2018

Master's thesis

# A Study on an Application of Extreme Learning Machine to Brain Decoding of Human Emotion induced by Visual Stimuli


1215099 Xiang LI

Advisor Associate Professor Shinichi YOSHIDA



2019 年 2 月 1 日



Informatics Course

Graduate School of Engineering, Kochi University of Technology

## Abstract

# A Study on an Application of Extreme Learning Machine to Brain Decoding of Human Emotion induced by Visual Stimuli

Xiang LI

Emotions combine people's feelings, thoughts, and behaviors. With the psychological reaction of the outside world's stimuli, our cerebral cortex will have physiological reactions. Artificial intelligence has been widely used in the field of emotion recognition.

Brain decoding has been studied and focused in neuroscience field. Brain decoding of human emotion has also been studied since 2014. In the course of our experiments, we try to use classifier to classify human emotion induced by visual stimuli using brain activity measured by functional Magnetic Resonance Imaging(fMRI). For brain decoding traditional classifiers such as support vector machine(SVM)and feature extraction methods have been employed. The recent development of neural network classifiers is remarkable. Hybrid neural network classifier of extreme learning machine(ELM) and deep convolutional neural networks(CNN) machine has been proposed and achieve high accuracy.

After the training of brain decoder, input brain activity images to the brain decoder. Use the trained brain decoder to predict the emotion of subjects is positive or negative. The traditional method in brain decoding is Support Vector Machine. In this thesis, we propose to use ELM to take the place of SVM to construct the brain decoder.The

ELM algorithm provides a unified model for all classification problems, inheriting the structural advantages of single-hidden-layer feedforward networks, and the computing speed is faster than the SVM algorithm. We choose the ELM algorithm to construct the brain decoder instead of SVM because ELM has simple structure and good learning efficiency.

In the MVPA-ELM method, we use ELM combine with MVPA method. Multi-voxel pattern analysis(MVPA) usually considers multi-voxel activation in the brain as a pattern in high-dimensional space, and uses pattern classification to decode the information contained in the activation pattern. In our experiment, analysis of fMRI data in the brain using multi-voxel pattern analysis typically involves the following steps: Feature selection, Pattern assembly, ELM Classifier training, cross-validation.

In the CNN-ELM method, we use the VGG16 CNN Model to extract features of fMRI dataset. Use fMRI dataset as the input to the CNN after preprocessing the experimental data to extract features. Then train the ELM using these features and finally use cross-validation to evaluate the accuracy of the CNN-ELM brain decoder model.

We could see the average accuracy of MVPA-ELM Model with tanh activation function has the best performance compare to other methods in the results. The average accuracy is around 0.688%.

***key words*** Brain Decoding, Multi-voxel pattern analysis, Extreme Learning Machine, Convolutional Neural Networks

# Contents

# List of Figures

# List of Tables

# Chapter 1

# Introduction

## 1.1 Research Background

Emotions combine people's feelings, thoughts, and behaviors. With the psychological reaction of the outside world's stimuli, our cerebral cortex will have physiological reactions. Artificial intelligence has been widely used in the field of emotion recognition. But it is obvious that it is still at the level of identifying and categorizing image and voice information. On the road of human self-exploration, it takes time to understand the emotions of human brain.

Lots of researchers has done some work on understanding of human emotion from different side. The emotional manifestation of people is multifaceted. Expressions, language, and movements can all serve as vectors for human expression of emotions. Many scholars analyze human emotions in different ways, including micro-expression analysis, language performance analysis and so on.

Brain decoding has been studied and focused in neuroscience field. Brain decoding of human emotion has also been studied since 2014[1][2]. In the course of our experiments, we try to use classifier to classify human emotion induced by visual stimuli using brain activity measured by functional Magnetic Resonance Imaging(fMRI)[3] [4] . For brain decoding traditional classifiers such as support vector machine(SVM)and feature extraction methods have been employed. The recent development of neural network classifiers is remarkable. Hybrid neural network classifier of extreme learning machine(ELM) and deep convolutional neural networks(CNN) machine has been proposed and achieve high accuracy[5].

## 1.2 Research Purpose and significance

Emotion plays an important role in the communication between people. In medicine, if we can know the emotional state of a patient, especially a patient with an expression disorder, we can use different treatments based on the patient's feeling. In the product development process, if we can identify the user's emotion transformation in the process of using the product, we can improve the product function and make the product more suitable for users. In various human-computer interaction(HCI) systems, if the system can recognize the emotion of the person, the interaction between the person and the machine will become more friendly. In the design of robots, if we did an emotion model for robots that may help them to have empathy with people and understand people ' s feelings.

Therefore, the analysis and recognition of emotions is an important interdisciplinary research topic in the fields of neuroscience, psychology, cognitive science, computer science and artificial intelligence. In this chapter, we briefly introduce the research background of Brain Decoding of human emotion induced by visual stimuli and the purpose and significance of the research. And after that is a brief description of the structure of the whole thesis.The structure of thesis is as follows: The second chapter is a literature review, which mainly introduces the research and contributions of other scholars in emotional recognition in recent years. And the research results of extreme learning machines that have achieved excellent results in classification algorithms in recent years. The third chapter is the research method, which introduces several research methods, Multi-variate Pattern Analysis, Extreme Learning Machine and Convolutional

Neural Networks. The fourth chapter of the experimental part, firstly introduced the experimental design, and then described the two different methods which we used in this paper:

a) MVPA-ELM combining Multi-Variate Pattern Analysis with ELM

b) CNN-ELM combining the Convolutional Neural Network method with ELM.

The results of two different methods were analyzed and discussed in this chapter. The fifth chapter is the conclusion and summary of the thesis.

# Chapter 2

# Related Work

## 2.1 Researches on Brain Decoding of Human Emotion

Brain decoding is developed since 2005[6]. Brain decoding system including two phase. Learning phase and predicting phase. Figure 2.1 shown the learning phase of brain decoding.

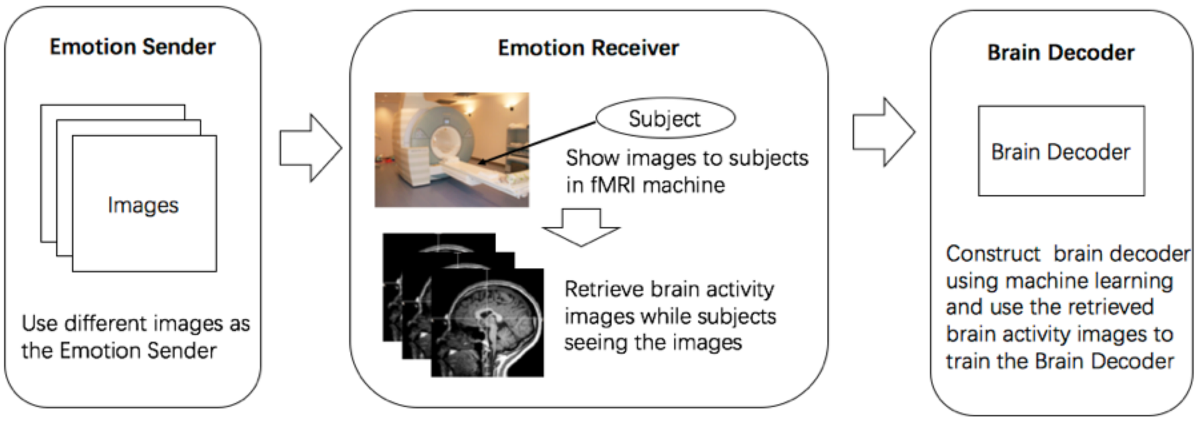


Fig. 2.1 Learning Phase of Brain Decoding

We use different images or videos as the emotion sender. During the experiment, show the images or videos to the subjects while they are in the functional Magnetic Resonance Imaging machine. Use functional Magnetic Resonance Imaging machine to retrieved their brain activity images while they seeing the images. Then construct a brain decoder using machine learning method and use the retrieved brain activity images to train the Brain Decoder.

After the learning phase is the predicting phase. The process is shown in Figure 2.2.

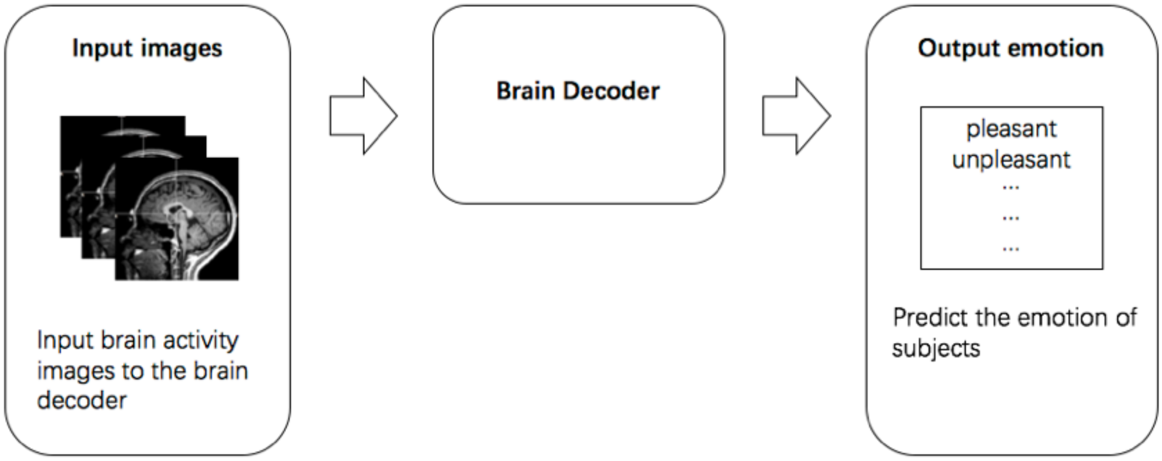


Fig. 2.2 Predicting Phase of Brain Decoding

After the training of brain decoder, input brain activity images to the brain decoder. Use the trained brain decoder to predict the emotion of subjects is positive or negative. The traditional method in brain decoding is Support Vector Machine (SVM). In this experiment, we proposed use Extreme Learning Machine to take the place of SVM to construct the brain decoder.

## 2.2 Researches on Extreme Learning Machine

Computer intelligence technology has been widely used in various fields. In the past few decades, artificial neural networks have fallen into a bottleneck in the applications that require highly real-time performance because their learning speeds are far from needs.

Support Vector Machine(SVM), Decision Tree(DT), K-Nearest Neighbors(KNN) and Back Propagation Neural Network(BPNN) are well-known algorithms in machine learning.

The Support Vector Machine algorithm was first proposed by Cortes C, Vapnik V[7]. SVM is a supervised learning model used for classification and regression analysis. The motivation of SVM is to find the maximum-margin hyperplane for two or more classes which the distance from it to the nearest data point on each side is maximized to classify the data.

The K-Nearest Neighbors(KNN) Algorithm has been proposed by Altman N S in 1999[8]. KNN is a non-parametric method used for classification and regression. Nonparametric regression is a set of techniques for estimating a regression curve without making strong assumptions about the shape of the true regression function.

Back Propagation has been put forward by Rumelhart D E[9]. BPNN is a feedforward neural network that backward errors during training and iteratively updates the network parameters using the Gradient Descent method until the loss value converges.

Huang, a professor at Nanyang Technological University in Singapore, proposed a new single hidden layer feedforward neural network called Extreme learning machine [10]. Extreme learning machine (ELM) is a single-hidden layer feedforward neural networks (SLFNs) which randomly chooses the input weights and analytically determines the output weights of SLFNs. The experimental results show that the new algorithm can produce best generalization performance in some cases and can learn much faster than traditional popular learning algorithms for feedforward neural networks [11] .

