Decoding of Visual Information Based on EEG Data

Qing Zhou¹*, Yadong Liu¹, Saisai Zhong¹, Jingsheng Tang¹ and Yaru Liu¹

¹ College of Intelligence Science and Technology, National University of Defense Technology, Changsha, Hunan, 410073, China

* Corresponding author’s e-mail: qzhoukn@163.com

Abstract. We perhaps fantasized about reading the mind of the human brain directly and use it in other ways. In this paper, we tried to analyze human brain activity to identify the feature space employed by humans for visual classification. The brain signals driven by visual information were classified, and the visual information classifier were trained, and the visual information were decoded by the brain signals. The 64 channels EEG data induced by 40 types of images were collected, and recurrent neural network (RNN) was used for extracting features of electroencephalogram (EEG) signals and predicting the types of EEG signals. The accuracy of decoding could reach 80.90%. The technology of decoding of visual information based on brain EEG data could be applied to the brain-computer interface (BCI).

1. Introduction

For decades, researchers have been trying to read people's minds by reading brain signals. Brain activity records contain information about categories such as visual objects and auditory objects, so EEG readings can be used to analyze the information the brain receives from the outside world. Understanding object-induced EEG has been one of the goals of brain-computer interface, which is a direct communication and control channel between human brain and computer or other electronic devices[1].

At present, many research groups are exploring the techniques and methods of decoding the EEG signals and some achievements have been achieved. Recently, thanks to advances in deep learning, researchers have had more choices in decoding brain signals. Deep learning methods have been deeply applied in many fields and achieved performance that traditional methods cannot achieve. Deep learning can be found in signal processing, image classification, scene analysis, speech translation and other fields. In 2015, Kaneshiro came up with the comprehensive approach, training a classifier to distinguish brain computer signals evoked by 12 different object classes with about 29% accuracy[2]. In 2016, Stober proposed a method using convolutional neural networks (CNN) to learn how to classify EEG recordings induced by audio, but the accuracy was merely 28%[3]. And in 2017, Spampinato found that when the subjects looked at different types of images, the evoked EEG signals would be very different, so that the visual information received from the outside could be interpreted by a series of technical means. As shown in Figure 1, the EEG signals induced by the images of pyramids and golf balls are different. Spampinato collected 32-channel EEG signals induced by 40 kinds of target objects, extracted visual features from EEG signals, and input them into the constructed architecture. The accuracy of classification reached 35%[4]. These methods have proved the potential of using deep learning for classification of brain signals, but low accuracy was not suitable for BCIs and other applications. There is still a long way to go to study the decoding of brain signals.
The purpose of this paper is to explore how to classify the EEG signals induced by different kinds of images so as to decode the corresponding image categories of EEG signals. Thanks to the development of deep learning, we can use deep learning network as classifier algorithm.

![Figure 1. Examples of brain signals evoked by visual stimuli of two different ImageNet object classes[4].](image)

2. **Method**

The first step of our approach was to acquire EEG signals recorded while a subject looks at images. Then, we trained an recurrent neural network to extract EEG features from raw EEG signals; the training process was supervised by the class of the images for which each input EEG sequences were recorded, and a classifier for EEG features was jointly trained in the process. The overview of the approach is shown in Figure 2.

![Figure 2. Overview of the proposed approach.](image)

2.1 **The EEG data acquisition**

Six subjects participated in the experiment, including five males and one female. The six subjects were homogeneous in terms of age, education and cultural background. During the experiment, the subjects stared at a computer screen display of 40 images which were the subset of ImageNet[5]. There were 50 images in each category, and each EEG data evoked by an image was taken as a sample. Each image was displayed for 0.5 seconds, and there was no time interval for same type of images. After each type of image was displayed, there was 10 seconds for the subjects to rest. EEG data was collected using the 64-channel active electrode, of which the 1 channel was the reference electrode, so data of 63 channels were used for classification in the experiments. The electrodes were arranged according to the position of the 10-20 system[6]. Amplifier number sampling frequency was set to 1KHz.
2.2 Data preprocessing

