Decoding Human Actions in Human Brain by Searchlight MVPA Method

Xiaolin Sun[a], Baolin Liu[b,*]

College of Intelligence and Computing, Tianjin Key Laboratory of Cognitive Computing and Application, Tianjin University, Tianjin 300350, P. R. China

[a]monica_sxl@foxmail.com, [b]aoishihanae@foxmail.com

*Corresponding author

Keywords: Machine learning, Support vector machine, Functional magnetic resonance imaging, Multi-voxel pattern analysis

Abstract: Humans can identify actions performed by others rapidly and accurately. Here we applied machine learning techniques in functional magnetic resonance imaging (fMRI) data to reveal the neural mechanism underlying such neural activity. We used searchlight multi-voxel pattern analysis method based on support vector machine classifier to identify the regions that can discriminate different actions significantly, so as to clarify the brain regions recruited for action identifications. Our results indicated that part of the regions in occipito-temporal and frontal cortex hold the ability to identify different human actions.

1. Introduction

There is an emerging trend in the use of machine learning classifiers for analyzing fMRI data. It has been shown that machine learning classifiers can be trained to decode the variables of interest from fMRI data, and so as to show the data contain information about them [1]. Humans can identify actions performed by others rapidly and accurately. Using machine learning algorithms with fMRI data to reveal the neural mechanism of the action perception in human brain could be feasible.

In the present study, human subjects were scanned while four categories of actions were displayed. The fMRI data was collected. After the preprocessing of the data, we used the searchlight multi-voxel pattern analysis(MVPA) [2] to find the brain regions that can classify different action categories significantly above the chance level. In each classification analysis, 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 (to be specific, action recognition in the present study), the region should successfully decode different mental states. The regions identified with significant above-chance accuracies could be interpreted as being involved in the perception. The aim of the current study was to find the regions.

2. Experimental procedure and methods

2.1 Experimental Procedure

Twenty-five healthy volunteers (age: M = 21.8 ± 1.6 years old, 10 females) with normal or corrected-to-normal vision were recruited in this study. All participants were right-handed, and reported no history of psychiatric or neurologic disorders. Prior to the experiment, all participants gave written informed consent and were compensated for their time after the experiment. Eight additional participants (finally fMRI data from 17 volunteers were used in this study) were excluded for further analyzes because of a poor task performance or excessive head motion.

A block design was adopted in our experiments. Stimuli included four categories of colored natural videos, which depicted jumping, skipping, walking and running from left to right or vice versa performed by 8 actors, with similar outdoor backgrounds and 360 pixels wide by 288 pixels tall [3]. The exemplar stimuli used in the experiment were illustrated in Fig.1 A.
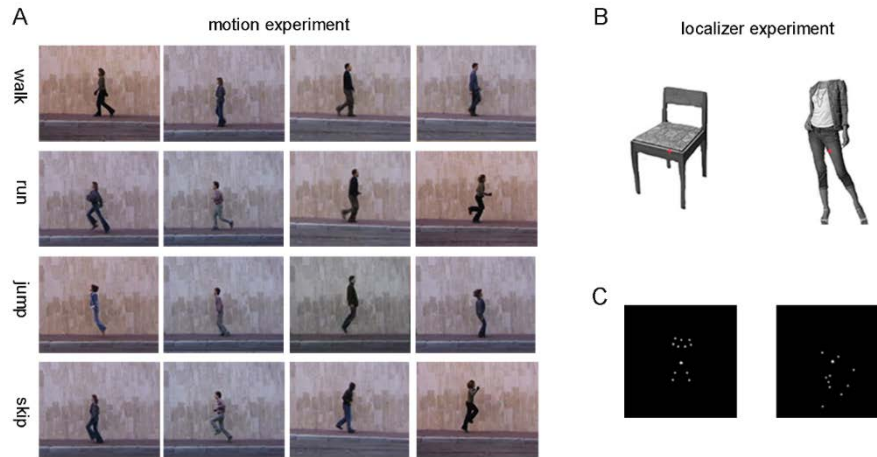
Figure 1. Exemplar stimuli used in motion experiment and functional localizer

A, Exemplar stimuli used in the motion experiment, including videos depicting four categories of human movements performed by 8 actors: walking, running, jumping and skipping, from left to right or vice versa; B, exemplar stimuli in functional localizer for EBA, including pictures of headless bodies and chairs; C, exemplar stimuli in functional localizer, including point-light biological motion videos, scrambled versions and the pictures of static frames from scrambled biological motions. The pictures shown here are the screenshots of the two kinds of videos.

