A Self-Supervised Voxel Shuffling Framework for Kernel-Based fMRI activation detection

Poster No:

2475

Submission Type:

Abstract Submission

Authors:

Chendi Han[1], Zhengshi Yang[1], Xiaowei Zhuang[1], Dietmar Cordes[1]

Institutions:

[1]Cleveland Clinic, Las Vegas, NV

First Author:

Chendi Han
Cleveland Clinic
Las Vegas, NV

Co-Author(s):

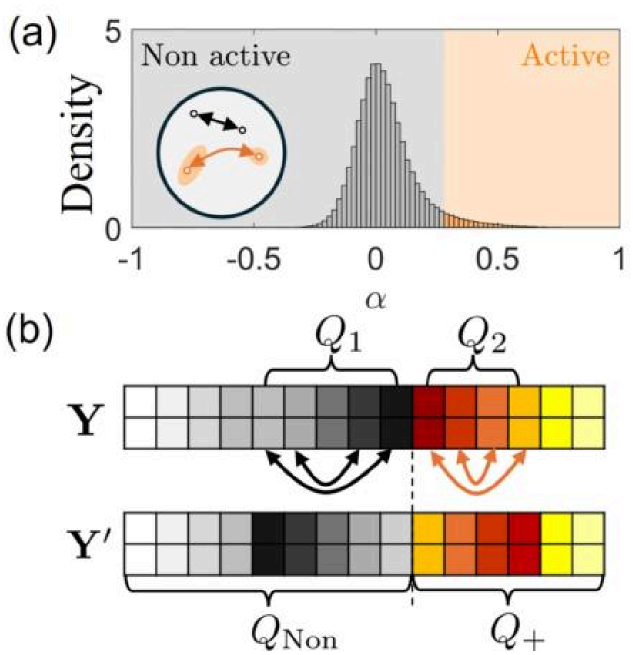
Zhengshi Yang
Cleveland Clinic
Las Vegas, NV
Xiaowei Zhuang
Cleveland Clinic
Las Vegas, NV
Dietmar Cordes
Cleveland Clinic
Las Vegas, NV

Introduction:

Kernel-based methods are powerful for dimension reduction in fMRI analysis. Previous methods used supervised optimization (Xifra-Porxas, 2021) or an additional dataset serving as negative pairs (Yang, 2018) to determine the unknown kernel mapping, which is less efficient and hard to expand to real datasets when there is no ground truth. Inspired by self-supervised learning in image processing (Chen, 2020), we propose a data augmentation method that could generate supervisory examples by voxel shuffling. By maximizing the similarity before and after augmentation, we show our method could reduce overfitting and increase accuracy even for complicated kernels. The results are validated using real fMRI datasets from two different sources for the activation detection problem.

Methods:

Suppose fMRI data with Y∈ R^{T*Q}, where T indicates the time and Q is the total number of voxels, given a certain design signal X_eff ∈ R^{T*1}. Kernel canonical correlation analysis (KCCA) solves this problem by mapping the original data to kernel space and maximizing the correlation (Hardoon, 2004). This process could involve unknown functions or regularization parameters. For example, the linear kernel contains one additional regularization parameter γ to reduce overfitting (Yang, 2018). Our goal is to find the augmented data Y', which could serve to determine the unknowns.
Figure 1 shows our proposed algorithm. Starting from arbitrary kernel mapping and performing KCCA, we can divide the output α into activated and non-activated clusters; we use Q_+=0.1Q and Q_Non=0.9Q to indicate the number of voxels in each cluster. Our shuffling algorithm involves switching the voxel locations inside each cluster along the decision boundary. Specifically, we select Q_1 voxels in the non-activated cluster and Q_2 voxels in the activated cluster. Within the Q_1 voxels, the voxel with the highest α is changed to the lowest α, and the same for the second highest/lowest voxels. The same shuffling is repeated for Q_2. The detailed choice for Q_1 and Q_2 is shown in Figure 1. Practically, we use two shufflings and take the average. The data after shuffling is indicated by Y', which is then input to the KCCA to get another activation pattern α'. The similarity is measured by treating α as ground truth, then using different thresholds to compute the Receiver Operating Characteristic (ROC) curve, with area under the curve (AUC) with False Positive Rate (FPR) smaller than 0.1 defined as similarity. We implement the surrogate optimization algorithm in MATLAB to maximize it.



(c)

Shuffling 1	$Q_1 = Q_2 = 0.5Q_+$
Shuffling 2	$Q_1 = Q_{\mathrm{Non}},\ Q_2 = Q_+$

Figure 1: Illustration of the voxel shuffling method. (a) Distribution of the activation $\boldsymbol{\alpha}$. Results do not change significantly for different data or kernels. Based on the output $\boldsymbol{\alpha}$ the top 10% of voxels are set to be activated. (b) Two clusters with numbers of voxels Q_1 and Q_2 are chosen around the boundary. Voxels inside each cluster are shuffled based on their $\boldsymbol{\alpha}$ value. (c) The choice for the Q_1 and Q_2 used in the paper.

