سیزدهمین کنفرانس مهندسی برق ایران
۲۲-۲۰ اردیبهشت ۱۳۸۴

Feature Space Analysis for Group Inference in fMRI Data

Hesamoddin Jahanian
Jahanian@ipm.ir
University of Tehran, Tehran, Iran
School of Cognitive Sciences, IPM, Tehran, Iran

Seyyed Mohammad Shams
msshams@ipm.ir
University of Tehran, Tehran, Iran
School of Cognitive Sciences, IPM, Tehran, Iran

Gholam Ali Hossein-Zadeh
ghzadeh@ut.ac.ir
University of Tehran, Tehran, Iran
School of Cognitive Sciences, IPM, Tehran, Iran

Hamid Soltanian-Zadeh
hamids@rad.hfh.edu
University of Tehran,Tehran, Iran
School of Cognitive Sciences, IPM, Tehran, Iran
Image Analysis Lab., Radiology Department, Henry Ford Health System, USA

Abstract: *Recently, fMRI multisubject analysis has emerged as a new topic to embrace the differences in brain response. State-of-the-art multisubject methods - like general linear model (GLM) - suffer from limited sensitivity. A novel method based on feature space fuzzy cluster analysis for fMRI group inference is introduced to overcome this limitation. In the proposed method, a brain tensor is obtained using the cross-correlation analysis of each subject. Then a feature space is constructed from the brain tensor. Fuzzy cluster analysis of the proposed feature space generates a map of "membership to active cluster". Statistical significance of the membership map is then assessed using randomization to derive the group inference activation map. This method is applied to experimental and simulated multi-subject fMRI data and results are compared to those of the GLM. We show that the proposed method detects more activated regions in analyzing experimental data and considerably more true positives (30-40%) at all false alarm rates in the simulation study. This means that the proposed method has higher detection sensitivity compared to GLM.*

Keywords: fMRI, group inference, feature space clustering, visual task.

1 Introduction

Functional magnetic resonance imaging (fMRI) allows evaluation of brain activity when a subject performs cognitive, sensory, or motor tasks, via measuring local variations of hemodynamic over time [1]. Conventional fMRI analyses methods analyze the data of every subject separately. But multisubject analysis of fMRI data deals about the differences or commonalities of brain response among different subjects. In single-subject analysis, it is assumed that each subject has a fixed (observed) activation and therefore discounts random variations from subject to subject [2] and only a scan-to-scan variance is considered. However, the variability in activation effects, from subject to subject, should be assessed. In multisubject fMRI studies, both variations are included and the resulting response reflects the pattern of response for a group of subjects. When the objective of the analysis is a case study, the *fixed-subject-effect* method is used. On the other hand, for deduction about the population, the *random-subject-effect* method is used. Some multi-level methods have been introduced for such analyses. These statistical model-based methods are based on General Linear Model (GLM) framework. In this framework, a statistical map is first derived for each subject. This process is a uni-variate analysis. Using the maps of all subjects, the "effects" and "standard errors" are then combined. Finally, the decision is made with the use of group

t-test [3] or similar statistics. The effects of similar voxels from different subjects are combined using a variety of approaches. In multi-level approaches, the effect and standard errors are merely used in group analysis level, resulting in a noticeable reduction in the sensitivity of analysis to time series [3].
Here, we propose a novel method based on feature space analysis of fMRI data to deal with this issue. Clustering method has been successfully used for single subject fMRI activation detection [4]. Here, by introducing feature space corresponding to the time series of each voxel along all subjects, we employ it for fMRI group analysis.

2 Materials

2.1 Experimental fMRI Data

A set of sensory-motor fMRI data is analyzed in this research. This set is provided by fMRI data center (http://www.fmridc.org). They were acquired during an event-related fMRI experiment in a 1.5 T scanner. During the experiments, 128 T2*-weighted volume images were acquired using asymmetric spin echo pulse sequence. Each volume image consisted of 16 slices and each slice was composed of 64×64 pixels. A set of anatomical images was also acquired from each subject, which consists of 128 sagital slices with 256×256 pixels. Eleven young non-demented subjects were selected from these data. Their functional images were motion corrected using the AFNI software package (Medical College of Wisconsin, Milwaukee, WI) [5]. Then their anatomical images were transferred to the standard space of Talairach and Tournoux and the resulted transform is used for spatial normalization of functional images in the AFNI software package. The anatomical images were used to localize the active regions in the AFNI software. For each volume of functional data, the sub sampling process produced a volume image with 54×64×50 voxels and voxel size of 3×3×3 mm. Linear drift and the mean component were removed from time series of each voxel.