In each scan, 8 stimulus blocks were included. Within each stimulus block, eight 2520-ms videos of the same movement performed by different actors were shown, alternating with 480-ms inter-stimulus interval accounting for a total time of 24 s for one block. After each block, participants were instructed to press a button at the response box within two seconds, indicating which motion they had seen. Each session started by a 10-s fixation-only baseline, and followed by a stimulus block. The same category of motion was presented twice in a pseudorandom order, with the constraint that the same block was not presented in succession. Each scan lasted 282 s, and each subject participated in 4 scans [3].

The functional and anatomical data were collected by a 3.0 T Siemens scanner in Yantai Hospital affiliated to Binzhou Medical University with a 20-channel head-neck coil. Foam pads and earplugs were used to reduce the head motion and scanner noise. T1*-weighted images for an anatomical localization were acquired using a three dimensional magnetization-prepared rapid-acquisition gradient echo (3D MPRAGE) sequence: repetition time (TR) = 1900 ms, echo time (TE) = 2.52 ms, voxel size = $1 \times 1 \times 1$ mm^3, matrix size = 256×256, flip angle (FA) = 9°). T2*-weighted images were acquired using an echo-planar image (EPI) sequence: TR = 2000 ms, TE = 30 ms, voxel size = $3.1 \times 3.1 \times 4.0$ mm^3, matrix size = 64×64, slices = 33, slices thickness = 4 mm, slices gap = 0.6 mm, FA = 90°).

2.2 Data Processing

Data preprocessing and statistical analysis were conducted using SPM8 (Wellcome Department of Imaging Neuroscience, London; http://www.fil.ion.ucl.ac.uk/spm). First, slice timing and head motion correction were performed for the functional data. The structural images (T1) were then segmented for normalization after co-registrating to the functional images. The spatial normalization parameters were applied to normalize the functional images into the Montreal Neurological Institute space (MNI), with a re-sampled voxel size of $3 \times 3 \times 3$ mm^3. The images in both functional and localizer runs were smoothed with a full-width at half-maximum = $4 \times 4 \times 4$ mm kernel to attenuate noise.

In search of the regions that can discriminate different actions, we conducted a whole-brain searchlight analysis [4]. Each time we centered a small spherical ROI (radius 6mm) around every voxel of the brain separately for each participant and then calculated the classification accuracy within each sphere. The resulting value was assigned to the central voxel of the sphere. A linear

support vector machine (SVM) classifier implemented by LIBSVM [5] was adopted. The classification accuracies were computed using leave-one-out cross validation. The individual accuracy maps for all of the subjects were finally entered into a one-sample t test to identify voxels yielding an index significantly above the chance level (0.25). A threshold of $p<0.05$ was adopted.

3. Experimental results

We found a series of brain regions that can discriminate different human actions above the chance level, which were summarized in Table 1. The location of the areas was illustrated in Fig. 2.

Table.1. Clusters identified in the searchlight MVPA analysis

|	Region	MNI coordinates			voxel number
		X	Y	Z	
1	Cerebelum_6_R (aal)	12	-60	-24	2159
2	Occipital_Inf_L (aal)	-51	-69	-12	1845
3	Cerebelum_Crus1_R (aal)	18	-81	-30	136
4	Lingual_L (aal)	-12	-102	-15	136
5	Cerebelum_6_L (aal)	-18	-57	-24	85
6	Cerebelum_9_R (aal)	6	-51	-48	76
7	Supp_Motor_Area_R (aal)	9	-3	45	50
8	Frontal_Inf_Orb_R (aal)	27	27	-24	39
9	Postcentral_R (aal)	42	-21	48	23
10	Precentral_L (aal)	-54	3	24	19
11	Olfactory_R (aal)	3	21	-6	18
12	Temporal_Sup_R (aal)	66	-12	0	14
13	Frontal_Inf_Tri_L (aal)	-48	27	18	14
14	Paracentral_Lobule_R (aal)	6	-30	78	14
15	Insula_L (aal)	-30	18	9	13
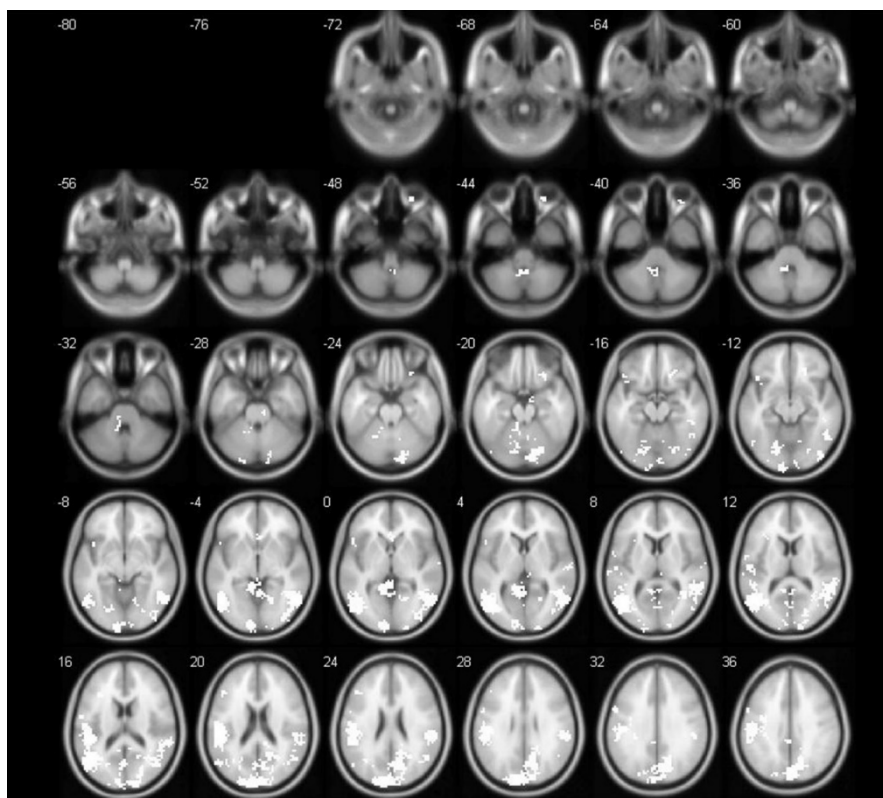
Figure 2. Clusters identified in the searchlight MVPA analysis