Results:

We use two datasets to test the method: the HCP dataset with 87 subjects performing a working memory task (Barch, 2018), and our in-house scans with 16 subjects performing both faces and pictures episodic memory tasks (Jin, 2012). In the HCP dataset, the design signal is arranged with contrasts targets minus non-targets. For our in-house scans for each subject and task, the contrasts encoding minus control and recognition minus control effectively give us 64 different combinations.

All the data are minimally preprocessed using the SPM12 package (J. Ashburner, 2008), normalized to the MNI atlas (F. F. Glasser, 2013), with Gaussian smoothing FWHM=4 mm and 7 spatial orientation filters for kernel methods with the same smoothing level (Yang, 2018). We set the criteria to be Gray Matter (GM) overlapping. Using gray matter as ground truth, different thresholds for α are used to evaluate the ROC curve. The results for two different datasets are shown in Figure 2 (a) and (b). Relationship between shuffling robustness and gray matter overlapping is observed.

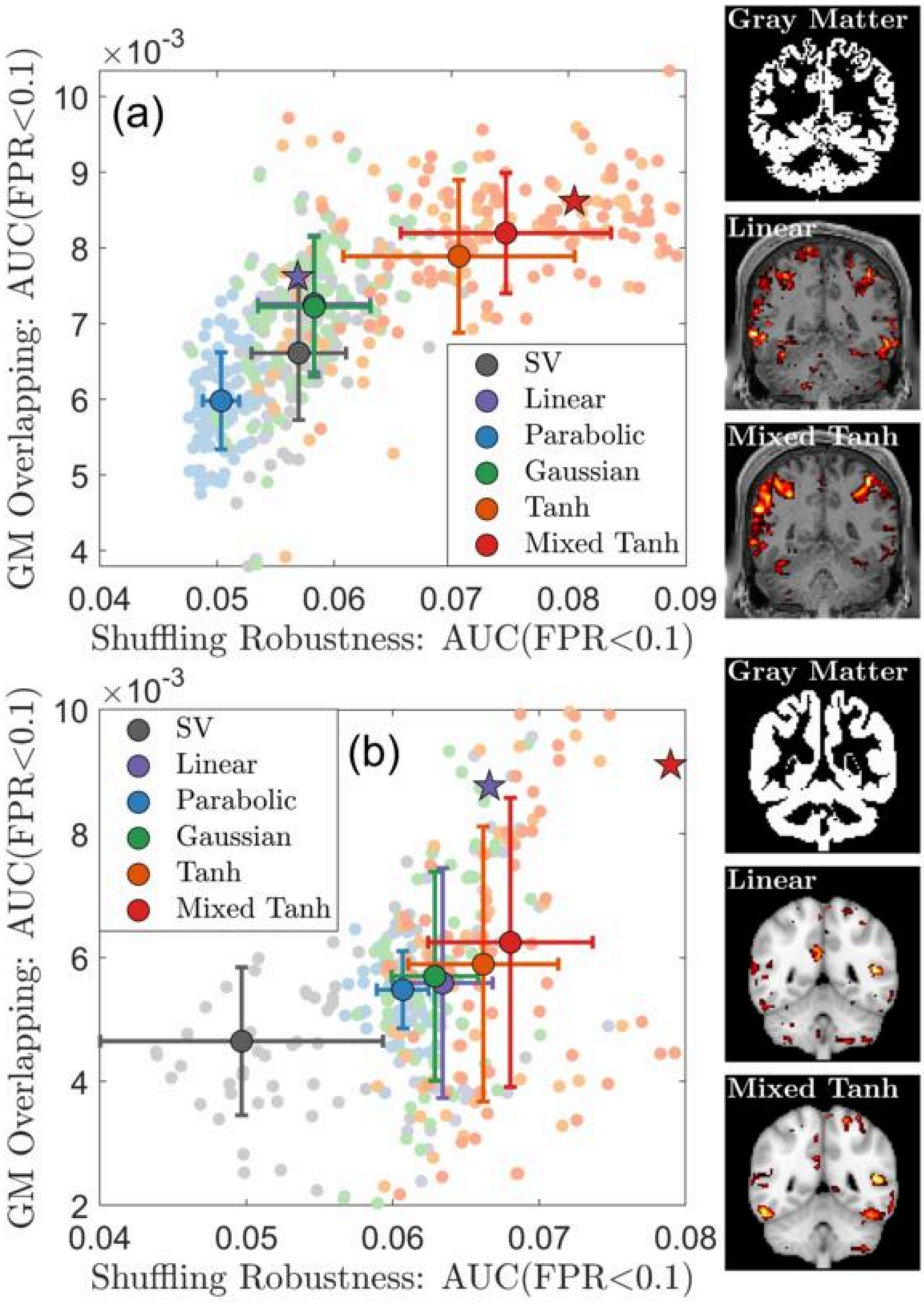


Figure 2: The relationship between shuffling robustness and gray matter overlapping for task fMRI. Each dot represents one subject analyzed using the specified methods. Panels (a) and (b) correspond to the HCP dataset and in-house scans, respectively. In each dataset, one subject was selected (indicated by a star), and the corresponding activation patterns are displayed on the right. Compared with SV, the mixed hyperbolic tangent kernel increased the AUC by 25.74% and 37.14% for HCP and in-house scans, respectively.

