Nanjing University of Aeronautics and Astronautics
College of Computer Science and Technology

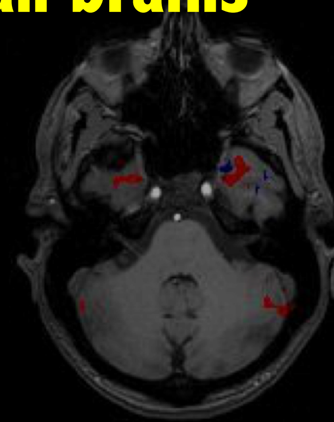
iBRAIN

Multi-Region Neural Representation

A novel model for decoding visual stimuli in human brains

Muhammad Yousefnezhad, Daoqiang Zhang

SIAM International Conference on Data Mining (SDM17)

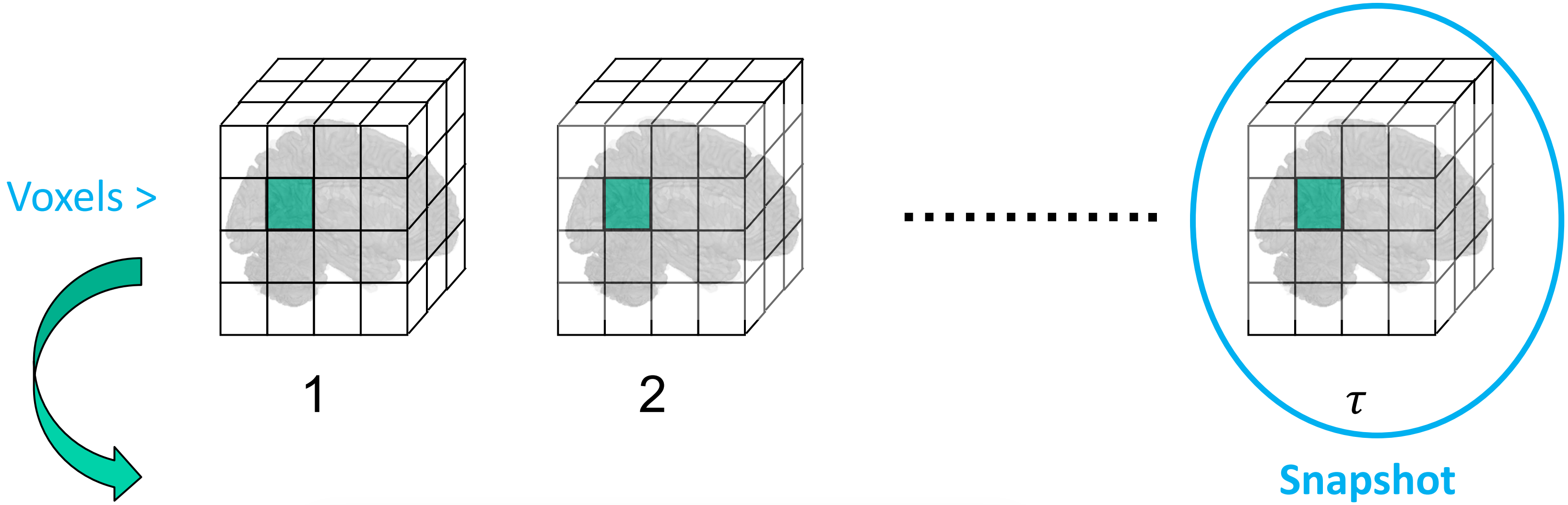


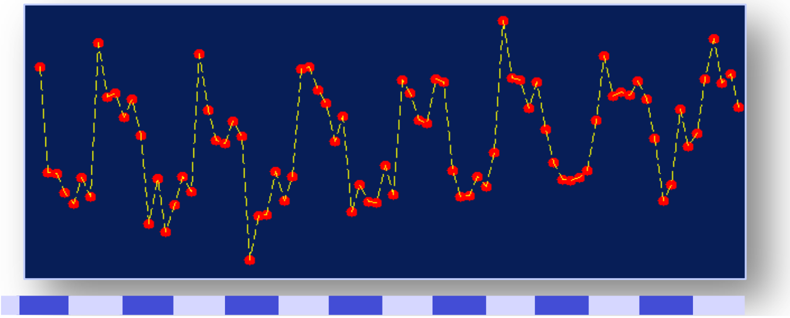
Presented by

Muhammad Yousefnezhad

2017 / March

Motivation





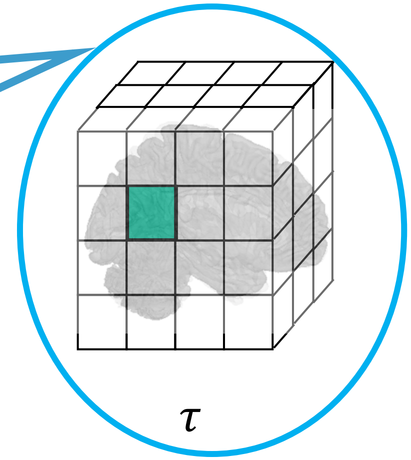
Blood Oxygen Level Dependent (BOLD) signals

Motivation

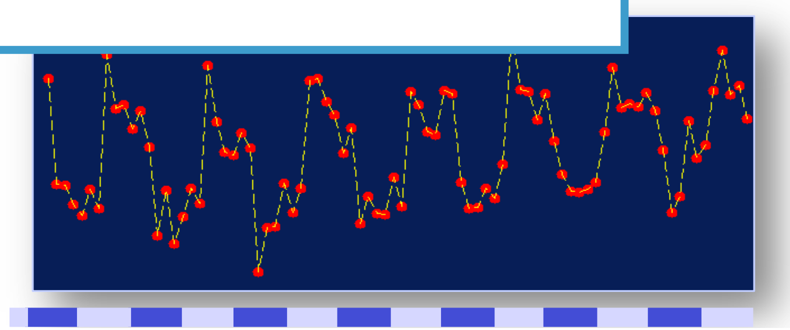
1. Selecting a set of **effective** snapshots rather than using whole of the **noisy and sparse** time series

