

The Application of kNN and SVM in the Decoding of fMRI Data

Fangyuan Ma and Junhai Xu*

School of Computer Science and Technology, Tianjin Key Laboratory of Cognitive Computing and Application, Tianjin University, Tianjin 300072, P.R. China

*Corresponding author

Abstract—For the decoding analysis of functional magnetic resonance imaging (fMRI) data, the appropriate method for feature selection and classification algorithm was a core issue. Given the high dimensionality of fMRI data in whole brain, the localization of regions of interest (ROI) usually was used to select voxels relevant to task, which was based on the physiological evidence. K-nearest neighbor (kNN) and linear support vector machine (SVM) were performed to classify the fMRI data. The result indicated that the method of ROI could indeed select voxels involved in the recognition task. The accuracy of ROI relevant to task was significantly higher than chance level and different stimuli could be decoded successfully. Additionally, by the comparison of kNN and SVM, the performance of SVM was better than that of kNN on the whole.

Keywords—functional magnetic resonance imaging; support vector machine; k-nearest neighbor

I. INTRODUCTION

Recent years, functional magnetic resonance imaging (fMRI) was widely used in the study of cognitive mechanism. As a non-invasive imaging technique, fMRI could measure the neural activity with high spatial resolution. Blood oxygen level dependent (BOLD) was a common MRI technique. Under external stimulus, the blood oxygen level in local areas would change depending on the stimulus condition, such as stimulus type or intensity. BOLD-fMRI could indirectly measure this change to produce MRI data. The MRI data could be used for the study of diseases and the function of brain. Previous neuroimaging studies have used brief point-light motion stimuli or human body stimuli to show that the occipito-temporal cortex plays a critical role at biological motion recognition. But now it remains unclear to what extent those regions are involved in the decoding of different movements.

After the fMRI data was pre-processed, multi-variate pattern analysis (MVPA) was applied to classify different activity patterns corresponding to different stimulus types [1]. This analysis method was driven by the data. The response evoked by one stimulus type in a region entered the classifier as input, and different stimulus type acted as labels. If one region contained some information about the recognition, the region should successfully decode different mental states.

The development of machine learning made it possible to apply algorithms to many fields. Many classification

algorithms were adopted by the decoding study of fMRI data, such as the classification for scenes or finger movements. The current study designed fMRI experiment and conducted classification analysis to explore whether the regions represent specific movement information. The procedure was mainly involved in feature selection and classification algorithm. The goal of feature selection was to select voxels the most relevant to the recognition task to reduce the dimension of data. Regions of interest was usually used to select voxels. The localizer experiment was performed to localize extrastriate body area (EBA) and posterior superior temporal sulcus (pSTS). At the same time, to confirm the rationality of feature selection, lateral ventricle (LV) which was irrelevant to the cognitive process was included in our analysis. About the selection of algorithm, the study utilized using k-nearest neighbor (kNN) and linear support vector machine (SVM) [2]. We aimed at making a comparison between the performances of classifiers to determine which of them was appropriate to the fMRI.

II. EXPERIMENTAL PROCEDURE AND METHODS

A. Experimental Procedure

Our experiment adopted a block design, including motion recognition task and localizer task. In the movement recognition task, subjects watched the videos of four kinds of movements performed by different persons. We used E-prime to display the stimulus. There were four types of movement, jump, run, skip and walk. Each scanning procedure consisted of eight block and began with a 10 s fixation-only baseline. Figure 1 shows the process in one block. In each block, eight 2520 ms videos of movement from the same category were presented with an inter-stimulus interval of 480 ms. The total time of one block was 24 s. After one block, subjects had to make a keypress to indicate the category of movement they had seen within 2 seconds. Each scanning session included eight blocks. Each session lasted 282 s. All subjects completed four sessions. The classification analysis was performed each subject separately. The fMRI scanner acquired one image every two seconds. Each session consisted of two blocks per class, with ten data points in each block (after two data points were removed). The total number of samples per class was $2 \times 10 \times 4 = 80$. Compared to previous decoding studies, this was enough to demonstrate the effectiveness of classifiers.

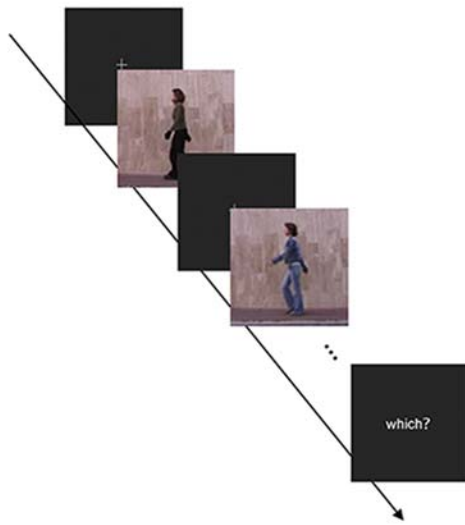


FIGURE 1. THE ILLUSTRATION OF EXPERIMENTAL PROCESS.

