

INTERNATIONAL JOURNAL OF COMPUTERS COMMUNICATIONS & CONTROL
ISSN 1841-9836, e-ISSN 1841-9844, 14(4), 475-488, August 2019.

Facial Expression Decoding based on fMRI Brain Signal

B. Cao, Y. Liang, S. Yoshida, R. Guan

Benchun Cao

1. Zhuhai Laboratory of Key Laboratory of Symbol Computation and Knowledge Engineering of Ministry of Education, Zhuhai College of Jilin University
Zhuhai, 519041, China
2. Key Laboratory of Symbol Computation and Knowledge Engineering of Ministry of Education, College of Computer Science and Technology,
Jilin University, Changchun, 130012, China
caobc15@mails.jlu.edu.cn

Yanchun Liang

1. Zhuhai Laboratory of Key Laboratory of Symbol Computation and Knowledge Engineering of Ministry of Education, Zhuhai College of Jilin University
Zhuhai, 519041, China
2. Key Laboratory of Symbol Computation and Knowledge Engineering of Ministry of Education, College of Computer Science and Technology,
Jilin University, Changchun, 130012, China
ycliang@jlu.edu.cn

Shinichi Yoshida

School of Information, Kochi University of Technology,
Kochi 782-8502, Japan
yoshida.shinichi@kochi-tech.ac.jp

Renchu Guan*

1. Zhuhai Laboratory of Key Laboratory of Symbol Computation and Knowledge Engineering of Ministry of Education, Zhuhai College of Jilin University
Zhuhai, 519041, China
 2. Key Laboratory of Symbol Computation and Knowledge Engineering of Ministry of Education, College of Computer Science and Technology,
Jilin University, Changchun, 130012, China
- *Corresponding author: guanrenchu@jlu.edu.cn

Abstract: The analysis of facial expressions is a hot topic in brain-computer interface research. To determine the facial expressions of the subjects under the corresponding stimulation, we analyze the fMRI images acquired by the Magnetic Resonance. There are six kinds of facial expressions: "anger", "disgust", "sadness", "happiness", "joy" and "surprise". We demonstrate that brain decoding is achievable through the parsing of two facial expressions ("anger" and "joy"). Support vector machine and extreme learning machine are selected to classify these expressions based on time series features. Experimental results show that the classification performance of the extreme learning machine algorithm is better than support vector machine. Among the eight participants in the trials, the classification accuracy of three subjects reached 70-80%, and the remaining five subjects also achieved accuracy of 50-60%. Therefore, we can conclude that the brain decoding can be used to help analyzing human facial expressions.

Keywords: Brain-computer interface, machine learning, extreme learning machine, fMRI, image processing.

1 Introduction

Brain is the control center of the entire body. It is of great importance and is an indispensable part of us. The brain-computer interface as one of the state-of-the-art issues in the field of science has attracted more and more attentions. If we can judge people's specific behaviors by analyzing the brain, we can understand the brain better [11]. Functional magnetic resonance imaging (fMRI) is a popular technique for studying brain in recent years [11, 18]. It can scan the brain to obtain images and shows the area of the voxel that is activated in the brain when the subject is stimulated by outside stimulus. Through the voxels we can determine which behaviors that the subject was doing are activated in the brain. Medical image processing is a hot issue in the image recognition field. More and more researchers have already invested into the study of medical image processing [12]. Using computers to help people identify medical images can reduce people's errors caused by overwork and improve the accuracy of judgment [12].

Recently, the problem of brain decoding has attracted broad attention. The facial expressions decoding is the main research direction for brain decoding. In daily human interactions, facial expression is one of the most intuitive feelings for humans to communicate. It is crucial for us to understand others' emotions. If the brain activity corresponding to the facial expression can be retrieved, this information can be used to analyze facial expressions. This can be achieved by Brain Computer Interface (BCI), which is a technique to obtain the internal signal of the brain through an external device connected to the brain [16]. At present, electroencephalogram (EEG) and functional magnetic resonance imaging (fMRI) have successfully completed the acquisition of non-invasive signals of brain-computer interface systems [14]. The imaging principle of fMRI is through the magnetic field inside of the machine to affects the activity of the human neuron. The activity of the neuron changes the blood flow and blood oxygen, and fMRI is imaged by measuring these changes. Because of its non-invasive and reproducible features, fMRI is a hot topic for medical researchers and experts. Studies [7] have pointed that fMRI not only can interpret simple movement instructions, but also can interpret complex advanced thinking activities that are related to certain regions of the brain. Therefore, fMRI technology is more suitable for decoding human facial expressions [6]. In particular, the research on the location of brain regions has been widely used.