2.2 Simulated fMRI Data

Two groups of simulated data were used in this study. The first group consists of 11 sets of simulated rest data, which is used for obtaining the histograms of parameters under null hypothesis (no activation in the group). The second group consists of 11 sets of simulated data that have some predefined active voxels.
Each set of the simulated data contains 172,800 time series with 124 points. For simulating the rest data (group 1), Gaussian noise was used with the mean value similar to real data sets, and variance of 2% of the mean value. In the second group, in addition to the Gaussian noise, activation was added to some voxels according to the spatial pattern depicted in Fig. 1. The contrasts of the activation regions varied as 1%, 1.25%, 1.5%, and 1.75% horizontally and the noise variances were selected randomly in the interval [4 10]. Simulated activation time-series consisted of 124 points, which have been obtained through convolving the stimulation pattern with the hemodynamic response function (HRF) and then adjusting the amplitude of the resultant signal to the desired contrast. The stimulation pattern was the same as that of the experimental data. The HRF was modeled according to the following Gamma function:

$$h(t;\tau,\sigma)=\begin{cases} e^{-t/\sqrt{\sigma.\tau}}\left(\frac{e.t}{\tau}\right)^{\sqrt{\tau/\sigma}} & t>0 \\ 0 & t<0 \end{cases} \qquad (1)$$

where τ shows the location of the peak and σ is related to the width of the peak [6]. In order to model HRF variations, parameters τ and σ were selected randomly within intervals [3 7] and [0.05 0.21], respectively. This process was applied to each voxel.

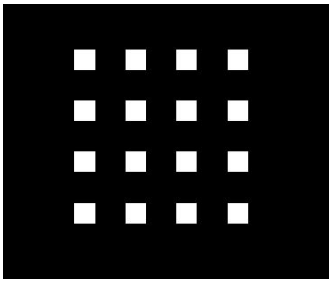


Fig. 1. The spatial pattern of activation in the simulated data.

3 Methods

The proposed method consists of the following 5 steps:

1) A Brain tensor is extracted for each brain voxel whose elements are the cross-correlation of that voxel time-series with the reference signal in different subjects (described in Section 3.1).
2) An overall feature space is obtained from the brain tensor (described in Section 3.2).

3) The Fuzzy c-means (FCM) algorithm is then applied to the overall feature space and the map of "membership degree to active cluster" is obtained (described in Section 3.3).
4) The statistical parametric map corresponding to the "membership degree" (obtained in previous step) is derived (described in Section 3.4).
5) The statistical parametric map is thresholded at the desired confidence level (p-value) (described in Section 3.5).

Figure 2 summarizes the proposed method.

3.1 Obtaining Brain Tensor

Here, we have used the cross correlation of each time-series and the reference signal, as the feature of that time-series. The reference signal is produced by convolving the stimulation pattern with the gamma function (1). In this step, for each voxel of the brain, we produce *n* features corresponding to *n* subjects. In other words, we have a tensor of size *n* for each voxel, where each element comes from one of the subjects.

3.2 Feature Space Extraction

We cannot apply the clustering algorithm directly to the brain tensor because each tensor element has the same importance as the others. But clustering algorithm assumes different importance for each tensor element because they would have different directions in brain tensor feature space. For example activation in subject 1 points to a direction in a feature space which is different from that of subject 6. Therefore in spite of their same importance they will be put in different clusters and this will mislead the algorithm in grouping the time series properly. Considering this fact, we should combine the obtained tensor elements so that each feature has the same contribution in the overall feature space.

We have used L2 norm to combine the elements of each tensor because it has the aforementioned properties. Thus, in this step, we obtain a feature space whose elements are the L2 norms of the tensor elements obtained in Section 3.1.

