EEG Classification Based on Sparse Representation and Deep Learning

Guangchun Gao¹, Lina Shang¹*, Kai Xiong¹, Jian Fang¹, Cui Zhang¹, Xuejun Gu²

ABSTRACT
For brain computer interfaces (BCIs) research, the classification of motor imagery brain signals is a major and challenging step. Based on the traditional sparse representation classification, a classification algorithm of electroencephalogram (EEG) based on sparse representation and convolution neural network is proposed by this paper. For the EEG signal, firstly, the features of the signal are obtained through the common spatial pattern (CSP) algorithm, and then the redundant dictionary with sparse representation is constructed based on these features. The sparse representation of the EEG signal is completed and the sparse features can be obtained. Finally, the sparse features are transformed into two dimensional signals, and the convolution neural network is used to complete the classification of EEG signals. Using the dataset downloaded from the website of BCI competition III (dataset IVa), for two types of EEG signals, the experiments show that the recognition accuracy of the method is over 80, and the recognition accuracy is better than that of the traditional SRC algorithm.

Key Words: Deep Learning, Sparse Representation Classification, EEG, Convolution Neural Network

Introduction
Brain research is the biggest challenge faced by mankind in natural science in this century, and electroencephalogram (EEG) signal processing is a typical research in this field (Wolpaw et al., 2002; Yong et al., 2008). EEG is a one-dimensional multichannel nonstationary signal which records the electrophysiological activity of the brain. In clinical medicine, the information processing of EEG signals not only provides an objective basis for the diagnosis of certain brain diseases, but also provides effective treatment for some brain diseases (Thornton, 2002; Goker et al., 2012). For a long time, doctors need to manually detect and analyze the waveform characteristics of electroencephalogram, with intensive labor and strong subjectivity. Therefore, the clinical requirements call for detecting and analyzing EEG signals quantitatively and automatically.

Due to the nonstationarity of the EEG signal and the influence of a large number of background waveforms and artifacts, the automatic detection and analysis of the pathological features is a challenging problem. At present, researchers have used algorithms in various fields to study the feature extraction and classification of EEG signals. The common feature extraction algorithms include: autoregressive model (AR model) (Argunsah and Cetin, 2010), power spectral density estimation (Seth et al., 2017), wavelet transform (Ocak, 2009), chaos method (Lv et al., 2004), common spatial pattern (CSP) (Blankertz et al., 2008; Ramoser et al., 2000), new descriptor (Kapoor et al., 2016), multi-dimensional statistical analysis (Weis et al., 2010) and so on. The frequently used classification methods include Fisher linear discriminant analysis (Harikumar, 2017),...
Bayesian method (Bashashati et al., 2016), BP neural network (Gao et al., 2012), support vector machine (Hortal et al., 2015) and so on (Roeva and Atanassova, 2016; Yang et al., 2012; Kaper et al., 2004; Wang et al., 2009; Deriche and Alani, 2001). But since EEG has strong nonstationarity and randomness, it is difficult to determine the representation and appropriate description.

Researchers are still seeking new signal analysis methods, and have completed the correct classification of EEG signals (Subasi adn Mishuhina (Suasi and Gursoy, 2010; Mishuhina and Jiang, 2018; Zhong et al., 2008). Sparse representation is a fast developing field by constructing sparse linear models, and good results in signal and image processing have been achieved (Candes et al., 2005; Candes et al., 2006; Donoho, 2006). After the Sparse Representation Classifier (SRC) is proposed (Wright et al, 2009), the classification of EEG signals based on sparse representation is also developing gradually (Shin et al., 2015; Shin et al., 2012; Zhou et al., 2012). However, they have to face the contradiction between dictionary size and algorithm recognition accuracy. For now, the deep learning method has been greatly developed and applied (Krizhevsky et al., 2012; Krahenbuhl et al., 2015; Cecotti and Grasser, 2011), and has achieved some breakthroughs in the field of computer vision and object recognition (An et al., 2014; Lu et al., 2017).

In order to solve the problems encountered in the sparse representation algorithm, sparse coefficients are classified by deep learning framework in the process of sparse representation classification in this paper. The classification algorithm combined with deep learning and sparse representation is proposed by the authors, and its application in EEG signals is discussed.