The brain information decoding technology refers to using fMRI and machine learning algorithms to estimate the subject's facial expressions under the stimulation of expression images according to the acquired brain signals. However, low recognition accuracy is a major problem in brain information decoding, which has limited this technology's fast spreading. The main task of this paper is to extract the brain image features that have been stimulated to show different states through the fMRI brain images analysis with higher classification accuracy. Comparing with the time series, we find out different stimulus labels corresponding to different brain image features.

The remainder of this paper is organized as follows. Section 2 introduces the fundamental concepts of related algorithms. Section 3 elaborates the data parsing process of fMRI data. Section 4 presents the specific implementation process of the experiment and the analysis of related experimental results. Finally, the conclusion is drawn in Section 5.

2 Related theory

To better understand our model, this section gives a detailed introduction and explanation about all the related algorithms including cross validation, support vector machine algorithm, extreme learning machine algorithm and brain decoding. For each algorithm, this section specifies its principles and implementation steps.

2.1 Cross validation algorithm

Cross-validation is an algorithm used to evaluate the performance of classifiers [1]. Its main idea is to divide data sets into training sets and verification sets. Training sets are mainly used to train classifiers and verification sets are used to verify the classification performance. According to the data partitioning, cross-validation methods are divided into three categories hold-out method [2], k-fold cross validation method [15] and leave-one-out cross validation method [10]. The k-fold cross-validation method divides the original data set into k parts, which is divided into k parts. Each of the k parts is used as verification set; the other k-1 parts are used as the training set. We get k models and take the arithmetic average as the accuracy of the model. The advantage of the k-fold cross-validation method is that the segmentation is reasonable so that the data set can be fully utilized, the probability of occurrence of the under-fitting problem is reduced and the training data is greatly increased, and the trained model can be more optimized. Because k-fold cross-validation can effectively avoid over learning and less learning and the result is convincing, we used k-fold cross-validation method.

2.2 Support Vector Machine

Support Vector Machine (SVM) is a very popular machine learning algorithm [17]. The original support vector machine was specifically designed to solve the linear divisibility problem. Introducing kernel functions, SVM can map low-dimensional data to high dimensions and solve the linear indivisibility problems. The purpose of SVM is to generate a hyperplane that divides the linearly separable data into two classes and make the point closest to this line as far away as possible from the hyperplane. The interval which is called the geometric interval should be as large as possible so that the data can be completely separated. We assume that the formula for the linear segmentation plane is $\mathbf{w}^T + b = 0$.

The geometric interval can be obtained by solving the following optimization problem:

$$f(x) = \text{Max}_{w,b} \frac{1}{2} \|w\|^2. \quad (1)$$

Subject to : $y^i(wx^i + b) \geq 1$

where the weight vector w and the offset value b represent the parameters of the classification plane. For nonlinear classification problems, SVM cannot classify them in the low-dimensional space and the kernel function solves this problem well. The kernel function can map the data from the low dimension to high dimension. With kernel functions, SVM can be used to classify nonlinear problems. There are different types of kernel functions such as sigmoid kernel, Gaussian kernel and tanh kernel.

2.3 Extreme Learning Machine

Extreme Learning Machine (ELM) is a neural network with a single hidden layer which randomly initializes weights and do not need to adjust parameters during training [9]. The ELM consists of three layers: the input layer, hidden layer, and output layer. The learning speed of ELM is fast because ELM has only one single hidden layer. ELM initializes weights randomly and there is no need to adjust offsets and weights in the training process.

Figure 1 shows a neural network with single hidden layer. If there are N groups input data for the neural network, it can be expressed as

$$(x_j, t_j) \quad (2)$$

where

$$x_j = [\mathbf{x}_{j1}, \mathbf{x}_{j2}, \dots, \mathbf{x}_{jn}]^T \in R^n$$

$$t_j = [\mathbf{t}_{j1}, \mathbf{t}_{j2}, \dots, \mathbf{t}_{jm}]^T \in R^m$$

we assume that the number of hidden nodes in the neural network is L .

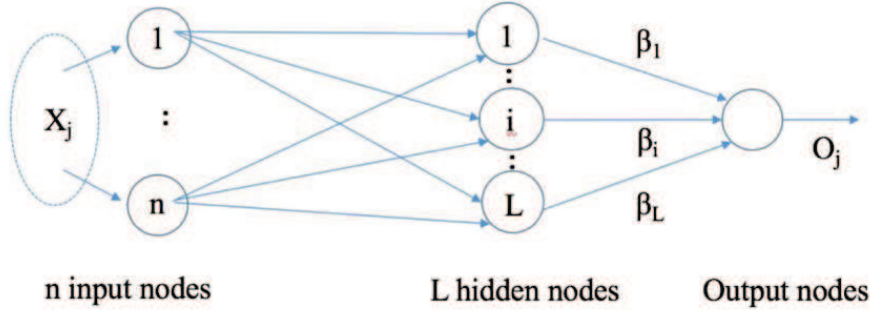


Figure 1: Single hidden layer neural network

Then the neural network can be expressed as

$$\sum_{i=1}^L \beta_i g(\omega_i x_j + b_i) = o_j, j = 1, 2, \dots, N. \quad (3)$$

where $g(x)$ denotes the activation function of the hidden layer node,

$\omega_i = [\omega_{i1}, \omega_{i2}, \dots, \omega_{in}]^T$ denotes the input weight of the i -th node on the hidden layer, and b_i denotes the offset of the i -th node on the hidden layer,