3.3 Fuzzy Clustering

The well known fuzzy C-means clustering algorithm is then applied to the features obtained in the previous step. In this paper, we use the cluster validity measure proposed by Fadili *et al.* in [7]. We also used the fuzziness index m=2 as

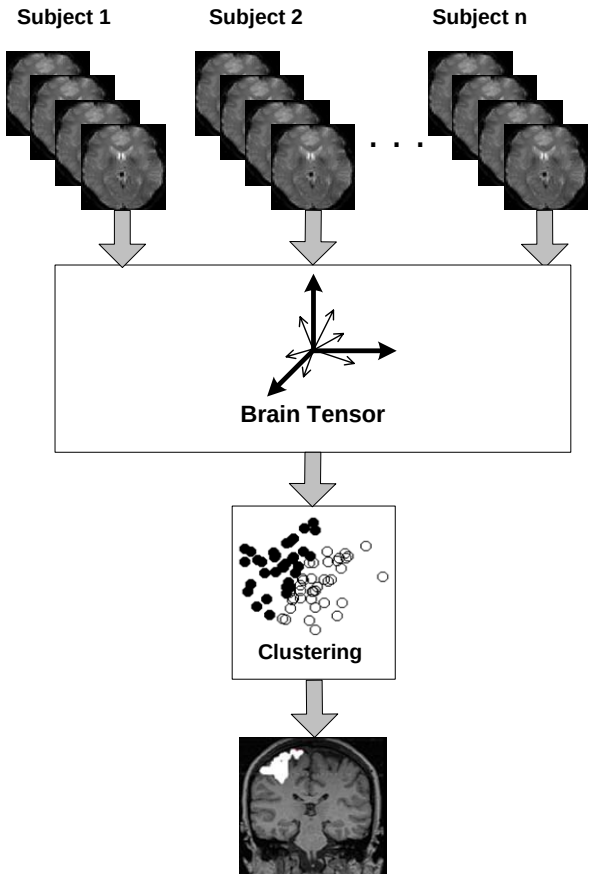


Fig. 2. Summary of the proposed method.

suggested in [7]. After FCM converges, the cluster with the most similar centroid to the reference pattern is chosen as the active cluster and the membership degrees to this cluster provide us with the map of "membership to active cluster". In the case of suggested feature space, this is the cluster with a centroid that has maximum value of L2 norm.

3.4 Statistical Significance Assessment

To assess the statistical significance of the obtained membership degree map, we need to have the pdf $f_u(u|H_0)$. In this paper, we use the method proposed by Jahanian, *et al.* [4] to find the statistical significances corresponding to active membership degrees. In this method, an empirical histogram is constructed using randomization. This histogram estimates the required pdf $f_u(u|H_0)$. Using this histogram, one finds a threshold corresponding to the desired p-value.

3.5 Obtaining Statistical Parametric Map

Thresholding the "membership to active cluster" map of brain voxels with the threshold obtained in previous step generates statistically meaningful

results. Especially, this step is needed when we intend to compare detection accuracy of different methods.

4 Experimental Results

Simulated and experimental fMRI data sets, described in Section 2, have been used to evaluate the proposed method. The proposed method was compared to GLM multi-subject analysis [3]. The comparison study should be done at the same false alarm rates. We computed the equivalent thresholds for the proposed method as described in Section 3.4. Since the proposed method is a single level analysis and GLM is a two level analysis method, we used the thresholds obtained from the simulated data to achieve similar conditions for both methods. Analysis of rest data produces realizations of its parameter under the null hypothesis, which can be used to construct an empirical histogram. This histogram is used for computing the thresholds of the GLM parameter for different false alarm rates.

The methods were applied to both simulated and experimental fMRI data sets. Fig. 3 shows the number of true positives in the simulated data set at different false alarm rates in the interval $\alpha \in [0.0001\ 0.001]$. Fig. 4 shows the activated regions detected by both methods at the false alarm rate of 0.0001. Results show that the proposed method provides improved detection sensitivity over the GLM method.

Figs. 5, 6 show activated regions in the experimental data set, superimposed on the anatomical MRI images. Single voxels were removed from the activation maps. Table 1 lists the activated regions detected by the proposed method and GLM. As shown in Table 1 and Figs. 5, 6, both methods detected activation in Occipital Cortex (BA 17, 18), Thalamus, and Precuneus which were also reported in previous fMRI investigations of visual task [8-11].

The proposed method also detected activations in Cerebellum, Superior Temporal Gyrus, Cingulate Gyrus (BA 24), Precentral Gyrus, Superior Frontal Gyrus and Cuenus, where GLM method did not detect any activation (see Fig. 5).