In order to localize EBA and pSTS, according to previous studies, subjects in the scanner watched the pictures of headless bodies and chairs. EBA was defined by the activation cluster in the posterior inferior temporal sulcus, in which headless bodies evoked stronger compared to chairs [3]. pSTS was defined as the activation cluster in the superior temporal sulcus using the contrast between point-light biological motion and the scrambled versions [4].

Seventeen healthy subjects (age: $M = 21.8 \pm 1.6$, 10 females) participated in our experiment. The functional and anatomical data were collected by a 3.0 T Siemens scanner in Yantai Hospital Affiliated to Binzhou Medical University with an eight-channel head coil. A three dimensional magnetization-prepared rapid-acquisition gradient echo (3D MPRAGE) sequence was used to acquire T1*-weighted images for an anatomical localization. T2*-weighted images were acquired using an echo-planar image (EPI) sequence.

B. Algorithms

KNN was a simple nonparametric method. Let $\mathbf{x} = x_1, \dots, x_n$ denote an fMRI pattern and $w_j, j = 1, \dots, c$, denote the different visual stimuli. kNN assumes that the fMRI pattern \mathbf{x} act as a point in R^n space. It defines the neighbors of \mathbf{x} in terms of a distance measure, such as Euclidean distance. The classifier would choose the most common class among the k training examples nearest to the pattern \mathbf{x} . Let X represent the set of all samples. If we have known all classes of X_i in X , the classifier would extend its spatial scale around \mathbf{x} until the space could contain k examples. Then \mathbf{x} was classified as the class which is the most frequent in the sample set. In other words, if most samples around \mathbf{x} belonged to w_j , \mathbf{x} was judged to class w_j .

As a multivariate method, SVM has great superiority in the pattern recognition of small sample and high-dimension data set. SVM just needs to find the support vector which depends on points at the border. Thus, SVM could probably gain relatively good performance. Now SVM has been employed to

the decoding study of fMRI data for the brain activity prediction and classification [5]. In our current study, we applied both SVM and kNN to the classification of fMRI data and made a comparison between their performances.

C. Data Processing

At first, SPM8 was used to process the data. The main procedure included preprocessing and statistical analysis. At the stage of preprocessing, slice timing and head motion correction were performed to the functional data. The structural images were segmented and co-registered to the functional data. Spatial normalizing was conducted to normalize the functional images into the Montreal Neurological Institute space. In addition, smoothing was also performed for the localizer data. Based on general linear model, first-level analysis was conducted to localize EBA and pSTS by the contrast between the response to headless bodies and chairs, and the contrast between point-light biological motion and the scrambled display. ROIs was defined at group level.

We used raw intensity for the classification analysis. The unsmoothed functional data were normalized by z-score. Two TRs were shifted to decrease the influence of hemodynamic lags. The LibSVM was used to extract the response activity of the three regions (<http://www.csie.ntu.edu.tw/~cjlin/libsvm/>). The imaging data was transformed to a matrix. The matrix represented the response of each voxel in one ROI at different time points. Subjects watched different movements at different time points. Thus, the class of each response vector at one time point could be calculated by the time of data acquisition. The dimensions of the matrix were the number of samples and the number of voxels in one ROI respectively. Data corresponded to four labels representing four movements according to the time points. Then the data was divided into eight groups and cross validation ($n = 8$) was performed. We used the data of seven groups to train a classifier and test it in the remaining group for each subject. Each group acted as the test set in turn. The final accuracy was the average of eight cross validations.

III. EXPERIMENTAL RESULTS

The accuracy of EBA, pSTS and LV from 17 subjects was averaged and showed in Figure II. The one-tailed t-test showed the classification accuracy of EBA and pSTS was significantly higher than the chance level of 0.25 ($p < 0.025$), which was not the case for LV. The average accuracy of SVM was significantly higher than kNN. In order to compare algorithms in detail, the accuracy of every subject was calculated.

We displayed the curve of accuracy of kNN and SVM in 17 subjects in Figure III. It showed the change of accuracy in EBA and pSTS. The accuracy of kNN was lower than that of SVM in about 14 of 17 subjects. Even in LV, Figure IV showed the performance of SVM was better than kNN.

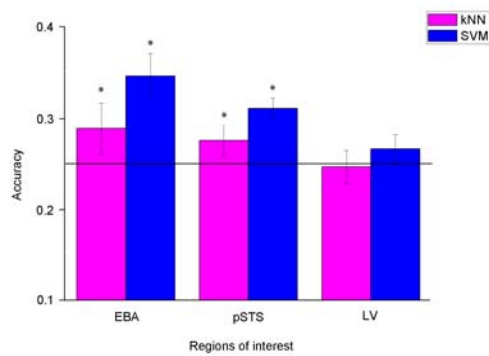


FIGURE II. THE AVERAGE ACCURACY OF 17 SUBJECTS IN ROIS.