$\beta_i = [\beta_{i1}, \beta_{i2}, \dots, \beta_{im}]^T$ represents the output weight of the i -th node of the hidden layer, $\omega_i \cdot \omega_j$ indicates the inner product of the weights ω_i and ω_j .

The goal of training single hidden layer neural network is to minimize the error between the output and the expected value of the network output layer $\min(\sum_{j=1}^N \|o_j - t_j\|)$.

Hereinto,

$$\sum_{i=1}^L \beta_i g(\omega_i x_j + b_i) = t_j, j = 1, 2, \dots, N.$$

It can be expressed in the form of matrix

$$\mathbf{H} \cdot \beta = \mathbf{T} \quad (4)$$

where \mathbf{H} represents the output matrix of the hidden layer node,

β represents the weight of the output of the hidden layer,

\mathbf{T} represents the target output of the output layer:

$$\mathbf{H}(\omega_1, \dots, \omega_L, b_1, \dots, b_L, x_1, \dots, x_N) = \begin{bmatrix} g(\omega_1 \cdot x_1 + b_1) & \cdots & g(\omega_L \cdot x_1 + b_L) \\ \vdots & \ddots & \vdots \\ g(\omega_1 \cdot x_N + b_1) & \cdots & g(\omega_L \cdot x_N + b_L) \end{bmatrix} \quad (5)$$

$$\beta = [\beta_1 \cdots \beta_L]^T \quad (6)$$

$$\mathbf{T} = [\mathbf{t}_1 \cdots \mathbf{t}_N]^T. \quad (7)$$

In order to optimize the neural network, we expect to obtain \widehat{w}_l , \widehat{b}_l and $\widehat{\beta}_l$, making

$$\|\mathbf{H}(\widehat{w}_l, \widehat{b}_l) \cdot \beta - \mathbf{T}\| = \min_{w,b,\beta} \|\mathbf{H}(\widehat{w}_l, \widehat{b}_l) \cdot \beta - \mathbf{T}\| \quad (8)$$

where $i = 1, 2, \dots, L$.

This is equivalent to minimizing the loss function:

$$\mathbf{E} = \sum_{j=1}^N \left\| \sum_{i=1}^L \beta_i g(\omega_i x_j + b_i) - t_j \right\|^2 \quad (9)$$

For solving this kind of problem, the gradient descent method is often used to find the optimal solution. However, using the gradient descent algorithm, we need to adjust the parameters of all hidden layers based on expert experience or optimization algorithms. To avoid this problem, we introduce extreme learning machine (ELM).

For ELM, it is no need to adjust the parameters and the offsets as long as the weights are initialized. The training process can be summarized as: $\mathbf{H} \cdot \beta = \mathbf{T}$, and we can get output weight $\widehat{\beta}$

$$\widehat{\beta} = \mathbf{H}^+ \cdot \mathbf{T} \quad (10)$$

where \mathbf{H}^+ represents the generalized inverse matrix of matrix \mathbf{H} . Huang et al. have proved that the norm of the solution $\widehat{\beta}$ is not only smallest but unique [8].

2.4 Brain decoding

The flowchart of brain information decoding is shown in Figure 2. First, by giving stimulus images and measuring brain activity data, the facial expressions of the subjects are recorded according to the time series. Then, we extract the feature vectors under this stimulus according to activated voxels. Finally, the obtained feature vectors are corresponding to the facial expressions through the time series to form a feature vector and label input classifier for counting the recognition accuracy.

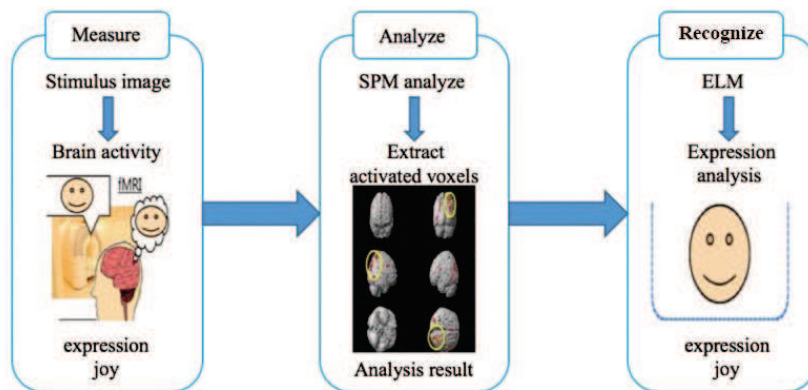


Figure 2: Brain information decoding processes

3 fMRI data analysis

For the preprocessing procedure, we use Statistical Parametric Mapping (SPM) [13] to analyze the obtained fMRI images, personal parsing process and group parsing process. The process of fMRI Data Analysis is shown in Figure 3.