The main contents of this paper, are SRC and deep learning. According to this structure, this paper is organized as follows: The next section introduces the sparse representation-based classification method. Section 3 gives the EEG classification algorithm combined with SRC and deep learning. The section 4 shows the experiment Results and discussion. Finally, section 6 gives the conclusions of the proposed method.

**Sparse representation-based classification**

For the EEG signals, a feature vector can be obtained by Common Spatial Pattern (CSP) as $v_i = [v_{i1}, v_{i2}, ..., v_{im}]^T \in R^m$, where $m$ is the sample dimension. If all the characteristic vector signals from different types of brain waves are put in $A$, the matrix $A$ can be written as the following form:

$$A = [v_1, v_2, ..., v_n] \in R^{mn}. \quad (1)$$

The document (Harikumar et al., 2016) declares that if the training data from the $i^{th}$ category are enough, the test sample $y$ from the same category can be shown as a linear combination of the training set associated with subject $i$:

$$y = \alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_n v_n \quad (2)$$

where $\alpha$ is the coefficient vector, and its elements are not all zero. By concatenating $A_i$, the dictionary matrix $A$ for all $k$ classes can be acquired as $i = 1, 2, ..., k$. The dictionary can be given as follows:

$$A = [A_i, A_2, \cdots, A_k] \in R^{mnk} \quad (3)$$

If the EEG feature signal $y$ is the tested signal, $y$ can be written as a linear combination of all $nk$ training data.

$$y = Ax = x_1 v_{11} + x_2 v_{12} + \cdots + x_n v_{1n} + x_2 v_{21} + \cdots + x_n v_{2n} \quad (4)$$

where $x = [v_{11}, v_{12}, ..., v_{kn}]^T \in R^{nk}$ are the coefficients vectors. In the ideal case, $x = [0, 0, ..., x_{l1}, x_{l2}, ..., x_{lm}, 0, 0, ..., 0]$ is a coefficient vector, which is mostly zero value except for those elements corresponding to the class of $i^{th}$, thus the corresponding class of EEG feature signals can be classified. The two types of sparse presentation classification operations are shown in Fig. 1.

![Figure 1. Sparse representation classification scheme](image-url)

---

**Figure 1.** Sparse representation classification scheme
Based on the vector \( x \) and the test signal \( y \), L1-norm minimization equation can be listed as follows:

\[
\hat{x}_i = \arg \min \|x\|_{1} \quad \text{subject to} \quad Ax = y
\]  \hspace{1cm} (5)

In the ideal case, when we got the estimated \( \hat{x}_i \), it should has nonzero elements corresponding to \( y \). Through analyzing the indices of the non-zero elements in \( \hat{x}_1 \), the class of \( y \) can be determined. However, because of the modeling limitations and noise, \( \hat{x}_1 \) is not exactly zero but is close to zero, and these small nonzero elements belong to different types which are shown in the Fig.2. To resolve this problem, the following equation will be calculated generally as:

\[
r_i(y) := \|y - A\delta_i(x)\|_2.
\]  \hspace{1cm} (6)

For each class \( i \), \( \delta_i(x) \) is obtained by nulling all the elements corresponding to the other class, then the class \( i \) can be determined by analyzing the residuals, that is:

\[
\text{class}(y) = \arg \min r_i(y)
\]  \hspace{1cm} (7)

Figure 2. Sparse coefficients

The accuracy of this classification method depends largely on the size of the dictionary. Facing this question, at the same time, we hope to use the advantages of sparse representation classification algorithm, and the deep learning algorithms are used for coefficients classification in this paper so that the categories of electroencephalogram signals can be identified. This method will be described in detail in the next section.

**Deep learning**

*Algorithm flow chart*

In general, the image classification method based on deep learning is used to train the pixel data of the image set as the input, and then learn a number of image expression vectors. The time complexity and computational complexity of this method are very high so that it takes a lot of manpower and material resources. In addition, the interpretation of the characteristics of learning is poor, that is, the expression of image content is still at the bottom level of visual features. For the algorithms proposed by this paper, the learning and training process mainly focuses on sparse feature vectors. As the algorithm structure is shown in Fig.3, the whole algorithm framework mainly consists of four parts: data preprocessing, feature extraction, sparse representation and deep learning. Since he data preprocessing and feature extraction part adopt the methods proposed by the literature (Shin et al., 2012), there is no more detailed description in this section. Convolution neural networks have good recognition effects on many recognition problems, such as handwritten font recognition, face recognition, traffic sign classification, pedestrian detection, image annotation, and behavior detection. Because of the excellent performance of convolution neural network (CNN) model in the field of image, in this paper, the sparse coefficient recognition is based on image method. Since the EEG feature vector signal is processed by SRC algorithm and the sparse coefficient is one-dimensional vector, the sparse coefficients must be transformed into a two-dimensional vector. In this paper, after data filling and expansion, the sparse coefficient vector can be represented as image signal shown in Figure 4, which can be seen as the input signal of the CNN.