2. Extracting **robust** features from the selected snapshots

3. Improving the **performance** of the generated cognitive model



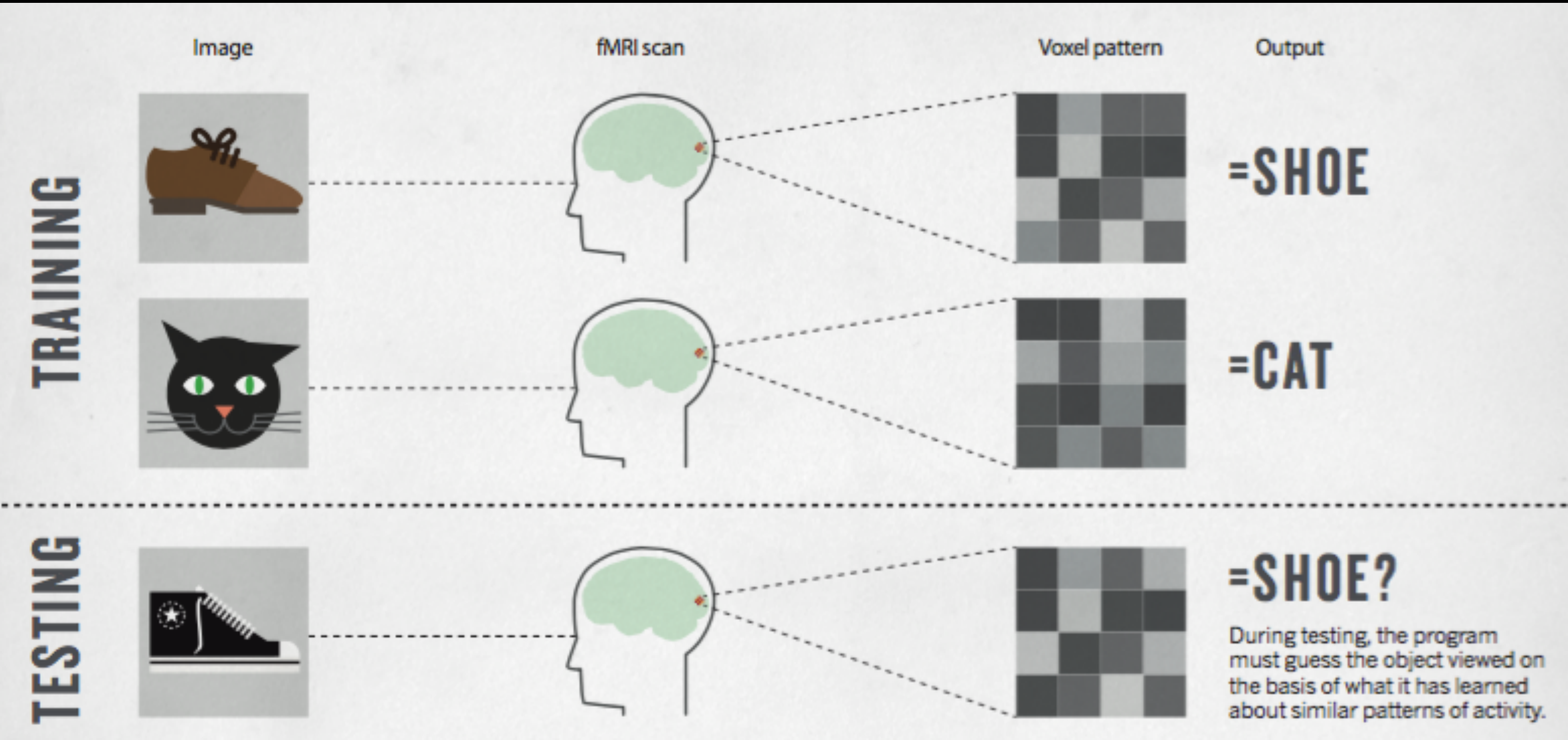
Snapshot



Blood Oxygen Level Dependent (BOLD) signals

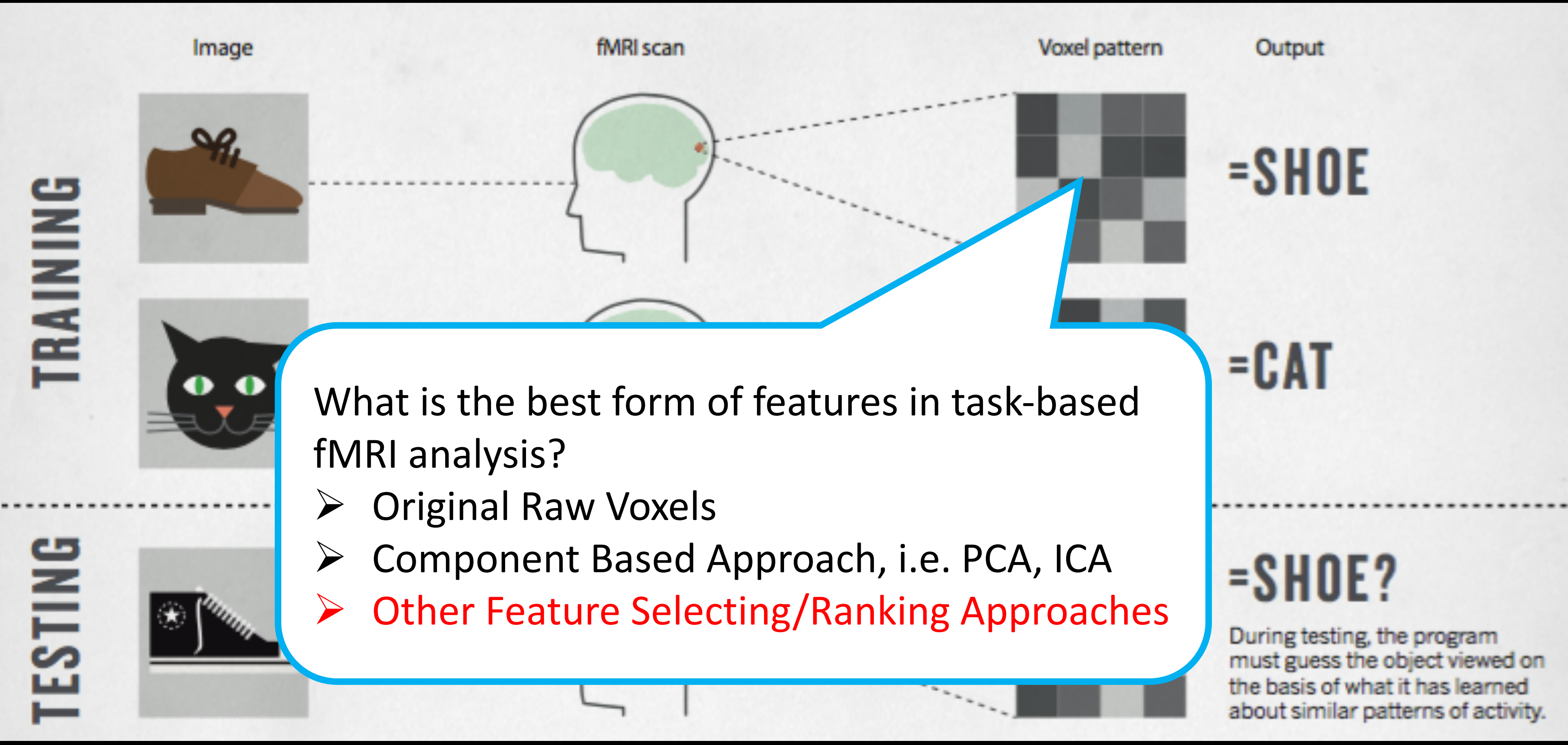
Basic Concepts

The Human Brain Decoding



Smith, Nature, 2013

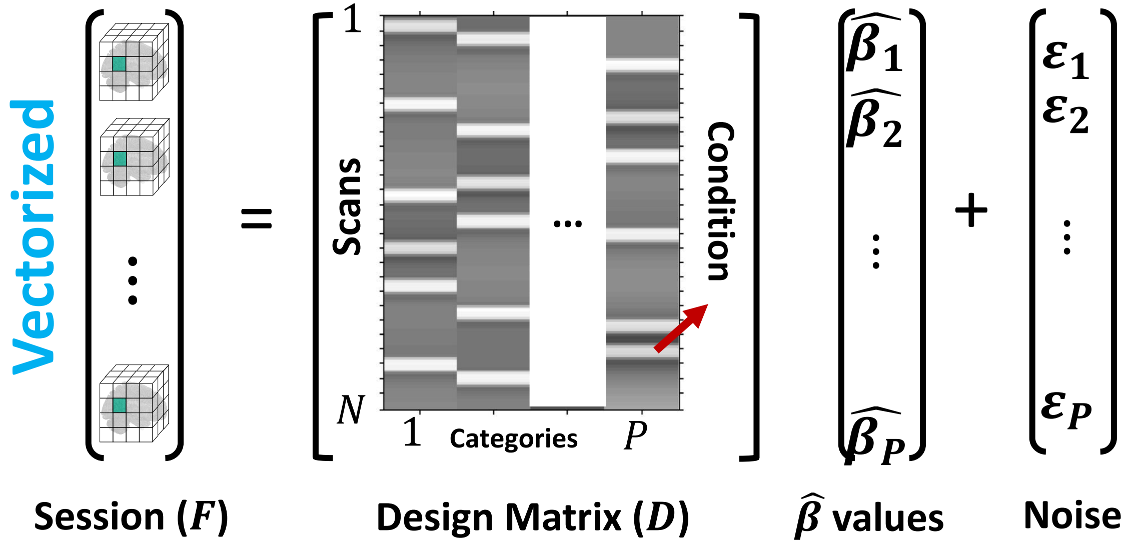
The Human Brain Decoding



What is the best form of features in task-based fMRI analysis?

- Original Raw Voxels
- Component Based Approach, i.e. PCA, ICA
- Other Feature Selecting/Ranking Approaches

Smith, Nature, 2013



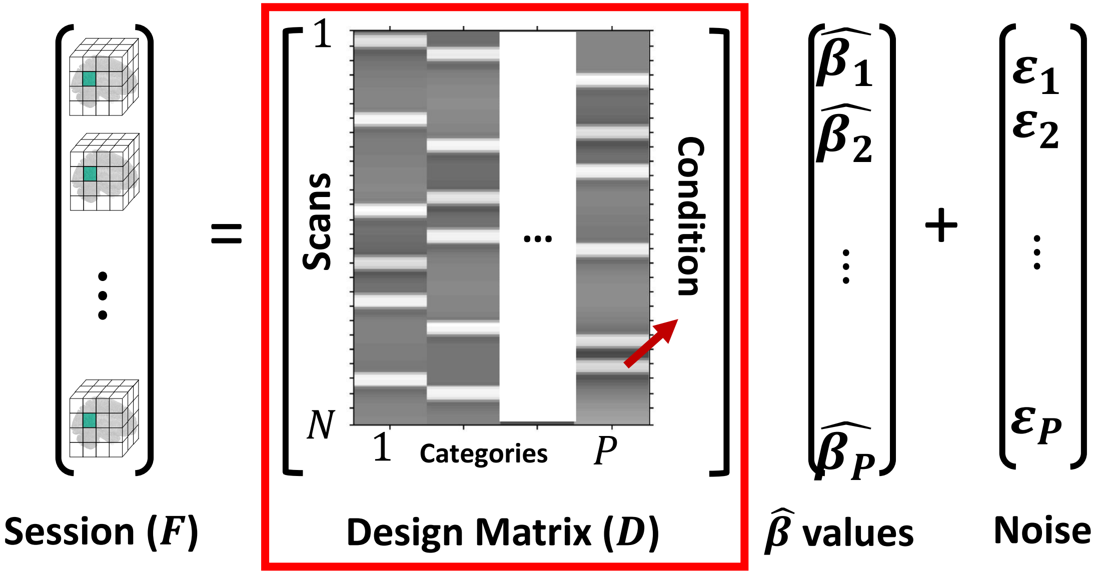
$$F = D\left(\hat{\beta}\right)^{\top} + \varepsilon$$

- ❑ This paper uses Generalized Least Squares (GLS) approach for estimating optimized solution:

$$\widehat{\beta} = \left(\left(\mathbf{D}^{\top}\mathbf{\Sigma}^{-1}\mathbf{D}\right)^{-1}\mathbf{D}^{\top}\mathbf{\Sigma}^{-1}\mathbf{F}\right)^{\top}$$

$$Var(\varepsilon) = \mathbf{\Sigma}\sigma^2 \neq \mathbb{I}\sigma^2$$

The first level analysis



$$F = D\left(\hat{\beta}\right)^{\top} + \varepsilon$$

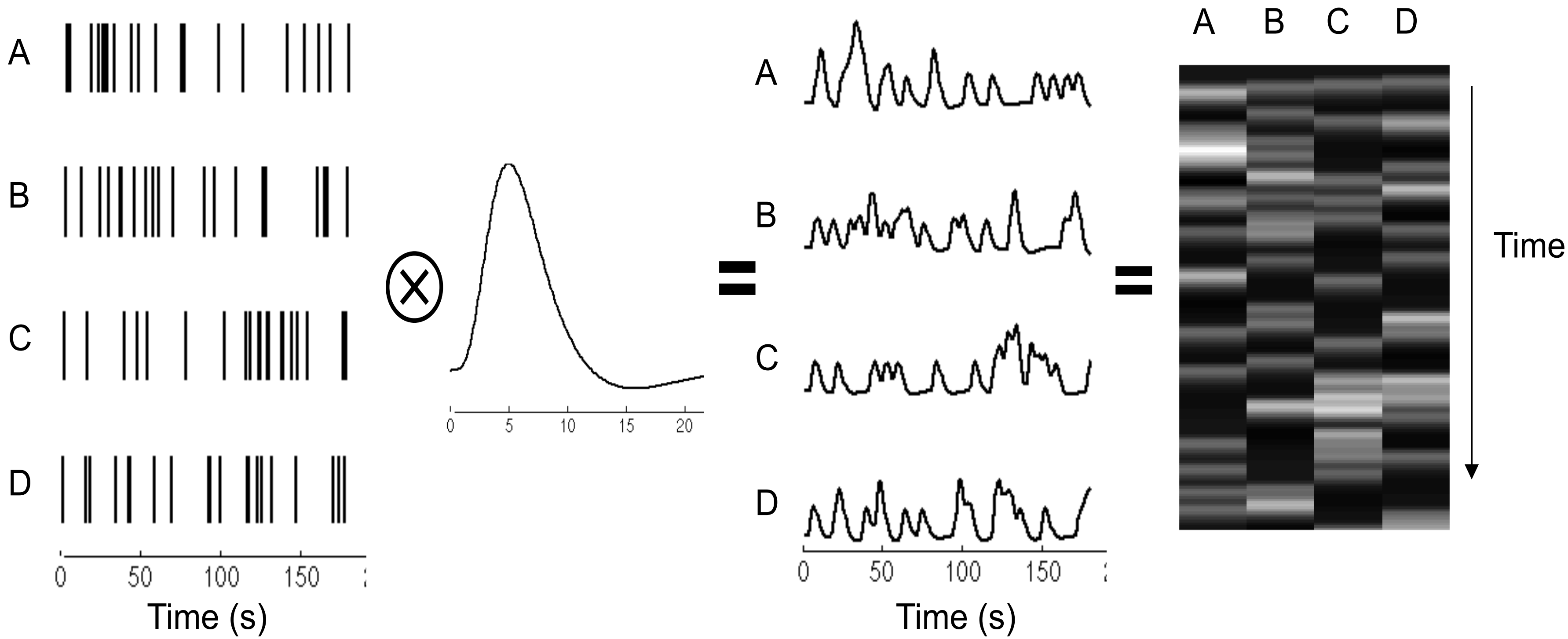
- ❑ This paper uses Generalized Least Squares (GLS) approach for estimating optimized solution:

$$\widehat{\beta} = \left(\left(\mathbf{D}^{\top}\boldsymbol{\Sigma}^{-1}\mathbf{D}\right)^{-1}\mathbf{D}^{\top}\boldsymbol{\Sigma}^{-1}\mathbf{F}\right)^{\top}$$

$$Var(\varepsilon) = \boldsymbol{\Sigma}\sigma^2 \neq \mathbb{I}\sigma^2$$

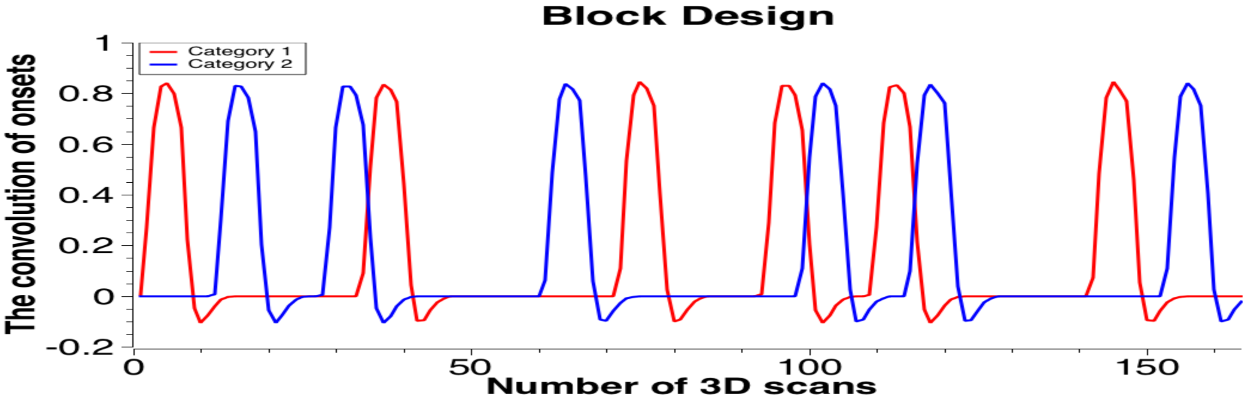
What is the Design Matrix?

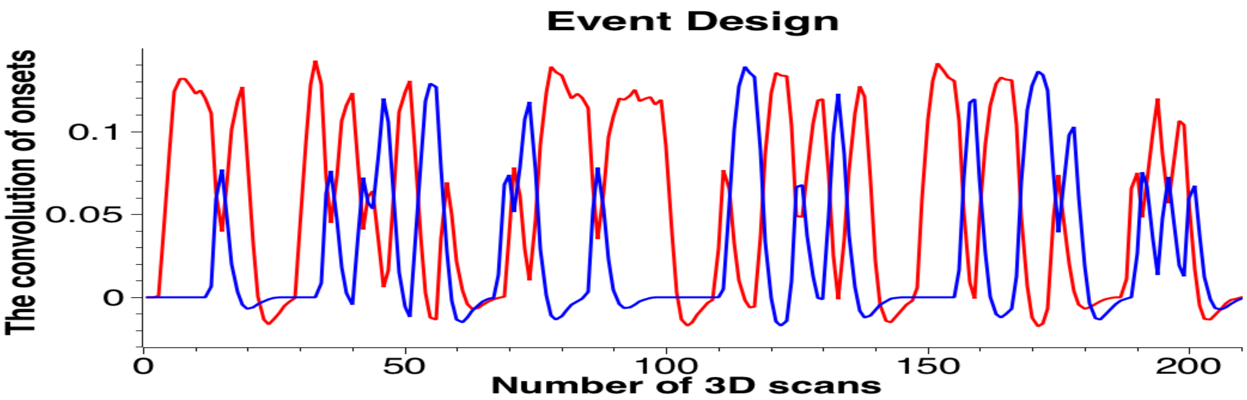
Indicator functions (Onsets) | Assumed HRF (Basis function) | Design Matrix (X^T) | Design Matrix (X)



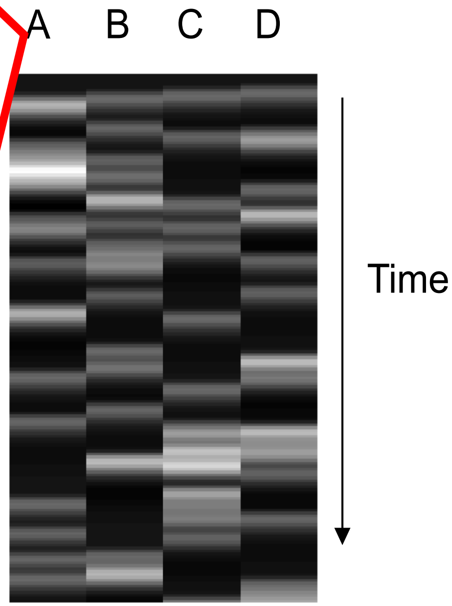
What is the Design Matrix?

Based on onsets, we have two types of design matrix:





Design Matrix (X)

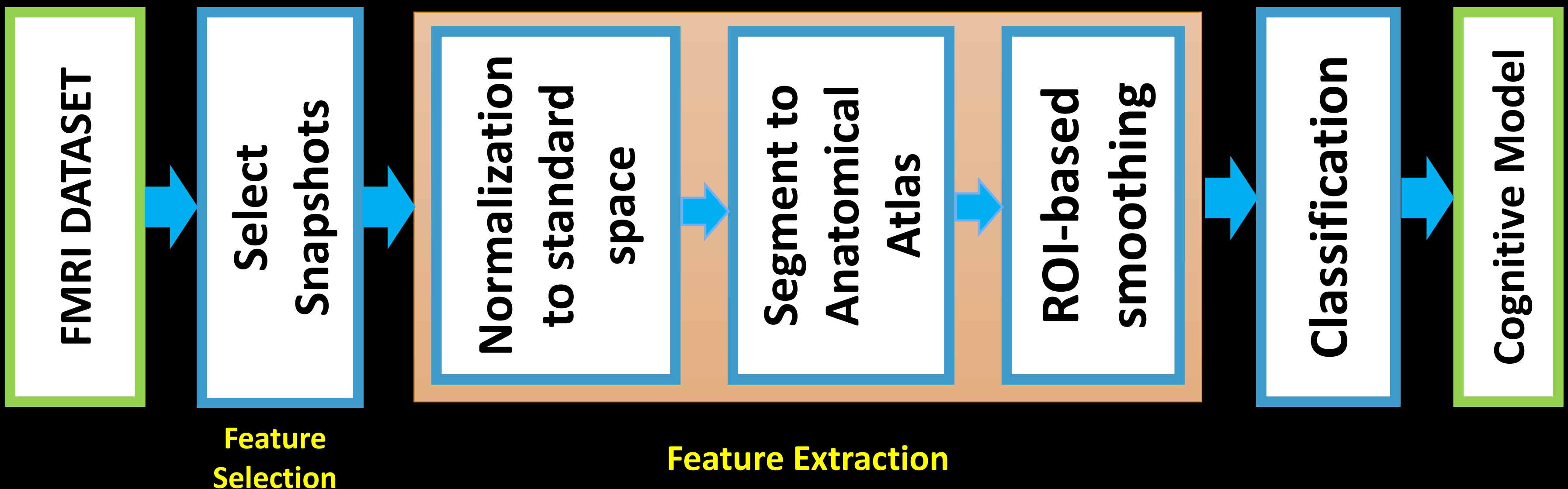


Multi-Region Neural Representation

Multi-Region Neural Representation

- The proposed method includes three steps:
  1. Snapshots Selection
  2. Feature Extraction
     - 2.1 Normalizing snapshots to standard space
     - 2.2 Segmenting the snapshots in the form of anatomical regions
     - 2.3 Removing noise in the level of ROIs.
  3. Ensemble Learning

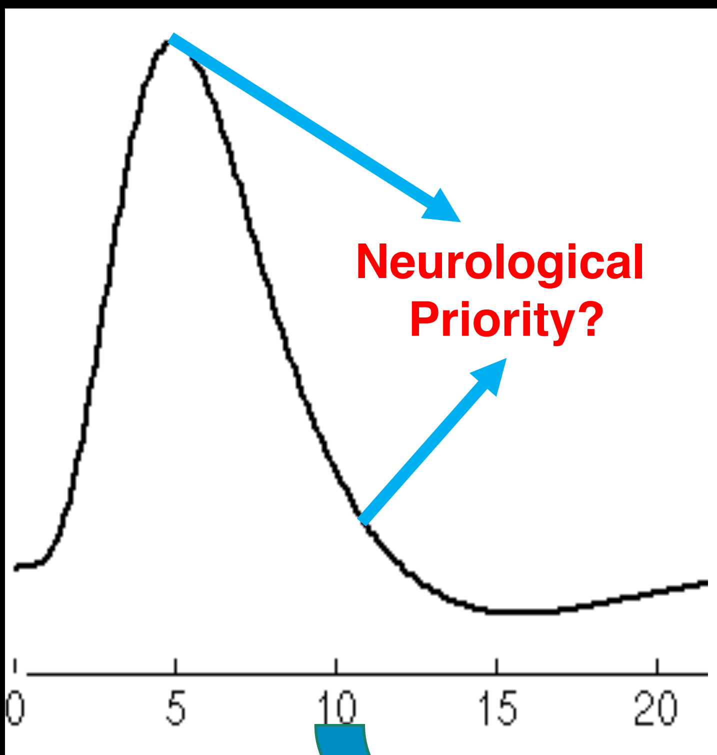
- The graphical pipeline of the proposed method:

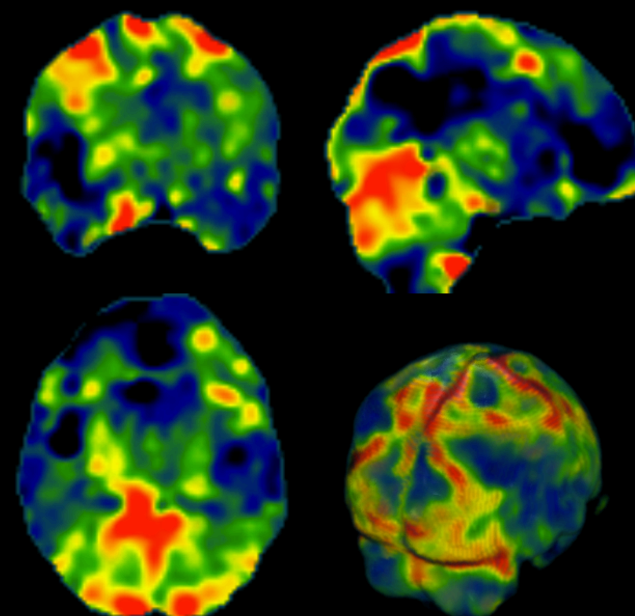


Snapshots Selection

- **The first key idea: Neurological Priority**

Which time point is better for ranking 3D images?