Both the ELM algorithm and the BPNN have a feedforward neural network architecture, the difference is that the two algorithms use different learning methods. BPNN uses the gradient descent method to learn by back propagation, which requires continuous iteration. ELM uses the method of randomly generating parameters to transform the iterative solution process into the solution process of linear equations, thus speeding up the solution and avoiding the possibility of falling into the local optimum. Compared with the SVM algorithm, the SVM algorithm is based on nonlinear mapping theory and uses kernel functions instead of high-dimensional space mapping. According to different classifications, two solution models are provided for the two classification and multi-classification problems. In contrast, the ELM algorithm provides a unified model for all classification problems, inheriting the structural advantages of Single-hidden layer feedforward networks, and the computing speed is faster than the SVM algorithm. The ELM algorithm has simple structure and good learning efficiency. We choose the ELM algorithm to construct the brain decoder instead of SVM.

# Chapter 3

# Research Methods

## 3.1 Multi-variate Pattern Analysis

When a person performs some kind of cognitive processing, the corresponding neuron activity in the brain is enhanced. The oxygen consumption in the area where these neurons are located increases, and the oxygen supply in the brain active area increases significantly and the increase is greater than the increase in oxygen consumption. The combined effect of the two is that the total amount of oxyhemoglobin in the brain's active area is significantly higher than that in other areas, and the deoxygenated hemoglobin content is relatively reduced. Blood oxygen level depend fMRI (BOLD-fMRI) uses the change of regional proton lateral relaxation time caused by changes in oxygenated hemoglobin and deoxygenated hemoglobin in brain activity area to get the brain activity image.

Haxby J V has proposed multi-voxel pattern analysis of fMRI data in 2001[12][13]. Before the advent of MVPA, researchers used a generalized linear model to analyze the data. Many important research results are based on this method, but this method has its limitations. The principle of GLM determines that it can only analyze the activity of each voxel in the brain separately, and cannot analyze the activation patterns of multiple voxels in the brain [14] . MVPA breaks the limitations of GLM and more sensitively detects subtle multi-voxel pattern changes in brain fMRI data, thereby inferring neural representation of specific cognitive states. This method treats the activation of multiple voxels in fMRI data as a pattern and decodes information related to experimental conditions from the activation mode of the voxel.

Norman K A and his team also did a lot of work on MVPA. Their paper introduced pattern-classification algorithms to multi-voxel patterns of functional MRI data. This multi-voxel pattern analysis approach as led to several impressive feats of mind reading. And they think MVPA methods constitute a useful new tool for advancing our understanding of neural information process [15].

Multi-voxel pattern analysis usually considers multi-voxel activation in the brain as a pattern in high-dimensional space, and uses pattern classification to decode the information contained in the activation pattern. Analysis of fMRI data in the brain using multi-voxel pattern analysis typically involves the following steps:

Step1: Feature selection.

Step2: Pattern assembly.

Step3: Classifier training.

Step4: cross-validation

### 3.1.1 Feature Selection

In the feature selection step, which voxels will be included in the analysis is decided. The spatial resolution of fMRI is relatively high, and fMRI whole brain data scanned by fMRI imaging equipment may contain tens of thousands of voxels. If each voxel is used as a feature for classification, the samples created using whole brain voxels will be ultra-high dimensional. In recent years, with the continuous improvement of spatial resolution, the total number of whole brains has reached hundreds of thousands. The magnitude of this dimension is clearly beyond many classifications.

The upper limit of the number of input features. Therefore, it is necessary to pre-select a subset of whole brain voxels or to calculate the weighted combination of voxels as new features before performing classification, and it is necessary to reduce the dimension of fMRI data.

Feature selection is to preselect the voxels, select a subset of voxels, use the subset to create samples and classify the patterns. One of the simplest feature selection methods is to limit the analysis to a specific brain area, that is, to select the voxels in the Region of Interest (ROI) for subsequent analysis. Another common method is to calculate the univariate/voxel-wise statistics. These methods are collectively referred to as univariate feature selection methods. For example, voxels that can effectively distinguish experimental conditions can be selected as features alone. In fact, univariate statistics commonly used in traditional fMRI analysis can be used for feature selection. Such as raw fMRI data, averaged fMRI data, searchlight and values from a Generalized Linear Model analysis.

In this experiment, we chose the GLM method to find out the active brain area in the cognitive task. Combined with the knowledge of the functions of various parts of the brain, the Region of Interest is selected and all the other voxels are removed. This method can reduce the dimension of the input data by 1-3 orders of magnitude.

### 3.1.2 Pattern assembly

Before training the classifier, you also need to consider how to create samples using fMRI data. A sample is typically a vector of multiple eigenvalues, each of which has a category label that identifies the category to which the sample belongs. Usually, each category has multiple samples, all samples and their corresponding category labels make up the data set, and pattern classification is performed on the data set. In the extraction of feature values, you can use the average of multiple TRs in a single trial to create a sample. We can use either TR in a trial to create a sample or the average of multiple trials of the same stimulus to design the pattern[17].

### 3.1.3 Classifier

After obtaining the sample data, we divide the sample data into training and testing machines, train the classifier with the training set, and test the accuracy of the classifier with the test set. The test set can be used to evaluate the correctness of the classifier based on the training set based on the assumption that the training set and the test set satisfy the properties of independent and identical distribution. Therefore, the allocation of training sets and test sets must be randomly assigned.Traditional Classifier of MVPA including Nearest neighbor, Support Vector Machines(SVM) and Neural Networks.

### 3.1.4 Cross-validation

Since the number of samples obtained from the experiment is limited, in order to ensure a good estimate of the experimental results, we choose cross-validation. Cross-validation can make good estimates of the accuracy of the classification while using as much training data as possible. We use the k-fold cross-validation[16] . The processing steps are as follows:

Step1: Divide the dataset into $k$ groups, each group must contain samples of all categories, and the quantity is balanced with each other.

Step2: Take the first group as the test set, and the rest as a training set to train a classifier.

Step3: Take the $i$ group as the test set, $i = 2, 3, ..., k$, repeat the previous step until the $k$ group is done.

Step4: Calculate the prediction accuracy.

## 3.2 Extreme Learning machine

In our experiment, we choose Extreme learning machine as the classifier of the brain decoder. Compared to BPNN and SVM，ELM has several salient features:

- Easy to use. You could adjust the network by tuning the number of hidden nodes.No other parameters need to be manually tuned except predefined network architecture.
- Very fast training speed. Most training can be completed in fast speed. Even the features is complex it could also finish the training in a fast speed like SVM.
- Higher performance. From researchers experiment,ELM could obtain better generalization performance than BP in most cases, and reach generalization performance similar to or better than SVM.
- Suitable for lots of activation functions. Almost all piecewise continuous can be used as activation functions. And fully complex functions can also be used as activation functions in ELM.

The Single-hidden layer feedforward neural network consists of three layers, an input layer, an output layer, and an implicit layer, each of which is composed of neurons. The external information is received through the input layer. Each neuron receives only the output of the previous layer and uses it as its own input value. The output layer can feed back information to the outside world. Figure 3.1 below is a schematic diagram of the neural network structure model. The input of the neural network in the figure is represented by a circle. The left layer of the network is called the input layer. All the nodes in the middle of the network are called the hidden layer. The layer on the right is called the output layer. Neurons in adjacent layers are connected by connection weights.

For $N$ arbitrary distinct samples $(\mathbf{x}_i, \mathbf{t}_i)$, where $\mathbf{x}_=[x_{i1}, x_{i2}, x_{i3}, x_{i4}, ..., x_{in}]^T \in \mathbf{R}^n$ and $t_i = [t_{i1}, t_{i2}, t_{i3}, ..., t_{im}]^T] \in \mathbf{R}^m$ , standard SLFNs with$\tilde{N}$hidden nods and activation

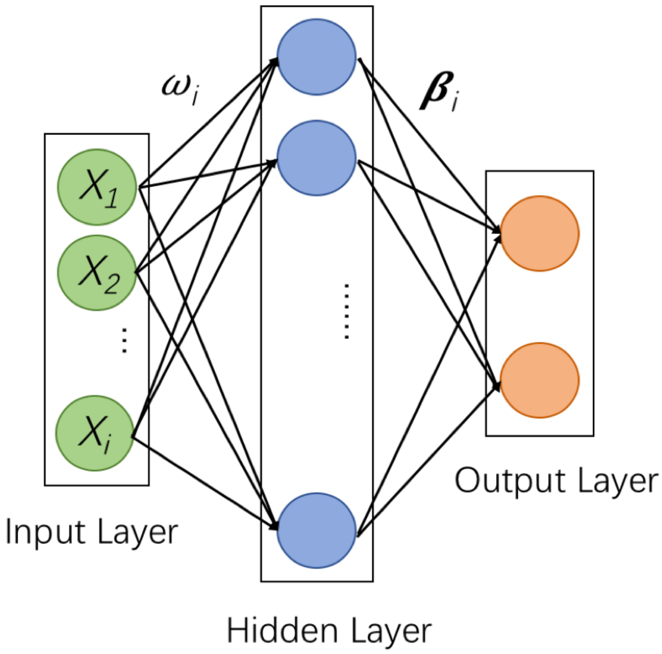


Fig. 3.1 Single-hidden layer feedforward neural network

function $g(x)$ are mathematically modeled as

$$\sum_{i=1}^{\tilde{N}} \beta_i g_i(\mathbf{x}_j) = \sum_{i=1}^{\tilde{N}} \beta_i g(\mathbf{w}_i \cdot \mathbf{x}_j + b_i) = \mathbf{o}_j, j = i, \cdots, N. \tag{3.1}$$

where $\mathbf{w}_= [w_{i1}, w_{i2}, w_{i3}, w_{i4}, ..., w_i n]^T$ is the weight vector connecting the $i$th hidden node and the input nodes, $\beta_= [\beta_{i1}, \beta_{i2}, \beta_{i3}, \beta_{i4}, ..., \beta_i n]^T$ is the weight vector connecting the $i$th hidden node and the output nodes, and $b_i$is the threshold of the $i$th hiddden node. The standard SLFNs with $\tilde{N}$ hidden nodes with activation function $g(x)$ can approximate these $N$ samples with zero error means that $\sum_{j=1}^{\tilde{N}} \|\mathbf{o}_j - \mathbf{t}_j\| = 0$, there exist $\beta_i, \mathbf{w}_i, \mathbf{b}_i$such that

$$\sum_{i=1}^{\tilde{N}} \beta_i g(\mathbf{w}_i \cdot \mathbf{x}_j + b_i) = \mathbf{t}_j, j = i, \cdots, N. \tag{3.2}$$

The N equations can be written compactly as

$$\mathbf{H}\beta = \mathbf{T} \tag{3.3}$$

where

$$\mathbf{H}(\mathbf{w}_1, \cdots, \mathbf{w}_{\tilde{N}}, b_1, \cdots, b_{\tilde{N}}, \mathbf{x}_1, \cdots, \mathbf{x}_{\tilde{N}}) = \begin{bmatrix} g(\mathbf{w}_1 \cdot \mathbf{x}_1 + b_1) & \cdots & g(\mathbf{w}_{\tilde{N}} \cdot \mathbf{x}_1 + b_{\tilde{N}}) \\ \vdots & \cdots & \vdots \\ g(\mathbf{w}_1 \cdot \mathbf{x}_N + b_1) & \cdots & g(\mathbf{w}_{\tilde{N}} \cdot \mathbf{x}_N + b_{\tilde{N}}) \end{bmatrix}_{N \times \tilde{N}}, \tag{3.4}$$

$$\beta = \begin{bmatrix} \beta_1^T \\ \vdots \\ \beta_{\tilde{N}}^T \end{bmatrix}_{\tilde{N} \times m} \quad \text{and} \quad \mathbf{T} = \begin{bmatrix} \mathbf{t}_1^T \\ \vdots \\ \mathbf{t}_N^T \end{bmatrix}_{N \times m}. \tag{3.5}$$

$\tilde{H}$is called the hidden layer output matrix of the neural network; the $i$th column of $\tilde{H}$ is the $i$th hidden node output with respect to inputs $x_1, x_2, x_3, x_4, ..., x_N$.

