number

The model is then evaluated and validated using a test dataset, while the evaluation that is calculated from confusion matrix is reported on the table 1. The significant result from this decoding method is achieved on the offline decoding, with the relaxed intention is highly distinguishable than the other intention.

Additionally, we reported different model properties using the same architecture to optimize the network and hyperparameter. Nonetheless, the deeper model does not achieve comparable result. It indicates that by adding or reducing the trainable parameter or network does not simply solve the decoding problem. Therefore, we further propose a deeper decoding method, different way than directly feeding the feature representation into neural network to explore feature connection over the EEG channels.

Table 1. Confusion matrix evaluation, PPV stand for positive predictive value, TPR for true positive rate, and k for Cohen's k-value. The results are respected to testing test for 900 trials from three different intention.

N=300	Right hand	Relaxed	Left hand	PPV
Right hand	267	8	23	0.89
Relaxed	6	272	21	0.90
Left hand	39	9	251	0.83
TPR	0.85	0.94	0.85	k=0.82

Ⅲ. Conclusion

In this paper, we studied a decoding method for motor imagery from three different intention. By comparing the true positive rate from three different intention, we observed the significant achievement on the relaxed state, hence the two other intentions have the same evaluation result. We achieved overall result performed by Cohen's k-value by 0.82.

ACKNOWLEDGMENT

본 연구는 과학기술정보통신부 및 정보통신기획평가원의 ICT 명품인재양성 사업의 연구결과로 수행되었음 (IITP-2021-2020-0-01821).

Reference

[1] Yulun Zhang, Yapeng Tian, Yu Kong, Bineng Zhong, Yun Fu. "Residual Dense Network for Image Super-Resolution". IEEE Converence on computer vision and pattern recognition, pp. 2472-2481. 2018.

[2] Huang Gao, Zhuang Liu, Laurens Van Der Maaten, Kilian Q.Weinberger. "Densely Connected Convolutional Networks". IEEE Converence on computer vision and pattern recognition, pp. 4700-4708. 2017.

[3] Kaiming He, Xiangyu Zhang, Shaoqing Ren, Jian Sun. "Deep Residual Learning for Image Recognition". IEEE Converence on computer vision and pattern recognition, pp. 770-778. 2016.