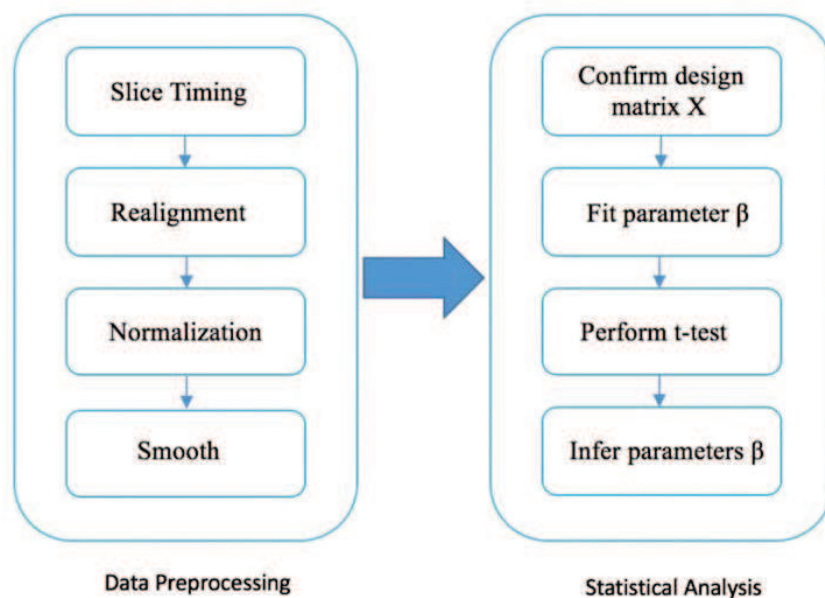


Figure 3: fMRI data analysis process

3.1 Data preprocessing

Before analyzing the data of the brain image, we need to preprocess the brain data and use the SPM software to complete the process. It mainly consists of four steps: Slice Timing, Realignment, Normalize, and Smooth.

Slice Timing. During the collection of brain images, because we cannot scan the entire brain image at a time, we scan brain in slice. In the process of scanning, there will be slight differences in the acquisition time of each slice; therefore we need to correct time difference in the acquisition slice to ensure that the acquisition time is the same for each slice. Slice timing is used to correct the difference in time between the slice and the slice in the brain image.

Realignment. Although we fix the head of the subjects, some subjects still had slight head shake during the experiment. The process of realignment is to unify all the head portraits of the subjects in the entire test process, and the purpose is to correct head shake. Assuming that subject's brain shaking category is within an acceptable category, then we correct it by some methods to make it close to the exact value, but if it is not in this category, we need to delete the portrait. In general, the range of head movements we examined was that the translation does not exceed 2.0 mm and the rotation does not exceed 2.0 degrees.

Normalization. Because different subjects have different brain sizes and shapes, we standardize the brain images of the subjects before carrying out the experiment in order to deal with them in a unified way. We place the brain images of the subjects with differences in a common space and use the same norms to describe the specific sites. For example, when there are multiple experimental participants with different head shapes, if we want to make an experiment for them

to select an area of the brain, we should develop a standard spatial brain template to achieve accurate positioning.

Smooth. Since fMRI contains a large amount of noise, to minimize the impact of noise and obtain images with less difference from the original data, we smooth the acquired image. In SPM software, smoothing is the process of deconvoluting the acquired brain image using a Gaussian function.

3.2 Statistical analysis process

Directly using the preprocessed data to data analysis will bring a lot of unnecessary problems. To enable the entire experiment to be proceed smoothly, we divide the preprocessed data into two areas, the interest region and other regions. In data analysis process, we extract feature vectors directly from the regions of interest (ROI) [4]. In the statistical analysis process, a very important statistical analysis model is used, namely the generalized linear model (GLM) [19].

GLM is an extension of the linear model. The independent variables should be continuous values rather than discrete integers or data, and the relationship between the expected values of the independent and dependent variables are linear. The independent variables of the generalized linear model are not only limited to continuous numerical data, but discrete integers are also satisfactory. The other difference is that we need to determine the connection function when modeling GLM.

In simple terms, GLM is based on the assumption that the test result (represented by Y) at each pixel is a linear combination of certain parameters X . These parameters are not only related to the specific brain regions, but are related to the experimental task and the experimental time. The matrix composed of these parameters is called the design matrix. It can be expressed as $Y = \beta X + \varepsilon$ where ε represents the error, we have changed the parameters of the solution after the transformation through the GLM. We originally asked for a statistical analysis of the variable Y , and now we are solving the parameter β .