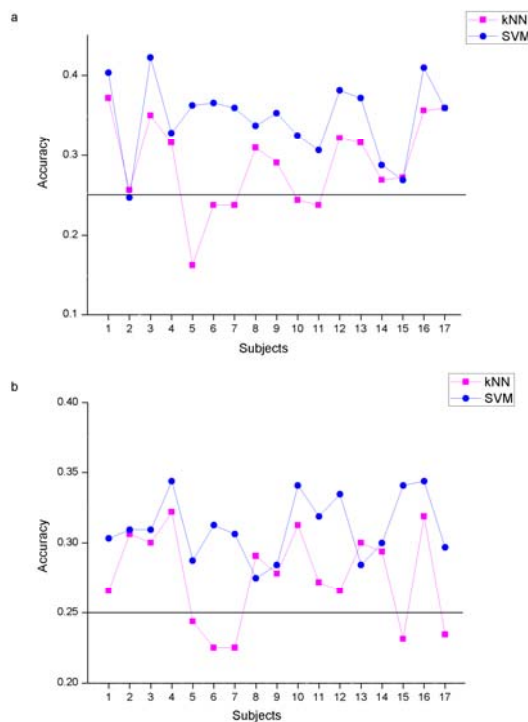


FIGURE III. THE CURVE OF ACCURACY IN EBA AND PSTS. (A) THE CURVE OF ACCURACY IN EBA. (B) THE CURVE OF ACCURACY IN PSTS.

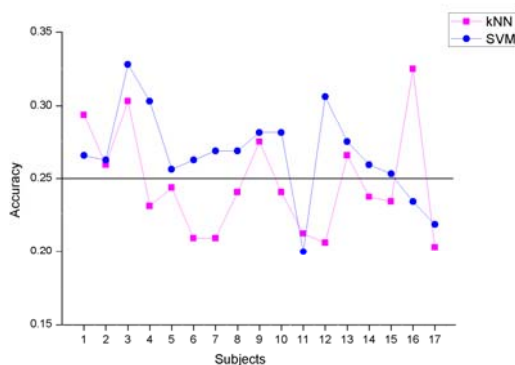


FIGURE IV. THE CURVE OF ACCURACY IN LATERAL VENTRICLE.

IV. DISCUSSIONS

Our result showed that the multivariate activity pattern in EBA and pSTS could successfully discriminate different movements. But the accuracy of LV was at chance level. The study indicated that both EBA and pSTS could represent motion information and they were indeed involved in the motion recognition task. This was consistent with previous investigations using the paradigm of fMRI adaptation [4, 6]. No previous studies has used the method of classification to investigate human movement recognition.

It has been proved that the area EBA was activated by human body form in static or dynamic displays, while pSTS was activated by the point-light biological motion. In our fMRI study, the discrimination of multi-voxel pattern in EBA and pSTS made clear their roles in human motion representation. In a recent study, the author and colleagues used walking direction to represent motion information to show that the activity pattern of pSTS carried motion information, whereas this was not the case for EBA [7]. The inconsistency could be caused by different measurement indexes. The correlation analysis was conducted in their study and the voxels overlapping with human middle temporal complex were removed from EBA, while we performed classification analysis in the whole EBA.

We also made a comparison between two algorithms, the performance of SVM was better than kNN. This indicated it was more appropriate to utilize SVM in the decoding of fMRI data than using kNN. The worse performance of kNN than SVM may result from its sensitivity to noise. Due to the high dimension of fMRI data, irrelevant features in the distance measure may result in poor performance. Additionally, although the size of our training data was small, SVM needed only the support vector defined by the samples in boarder and was not influenced by it.

ACKNOWLEDGMENT

We are grateful to professor Grossman for offering us the matlab code and materials of point-light biological motion.

REFERENCES

- [1] Allefeld C, Haynes JD. Multi-voxel Pattern Analysis A2 - Toga, Arthur W[M]. Brain Mapping. Waltham; Academic Press. 2015: 641-646.
- [2] Pereira F, Mitchell T, Botvinick M. Machine learning classifiers and fMRI: A tutorial overview[J]. NeuroImage, 2009,45(1, Supplement 1): S199-S209.
- [3] Downing PE, Jiang Y, Shuman M, Kanwisher N. A cortical area selective for visual processing of the human body[J]. Science (New York, NY), 2001,293(5539): 2470-2473.
- [4] Grossman ED, Jardine NL, Pyles JA. fMR-Adaptation Reveals Invariant Coding of Biological Motion on the Human STS[J]. Frontiers in Human Neuroscience, 2010,4: 15.
- [5] Wang Z, Childress AR, Wang J, Detre JA. Support vector machine learning-based fMRI data group analysis[J]. NeuroImage, 2007,36(4): 1139-1151.
- [6] Kable JW, Chatterjee A. Specificity of Action Representations in the Lateral Occipitotemporal Cortex[J]. Journal of Cognitive Neuroscience, 2006,18(9): 1498-1517.
- [7] Vangeneugden J, Peelen MV, Tadin D, Battelli L. Distinct Neural Mechanisms for Body Form and Body Motion Discriminations[J]. The Journal of Neuroscience, 2014,34(2): 574-585.