In order to exclude any possible interference from the previously shown image, the EEG signals of the first 40ms of each sample were discarded. At the same time, the EEG signals induced by some images are less than 500ms because of the errors in the acquisition equipment, while the neural network required that the input signals have the same dimension. Therefore, for the samples less than 480ms, we choose to discard them. After data screening, the original 12,000 samples were reduced to 11,883. A notch filter with a band resistance range of 49~51Hz was used to eliminate the 50Hz power frequency interference in the signal. The dataset also used a second-order band-pass butterworth filter with a cut-off frequency of 14~71hz above and below to filter out cognitive information unrelated to visual perception. Meanwhile, the data of channel 63 minus the corresponding mean of each channel was centralized, so that the data of each channel was 0 mean. After centralized processing, the data of each channel was divided by its standard deviation, so that the distribution of new data was close to the gaussian distribution.

2.3 The network structure

In this paper, the method of deep learning is used to classify EEG data. Considering that EEG signal is a time series signal, while common RNN structure is prone to gradient dispersion or gradient explosion when processing a long time series, long-term Memory (LSTM) structure is adopted to extract the characteristics of EEG signal for classification[7].

EEG signals are sent to the network in time series. The output of the n layer LSTM serves as the input of the n+1 layer. Different from ordinary LSTM, an additional output layer is added after the last layer of LSTM, and the feature of the EEG is the output of the layer which is connected with the classifier. The classifier is softmax classifier, which can output the classification accuracy of 40 types of images decoding. After training, the encoder can generate EEG features from the input EEG sequence, and the classification network is used to predict the image class of the input EEG feature representation. The network structure is shown as Figure 3.

3. Results

A The EEG data set was divided into training set, validation set and test set with respective fractions 80%, 10%, 10% in order to train and verify the network.

In LSTM networks, the number of looping layers (LSTM Layer) and the hidden state (LSTM Size) are two important parameters, and they affects the learning ability of the network, thus affecting the final result. This paper mainly evaluated the influence of these two parameters on network classification performance. For the convenience of expression, LSTM(Layer, Size) is used in this paper to represent the network structure under a specific parameter. Table 1 reports the classification results under several sets of parameters. Where, Max VA (Maximum Validation Accuracy) is the Maximum Accuracy of the decoding network on the verified data set in the training iteration process, and TA at Max VA (Test Accuracy at Maximum Validation Accuracy) is the decoding accuracy of the network model with the best accuracy on the test set data.

The results in Table 1 show that under the LSTM(4,128) structure, the EEG classification effect is the best, and the test set accuracy is 87.10% under the maximum verification accuracy. When the loop
Table 1. Classification results under several sets of parameters.

| LSTM Layer | LSTM Size | Max VA   | TA max VA |
|------------|-----------|----------|-----------|
| 1          | 64        | 77.77%   | 76.31%    |
| 1          | 128       | 79.04%   | 78.07%    |
| 2          | 128       | 82.16%   | 80.90%    |
| 4          | 128       | 87.35%   | 87.10%    |
| 8          | 160       | 2.97%    | 2.27%     |

Level increases to 8 layers and the implied state size is 160, the network cannot show any classification ability. We consider that under the LSTM(16,120) structure, the number of network parameters is too large, and the data set size in this paper is not enough to optimize the network. Therefore, the classification performance of the network is not good.

In order to explore the differences in EEG decoding ability induced by images of different categories, the EEG of 40 types of images were classified. The classification accuracies of each category is shown in Figure 4.

As can be seen from Figure 4, the network constructed in this paper had great differences in the classification results of different images. Although the final classification results could reach 80%, the classification accuracy of some images was less than 50%, indicating that the EEG signal characteristics were indistinguishable from the EEG signals of other images. However, the classification accuracy of 9 types of images exceeded 80%, which could be used as the stimulation images for online experiments.

The reason why we chose 0.5 seconds for each image display was that the rapid display frequency was easy to cause damage to the eyes of the subjects, and at the same time, we also wanted to collect as much experimental data as possible. The second reason was that the subjects felt tired due to the excessively long experimental time, and it was difficult to have a good experimental effect. Studies have shown that the feature extraction of object recognition by humans takes place between the initial 50~120ms. In order to evaluate the impact of different time stages on classification results, we divided the EEG signals collected in 500ms into multiple time periods as input data, and the model used was LSTM (2, 128). The classification accuracy is shown in Table 2. There was a certain gap in the accuracy after the decoding of EEG signals in different time periods. Similar to the above research results, the best classification result was the initial 40~180ms, which could achieve 89% classification accuracy, and the effect was significantly improved compared with other time periods.