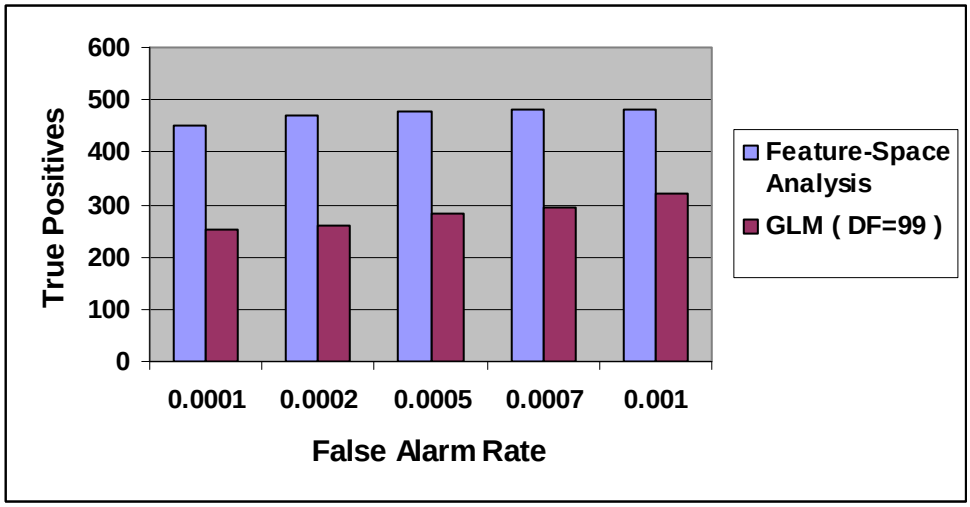
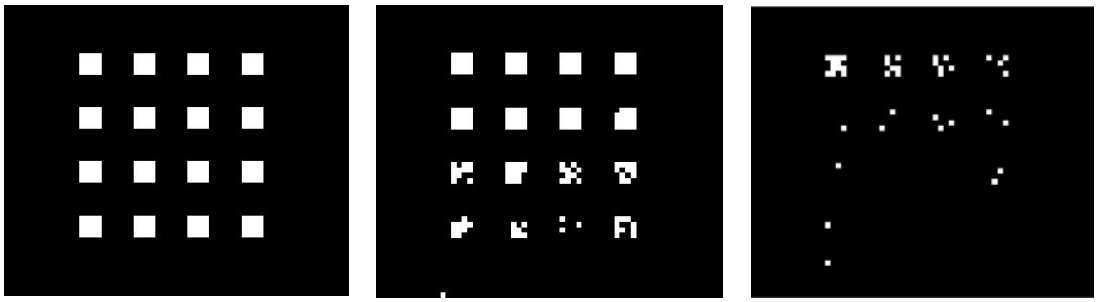


Fig. 3. True detected active voxels (true positives) in the simulated data for GLRT and GLM methods.



(a) (b) (c)

Fig. 4. The spatial pattern of activation in simulated data (a), Activated areas detected by feature space analysis (b), and GLM method (c) at false alarm rate of 0.0001.