Figure 3. The algorithm flow chart proposed by this paper
In the CNN model, nonlinear functions are introduced as the activation functions so that it can approximate any function, rather than the linear combination of input, and realize the meaning of deep network. In this paper, ReLU (Rectified Linear Units) function is used as the activation function. The ReLU function is a nonsaturated nonlinear function, and the formula is as follows:

\[
f(x) = \begin{cases} 
0, & x < 0 \\
 x, & x > 0 
\end{cases}
\]

### Results

#### Dataset

Because the Berlin dataset produced in the BCI competition is widely used in the BCI field for the analysis of EEG signal processing, it was also adopted by this paper. Although there are five different healthy subjects (aa, al, av, aw, and ay) records, only two corresponding to the right hand (R) and right foot (F) were used. In this paper, they are named Class-1 and Class-2. Each EEG signal has 118 channels, one of which is described in Fig. 6. After the EEG signals were sampled down to 100 Hz, a timed trial procedure was shown in Fig. 7.

### Data processing

For further analysis, EEG waveforms need to be segmented. According to Figure 3, after the cue appeared, the 1-2s time samples were used in all of the experimental trials so that 140 experimental samples can be achieved for each type of EEG signal. To reduce the interference from other sources such as electrooculograms (EOGs) and electromyograms (EMGs), 8 to 15Hz bandpass filters was applied in this paper. Because CSP is a widely used and powerful signal processing technique that is suitable for two
classes (conditions) of multi-channel EEG-based BCIs, this paper adopts the CSP method proposed by the document. When the number of CSP filters is set as 32, after filtering operation, training and testing EEG samples can be converted to 32 CSP eigenvalues, which can be used for data classification.

![Figure 7. The timed trial procedure](image)

**Classification**

For the EEG signals of Class-1 and Class-2, each class can get 140 groups of 32 eigenvalues after the above data processing. Based on the set number of training samples, the redundant dictionaries of sparse classification algorithm were constructed by using the CSP eigenvalues obtained by Class-1 and Class-2. Then, the test samples were selected randomly, and the classification of EEG signals was completed based on the classification algorithm described in section 2. The classification results are drawn in Fig.8. For the sparse coefficients, if the algorithm proposed by this paper is used, the result of classification is shown in Fig. 8. When the number of training samples is 112, the change process of accuracy and loss in depth learning process is shown in Fig.9.

From Fig.8, it can be seen that the classification result of the proposed algorithm is better than that of the sparse representation classification algorithm. Although the classification accuracy of the proposed algorithm decreases as the number of training samples decreases, as long as the training samples are good, the accuracy of recognition is less affected. In Fig. 10, the recognition rate of two dimensional signals of Class-1 and Class-2 EEG sparse features is shown. In this paper, the total number of EEG training trials is 280. For the training process of deep learning, the number of samples is still insufficient. If the number of training samples is sufficient, the accuracy of classification will be further improved.

![Figure 8. Average accuracy of SRC with different numbers of training signals](image)

![Figure 9. Precision and loss change of training process in deep learning](image)

**Figure 10. Recognition rate of class-1 and class-2 EEG signals**

**Conclusions**

Because of the complexity of EEG signals, the classification of EEG is a difficult problem. In this paper, an algorithm of EEG classification based on sparse representation and convolution neural network is proposed, and the experimental results verify the feasibility of the algorithm. In the future, we will optimize the redundant dictionary in sparse representation based on the improvement of recognition accuracy of the deep learning algorithm.
Acknowledgments
This work was supported by the teacher's scientific research funds of Zhejiang University City College (J-15021), the scientific research project of Zhejiang Education Department (Y201635575), and the teachers scientific research funds of Zhejiang University City College (JYB18003).

References


Wright J, Yang AY, Ganesh A, Sastry SS, Ma Y. Robust face recognition via sparse representation. IEEE Transactions

794