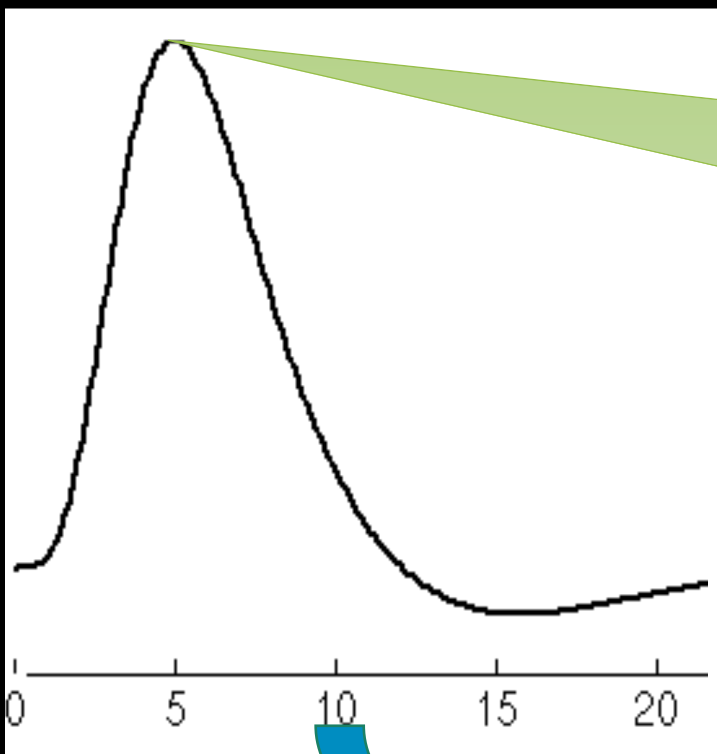


Each time point shows a 3D snapshots

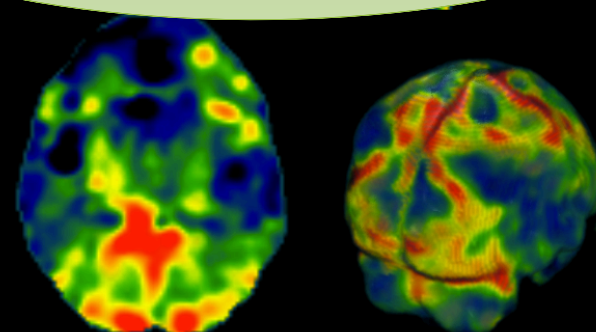
Snapshots Selection

- The first key idea: Neurological Priority

Which time point is better for ranking 3D images?



Looking for local maximums in the design matrix



Each time point shows a 3D snapshots

Definition of snapshots

Onsets (time points): $\boldsymbol{S} = \{\boldsymbol{S_1}, \cdots, \boldsymbol{S_i}, \cdots, \boldsymbol{S_p}\}$

removing soft noise

Design Matrix: $\boldsymbol{D} = \{\boldsymbol{d_1}, \cdots, \boldsymbol{d_i}, \cdots, \boldsymbol{d_p}\}$

Gaussian Kernel: $\widehat{\mathbf{G}} = \left\{ \exp\left(\frac{-\widehat{\mathbf{g}}^2}{2\sigma_G^2}\right) \middle| \widehat{\mathbf{g}} \in \mathbb{Z} \text{ and } -2\lceil\sigma_G\rceil \leq \widehat{\mathbf{g}} \leq 2\lceil\sigma_G\rceil \right\}$, $\mathbf{G} = \frac{\widehat{\mathbf{G}}}{\sum_j \widehat{\mathbf{g}}_j}$

Smoothed Design Matrix: $\phi_i = \mathbf{d}_i * \mathbf{G} = (\mathbf{S}_i * \mathbf{H}) * \mathbf{G}$, $\mathbf{\Phi} = \{\phi_1, \phi_2, \ldots, \phi_p\}$

The local maximum points: $\mathbf{S}_i^* = \left\{ \underset{\mathbf{S}_i}{\arg} \quad \phi_i \middle| \frac{\partial \phi_i}{\partial \mathbf{S}_i} = 0 \text{ and } \frac{\partial^2 \phi_i}{\partial \mathbf{S}_i \mathbf{S}_i} > 0 \right\}$

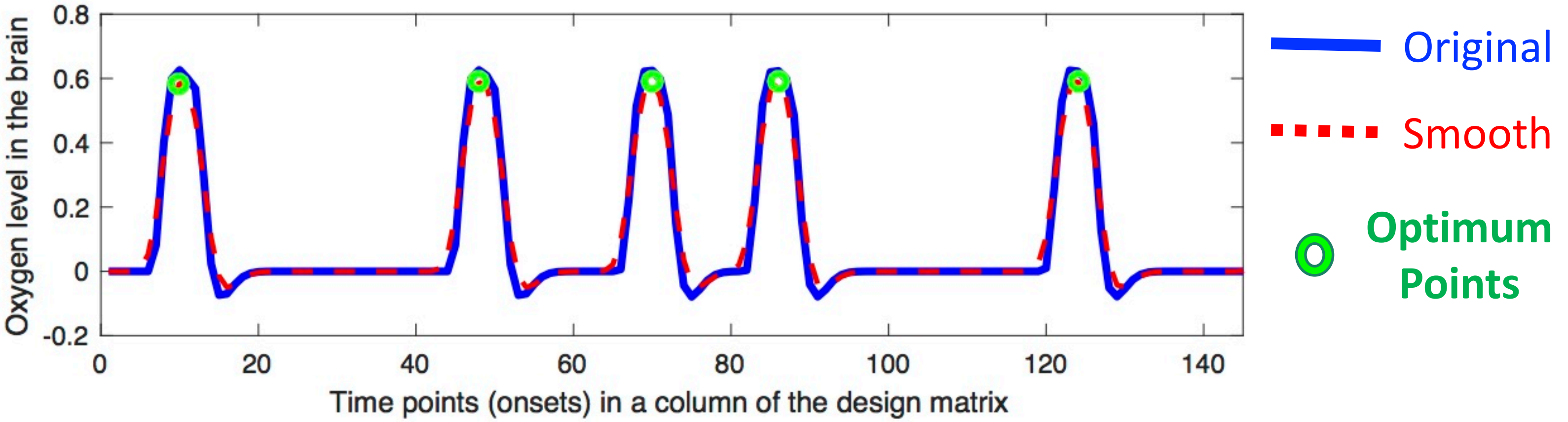
The set of snapshots can be formulated as follows:

$$\widehat{\mathbf{\Psi}} = \{\mathbf{f}_j^\intercal \mid \mathbf{f}_j^\intercal \in \mathbf{F}^\intercal \text{ and } j \in \mathbf{S}^*\} = \{\widehat{\psi_1}, \widehat{\psi_2}, \ldots, \widehat{\psi_k}, \ldots \widehat{\psi_q}\} \in \mathbb{R}^{m \times q}$$