We can get brain activation map by fitting statistics to β . The statistical inference procedure for parameter β is as follows:

Step 1: Confirm the accuracy of design matrix X .

Step 2: The least squares method is used to fit the parameter β , and make the error sum $\sum_{i=1}^N \varepsilon^2$ reach the minimum.

Step 3: Perform t-test or F-test on parameter β .

Step 4: According to t test or F test statistical results to infer parameters β .

4 Experiment and discussion

In our experiment, the fMRI image of the subject under stimulation was first obtained by Magnetic Resonance. Then the fMRI image analysis software SPM was used to analyze the fMRI images of different subjects stimulated by different expression images (anger and joy) respectively. The activated brain voxel area was obtained under the corresponding stimulus. This experiment allows participants to maintain a consistent facial expression and the stimulus images they see. The voxel area that is activated by the time series corresponds to the stimuli portrait. The obtained active voxel area was used as the feature vector and the facial expressions of the subjects were used as labels, which are input into the classifier to evaluate the performance of the algorithm.

Paul Ekman claims that there are six basic emotions in human emotion [5]. They are "anger", "disgust", "sadness", "happiness", "joy", and "surprise". In our case, anger and joy will be used to predict.

4.1 Experimental data

There are 8 subjects participated in the experiment, where 5 were males and 3 were females. All the subjects were from 21 to 22 years old and in good health to avoid the effects of gender and age on the experimental results. The safety of the fMRI equipment was explained to the subjects before the experiment and a confidentiality agreement on the experimental results was signed. The stimuli used in this study are anger and joy images. To emphasize facial features, the clothing and background of the stimuli are filled with gray to reduce the impact on the experimental results.

In the experimental design, 216 seconds brain activity was imaged for each subject. Eight brain scans were performed for each subject and 72 brain activity slices were obtained each time during this process. To reduce bold signal, we show 3 black background scans between the stimuli to stabilize the brain. The stimulating images randomly presented 64 facial portraits, including 32 "joy" facial expressions and 32 "angry" facial expressions. The specific form is shown in Figure 4.



Figure 4: Brain scan time

This following step is mainly carried out through three steps. First, different facial expression images were randomly presented to stimulate the subjects, and the subjects were asked to make corresponding facial expressions with the images, mainly including two expression images of anger and joy. Then the brain data of subjects was scanned by MRI to images. The obtained brain images are analyzed using SPM to obtain useful information. Finally, the data obtained by the experiment was input to the classifier.

4.2 Data preprocessing and analysis

The original image taken from the fMRI device is a DICOM format image. To use SPM for data analysis, we need to convert the image to the desired analysis format. This experiment uses the MRIConvert software to convert the image format. In the process of analyzing the brain images of the subjects, we used SPM software to identify the brain activation voxels. The personal analysis process is as follows: 1. Input preprocessed brain activity data into generalized linear model (GLM). 2. Use the design matrix created to calculate the regression coefficient β of the GLM model 3. Perform T-test on the subject's brain image to determine which brain voxels are activated 4. Analyze the brain regions of activated voxels belongs from the display image (reference Brodmann region system [20]).

When the stimulus image is joy or anger, the analysis results for each subject's fMRI image are presented in Table 1 and Table 2, respectively. The pulse repetition time (TR) of the subjects A, B, C, D was 3090ms, and E, F, G, H was 3480ms. The pulse repetition time refers to the interval where the brain is scanned using an MRI.

The group analysis is to analyze the information collected by all subjects as a whole. In the group analysis, the contrast file obtained by personal analysis of the subjects was used to analyze the whole group of subjects. The results of eight subjects group analysis are shown in Figure 5. When the stimulus image is "joy", the activated voxels are located in the lower parietal regions

Table 1: Brain image analysis results on the stimulus of joy

Subject	P value	T value	Activated Brodmann region
A	0.01	2.89	17-19,37,39
B	0.001	4.05	17-19,7
C	0.001	5.11	9-19,23-40,44-47
D	0.001	4.15	9-12
E	0.001	4.82	9-12
F	0.0001	9.51	17-19,37
G	0.001	3.70	9,46
H	0.001	4.00	9-12,23-27

Table 2: Brain image analysis results on the stimulus of anger

Subject	P value	T value	Activated Brodmann region
A	0.01	3.02	9-12,17-19
B	0.001	4.11	17-19
C	0.001	4.04	37,20,21
D	0.001	4.34	7-16,20-22,39-47
E	0.001	5.92	9-12
F	0.0001	3.72	Lateral geniculate body, posterolateral nucleus
G	0.001	3.59	9,11-16.25-33,34-36,46,47
H	0.001	4.02	9-12,45-47

of the 40 and 41 regions of the Brodmann region, mainly related to perception, vision, reading and language-functional area.