Traditionally, inorder to train an SLFN, we wish to find specific $\hat{\mathbf{w}}_1, \hat{b}_1, \hat{\beta}(i = 1, 2, ..., \tilde{N})$ such that

$$\left\| \mathbf{H}(\hat{\mathbf{w}}_1, \cdots, \hat{\mathbf{w}}_{\tilde{N}}, \hat{b}_1, \cdots, \hat{b}_{\tilde{N}})\hat{\beta} - \mathbf{T} \right\| = \min_{\mathbf{w}_i, b_i, \beta} \left\| \mathbf{H}(\mathbf{w}_1, \cdots, \mathbf{w}_{\tilde{N}}, b_1, \cdots, b_{\tilde{N}})\beta - \mathbf{T} \right\| \tag{3.6}$$

which is equivalent to minimizing the cost function

$$E = \sum_{j=1}^{N} \left( \sum_{i=1}^{\tilde{N}} \beta_i g(\mathbf{w}_i \cdot \mathbf{x}_j + b_i) - \mathbf{t}_j \right)^2. \tag{3.7}$$

So, when the $\mathbf{w}$ and $b$ has been chosen randomly before training , the $\beta$ could be solved by

$$\min_{\beta} \| H\beta - T' \| \tag{3.8}$$

the solution is

$$\hat{\beta} = \mathbf{H}^{+}T' \tag{3.9}$$

where $\mathbf{H}^{+}$ is the Moore-Penrose Generalized inverse of matrix $\mathbf{H}$ [18].

Thus, Huang proposed a single-hidden layer feedforward neural network Extreme Learning Machine.

## 3.3 Convolutional Neural Networks

Convolutional neural networks are multi-layer neural networks that excel at dealing with related machine learning problems of images, especially large images[21].

Convolutional neural networks have succeeded in reducing the dimension of data-intensive image recognition problems and ultimately enabling them to be trained. CNN was first proposed by Yann LeCun and applied to handwritten font recognition (MINST)[20]. There is Convolution layers, pooling layers and fully connected layers. There are multiple famous CNN model such as AlexNet[21] , ResNet[22] , GoogleNet [23]and so on. Convolutional neural networks structure is shown in Figure 3.2:

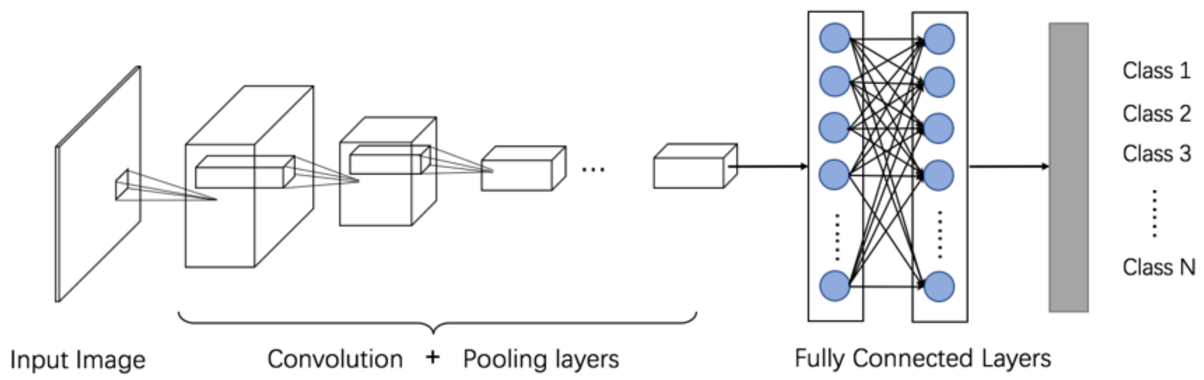


Fig. 3.2 Structure of Convolution neural networks

### 3.3.1 Convolution

Convolution Layer is the layer to extract features form an input image. It apply a convolution operation to the input and pass the result to the next layer. Convolution operation is a mathematical operation that takes two inputs such as image matrix and a filter or kernel.

### 3.3.2 Pooling

Pooling layers section would reduce the number of parameters when the images are too large. There are different types of pooling: Max Pooling, Average Pooling and Sum Pooling.

# Chapter 4

# Experiments

## 4.1 Experiments Design

### 4.1.1 Experiment Dataset

We selected 24 images which could make people feel pleasant and 24 images which could make people feel unpleasant from Open Access Series of Imaging Studies(OASIS) and show these images to subjects. OASIS is a project that making freely available brain neuroimaging datasets. The selected images which could make subjects feel pleasant including healing images such as lovely animals, beautiful views and smiling faces and so on. The selected images which could make subjects feel unpleasant including injured animals, crying faces blooding knife and so on. We shown these images to subjects while they are in the fMRI Machine. Used fMRI machine to retrieved their brain activity images while they seeing the images.

### 4.1.2 Subjects

There are 6 subjects (4 males and 2 females). The fMRI machine was used to collect the stimuli of the subject's brain. The safety of the fMRI experimental device was informed before the experiment. The experiments were started after obtaining the consent of the subjects.

### 4.1.3 Experiment design

SIEMENS MAGNETOM Verio 3T is used as fMRI device located in Kochi University of Technology, Japan. TR (Repetition time) is 3 seconds. The sequence is ep2d (2-dimensional EPI).

The experimental design is shown in Figure 6.6.

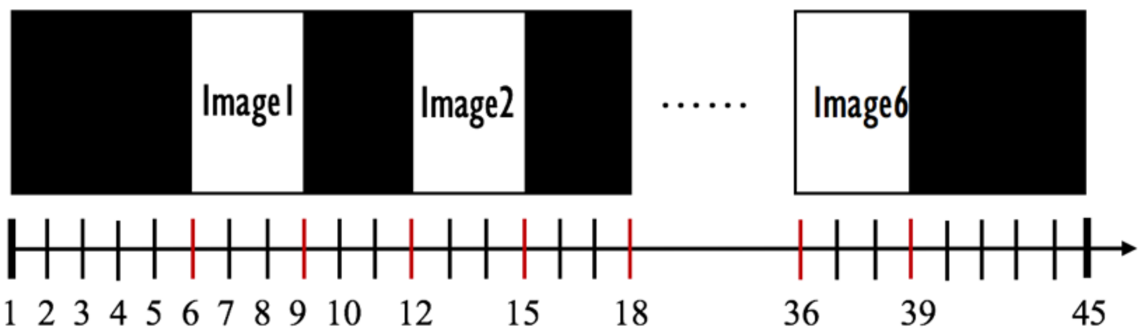


Fig. 4.1 Experimental Design

Each session lasted 132 seconds (44 scans). Each experiment started playing the picture 15 seconds after the start of imaging. Put the pleasant/unpleasant image on a black background and the time gap between each image is 9 seconds (3 scans).

Each subject do the experiment for 8 times. The order of the pleasant/unpleasant images were presented in each session is shown in Table 4.1 with pleasant images use "◯" and unpleasant images use "×". The order is same to the male subjects and female subjects.

Table 4.1 Experiment Session Design

| | Image1 | Image2 | Image3 | Image4 | Image5 | Image6 |
|---|---|---|---|---|---|---|
| Session1 | × | ◯ | × | ◯ | ◯ | × |
| Session2 | ◯ | × | ◯ | × | × | ◯ |
| Session3 | ◯ | × | ◯ | ◯ | × | × |
| Session4 | × | × | ◯ | × | ◯ | ◯ |
| Session5 | × | ◯ | ◯ | × | × | ◯ |
| Session6 | ◯ | ◯ | × | × | × | ◯ |
| Session7 | ◯ | × | ◯ | ◯ | × | × |
| Session8 | × | ◯ | × | ◯ | × | ◯ |

## 4.2 MVPA-ELM Brain Decoder Mode

In the MVPA-ELM Brain Decoder Model we use ELM combine with MVPA method to construct a brain decoder. The structure of MVPA-ELM Brain Decoder Model is shown in Figure 4.2. As the figure shown, there are 6 steps of our model.

Step1: Get Experiment Dataset.

Step2: fMRI dataset preprocessing.

Step3: Use GLM method define ROI.

Step4: Pattern assembly.

Step5: Use ELM to train the dataset and get the ELM classifier.

Step6: Cross-validation to evaluate MVPA-ELM model.

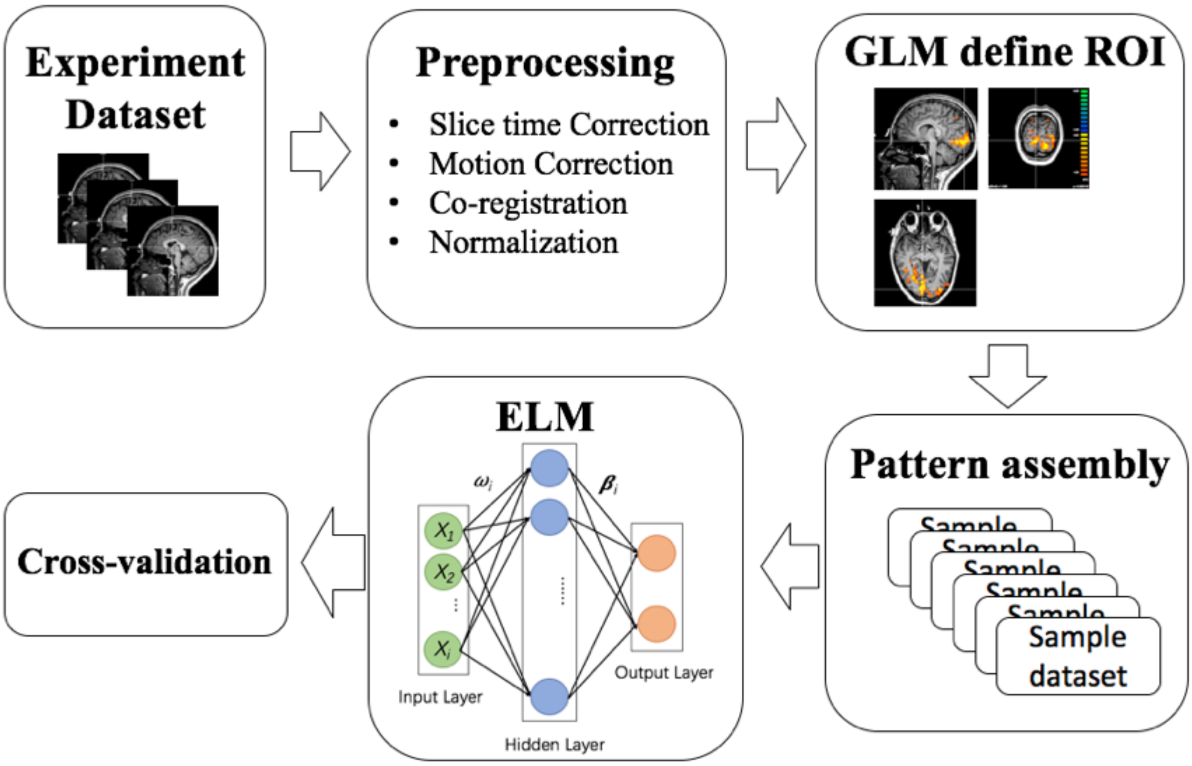


Fig. 4.2 MVPA-ELM Brain Decoder Model

### 4.2.1 fMRI Data Preprocessing

fMRI data preprocessing including: slice time correction, motion correction co-registration and normalization. After preprocessing we could get 3D Brain Images as Figure 4.3 shown.