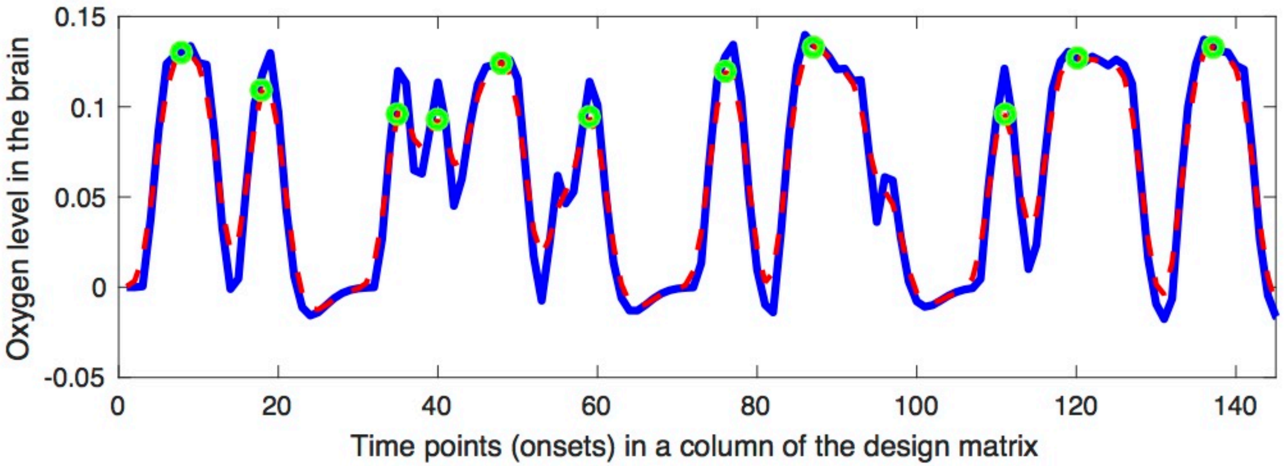
\# of conditions

\# of voxels

- **Block Based Example:**



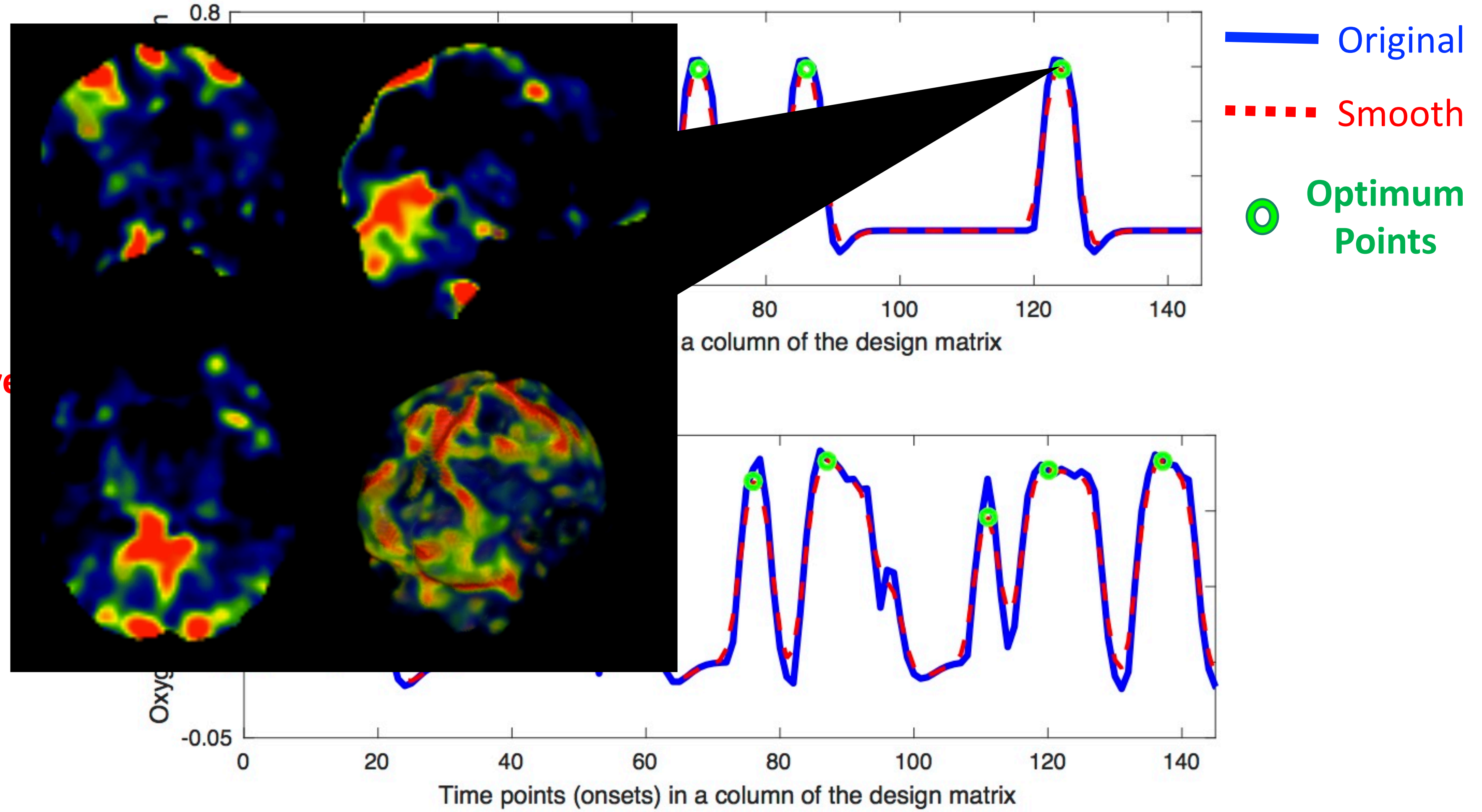
- **Event Based Example:**



Definition of snapshots (examples)

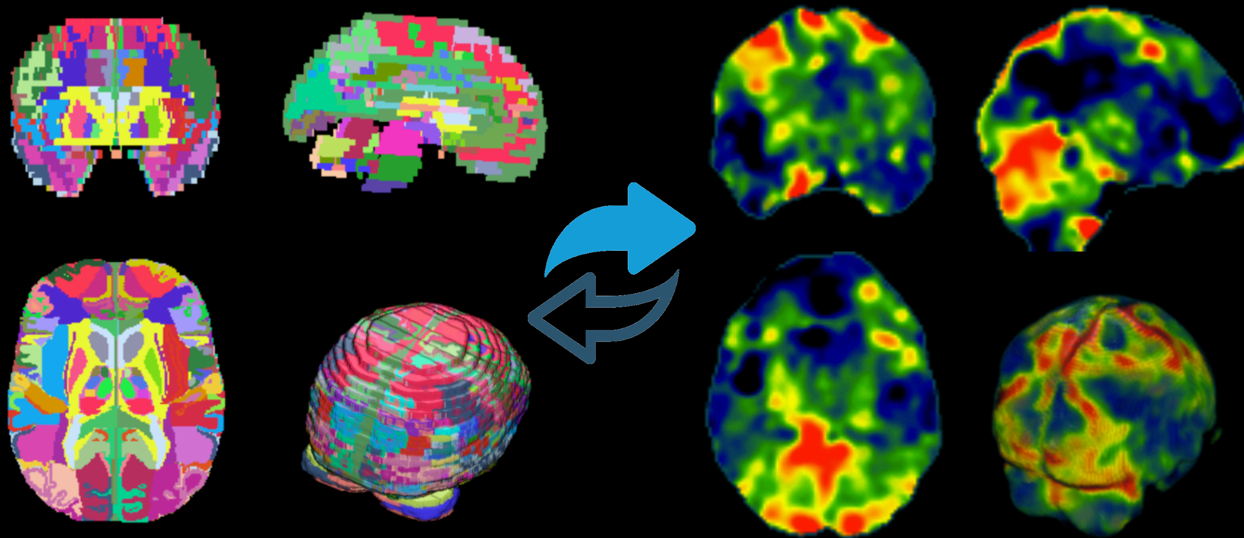
❑ **Block Based Example:**

❑ **Eve**



Feature Extraction

- The second key idea is extracting the features of snapshots based on an anatomical atlas for removing noise and sparsity and improving performance of learning.
- Three steps:
  - ✓ Normalizing snapshots to standard space
  - ✓ Segmenting the snapshots in the form of anatomical regions
  - ✓ Removing noise in the level of ROIs.



Anatomical Atlas   Functional Activities

- For reducing the time complexity, this paper uses β values for each category of stimuli to find a transformation matrix for mapping snapshots from the original space to the standard space.

$$\mathbf{T}_i: \quad \widehat{\beta}_i \in \mathbb{R}^m \quad \rightarrow \quad \beta_i \in \mathbb{R}^n$$

Original Space — Standard Space

- **Transformation:** $\mathbf{T}_i = \arg\min(NMI(\widehat{\beta}_i, \mathbf{Ref}))$
- **Snapshot Mappings:** $\mathbf{T}_j^*: \widehat{\psi}_j \in \mathbb{R}^m \rightarrow \psi_j \in \mathbb{R}^n \implies \psi_j = \left((\widehat{\psi}_j)^{\mathsf{T}} \mathbf{T}_j^* \right)^{\mathsf{T}}$
- **Applying non-zero correlations to snapshots:** $\mathbf{\Theta}_j = \psi_j \circ \beta_j^*$

where:

$$(\mathbf{T}_j^*, \beta_j^*) = Select(\widehat{\psi_j}, \mathbf{T}, \beta) = \{(\mathbf{T}_i, \beta_i) \mid \mathbf{T}_i \in \mathbf{T},\ \beta_i \in \beta\ ,\ \widehat{\psi_j} \text{ is belonged to the } i-th \text{ category} \implies \widehat{\psi_j} \propto \beta_i \propto \mathbf{T}_i\}$$

Step 2: Segmenting the snapshots in the form of anatomical regions

- The basic assumption is that the voxels belong to an anatomical regions must behave in unison for a each unique task.

- Anatomical Atlas: $\mathbf{A} \in \mathbb{R}^n = \{\mathbf{A}_1, \mathbf{A}_2, \dots, \mathbf{A}_\ell, \dots, \mathbf{A}_L\},$

$$\cap_{\ell=1}^{L} \{\mathbf{A}_\ell\} = \emptyset, \ \cup_{\ell=1}^{L} \{\mathbf{A}_\ell\} = \mathbf{A}$$

- A segmented snapshot based on the $i-th$ region can be denoted as follows:

$$\mathbf{\Theta}_{(j,\ell)} = \{\theta_j^k \mid \theta_j^k \in \mathbf{\Theta}_j \text{ and } k \in \mathbf{A}_\ell\}$$

- The automatically detected active regions can be also defined as follows:

$$\mathbf{\Theta}_j^* = \left\{\mathbf{\Theta}_{(j,\ell)} | \mathbf{\Theta}_{(j,\ell)} \subset \mathbf{\Theta}_j \text{ and } \sum_{\theta_{(j,\ell)}^k \in \mathbf{\Theta}_{(j,\ell)}} |\theta_{(j,\ell)}^k| \neq 0\right\}$$

Step 3: Removing noise in the level of ROIs

- This paper smooths voxels belong to each anatomical region.

- A Gaussian kernel for each anatomical region can be defined as follows:

$$\sigma_\ell = \frac{N_\ell^2}{5N_\ell^2 \log N_\ell}$$

# of voxels in $\ell - th$ region

$$\widehat{\mathbf{V}}_\ell = \left\{ \exp\left(\frac{-\widehat{\mathbf{v}}^2}{2\sigma_\ell}\right) \middle| \; \widehat{\mathbf{v}} \in \mathbb{Z} \text{ and } -2\lceil\sigma_\ell\rceil \leq \widehat{\mathbf{v}} \leq 2\lceil\sigma_\ell\rceil \right\}$$

$$\mathbf{V}_\ell = \frac{\widehat{\mathbf{V}_\ell}}{\sum_j \widehat{\mathbf{v}}_j}$$

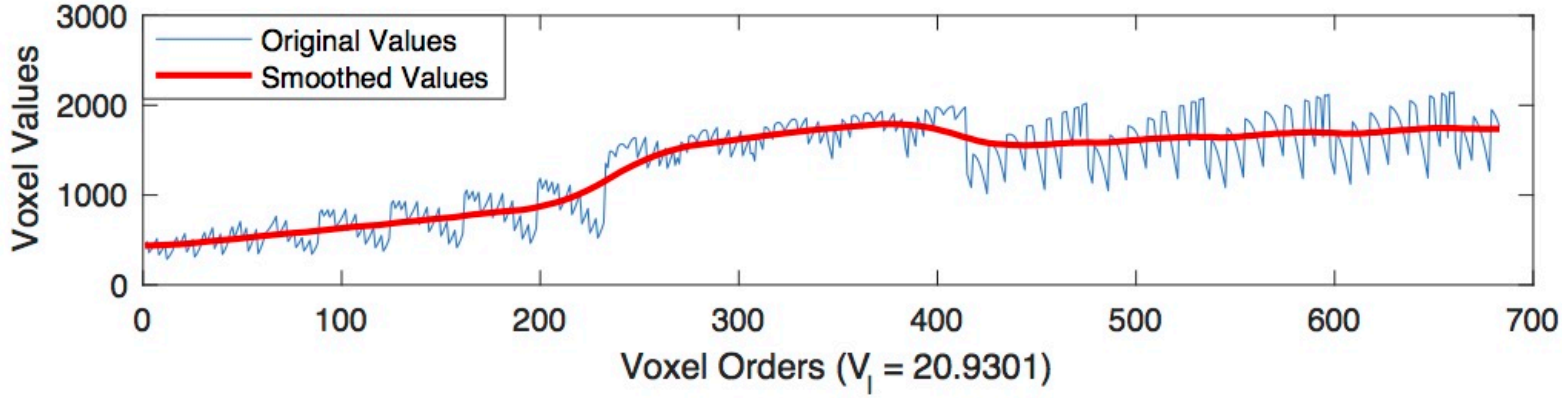
- The smoothed version of the $j - th$ snapshot can be defined as follows:

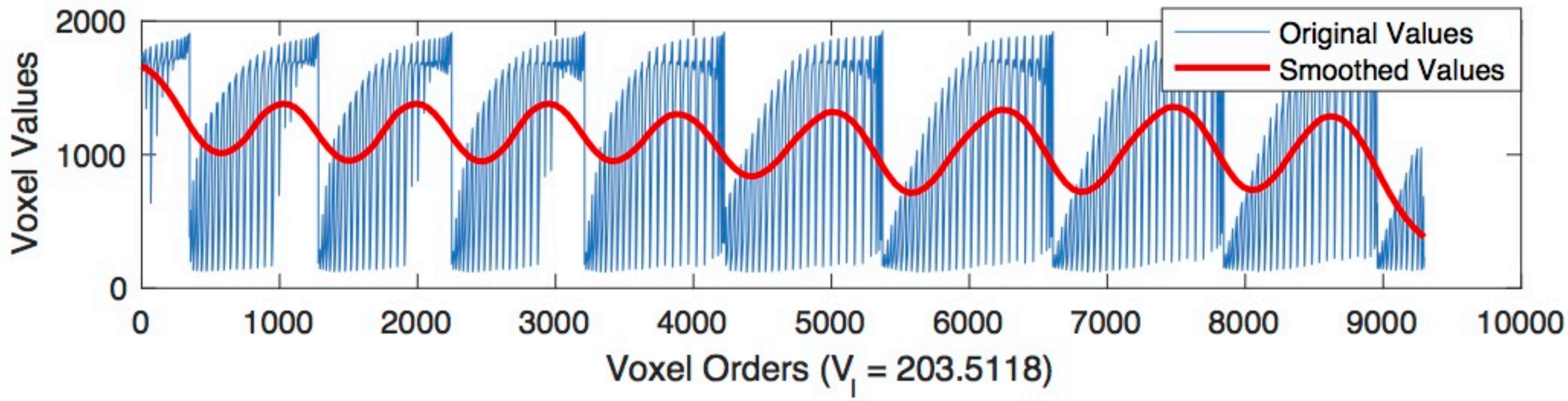
$$\forall \ell = L1 \dots L2 \rightarrow \mathbf{X}_{(j,\ell)} = \mathbf{\Theta}_{(j,\ell)} * \mathbf{V}_\ell,$$
$$\mathbf{X}_j = \{\mathbf{X}_{(j,L1)}, \dots, \mathbf{X}_{(j,\ell)}, \dots \mathbf{X}_{(j,L2)}\}$$

where L1 and L2 are the first and the last active regions in the $j - th$ snapshot

Feature Extraction (examples)

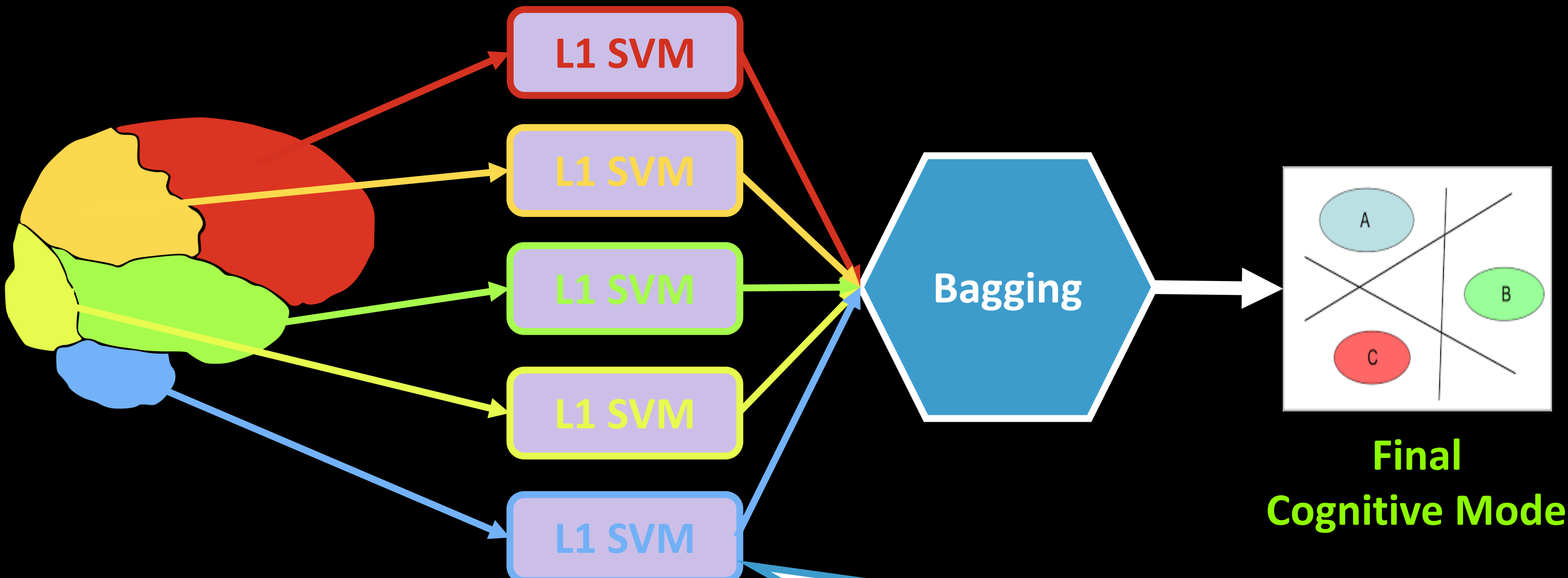
- Voxels belong to a unique anatomical region are smoothed as follows:





Learning: Cognitive Model

- The third key idea is training an efficient classifier by using an ensemble approach
- For each anatomical region, we use L1-SVM classifier.



$$\eta_\ell: \quad \min_{\mathbf{W}_\ell} \quad C \sum_{j=1}^{\tau} \max(0, 1 - y_j \mathbf{X}_{(j,\ell)} \mathbf{W}_{(j,\ell)}) + \|\mathbf{W}_\ell\|_1$$

Empirical Studies

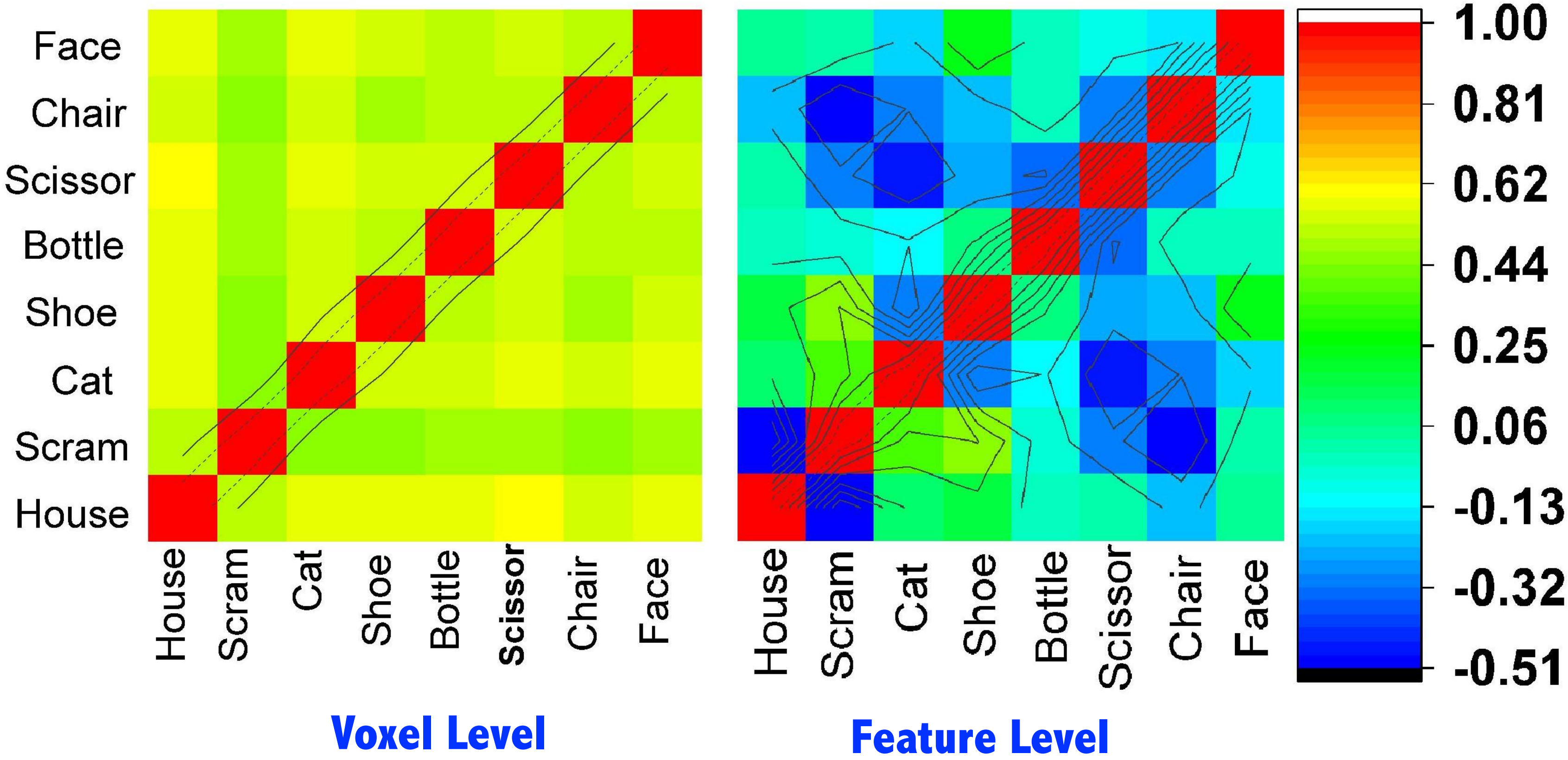
Datasets

Title	ID	U	p	t	X	Y	Z
Visual Object Recognition	DS105	71	8	121	79	95	79
Word and Object Processing	DS107	98	4	164	53	63	52
Multi-subject, multi-modal	DS117	171	2	210	64	61	33

- ❑ **U is the number of subject**
- ❑ **p denotes the number of visual stimuli categories**
- ❑ **t is the number of scans in unites of TRs (Time of Repetition)**
- ❑ **X, Y, Z are the size of 3D images**