The clusters with significant classification accuracy were listed. A threshold of p<0.05 was adopted. For each cluster, only the region showing the minimum p value (maximum t value) is listed. MNI coordinates (x, y, z) are indicated.

4. Discussion

In the present study, we used machine learning algorithm to decode the fMRI data. Our results indicated that part of the brain regions were involved in the action perception. The regions largely located at the occipital and temporal cortex, which is consistent with a large number of previous work [6-8]. For review, the work by Lingnau and Downing [9] systematically clarified the function of lateral occipito-temporal cortex in action processing. Besides, part of the regions in frontal lobe was also identified. It is widely acknowledged that, the involvement of frontal and parietal areas in action perception is related to the mirror neuron theory, also referred to as the human mirror neuron system, MNS. A recent study found that PD patients are impaired in perception of human movements, which is considered to be directly related to the impaired abilities in motor execution. The results of the study support the hypothesis that motor system may play a causal role in visual movement perception [10]

Besides, to make the results more reliable, it is feasible to apply various classifiers in the data, such as Logistic Regression (LR) and Gaussian Naive Bayes (GNB). A comparison between the performances of different classifiers is also valuable. Moreover, to reveal the neural mechanisms of different cognitive activities, more novel methods could be applied in neuroimaging data.

Acknowledgments

References

[1] Pereira, F. and M.M. Botvinick, Machine learning classifiers and fMRI: A tutorial overview. Neuroimage, 2009. 45(1): p. S199-S209.

[2] Norman, K.A., et al., Beyond mind-reading: multi-voxel pattern analysis of fMRI data. Trends in Cognitive Sciences, 2006. 10(9): p. 424-430.

[3] Ma, F., et al., Investigating the neural basis of basic human movement perception using multi-voxel pattern analysis. Experimental Brain Research, 2018. 236(3): p. 1-12.

[4] Nikolaus, K., G. Rainer, and B. Peter, Information-based functional brain mapping. Proceedings of the National Academy of Sciences of the United States of America, 2006. 103(10): p. 3863-3868.

[5] Chang, C.-C. and C.-J. Lin, LIBSVM: A library for support vector machines. Vol. 2. 2007.

[6] Hafri, A., J.C. Trueswell, and R.A. Epstein, Neural representations of observed actions generalize across static and dynamic visual input. Journal of Neuroscience, 2017. 37(11): p. 3056-3071.

[7] Tucciarelli, R., et al., MEG Multivariate Analysis Reveals Early Abstract Action Representations in the Lateral Occipitotemporal Cortex. Journal of Neuroscience the Official Journal of the Society for Neuroscience, 2015. 35(49): p. 16034.

[8] Urgen, B.A., S. Pehlivan, and A.P. Saygin. Representational similarity of actions in the human brain. in International Workshop on Pattern Recognition in Neuroimaging. 2016.

[9] Lingnau, A. and P.E. Downing, The lateral occipitotemporal cortex in action. Trends in Cognitive Sciences, 2015. 19(5): p. 268-277.

Kloeters, S., et al., Impaired perception of human movements in Parkinson's disease. Behavioural Brain Research, 2017. 317: p. 88-94.