When the stimulus image is "angry", the significance level is set to $p < 0.01$. At this time, the activated voxels are located in the inner and outer occipital regions of the 19 region Brodmann area. This area is mainly visual, color vision, and movement-related areas. It can be seen from the Figure 5 that amygdala is not completely activated. According to the relevant analysis, as the amygdala is the brain part on the facial expression recognition [3], the amygdala of the test was only activated a little. Furthermore, we can see that the visual and other areas are more fully activated.

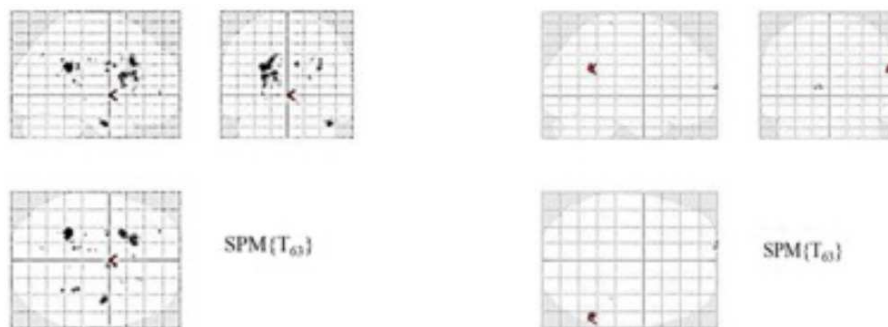


Figure 5: Brain activity state (joy, anger) of group analysis

4.3 Classification results comparison

According to the above discussion, we have already analyzed the fMRI image of the subject through SPM software and obtained the feature vector. According to the time series record in the experiment, we can associate the feature vector with the facial expression. Thus, a data containing feature vector and feature label is generated. To achieve high recognition accuracy, the selection of the classifiers is of great importance.

We classify the obtained data using SVM and ELM, respectively. In this comparison experiment, the number of ELM hidden layer nodes is 1000 nodes and the Sigmoid function is the activation function. And the C value of SVM is 0.5 and the Sigmoid function is selected as the kernel function, we use the grid search algorithm to find the optimal super parameters for the sigmoid function $K(x, y) = \tanh(\gamma \cdot x^t \cdot y + g)$ in order to get the optimized SVM model. In the parameter selection, we respectively let $\gamma = (1, 2, 3, 4)$ and $g = (0.2, 0.4, 0.6, 0.8)$. According to the experimental results, we get that the function is optimal when $\gamma = 2$ and $g = 0.6$. From the results it can be found that no matter what the data obtained in the stimulation of the joy images or anger images, the classification result of the ELM is better than SVM. Figure 6 shows the comparison of the classification accuracy under the stimulation of the joy images. Figure 7 shows the comparison of the classification accuracy under the stimulation of the anger image. Therefore, we use ELM to classify the obtained fMRI images.

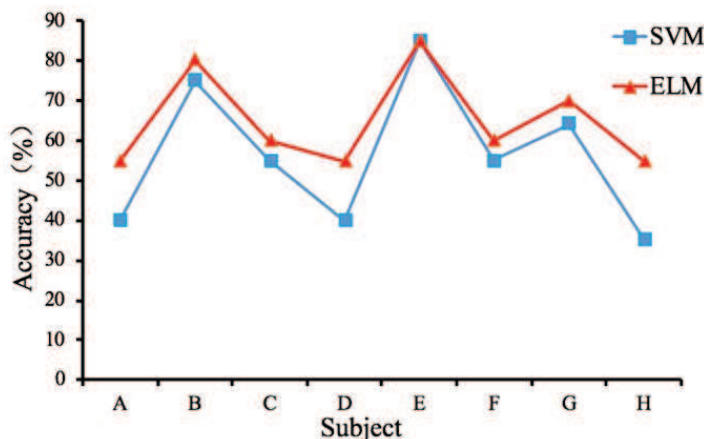


Figure 6: Performance comparison of SVM and ELM for joy images stimulation

Because the number of hidden layer nodes and the selection of activation function in the ELM directly affect the classification performance of the entire ELM, To pursue the highest accuracy, we need to determine the optimize number of hidden layer nodes and activation function.