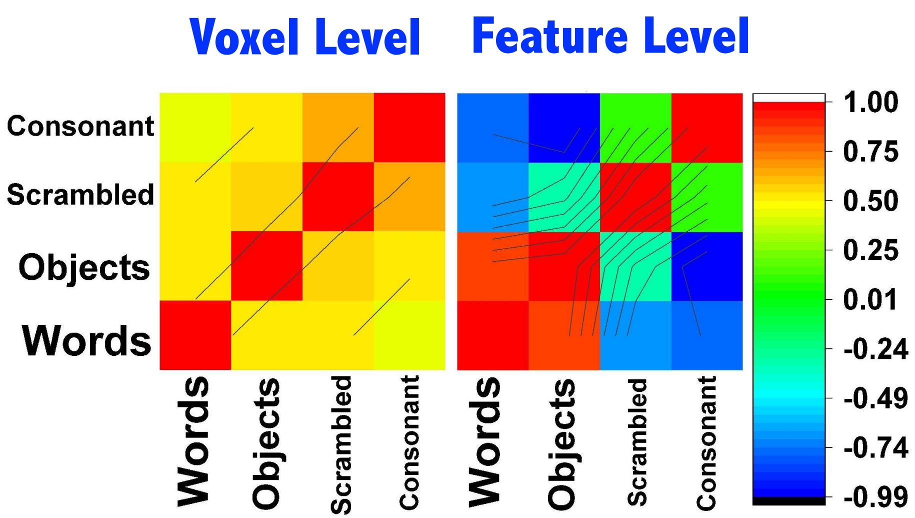
➢ **Provided by www.openfmri.org**

Correlation Analysis

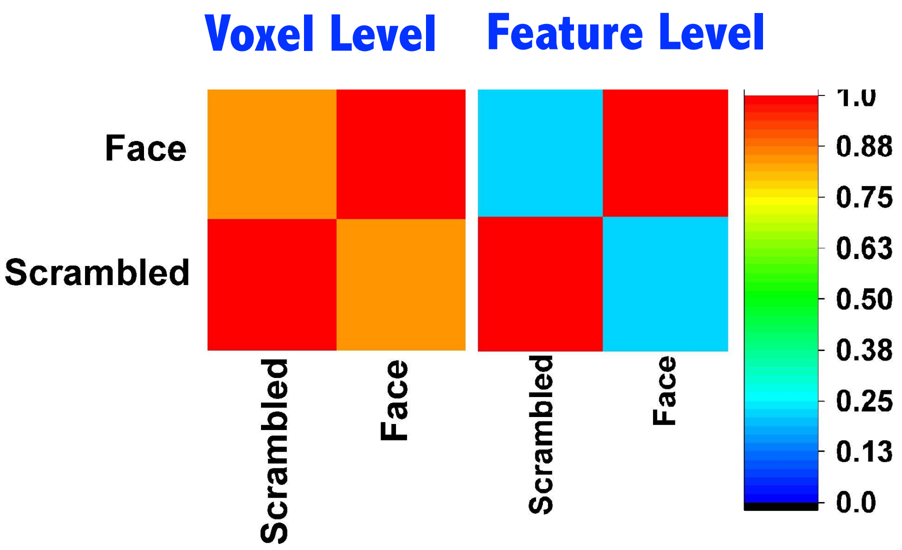


Visual Object Recognition (DS105)

Correlation Analysis



Word and Object Processing (DS107)



Multi-subject, multi-modal (DS117)

Classification Analysis

Table 1: Accuracy of binary predictors

Data Sets	SVM	Graph Net	Elastic Net	L1-Reg. SVM	Osher et al.	Proposed method
DS105: Objects vs. Scrambles	71.65±0.97	81.27±0.59	83.06±0.36	85.29±0.49	90.82±1.23	**94.32±0.16**
DS107: Words vs. Others	82.89±1.02	78.03±0.87	88.62±0.52	86.14±0.91	90.21±0.83	**92.04±0.09**
DS107: Consonants vs. Others	67.84±0.82	83.01±0.56	82.82±0.37	85.69±0.69	84.54±0.99	**96.73±0.19**
DS107: Objects vs. Others	73.32±1.67	77.93±0.29	84.22±0.44	83.32±0.41	**95.62±0.83**	93.07±0.27
DS107: Scrambles vs. Others	83.96±0.87	79.37±0.82	87.19±0.26	86.45±0.62	88.1±0.78	**90.93±0.71**
DS117: Faces vs. Scrambles	81.25±1.03	85.19±0.56	85.46±0.29	86.61±0.61	**96.81±0.79**	96.31±0.92
ALL: Faces vs. Others	66.27±1.61	68.37±1.31	75.91±0.74	80.23±0.72	84.99±0.71	**89.99±0.31**
ALL: Objects vs. Others	75.61±0.57	78.37±0.71	76.79±0.94	80.14±0.47	79.23±0.25	**92.44±0.92**
ALL: Scrambles vs. Others	81.92±0.71	81.08±1.23	84.18±0.42	88.23±0.81	90.5±0.73	**95.39±0.18**

Table 2: Area Under the ROC Curve (AUC) of binary predictors

Data Sets	SVM	Graph Net	Elastic Net	L1-Reg. SVM	Osher et al.	Proposed method
DS105: Objects vs. Scrambles	68.37±1.01	70.32±0.92	82.22±0.42	80.91±0.21	88.54±0.71	**93.25±0.92**
DS107: Words vs. Others	80.76±0.91	77.91±1.03	86.35±0.39	84.23±0.57	87.61±0.62	**91.86±0.17**
DS107: Consonants vs. Others	63.84±1.45	81.21±0.33	80.63±0.61	84.41±0.92	81.54±0.31	**94.03±0.37**
DS107: Objects vs. Others	70.17±0.59	76.14±0.49	81.54±0.92	80.92±0.28	**94.23±0.94**	92.14±0.42
DS107: Scrambles vs. Others	80.73±0.92	77±1.01	85.79±0.42	83.14±0.47	82.23±0.38	**87.05±0.37**
DS117: Faces vs. Scrambles	79.36±0.33	83.71±0.81	83.21±1.23	82.29±0.91	94.08±0.84	**94.61±0.71**
ALL: Faces vs. Others	61.91±1.2	65.04±0.99	74.9±0.61	78.14±0.83	83.89±0.28	**91.05±0.12**
ALL: Objects vs. Others	74.19±0.92	77.88±0.82	73.59±0.95	79.45±0.77	75.61±0.89	**89.24±0.69**
ALL: Scrambles vs. Others	79.81±1.01	80±0.49	82.53±0.83	88.14±0.91	88.93±0.71	**92.09±0.28**

Classification Analysis

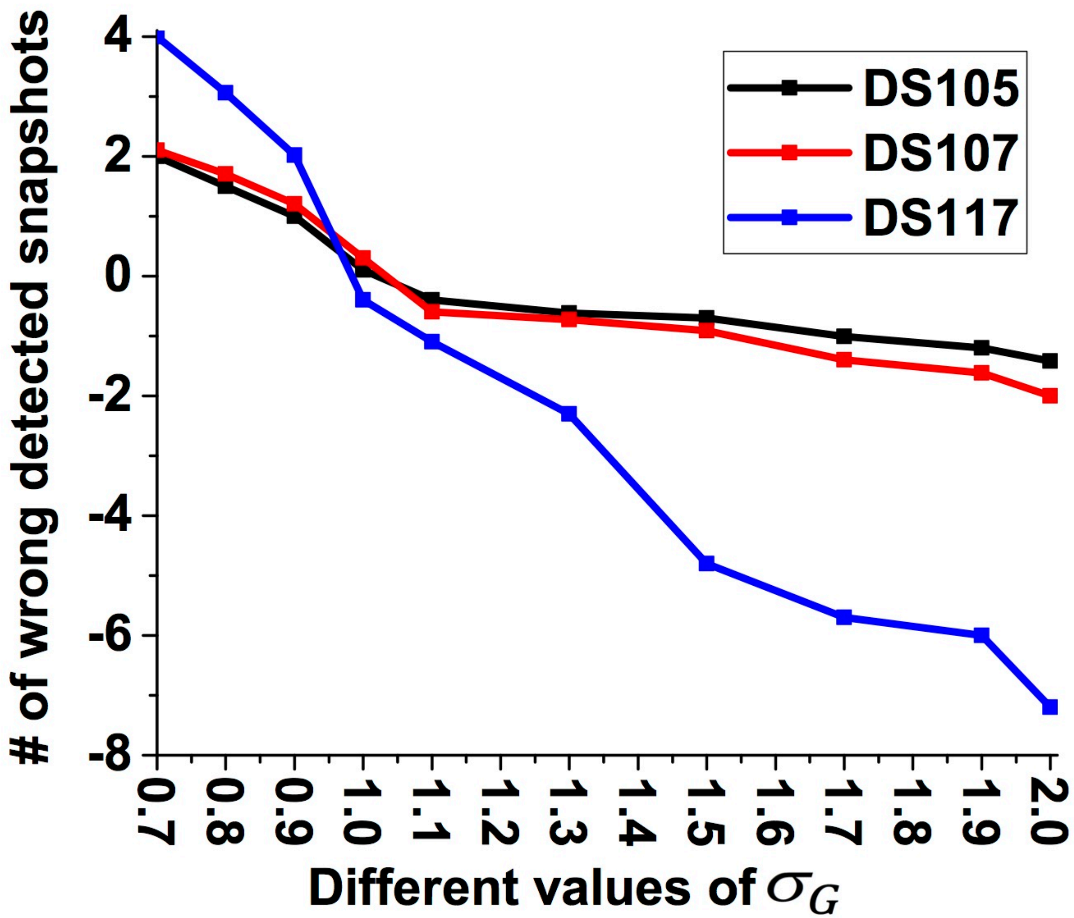
Table 1: Accuracy of binary predictors

Data Sets	SVM	Graph Net	Elastic Net	L1-Reg. SVM	Osher et al.	Proposed method
DS105: Objects vs. Scrambles	71.65±0.97	81.27±0.59	83.06±0.36	85.29±0.49	90.82±1.23	**94.32±0.16**
DS107: Words vs. Others	82.89±1.02	78.03±0.87	88.62±0.52	86.14±0.91	90.21±0.83	**92.04±0.09**
DS107: Consonants vs. Others	67.84±0.82	83.01±0.56	82.82±0.37	85.69±0.69	84.54±0.99	**96.73±0.19**
DS107: Objects vs. Others	73.32±1.67	77.93±0.29	84.22±0.44	83.32±0.41	**95.62±0.83**	93.07±0.27
DS107: Scrambles vs. Others	83.96±0.87	79.37±0.82	87.19±0.26	86.45±0.62	88.1±0.78	**90.93±0.71**
DS117: Faces vs. Scrambles	81.25±1.03	85.19±0.56	85.46±0.29	86.61±0.61	**96.81±0.79**	96.31±0.92
ALL: Faces vs. Others	66.27±1.61	68.37±1.31	75.91±0.74	80.23±0.72	84.99±0.71	**89.99±0.31**
ALL: Objects vs. Others	75.61±0.57	78.37±0.71	76.79±0.94	80.14±0.47	79.23±0.25	**92.44±0.92**
ALL: Scrambles vs. Others	81.92±0.71	81.08±1.23	84.18±0.42	88.23±0.81	90.5±0.73	**95.39±0.18**