### 4.2.2 Region of Interest

The voxel selection method based on the region of interest is based on anatomical localization or functional localization, defining a region of interest (ROI), and only selecting voxel components in the ROI to form a feature vector, which can greatly reduce the dimension of the sample. For the subsequent analysis and processing, especially the training of the classifier, providing great convenience.

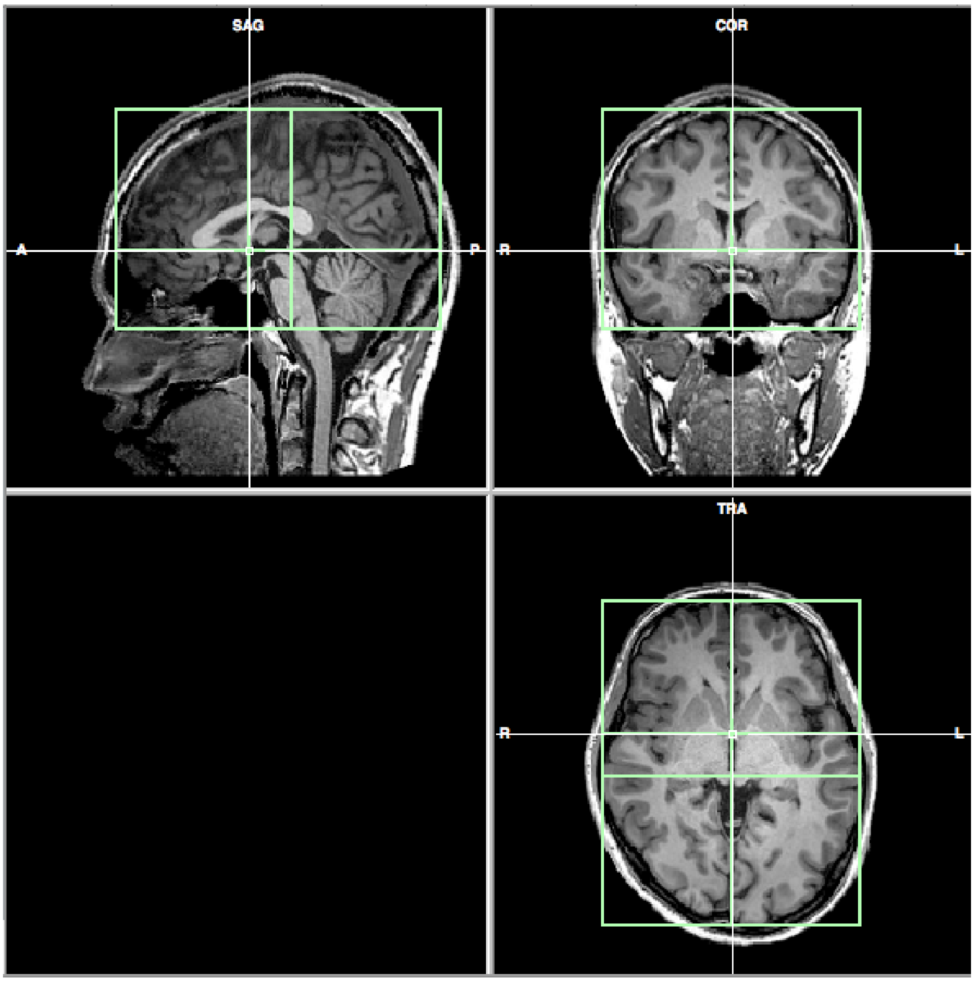


Fig. 4.3 3D Brain Image

### 4.2.3 Feature selection

A typical fMRI database contains Blood Oxygenation level dependent signal time courses recorded at multiple voxels in the brain. In this part we using GLM method to define the Region of Interes[24]. In order to map the cerebral areas involved in a given cognitive function, the BOLD signal at each voxel is analyzed. using the general linear model (GLM) approach to reveal activated brain areas by searching for linear correlations between the fMRI time course and a reference model[25].

Figure4.4 shows the response of one voxel in the whole time course. The yellow part is the time course which the subject see the picture which make them feel unpleasant and the green part is the time course of they feel pleasant.

From the left figure in Figure 4.4 we could see the main Brain activation areas of our experiment is mainly around 17,18 and 19 areas of Human Brodmann areas [26].

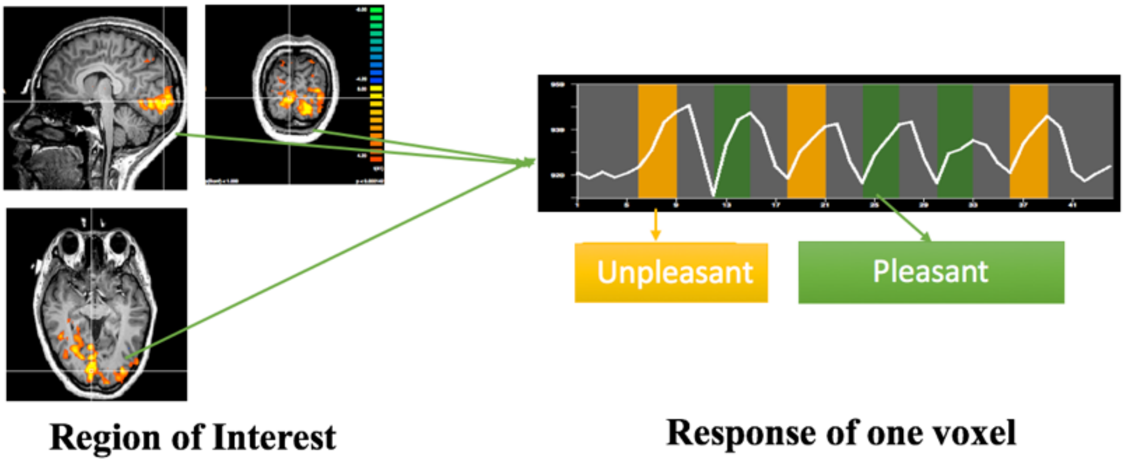


Fig. 4.4 Use GLM to define ROI

### 4.2.4 Classification and cross-validation

After feature selection, we use the features to train the ELM classifier and use cross-validation to evaluate the ELM model.

## 4.3 CNN-ELM Brain Decoder Model

In the CNN-ELM Brain Decoder, we combine Convolution neural network with ELM to construct a brain decoder. The structure of MVPA-ELM Brain Decoder Model is shown in Figure 4.5.

As the figure shown, there are 5 steps of our model.

Step1: Get Experiment Dataset.

Step2: fMRI dataset preprocessing.

Step4: Pattern assembly.

Step5: Use CNN to train the dataset and get the CNN classifier.

Step6: Cross-validation to evaluate CNN-ELM model.

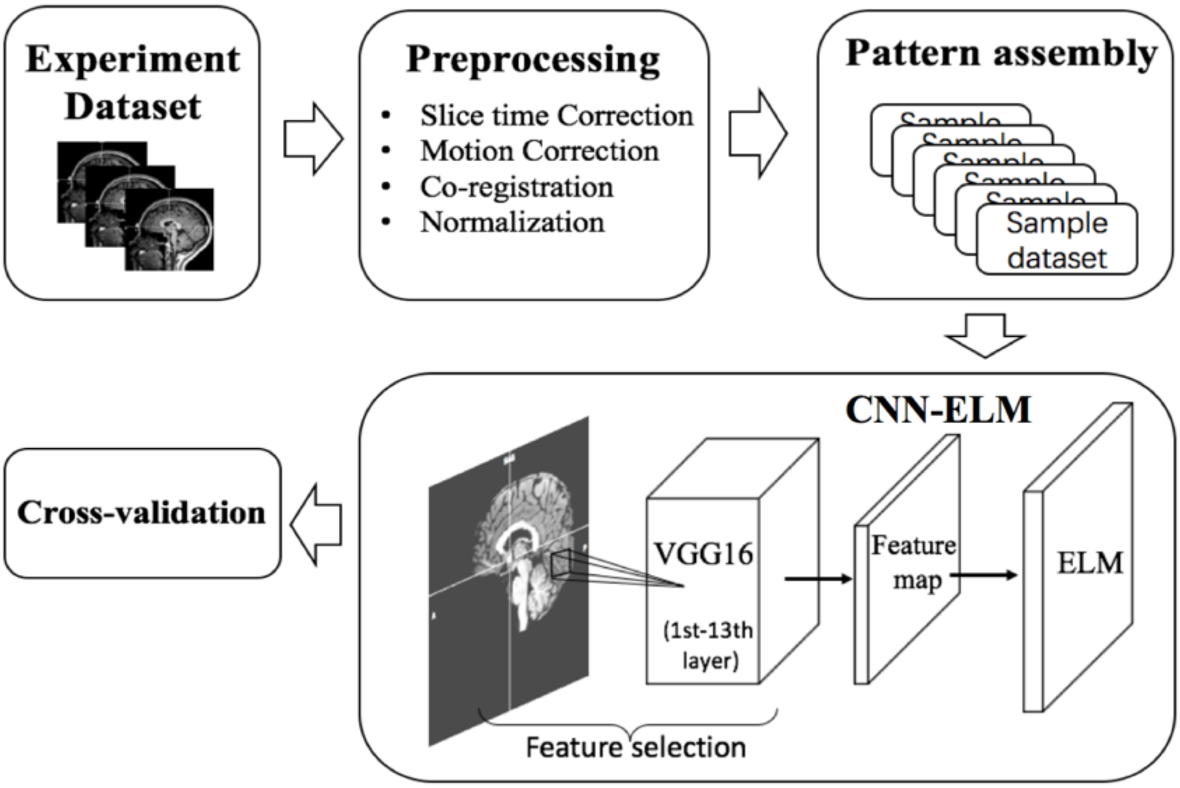


Fig. 4.5 CNN-ELM Brain Decoder Model

### 4.3.1 Feature selection

In the CNN-ELM method, we use the VGG16 CNN Model to extract features of fMRI dataset[27]. The network of CNN-ELM is shown in Figure 4.6. We use fMRI dataset as the input to the CNN after preprocessing the experimental data to extract features. 1st to 13th convolution layers of VGG16 is used to convolution and pooling to get the features. The input fMRI image size is $512 \times 512 \times 3$. After convolution we could obtain a $79 \times 79 \times 3$ size feature map. Then put the feature into Extreme learning machine.

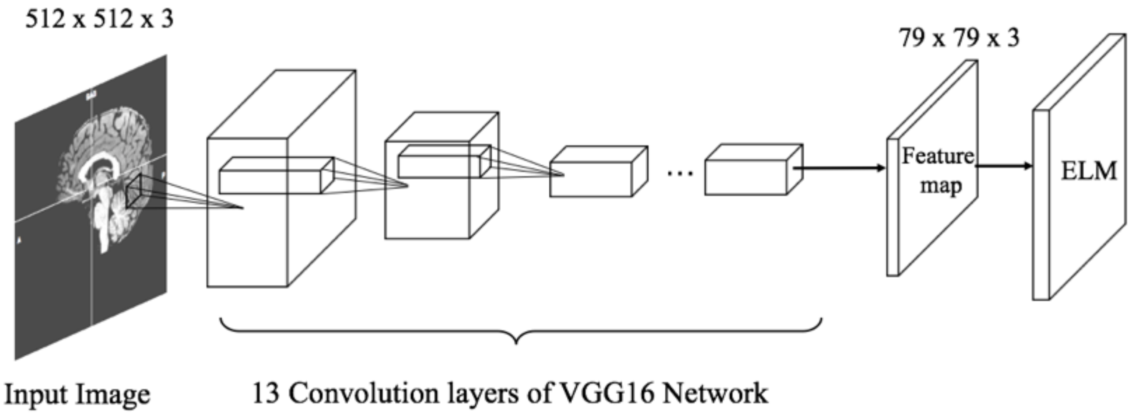


Fig. 4.6 CNN-ELM network

### 4.3.2 Classification and cross validation

After using CNN model to extract features we train the ELM using these features and finally use cross-validation to evaluate the accuracy of the CNN-ELM brain decoder model.