Figure 8 shows the effect of the active functions and the number of hidden nodes on the classification accuracy of ELM for the joy images. From Figure 8 we can see that the overall classification performance of the Sigmoid function is higher than the tanh function. Although using 2500 hidden nodes, the former function is not good as tanh, it does not affect the overall trend. Therefore, we use the Sigmoid function as the activation function of the classifier ELM. From the curve of the Sigmoid function, we can see that the number of nodes in the hidden layer ranges from 500 to 4000 and every 500 nodes are located in one interval. We can see that the classification performance reaches the best when the number of nodes is 1000. Therefore, the number of hidden layer nodes of the ELM classifier is chosen as 1000 nodes. In summary,

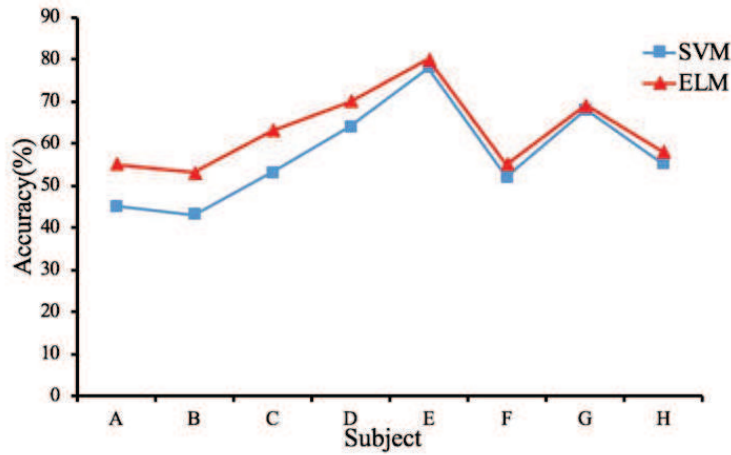


Figure 7: Performance comparison of SVM and ELM for anger images stimulation

the amount of hidden layer nodes is 1000 and the Sigmoid function is selected as the activation function.

4.4 Discussion

We use cross-validation to evaluate the classification performance of ELM. It not only enables ELM to make more effective predictions, but also reduces the chance of overfitting. The k fold cross-validation is used to evaluate the performance of the ELM. In our experiment, let k be 10, that is, we divide the data into 10 groups and use 9 groups each time as the training set to train the ELM, the other group as test set. The average accuracy of the ten groups of classification is the classification accuracy of the classifier. In this experiment, the classification accuracy of each subject is shown in Table 3.

Table 3: The accuracy of ELM on each subject data

Subject	Classification accuracy(joy)	Classification accuracy (anger)
A	55%	55%
B	80%	55%
C	60%	62%
D	53%	70%
E	88%	80%
F	60%	58%
G	68%	70%
H	55%	60%

Table 3 lists the classification accuracy of the brain images acquired by fMRI equipment from the 8 subjects under the effect of cross-validation and ELM classifiers. It can be seen that the accuracy of 3 subjects (subjects B, E, G) reached 70-80% and the remaining 5 subjects reached 50-60%. According to these results, it can be found that facial expressions can be analyzed by analyzing fMRI brain images. Among these subject's results, the accuracy of subject A is low, and the reason for the low accuracy rate may be due to noise. It may also be caused by the

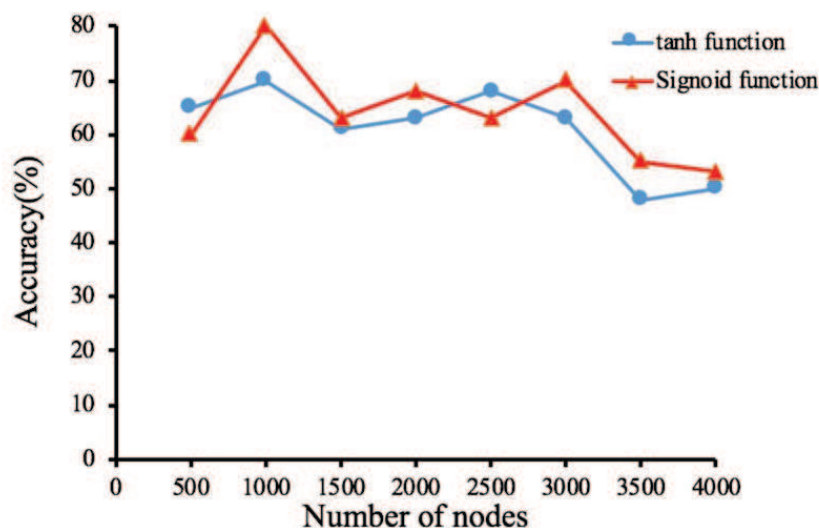


Figure 8: Effect of the activation functions and the amount of hidden layer nodes

insensitivity of the subject to the stimulating image.

5 Conclusion

In this paper, the machine learning algorithm ELM is applied to the classification of fMRI data. From the results of personal analysis, we can see that the accuracy of the recognition of 3 subjects (B, E, and G) in the 8 subjects reached 70-80%, and the accuracy of the remaining 5 subjects also reached 50-60%. It can be seen that using the brain image obtained by fMRI we can analyze people's facial expressions. From the results we can see that the activated voxels are usually located in the brain regions with visual, ideological, and cognitive functions. This article also provides the proof that we can analyze the facial expressions of the subjects by analyzing the fMRI images. It should be pointed out that in the near future, the stimulatory images may not be limited to still images. It is possible to study by using video to stimulate subjects and then observe the experimental results.