Table 2: Area Under the ROC Curve (AUC) of binary predictors

Data Sets	SVM	Graph Net	Elastic Net	L1-Reg. SVM	Osher et al.	Proposed method
DS105: Objects vs. Scrambles	68.37±1.01	70.32±0.92	82.22±0.42	80.91±0.21	88.54±0.71	**93.25±0.92**
DS107: Words vs. Others	80.76±0.91	77.91±1.03	86.35±0.39	84.23±0.57	87.61±0.62	**91.86±0.17**
DS107: Consonants vs. Others	63.84±1.45	81.21±0.33	80.63±0.61	84.41±0.92	81.54±0.31	**94.03±0.37**
DS107: Objects vs. Others	70.17±0.59	76.14±0.49	81.54±0.92	80.92±0.28	**94.23±0.94**	92.14±0.42
DS107: Scrambles vs. Others	80.73±0.92	77±1.01	85.79±0.42	83.14±0.47	82.23±0.38	**87.05±0.37**
DS117: Faces vs. Scrambles	79.36±0.33	83.71±0.81	83.21±1.23	82.29±0.91	94.08±0.84	**94.61±0.71**
ALL: Faces vs. Others	61.91±1.2	65.04±0.99	74.9±0.61	78.14±0.83	83.89±0.28	**91.05±0.12**
ALL: Objects vs. Others	74.19±0.92	77.88±0.82	73.59±0.95	79.45±0.77	75.61±0.89	**89.24±0.69**
ALL: Scrambles vs. Others	79.81±1.01	80±0.49	82.53±0.83	88.14±0.91	88.93±0.71	**92.09±0.28**

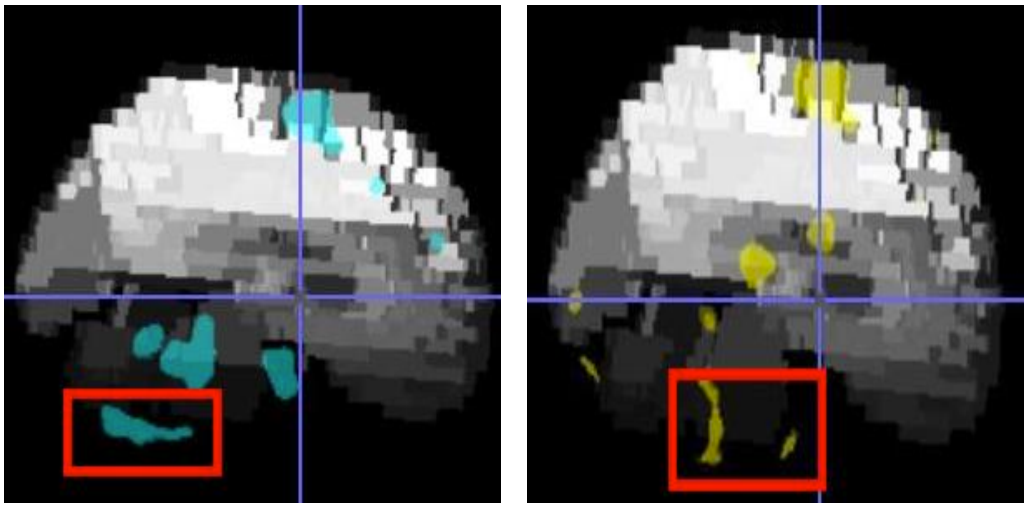
Parameters Analysis: σ_G for smoothing design matrix

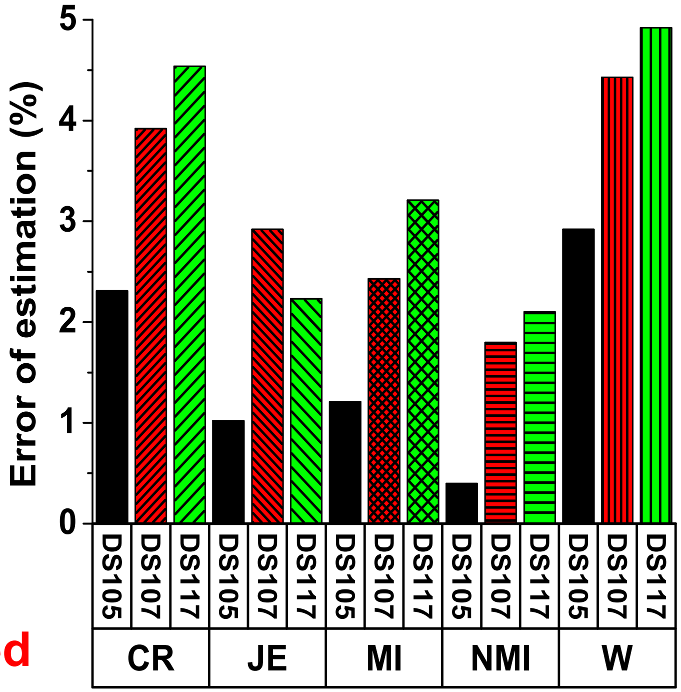


- ❑ $0 < \sigma < 1$ can create design matrix, which is sensitive to small spikes.
- ❑ $\sigma > 1$ can increase the level of smoothness that can remove some weak local maximums, especially in the event-related data sets.

$$\widehat{\mathbf{G}} = \left\{ \exp\left(\frac{-\widehat{\mathbf{g}}^2}{2\sigma_G^2}\right) \middle| \widehat{\mathbf{g}} \in \mathbb{Z} \text{ and } -2\lceil\sigma_G\rceil \leq \widehat{\mathbf{g}} \leq 2\lceil\sigma_G\rceil \right\} \Rightarrow \phi_i = \mathbf{d}_i * \mathbf{G} = (\mathbf{S}_i * \mathbf{H}) * \mathbf{G}$$

Parameters Analysis: Normalization Objective Functions

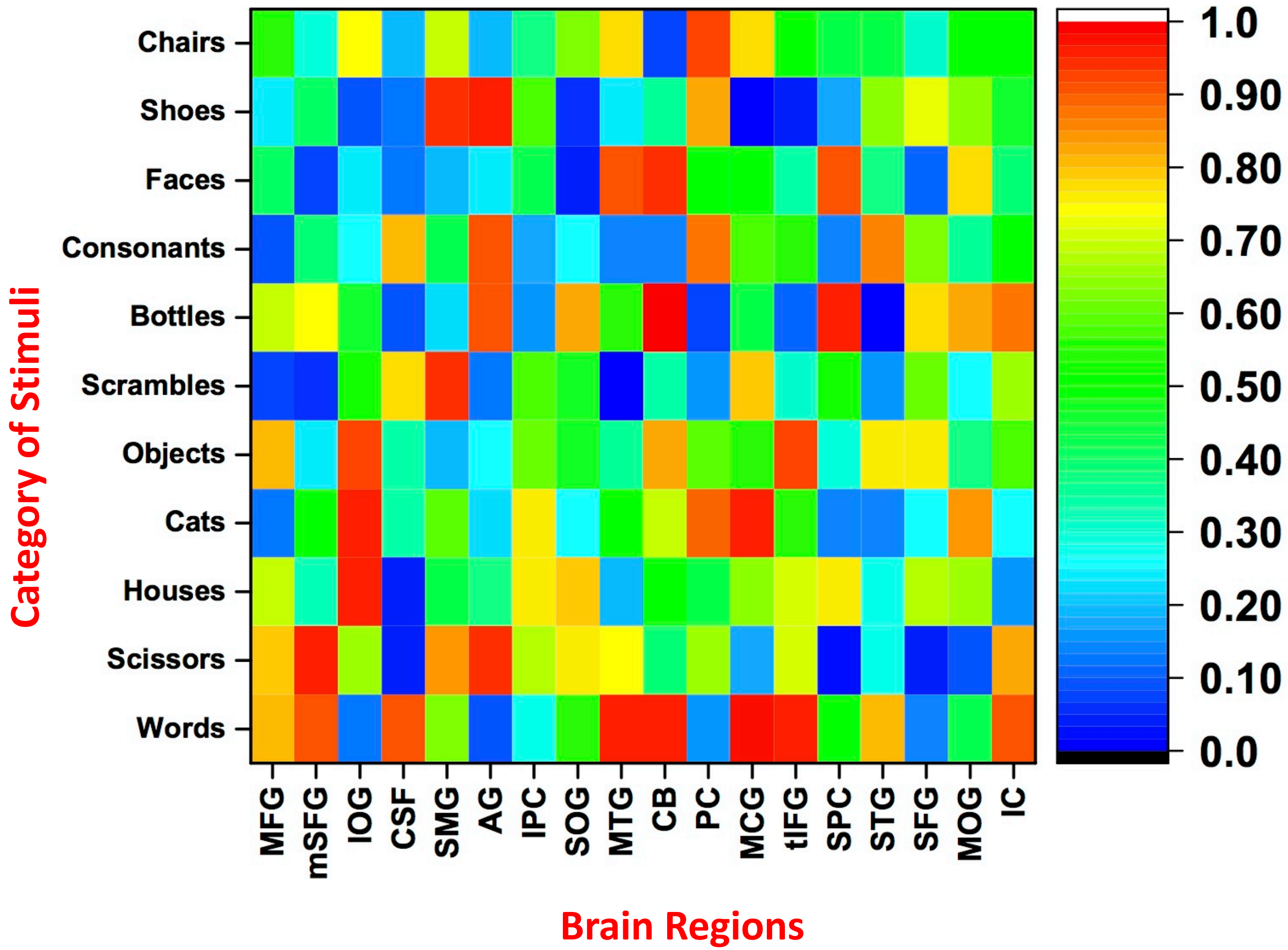




The error of registration (normalization): the red rectangles illustrate the error areas

$$\mathbf{T}_i = \arg\min(NMI(\widehat{\beta}_i, \mathbf{Ref}))$$

Regions of Interests (ROIs) Analysis



Future Works

Conclusion

- This paper proposes Multi-Region Neural Representation as a novel feature space for decoding visual stimuli in the human brain.

- Experimental studies on 4 visual categories (words, objects, consonants and nonsense photos) clearly show the superiority of our proposed method in comparison with state-of-the-art methods.

- In future, we plan to apply the proposed method to different brain tasks such as risk, emotion and etc.

Thank You

Q & A

For more details, contact:

myousefnezhad@nuaa.edu.cn

myousefnezhad@outlook.com

https://myousefnezhad.github.io