The 64 electrode was adopted in this paper. Although the EEG information collected by 64 channels was more abundant, more electrodes increased the time of experimental preparation in actual operation. Different brain regions have different functions, and it is difficult to evaluate the influence of different brain regions on the decoding. Therefore, the decoding results of EEG signals collected by electrodes distributed in various brain regions were compared by reducing the input data channels.

In this paper, EEG signals in the whole brain area were split into frontal lobe, occipital lobe and parietal lobe, and the LSTM model mentioned above was used for regional training. The results of three
Table 2. Classification accuracy using different portions of EEG signal data.

| Visualization time | Max VA | TA max VA |
|--------------------|--------|-----------|
| 40~480ms           | 82.16% | 80.90%    |
| 40~180ms           | 79.04% | 89.82%    |
| 180~320ms          | 91.03% | 84.02%    |
| 320~480ms          | 77.92% | 76.86%    |

Table 3. Classification accuracy using different portions of EEG signal data.

| Whole brain area | frontal lobe | occipital lobe and parietal lobe |
|------------------|--------------|----------------------------------|
| Accuracy         | 80.90%       | 75.35%                           |
|                  |              | 83.17%                           |

Figure 5. Decoding accuracy of data in different brain region.

Brain region image decoding are shown in Table 3 and Figure 5. Compared with EEG signals from the whole brain, the decoding accuracy of EEG signals in the frontal lobe decreased, while that in the parietal and occipital lobe increased. We consider that this is mainly because the occipital lobe is the center of the visual cortex, which responds more strongly to visual induction. Therefore, it is best to use the signals of the parietal and occipital regions to decode the visual induction.

It can be seen from Figure 5 that, when the accuracy of decoding images of the some categories of one brain region was not high, so the accuracy of the other two brain regions was also low, and the accuracy of the three brain regions was consistent. We believe that these categories of image-induced EEG signals are relatively indistinguishable. In order to get better experimental results, we should try to avoid using the image with low accuracy of decoding when using the online experiment of decoding.

4. Conclusions

In this paper, we explore the decoding of EEG signals of visual information, and build LSTM model to classify the EEG signals of 64 channels and different brain regions. By selecting EEG signals in appropriate time period, the classification accuracy can reach 89.82%. Multiple experiments have proved that the network is stable. However, there are great differences between different categories of image-induced EEG signal in the accuracy of decoding. After channel compression, the decoding effect of EEG signals on the parietal and occipital lobes is better than that in the whole brain and frontal lobes.
In the future work, we plan to design a new experimental paradigm of BCI, explore the real-time performance of image decoding, and apply it to online experiments.

Acknowledgments
This work was supported by the National Key Research and Development Program, under Grant 2017YFB1002505. The authors gratefully acknowledge this support.

References
[1] Devlaminck, D., Wyns, B., Boullart, L. (2009) Brain-Computer Interfaces: from theory to practice. the european symposium on artificial neural networks, 415-424.
[2] Simanova, I., Gerven, M. V., Oostenveld, R. (2012) Identifying Object Categories from Event-Related EEG: Toward Decoding of Conceptual Representations. Plos One, 5(12):e14465.
[3] Stober, S., Sternin, A., Owen, A. M. (2015) Deep Feature Learning for EEG Recordings. Neural and Evolutionary Computing.
[4] Spampinato, C., Palazzo, S., Kavasidis, I. (2017) Deep learning human mind for automated visual classification. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 6809-6817.
[5] Bigdelyshamlo, N, Vankov, A, Ramirez, R. R. (2008) Brain Activity-Based Image Classification From Rapid Serial Visual Presentation. IEEE Transactions on Neural Systems and Rehabilitation Engineering, 16(5): 432-441.
[6] Klem, G. H., Lüders, H. O., Jasper, H. (1999)The ten-twenty electrode system of the International Federation . Electroencephalogr Clin Neurophysiol, 52(3):3-6.
[7] Hochreiter, S., Schmidhuber, J. (1997) Long short-term memory. Neural computation, 9(8):1735-1780