Slide 1

Kernel Canonical Correlation Analysis (KCCA) is commonly used as a global method for fMRI activation analysis. Although KCCA has the advantage of characterizing global relationships in a single step, it also increases the risk of overfitting. As a result, current methods usually rely on additional datasets for supervision to determine hyperparameters. In this presentation, we propose a self-supervised learning algorithm that can detect global relationships without requiring additional datasets.

Slide 2

We declare that there is no conflict of interest. Part of the data used in this presentation is publicly accessible. All studies involving patient information are properly protected. This work is supported by the NIH.

Slide 3

Here we summarize different correlation-based methods for task-fMRI activation analysis. The general linear model with Gaussian smoothing is commonly performed on each voxel, but spatial smoothing reduces specificity and may cause spatial blurring. More advanced methods, such as local canonical correlation analysis, address this by introducing different spatially oriented filters to avoid spatial blurring. To incorporate nonlinear relationships and adaptive spatial smoothing, CCA with spatial constraints or deep CCA has been proposed. These methods all fall under the category of local methods, as correlations are computed between each voxel and its neighborhood. In contrast, Kernel Canonical Correlation Analysis can detect global relationships. The revised version using nonlinear kernels further extends activation detection power to more general types of relationships between the task-fMRI signal and the target signal.

Slide 4

In the upper row, we show the steps for KCCA. Both fMRI signals and the design signal are mapped to the kernel space using linear or nonlinear kernels. After establishing correlation in the kernel space, a back-reconstruction algorithm is used to extract activation patterns. In the lower row, we show some commonly used kernels. Mathematically, there is no restriction on the choice of kernel. This raises a general question of how to select the nonlinear mapping in KCCA. Due to the high dimensionality of fMRI signals, KCCA will always identify relationships in the kernel space. Properly defining the hyperparameters can increase the power of KCCA and help avoid overfitting.

Slide 5

To determine the kernel mapping, the traditional approach is to compare KCCA results on task-fMRI and resting-state fMRI signals. We find that this method is not sufficient to prevent overfitting as the kernel becomes nonlinear. Instead of introducing another dataset, we propose a self-supervised learning framework. The assumption is that a good method will maintain prediction accuracy after a properly designed shuffling algorithm. In machine learning, this approach is considered self-supervised learning, as the data itself is used to generate augmented samples. Specifically, starting from fMRI data labeled with Y and a specific kernel with certain hyperparameters, we perform KCCA to obtain the activation pattern (alpha). After voxel shuffling, we obtain the augmented data Y′, and then use the same kernel and hyperparameters to obtain another activation pattern (alpha′). The similarity between alpha and alpha′ is then evaluated. For voxel shuffling, we aim to perform data augmentation that preserves spatial correlation. In fMRI activation analysis, activated voxels tend to cluster together, as do non-activated voxels. We select voxels near the decision boundary—specifically, selecting Q1 voxels from the non-activated region and Q2 from the activated region, and then for each of them, reversing voxel order based on the alpha values. In practice, we perform multiple shufflings and average the results to obtain a stable outcome.

Slide 6

Once the data augmentation method is defined, we establish our objective function to maximize shuffling robustness. After identifying the optimal hyperparameters, we test the activation pattern using an additional metric. In previous studies, this metric is commonly the area under the curve (AUC). The difference here is that for simulated data, we compare the activation pattern to the ground truth, while for task-fMRI, we calculate how many activated voxels overlap with gray matter. The key assumption in this study is that activation patterns more robust to shuffling are also more accurate.

Slide 7

We first implemented our algorithm on two simulated datasets. Both datasets use resting-state data as noise, with signals added to specific regions. We then applied KCCA to obtain the activation pattern with maximum shuffling robustness. We observe that the relationship between shuffling robustness (x-axis) and accuracy (y-axis) generally forms a linear trend. Bounded kernels such as the hyperbolic tangent and mixed hyperbolic tangent show better results compared to unbounded kernels.

Slide 8

Next, we examine the same relationship in task-fMRI data, where ground truth is lacking. A similar relationship appears. The only difference is that for task-fMRI data, we use the overlap between activated voxels and gray matter as the evaluation metric.

Slide 9

To conclude, in contrast to previously used hyperparameter optimization algorithms in linear KCCA analysis, we find that voxel shuffling algorithms can optimize hyperparameters for more general types of linear and nonlinear kernels. Furthermore, this method does not require additional resting-state fMRI data for supervision and relies purely on shuffling robustness, making it broadly applicable to a variety of tasks.

We thank the NIH for research support.

Thank you for your attention.