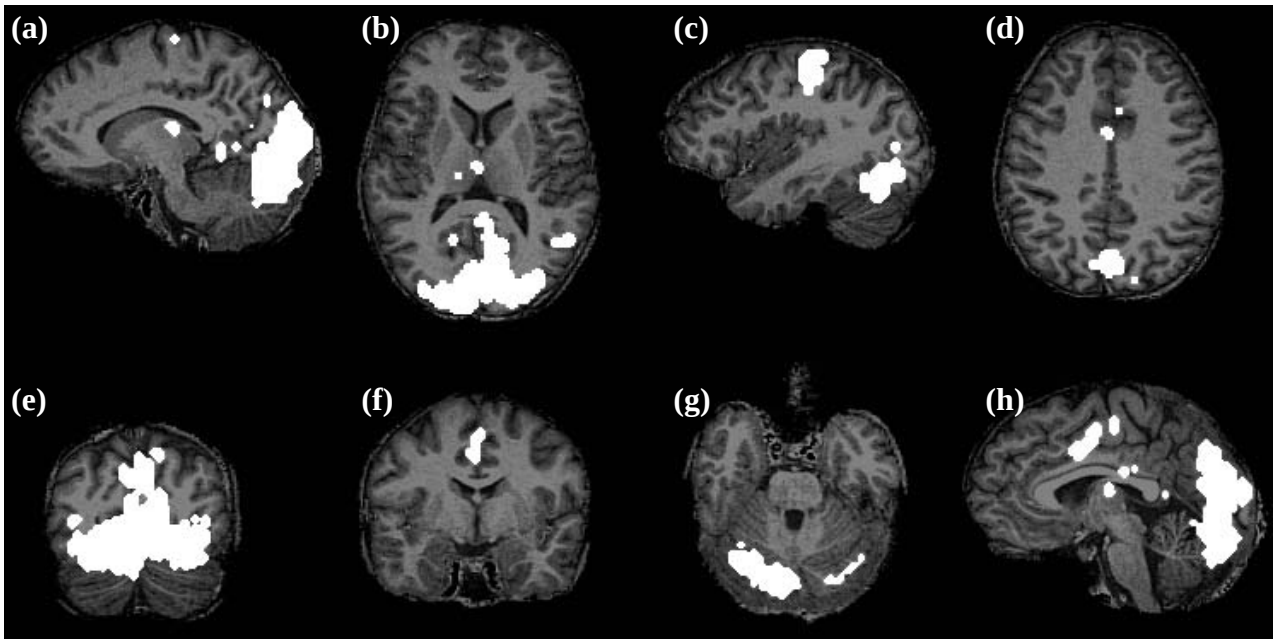


Fig. 5. Different views of brain activated areas, detected by applying feature space analysis on the visual task group data. The activated areas are overlaid on the high resolution anatomical images. Activation is detected in: a) Occipital Cortex and Superior Frontal Gyrus; b) Superior Temporal Gyrus and Thalamus; c) Precentral Gyrus and Cerebellum; d) Precuneus; e) Cuneus; f) Cingulated Gyrus; g) Cerebellum; h) Occipital Cortex, Thalamus and Cingulate Gyrus.

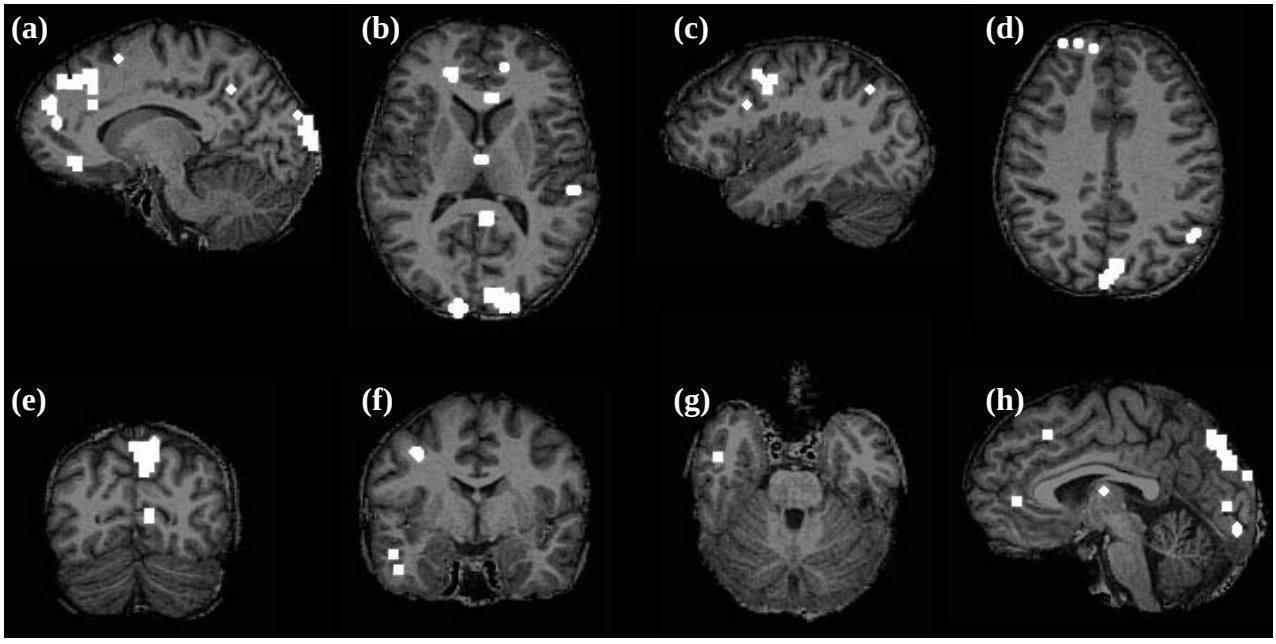


Fig. 6. Activated areas detected by GLM, in different views of brain (corresponding to the views of Fig. 5). The activated areas are overlaid on the high resolution structural images. Images show areas in: a) Occipital Cortex (*detected*) and Superior Frontal Gyrus (*not detected*) ; b) Superior Temporal Gyrus (*not detected*) and Thalamus (*detected*); c) Precentral Gyrus (*not detected*) and Cerebellum (*not detected*); d) Precuneus (*detected*); e) Cuneus (*not detected*); f) Cingulated Gyrus (*not detected*); g) Cerebellum (*not detected*); h) Occipital Cortex (*detected*), Thalamus (*detected*) and Cingulate Gyrus (*not detected*).