## 4.4 Results Analysis

### 4.4.1 MVPA-ELM Brain Decoder Model

After Cross-Validation we got the average accuracy of MVPA-ELM model. Figure 4.7 shows the average experimental accuracy when the number of hidden nodes of ELM changes. We use tanh and sigmoid as activation function. Blue line is the results of using tanh activation function and orange line is sigmoid activation function.

From the results we could find when the number of hidden nodes of ELM is 2000-4000, the average accuracy rate remains at stable state around 65%. For sigmoid acti-

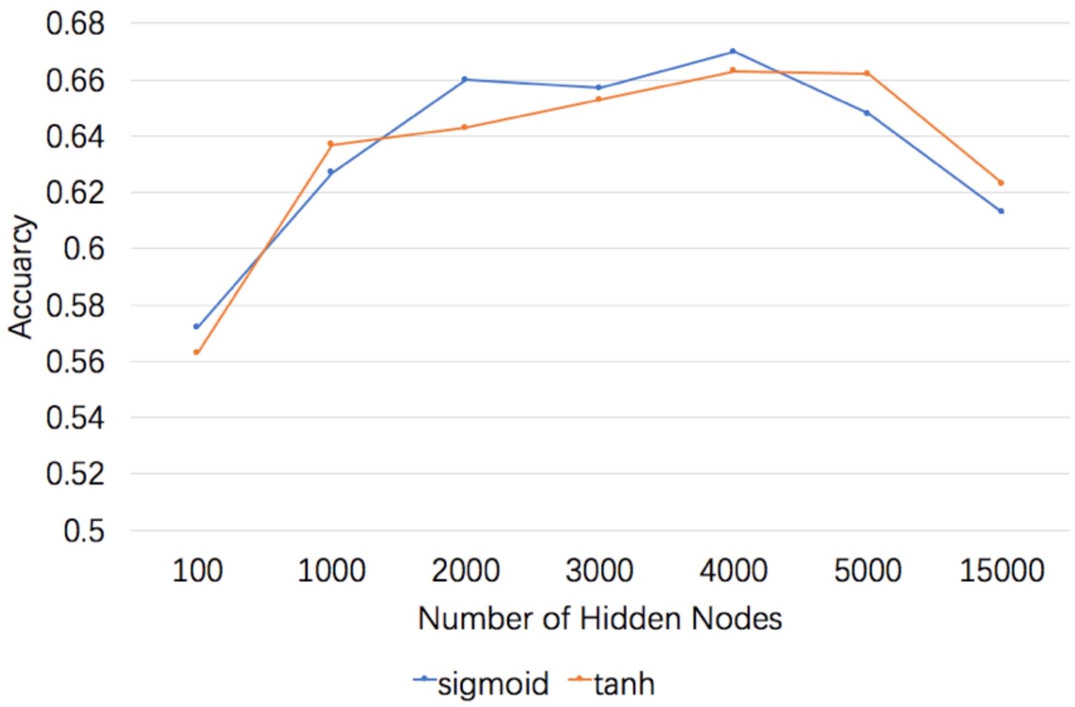


Fig. 4.7 Average accuracy of MVPA-ELM with different Number of hidden nodes

vation function the highest accuracy is 0.67 when the number of hidden nodes is 4000 and the lowest accuracy is 0.56 when the number of hidden nodes is 100. For tanh activation function the highest accuracy is 0.66 when the number of hidden nodes is 5000 and the lowest accuracy is 0.57 when the number of hidden nodes is 100. The tanh activation function has stable accuracy than sigmoid activation function. But when the number of hidden nodes is 2000 to 4000 sigmoid could get better results. The results shows different activation function didn't effect the accuracy too much. But the number of hidden nodes has a strong connection with the classify accuracy. Sigmoid activation function has better performance when the number of hidden nodes is between 2000 to 4000.

Figure 4.8 shows the MVPA-ELM Brain Decoder Model's accuracy of subject 1-6. We could find out the accuracy of the brain decoder have strong connection with the original dataset. Blue part is the results of using tanh activation function and orange

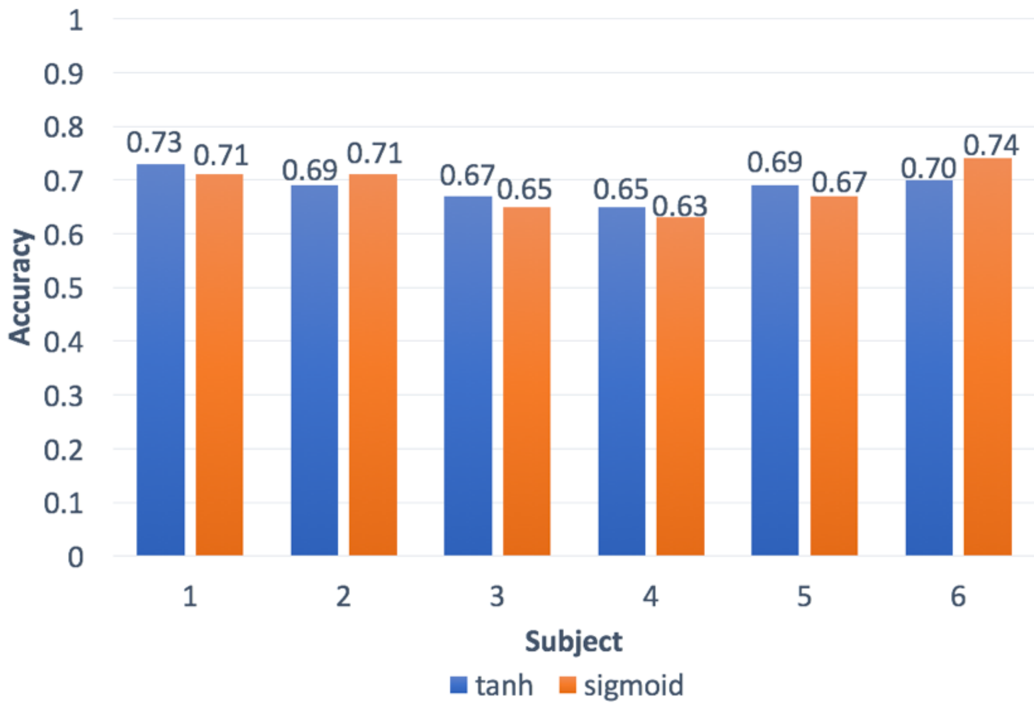


Fig. 4.8 MVPA-ELM accuracy of different subjects

part is sigmoid activation function.

In this experiment, using GLM method to get the ROI could help us to find the activation brain area makes the difference of different subjects could be seen easily. The brain decoder could obtain better accuracy from Subject 1 and Subject 6 than other subjects. Subject 3 and Subject 4 did not obtain as good results as others.

### 4.4.2 CNN-ELM Brain Decoder Model

The result of average accuracy of CNN-ELM with different Number of hidden nodes is shown in Figure 4.9. From the figure we could see the CNN-ELM Model could get good performance when the number of hidden nodes is between 1000-2000. The highest accuracy is 0.587 when the number of hidden nodes is 1000. From the result we could find when the number of hidden nodes is around 100 to 1300 and 2500 to 4000 tanh activation function has better performance. When the number of hidden nodes is 1300

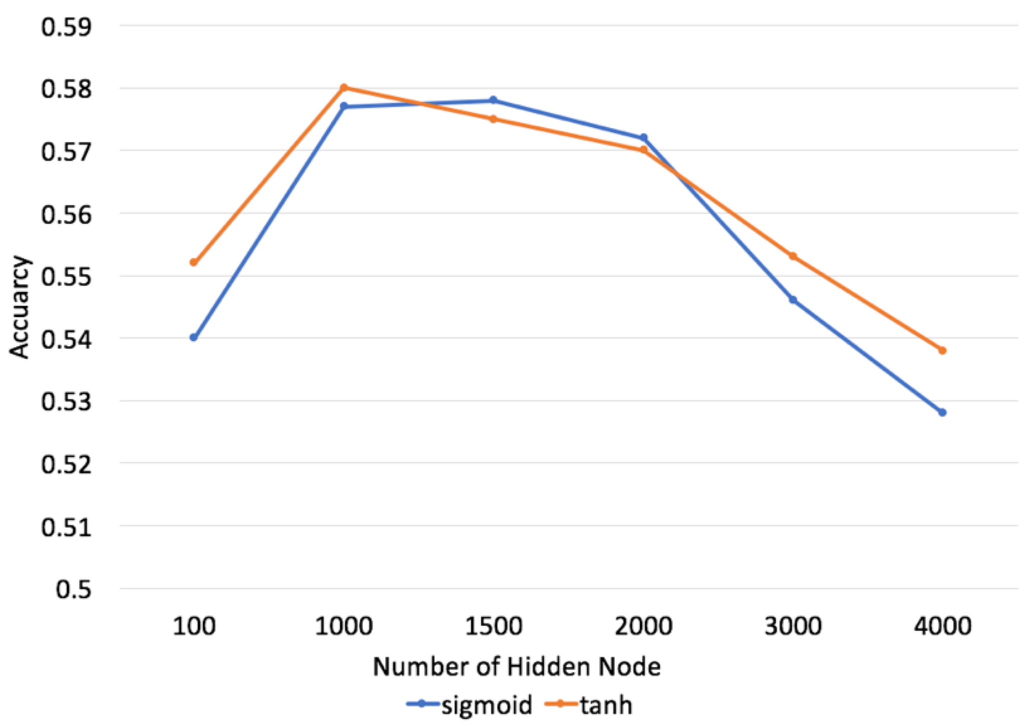


Fig. 4.9 Average accuracy of CNN-ELM with different Number of hidden nodes

to 2500 sigmoid activation function has better performance. The results shows different hidden nodes make a big difference of the accuracy.

From Figure 4.7 to Figure 4.10 we could see compare to CNN-ELM Brain Decoder Model, MVPA-ELM has better performance.

Figure 4.10 shows CNN-ELM Brain Decoder Model's accuracy of subject 1-6. From this results, the difference of different subject is not obvious. Using CNN method to select the feature is a new method in the experiments of fMRI dataset. There is no ROI that may makes the accuracy of different subjects did not make a big difference. Compare to Figure 4.8 we could ind out the MVPA-ELM Brain Decoder Model has higher accuracy and both two activation function has almost same performance in both two Models.