6 Acknowledgement

This work was supported in part by the National Natural Science Foundation of China (61572228, 61602207), the Jilin Provincial Key Laboratory of Big Data Intelligent Computing (20180622002JC), the Key Technological Research Projects in Jilin Province under Grant 20190302107GX, the Special Research and Development of Industrial Technology of Jilin Province under Grant 2019C053-7, Guangdong Premier Key-Discipline Enhancement Scheme (2016GDYSZDXK036), and Guangdong Key-Project for Applied Fundamental Research (2018KZDXM076).

Bibliography

- [1] Arlot, S.; Celisse, A.(2010). A survey of cross-validation procedures for model selection, *Statistics Survey*, 4, 40-79, 2010.

-
- [2] Blum, A.; Kalai, A.(1999). Beating the hold-out:bounds for K-fold and progressive cross-validation, *Proceedings of the twelfth annual conference on Computational learning theory. ACM*, 203-208, 1999.
- [3] Davis, M; Whalen, P. J.(2001). The amygdala: vigilance and emotion, *Mol Psychiatry*, 6(1), 13-34, 2001.
- [4] Desikan, R.S.; Segonne, F; Fischl, B.(2006). An automated labeling system for subdividing the human cerebral cortex on MRI scans into gyral based regions of interest, *Neuroimage*, 31(3), 968-980, 2006.
- [5] Ekman, P.(1992). An argument for basic emotions, *Cognition & Emotion*, 6(3-4), 169-200, 1992.
- [6] Forman, S D; Cohen, J.D.; Fitzgerald, M. et al.(1992). Improved Assessment of Significant Activation in Functional Magnetic Resonance Imaging (fMRI): Use of a Cluster-Size Threshold, *Magnetic Resonance in Medicine*, 3(5), 636-647, 2010.
- [7] Gilead, M.; Liberman, N.; Maril, A.(2013). The language of future-thought: An fMRI study of embodiment and tense processing, *Neuroimage*, 65(2), 267-279, 2013.
- [8] Huang, G.B.; Zhu, Q.Y.; Siew, C.K.(2004). Extreme learning machine: a new learning scheme of feedforward neural networks, *IEEE International Joint Conference on Neural Networks*, 2, 985-990, 2005.
- [9] Huang, G. B.; Zhu, Q. Y.; Siew, C.K.(2006). Extreme learning machine: Theory and applications, *Neurocomputing*, 70(1), 489-501, 2006.
- [10] Kearns, M; Ron, D.(1997). Algorithmic stability and sanity-check bounds for leave-one-out cross-validation, *Neural Computation*, 11(6), 1427-1453, 1997.
- [11] Logothetis, N.K.; Pauls, J.; Augath, M.(2001). Neurophysiological investigation of the basis of the fMRI signal, *Nature*, 412(6843), 150-157, 2001.
- [12] Lehmann, T.M; Gonner, C.; Spitzer, K.(1999). Survey: interpolation methods in medical image processing, *IEEE Transactions on Medical Imaging*, 18(11), 1049-1075, 1999.
- [13] Litvak, V.; Mattout, J.; Kiebel, S. et al. (2011). EEG and MEG Data Analysis in SPM8, *Computational Intelligence and Neuroscience*, 2011(3), 852961, 2011.
- [14] Michel, C.M; Murray, M.M.; Lantz, G.(2004). EEG source imaging, *Clinical Neurophysiology*, 115(10), 2195-2222, 2004.
- [15] Rodriguez, J.D.; Perez, A.; Lozano, J.A.(2010). Sensitivity Analysis of k-Fold Cross Validation in Prediction Error Estimation, *IEEE Computer Society*, 32(3), 569-575, 2010.
- [16] Schalk, G.; Mcfarland, D.J.; Hinterberger, T.(2004). BCI2000: a general-purpose brain-computer interface (BCI) system, *IEEE Transactions on Bio-medical Engineering*, 51(6), 1034-1043, 2004.
- [17] Suykens, J.A.K.(1999). Least squares support vector machine classifiers: a large scale algorithm[C]. *European Conference on Circuit Theory and Design*, 1999, 839-842, 1999.
- [18] Tugui, A. (2014). GLM Analysis for fMRI using Connex Array *International Journal of Computers Communications & Control*, 9(6), 768-775, 2014.

- [19] Ziegel, E.R.(2012). An Introduction to Generalized Linear Models[J]. *Technometrics*, 44(4), 406-407, 2012.
- [20] Zilles, K; Amunts, K.(2010). Centenary of Brodmann's map—conception and fate, *Nature Reviews Neuroscience*, 11(2), 139-145, 2010.