Table 1: Activated areas detected by two analysis methods.

Activated Area	GLM	Feature Space Analysis
Occipital Cortex	✔	✔
Thalamus	✔	✔
Precuneus	✔	✔
Precentral Gyrus	--	✔
Superior Frontal Gyrus	--	✔
Cerebellum	--	✔
Superior Temporal Gyrus	--	✔
Cingulate gyrus	--	✔
Cuenus	--	✔

5 Discussion and Conclusion

A novel method, based on feature space clustering for activation detection in group fMRI data, is presented. This method is applied to simulated and experimental fMRI data.

The use of cross-correlation for obtaining brain tensor benefits the method with decreasing the size of data and speeding up next analysis stages. In addition, clustering the raw fMRI time series may lead to stability problems and the risk of clustering on noise rather than on the activation because of the poor SNR of fMRI signal.

The comparison between the results of the proposed method and GLM on the simulated data shows that the proposed method detects more active voxels (about 30-40%) at all false alarm rates. It means that the proposed method has considerably improved detection sensitivity over GLM.

Also, the proposed method detects more significant activated regions in experimental data compared to the GLM method. The proposed method detected activated regions in Cerebellum, Superior Temporal Gyrus, Cingulate Gyrus (BA 24), Precentral Gyrus, Superior Frontal Gyrus and Cuenus, where GLM method failed to detect them.

References

[1] S. Ogawa, R. Menon, D. Tank, et al. "Functional brain mapping by blood oxygenation level-dependent contrast magnetic resonance imaging. A comparisio of signal characteristics with a biophysical model," *Biophys J*, vol. 64, pp. 803-812, 1993.

[2] K. J. Friston, A. P. Holmes, C. J. Proce, C. Buchel and K. J. Worsely, "Multisubject fMRI studies and conjuction analyses," *NeuroImage*, vol. 10, pp. 385-396, 1999.

[3] J. Worsley, H. Liao, J. Aston, V. Petre, H. Duncan, F. Morales, and C. Evans, "A general statistical analysis for fMRI data," *NeuroImage*, vol. 15, pp. 1–15, 2002.

[4] H. Jahanian, G. A. Hossein-Zadeh, H. Soltanian-Zadeh, B. A. Ardekani, "Controlling the False Positive Rate In Fuzzy Clustering Using Randomization: Application to fMRI Activation Detection," *Magnetic Resonance Imaging,* Vol. 22, pp. 631-638, 2004.

[5] R. Cox, J. Hyde, "Software tools for analysis and visualization of fMRI data," *NMR Biomed*, vol 10, pp. 171-178, 1997.

[6] G. A. Hossain-zadeh, B. Ardekany, and H. Soltanian-zadeh, **"**A signal subspace approach for modeling the hemodynamic response function in fMRI," *Magnetic Resonance Imaging,* vol. 21, pp. 843-853, 2003.

[7] M. Fadili, S. Ruan, D. Bloyet, "On the number of clusters and the fuzziness index for unsupervised FCA application to BOLD fMRI time series," *Medical Image Analysis*, vol. 5(2), pp. 55-67, 2001.

[8] M. S. Beauchamp, L. Petit, T. N. Ellmoro, J. Injeholm, and V. oxby, "A parametric fMRI for overt and covert shifts of visual spatial attentions," *NeuroImage*, vol. 14, pp. 310-321, 2001.

[9] L. Pesoa, E. Gutierrez, and P. A. Bordehini, "Neural correlates of visual working memory: fMRI amplitude predicts task performance," *Neuron*, vol. 35, pp. 975-987, 2002.

[10] B. J. Casey, et al, "Reproducing of fMRI results across for institutions using spatial working memory task," *NeuroImage*, vol. 8, pp. 249-261, 1998.

[11] D. Ress, B. T. Backus, and D. J. Heeger, "Activity in primary visual cortex predicts performance in a visual detection task," *Nature*, vol. 3, no. 9, pp. 940-945, 2000.