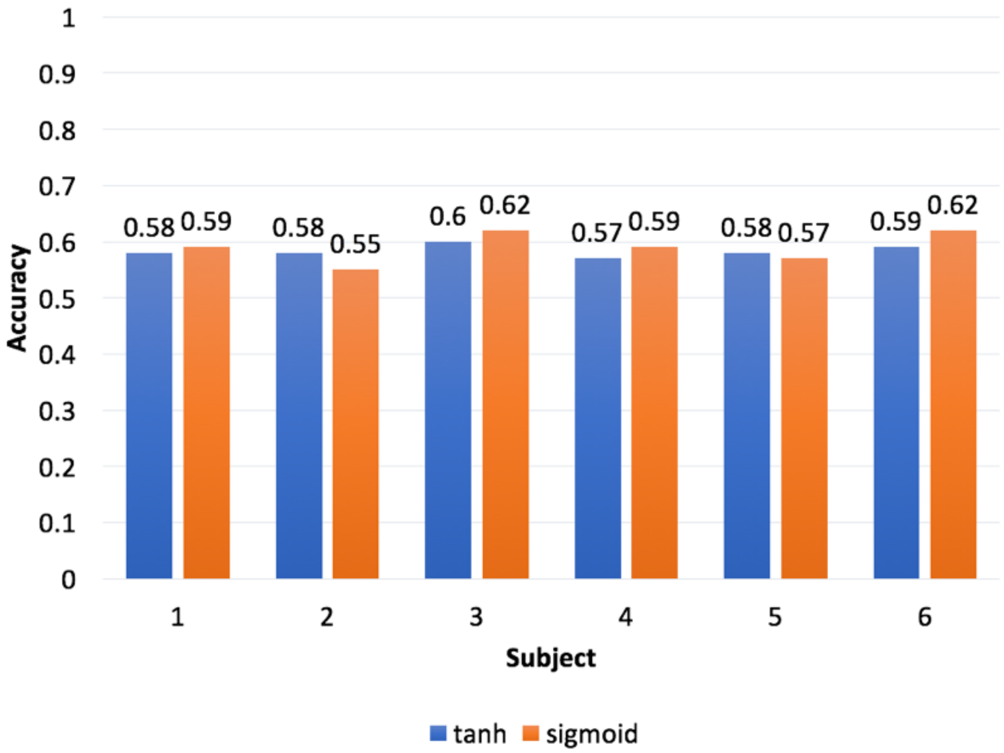


Fig. 4.10 CNN-ELM accuracy of different subjects

### 4.4.3 Comparison

The result of average accuracy of MVPA-ELM Model is shown in Table 1. Including ELM with tanh activation function, ELM with sigmoid activation function and ELM with CNN model. We also use the SVM with linear kernel, SVM with RBF kernel and SVM with CNN model as the classifier to construct brain decoder. The result comparison shows base on our fMRI dataset , ELM(tanh), ELM(sigmoid), and SVM(RBF) have good performance on the decoding of human emotion. From Table 1 we could see the accuracy of MVPA-ELM Model with tanh activation function has the best performance compare to other methods.

The result is shown in Table 1. we could see the average accuracy of ELM Model with tanh activation function has the best performance compare to other methods. The

Table 4.2 Accuracy of ELM Model, (t):tanh, (s):sigmoid, (C):CNN(L):linear, (R):RBF

| | ELM(t) | ELM(s) | ELM(C) | SVM(L) | SVM(R) | SVM(C) |
|---|---|---|---|---|---|---|
| 1 | **0.73** | 0.71 | 0.60 | 0.65 | 0.69 | 0.60 |
| 2 | 0.69 | **0.71** | 0.57 | 0.54 | 0.66 | 0.66 |
| 3 | 0.67 | 0.65 | 0.62 | 0.51 | **0.70** | 0.66 |
| 4 | 0.65 | 0.63 | 0.59 | 0.58 | **0.68** | 0.66 |
| 5 | **0.69** | 0.67 | 0.58 | 0.62 | 0.62 | 0.66 |
| 6 | 0.70 | **0.74** | 0.62 | 0.61 | 0.65 | 0.66 |
| Avg. | **0.688** | 0.685 | 0.597 | 0.585 | 0.667 | 0.65 |

average accuracy is around 0.688%.

# Chapter 5

# Conclusion

The visual system is the main way for the human brain to obtain information from the objective world. Its complexity and precision are beyond the reach of existing science and technology. Research on visual information processing mechanism has always been an important direction in brain science research. Functional magnetic resonance imaging technology has greatly promoted the research of visual information processing mechanism by virtue of its high spatial resolution and relatively moderate temporal resolution.

As the most popular visual information decoding method, MVPA method further promotes the research of visual information decoding. According to the technical characteristics and data characteristics of fMRI, this thesis studies the voxel selection method and feature extraction method in multi-voxel mode decoding technology. Combine the MVPA method with Extreme Learning Machine. It is the first time for analyzing the fMRI image by combining ELM and CNN.

The main research work includes:

1. MVPA-ELM Brain Decoder Model. Using MVPA-ELM to construct the brain decoding model, using the MVPA research method to introduce the ELM algorithm in machine learning to classify and process fMRI images.

2. CNN-ELM Brain Decoder Model. The brain decoding model was constructed with CNN-ELM. The convolutional neural network was used to extract the eigenvalues of fMRI, and the extreme learning machine was used as the classifier of CNN.

3. Using the ELM algorithm to construct the brain decoding model while using the SVM method to construct the brain decoding model. Compare and analyze the experimental results of several different models.

The research results shows both MVPA-ELM and SVM method could get good performance in the Brain Decoding of Human Emotion induced by Visual Stimuli. The accuracy is improved from 66.7% to 68.8%.Although decoding accuracy of CNN-ELM is a little lower than MVPA-ELM and SVM method and did not achieve the predicted performance.I think there are multiple reasons such as the fMRI Dataset is too small to CNN model , the feature of fMRI is too complex.

# Acknowledgement

Entering Kochi University of Technology to follow Professor Yoshida Shinichi is a turning point in my research career. I would like to thank Prof. Yoshida for his help in my research during my two-year master's degree. In Yoshida lab I started to the research about brain decoding.

My main research is to use machine learning to process human brain fMRI images for human brain decoding. During the research process, Professor Yoshida gave me a lot of valuable suggestions in the research and gave me careful guidance and help in the research. Especially in the process of writing my master's thesis. I have got a lot of valuable opinions from Prof. Yoshida . Prof. Yoshida has a deeply research on machine learning algorithms. In the time I live at the Kochi University of Technology, I further studied the knowledge of machine learning and successfully completed the experimental part of this thesis.

Another person I would like to thank is Prof. Ren Xiangshi of HEC. Because of him, I was able to study for my master's degree at Kochi University of Technology. I would like to thank Prof. Ren and all the teachers who have helped me in KUT for their help.

During my studies at Kochi University of Technology, my classmates were very friendly to me as an international student and gave me a lot of help in my study and life. I would like to thank all the students who have helped me.

I am grateful to my parents for their spiritual encouragement during my master's degree. They gave me the motivation to continue my study.

Finally, I would like to thank Kochi University of Technology for providing me with a good learning opportunity to enter a research team in different countries, to reach more scientific research fields, to increase my knowledge and experience, and to make me more scientific research.

# References


[1] N. Koike, H. Takahashi, and S. Yoshida, "Decoding Emotion by Deep Brain Voxel Selection using fMRI, " IDHF,2014

[2] Yoshida S. Decoding of emotional visual stimuli using fMRI brain signal[C]//Computer and Information Science (ICIS), 2016 IEEE/ACIS 15th International Conference on. IEEE, 2016: 1-4.

[3] Logothetis N K, Pauls J, Augath M, et al. Neurophysiological investigation of the basis of the fMRI signal[J]. Nature, 2001, 412(6843): 150.

[4] Fusar-Poli P, Placentino A, Carletti F, et al. Functional atlas of emotional faces processing: a voxel-based meta-analysis of 105 functional magnetic resonance imaging studies[J]. Journal of psychiatry & neuroscience, 2009.

[5] Matsuo T., Yoshida S., Improvement of Convolutional Neural Network using Extreme Learning Machine as Full-connected Layers, 4th InternationalWorkshop on Advanced Computational Intelligence 1 and Intelligent Informatics, 2015.

[6] Kamitani Y, Tong F. Decoding the visual and subjective contents of the human brain[J]. Nature neuroscience, 2005, 8(5): 679.

[7] Cortes C, Vapnik V. Support-vector networks[J]. Machine learning, 1995, 20(3): 273-297.

[8] Altman N S. An introduction to kernel and nearest-neighbor nonparametric regression[J]. The American Statistician, 1992, 46(3): 175-185.

[9] Rumelhart D E, Hinton G E, Williams R J. Learning representations by back-propagating errors[J]. nature, 1986, 323(6088): 533.

[10] Huang G B, Zhu Q Y, Siew C K. Extreme learning machine: a new learning scheme of feedforward neural networks[J]. Neural networks, 2004, 2: 985-990.

[11] Huang G B, Zhu Q Y, Siew C K. Extreme learning machine: theory and applications[J]. Neurocomputing, 2006, 70(1-3): 489-501.

[12] Haxby J V, Gobbini M I, Furey M L, et al. Distributed and overlapping representations of faces and objects in ventral temporal cortex[J]. Science, 2001, 293(5539): 2425-2430

[13] Haxby J V. Multivariate pattern analysis of fMRI: the early beginnings[J]. Neuroimage, 2012, 62(2): 852-855.

[14] Liang K Y, Zeger S L. Longitudinal data analysis using generalized linear models[J]. Biometrika, 1986, 73(1): 13-22.

[15] Norman K A, Polyn S M, Detre G J, et al. Beyond mind-reading: multi-voxel pattern analysis of fMRI data[J]. Trends in cognitive sciences, 2006, 10(9): 424-430.

[16] Kohavi R. A study of cross-validation and bootstrap for accuracy estimation and model selection[C]//Ijcai. 1995, 14(2): 1137-1145.

[17] Pereira F, Mitchell T, Botvinick M. Machine learning classifiers and fMRI: a tutorial overview[J]. Neuroimage, 2009, 45(1): S199-S209.

[18] Tamura S, Tateishi M. Capabilities of a four-layered feedforward neural network: four layers versus three[J]. IEEE Transactions on Neural Networks, 1997, 8(2): 251-255.

[19] Krizhevsky A, Sutskever I, Hinton G E. Imagenet classification with deep convolutional neural networks[C]//Advances in neural information processing systems. 2012: 1097-1105.

[20] LeCun Y, Bottou L, Bengio Y, et al. Gradient-based learning applied to document recognition[J]. Proceedings of the IEEE, 1998, 86(11): 2278-2324.

[21] Krizhevsky A, Sutskever I, Hinton G E. Imagenet classification with deep convolutional neural networks[C]//Advances in neural information processing systems.

2012: 1097-1105.

[22] He K, Zhang X, Ren S, et al. Deep residual learning for image recognition[C]//Proceedings of the IEEE conference on computer vision and pattern recognition. 2016: 770-778

[23] Szegedy C, Liu W, Jia Y, et al. Going deeper with convolutions[C]//Proceedings of the IEEE conference on computer vision and pattern recognition. 2015: 1-9.

[24] Johnston S J, Boehm S G, Healy D, et al. Neurofeedback: A promising tool for the self-regulation of emotion networks[J]. Neuroimage, 2010, 49(1): 1066-1072.

[25] Monti M M. Statistical analysis of fMRI time-series: a critical review of the GLM approach[J]. Frontiers in human neuroscience, 2011, 5: 28.

[26] Brodmann K. Brodmann's: Localisation in the cerebral cortex[M]. Springer Science & Business Media, 2007.

[27] Simonyan K, Zisserman A. Very deep convolutional networks for large-scale image recognition[J]. arXiv preprint arXiv:1409.1556, 2014.