97.14% for HCP and in-house scans, respectively.

Conclusions:

The main finding of the study is that we have found a robust way to do data augmentation. We find that by maximizing the similarities before and after augmentation, we can determine the best kernel mapping without using the spatial ground truth or additional datasets.

This study was funded by NIH-R01AG071566-02 and NIH-P20GM109025-08.

Modeling and Analysis Methods:

Activation (eg. BOLD task-fMRI) [1]
Methods Development

Novel Imaging Acquisition Methods:

BOLD fMRI [2]

Keywords:

FUNCTIONAL MRI

[1|2]Indicates the priority used for review

Abstract Information

By submitting your proposal, you grant permission for the Organization for Human Brain Mapping (OHBM) to distribute your work in any format, including video, audio print and electronic text through OHBM OnDemand, social media channels, the OHBM website, or other electronic publications and media.

I accept

The Open Science Special Interest Group (OSSIG) is introducing a **reproducibility challenge** for OHBM 2025. This new initiative aims to enhance the reproducibility of scientific results and foster collaborations between labs. Teams will consist of a “source” party and a “reproducing” party, and will be evaluated on the success of their replication, the openness of the source work, and additional deliverables. Click here for more information. Propose your OHBM abstract(s) as source work for future OHBM meetings by selecting one of the following options:

I do not want to participate in the reproducibility challenge.

OHBM is now accepting Abstract Submissions for Software Demonstrations in addition to traditional posters. Please indicate below if you would like to submit for a Software Demonstration. There are a limited number of Software Demonstrations available so this may result in your abstract being accepted without the demonstration portion. My abstract is being submitted as a Software Demonstration.

No

Would you accept an oral presentation if your abstract is selected for an oral session?

Yes

Please indicate below if your study was a "resting state" or "task-activation" study.

Task-activation

Healthy subjects only or patients (note that patient studies may also involve healthy subjects):

Patients

Was this research conducted in the United States?

Yes

Are you Internal Review Board (IRB) certified? Please note: Failure to have IRB, if applicable will lead to automatic rejection of abstract.

Yes, I have IRB or AUCC approval

Were any human subjects research approved by the relevant Institutional Review Board or ethics panel? NOTE: Any human subjects studies without IRB approval will be automatically rejected.

Yes

Were any animal research approved by the relevant IACUC or other animal research panel? NOTE: Any animal studies without IACUC approval will be automatically rejected.

Not applicable

Please indicate which methods were used in your research:

Functional MRI

For human MRI, what field strength scanner do you use?

3.0T

Provide references using APA citation style.

Xifra-Porxas, Alba, et al. (2021). Estimating brain age from structural MRI and MEG data: Insights from dimensionality reduction techniques. NeuroImage, 231: p. 117822.
Yang, Z., et al. (2018). 3D spatially-adaptive canonical correlation analysis: Local and global methods. Neuroimage, 169: p. 240-255.
Chen, Ting, et al. (2020). A simple framework for contrastive learning of visual representations. International conference on machine learning. PMLR,p. 1597-1607.
Hardoon, D.R., et al. (2004). Canonical correlation analysis: An overview with application to learning methods. Neural computation. 16: p. 2639-2664.
Barch, D.M., et al. (2013) Function in the human connectome: task-fMRI and individual differences in behavior. Neuroimage, 80: p. 169-189.
Jin, Mingwu, et al. (2012). A preliminary study of functional abnormalities in aMCI subjects during different episodic memory tasks. Magnetic Resonance Imaging. 2012. 30.4: p. 459-470.
J. Ashburner, (2008). Computational anatomy with the spm software, Magnetic resonance imaging. 27: p. 1163–1174

M. F. Glasser et al. (2013) The minimal preprocessing pipelines for the human connectome project. Neuroimage, 80: p. 105–124
Yang, Zhengshi, et al. (2020). A robust deep neural network for denoising task-based fMRI data: An application to working memory and episodic memory. Medical Image Analysis. 60: p. 101622.
Cordes, Dietmar, et al. (2012). Optimizing the performance of local canonical correlation analysis in fMRI using spatial constraints. Human brain mapping. 33.11: p. 2611-2626.

UNESCO Institute of Statistics and World Bank Waiver Form

I attest that I currently live, work, or study in a country on the UNESCO Institute of Statistics and World Bank List of Low and Middle Income Countries list provided.

No