# Chapter 6

# Appendex

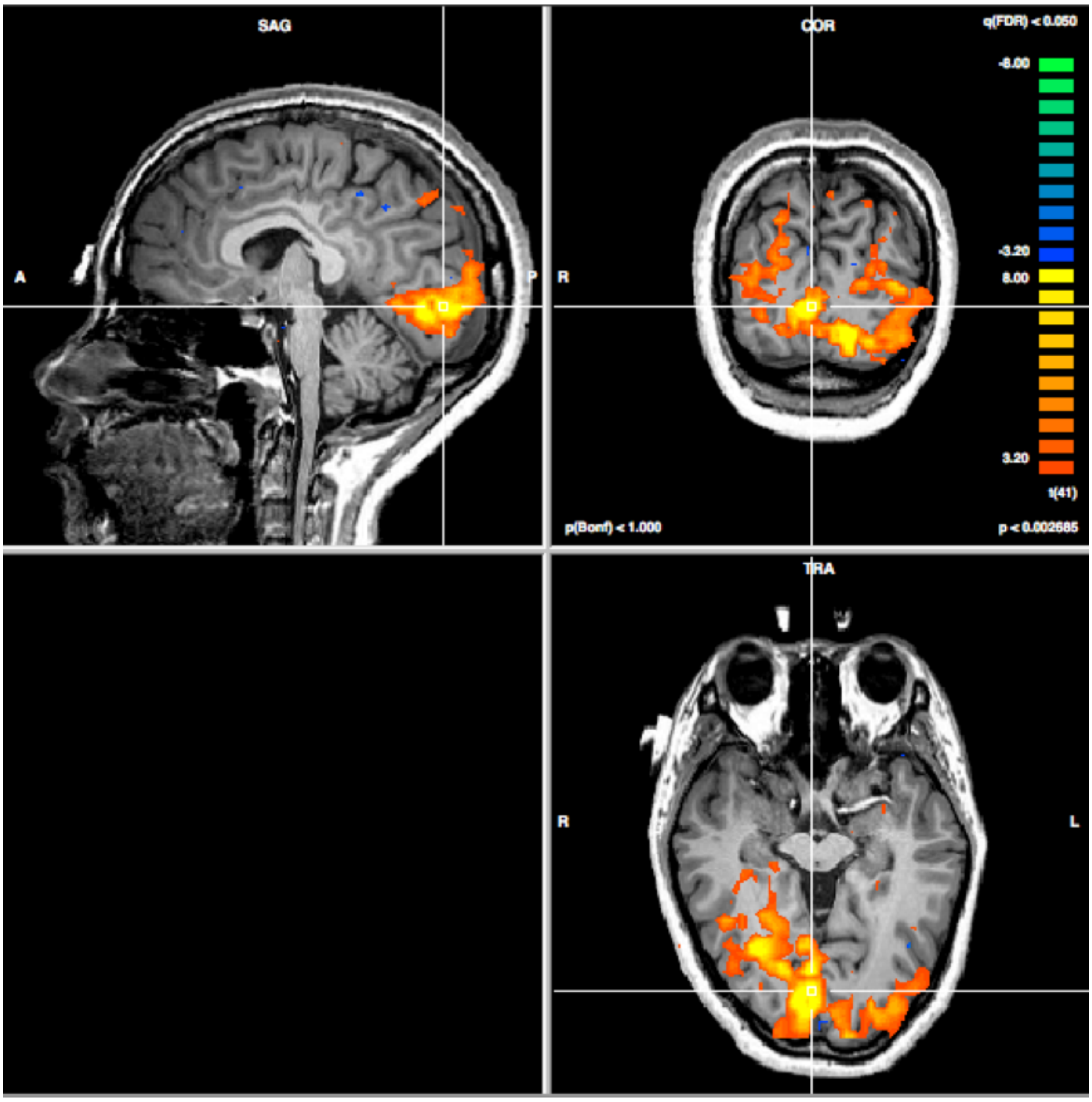


Fig. 6.1 ROI of Subject 1 using GLM

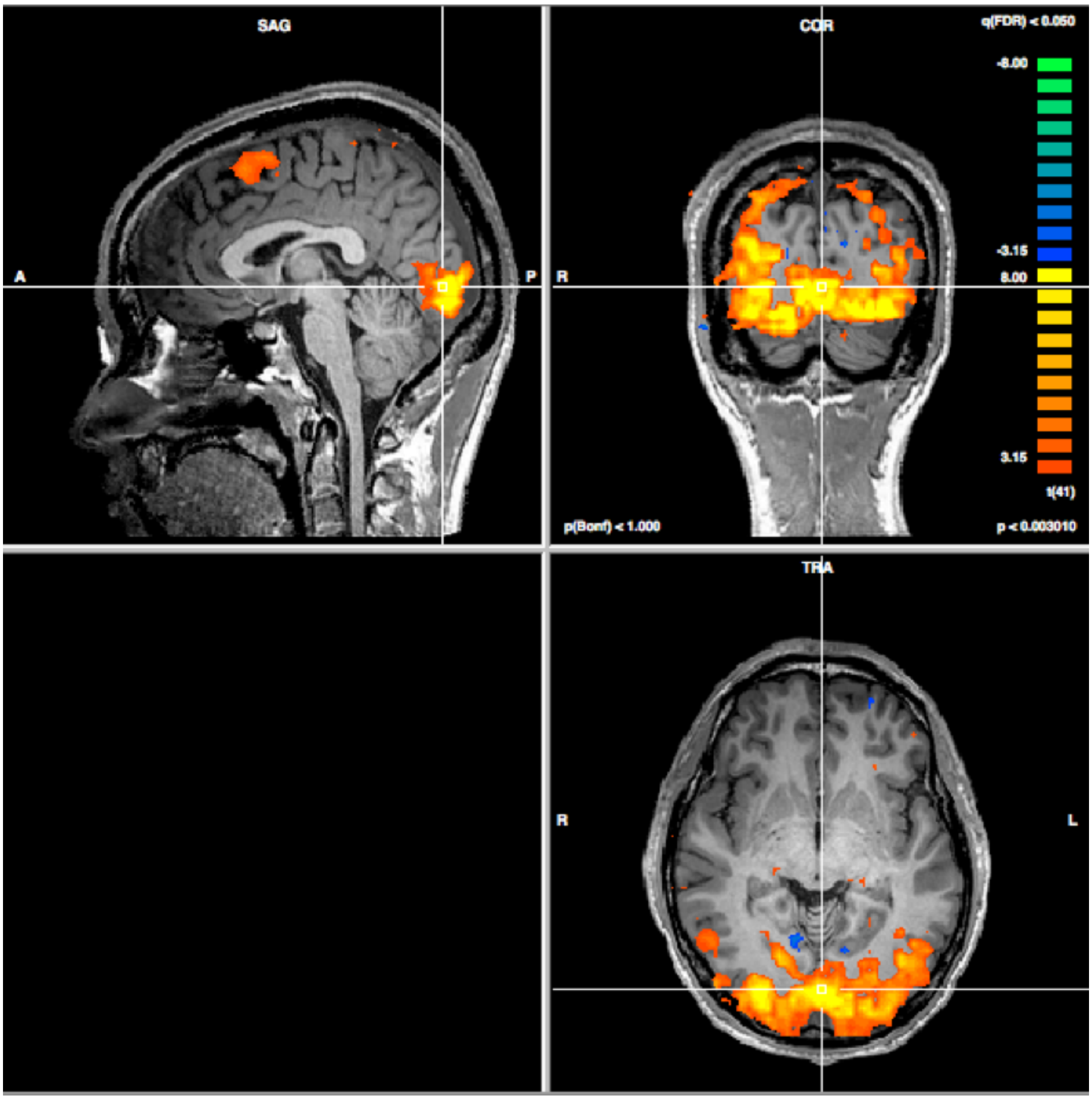

Fig. 6.2 ROI of Subject 2 using GLM

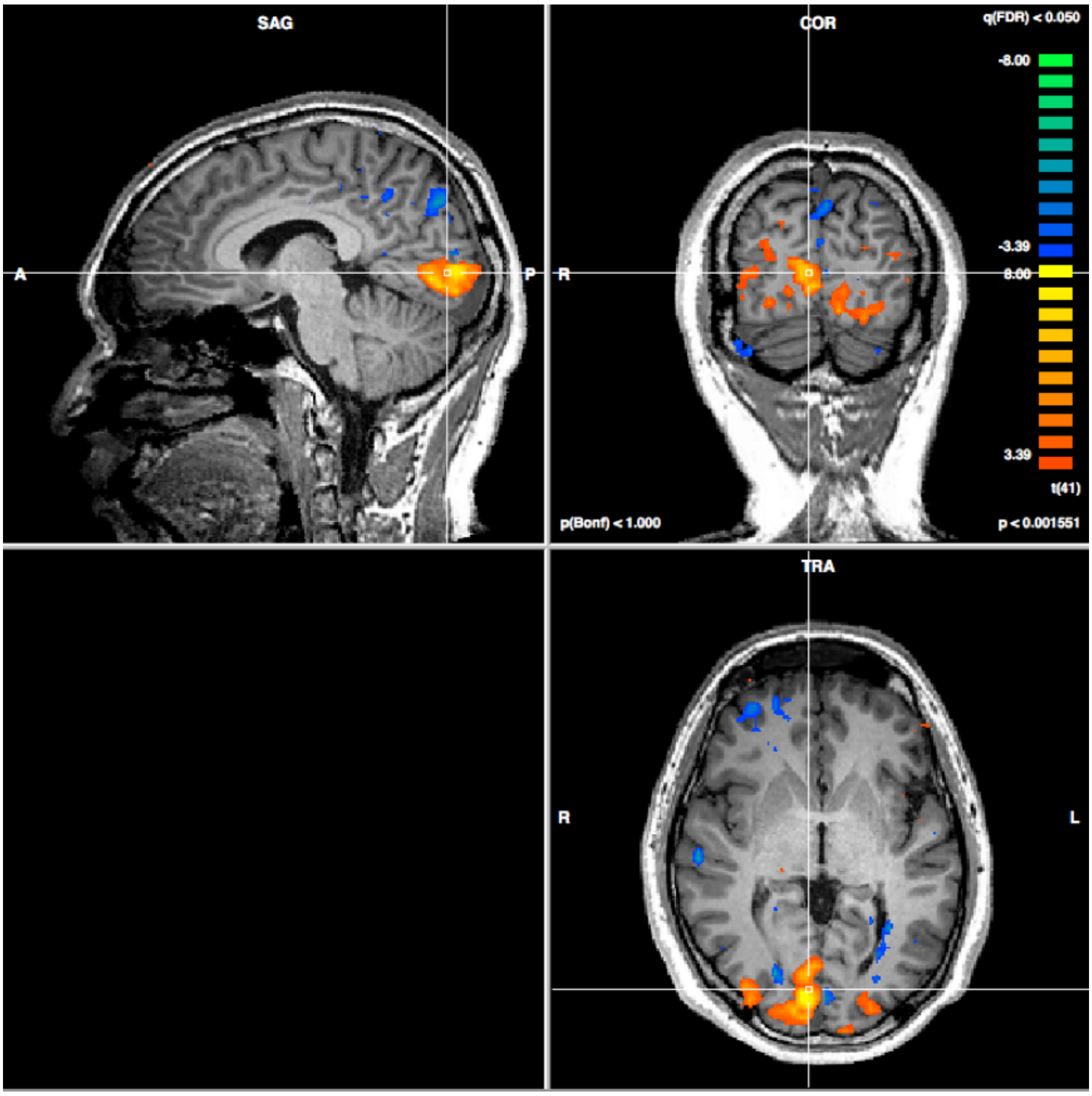


Fig. 6.3 ROI of Subject 3 using GLM

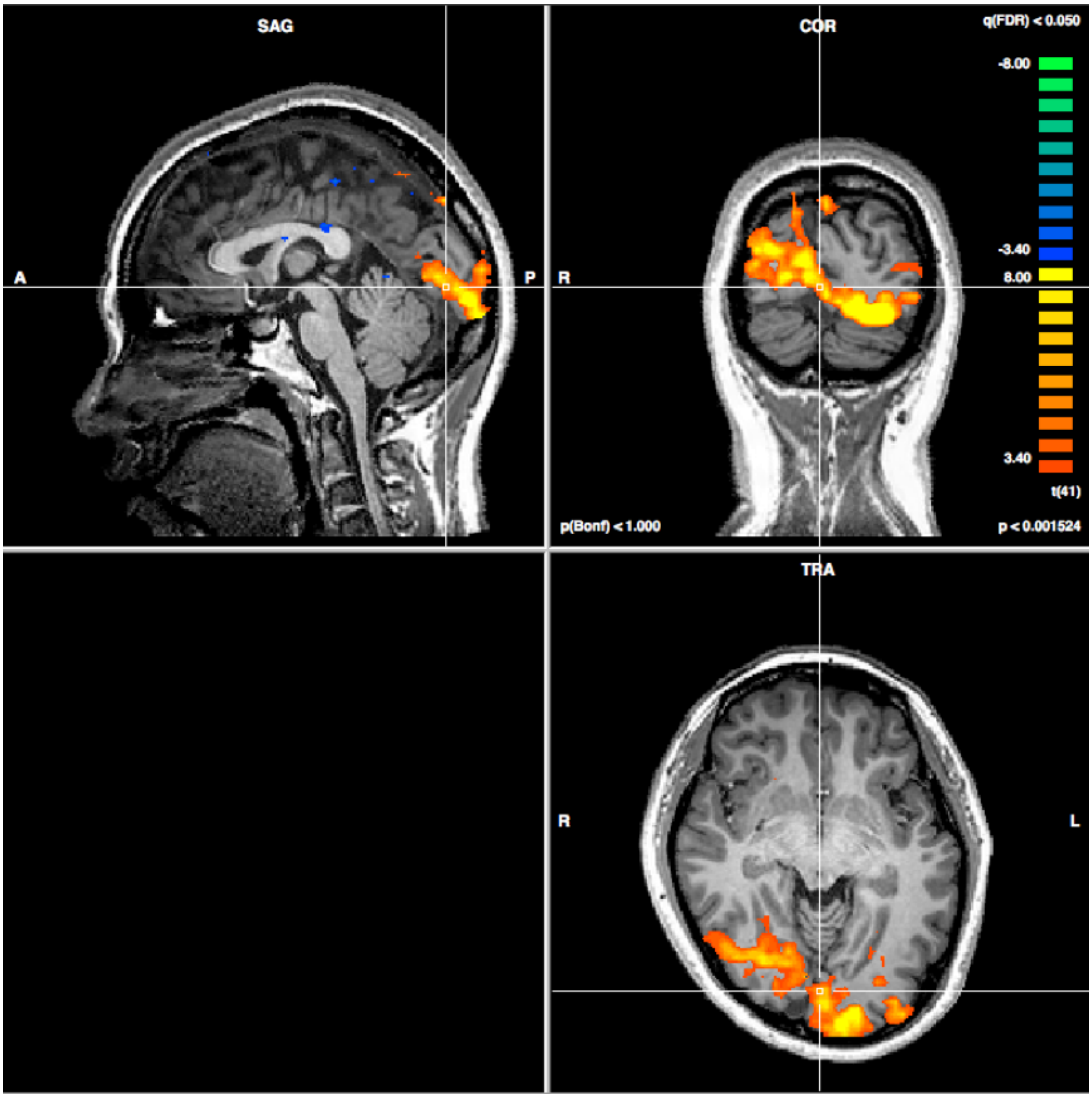


Fig. 6.4 ROI of Subject 4 using GLM

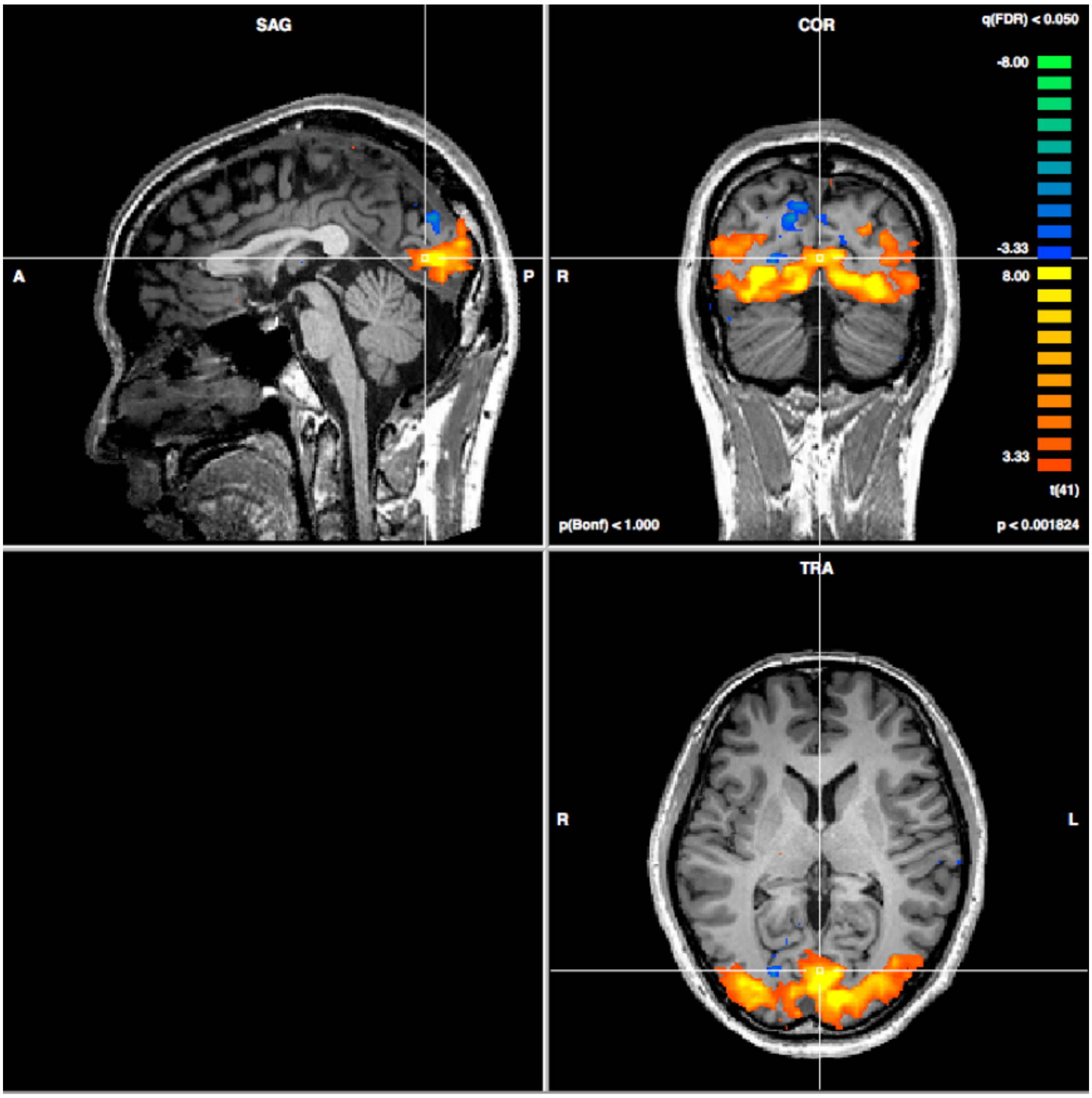


Fig. 6.5 ROI of Subject 5 using GLM

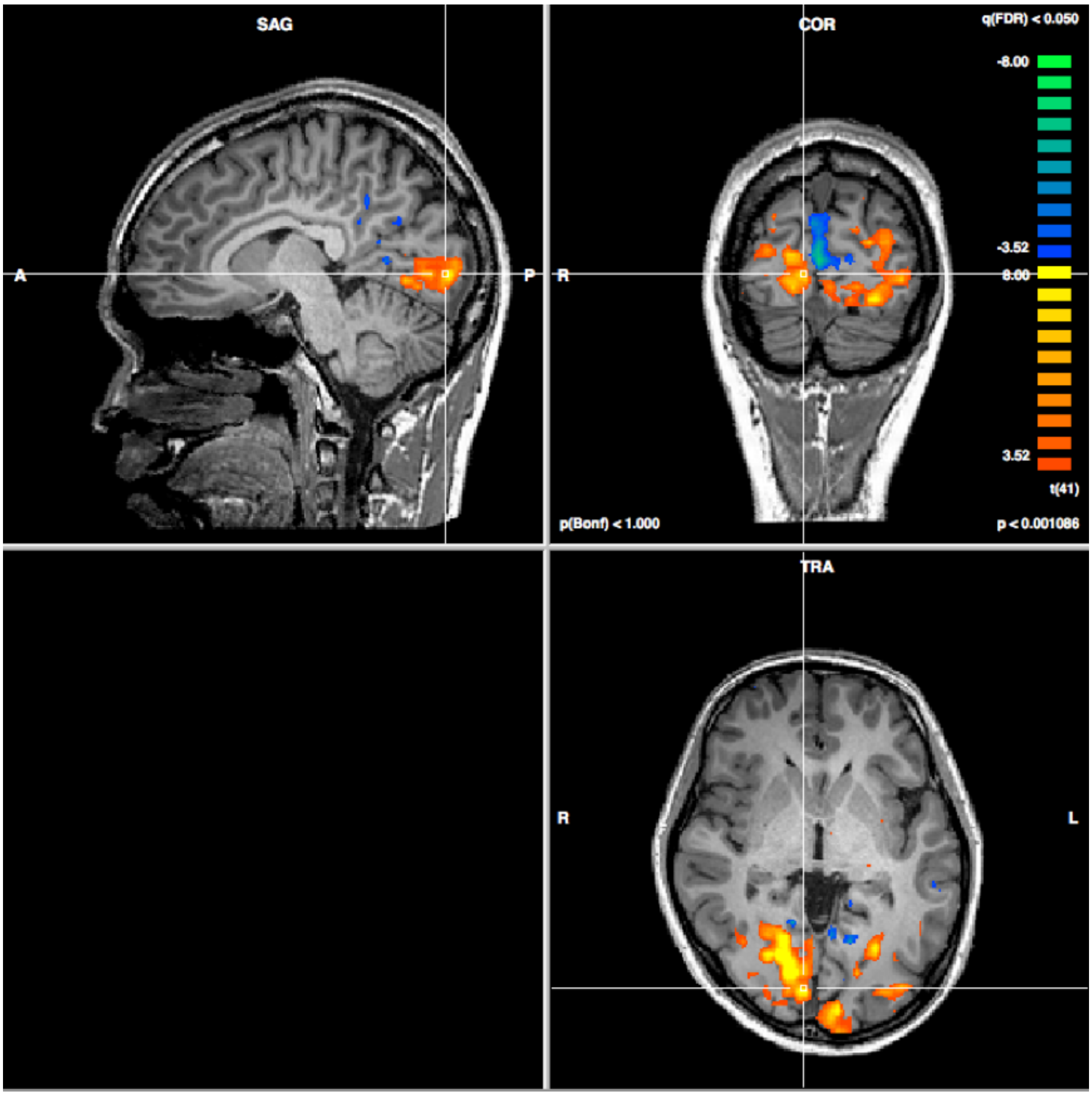


Fig. 6.6 ROI of Subject 6 using GLM