Multi-Region Neural Representation
A novel model for decoding visual stimuli in human brains

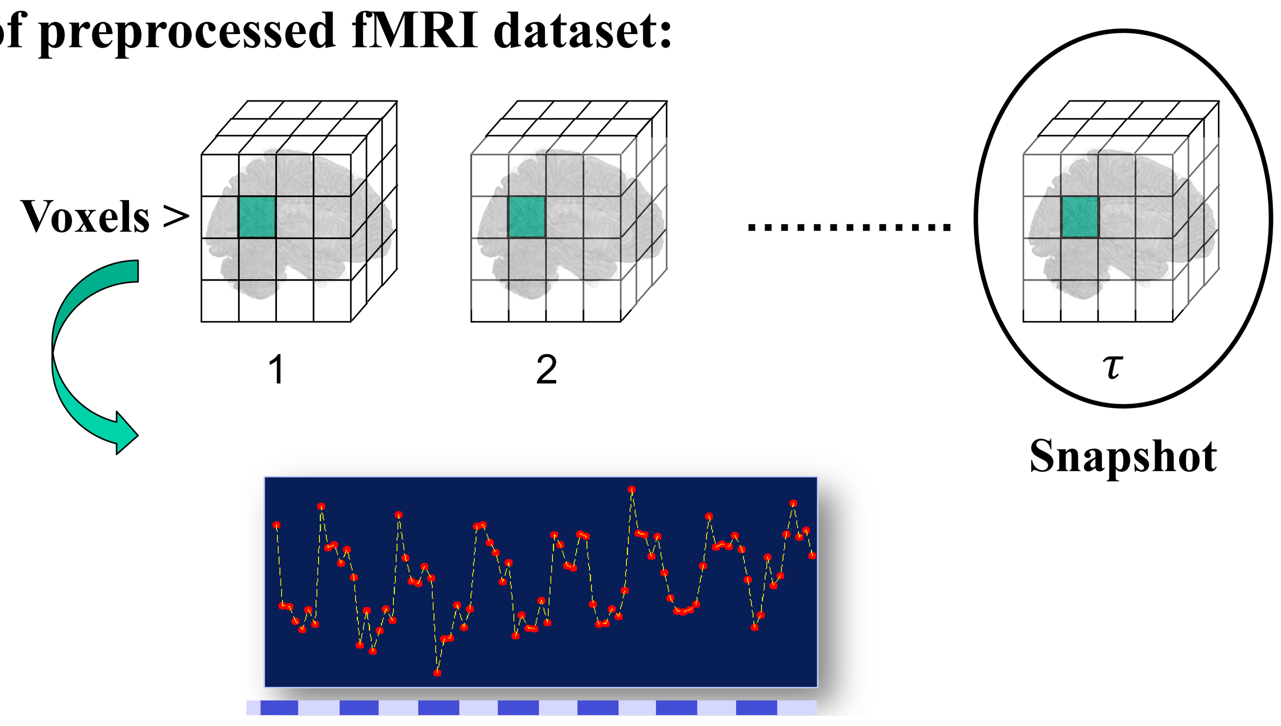
iBRAIN

Muhammad Yousefnezhad, Daoqiang Zhang

School of Computer Science and Technology, Nanjing University of Aeronautics and Astronautics, Nanjing, China

MOTIVATION

- **A session of preprocessed fMRI dataset:**



Blood Oxygen Level Dependent (BOLD) signals

Goals

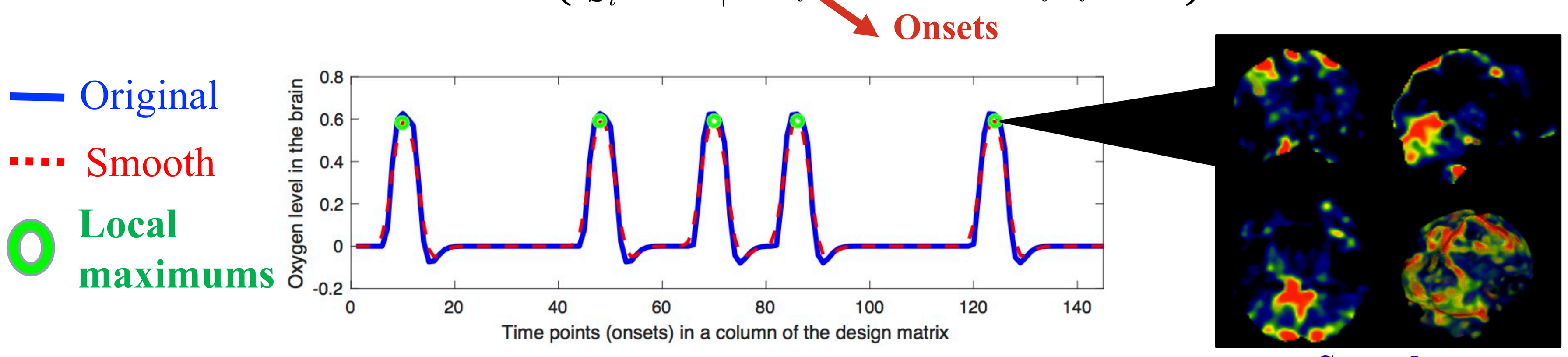
1. **Selecting a set of effective snapshots rather than using whole of the noisy and sparse time series.**
2. **Extracting robust features from the selected snapshots.**
3. **Improving the performance of the generated cognitive model.**

METHOD

The proposed method is applied in three stages: firstly, snapshots of brain image (each snapshot represents neural activities for a unique stimulus) are selected by finding local maximums in the smoothed version of the design matrix. Then, features are generated in three steps, including normalizing to standard space, segmenting the snapshots in the form of automatically detected anatomical regions, and removing noise by Gaussian smoothing in the level of ROIs.

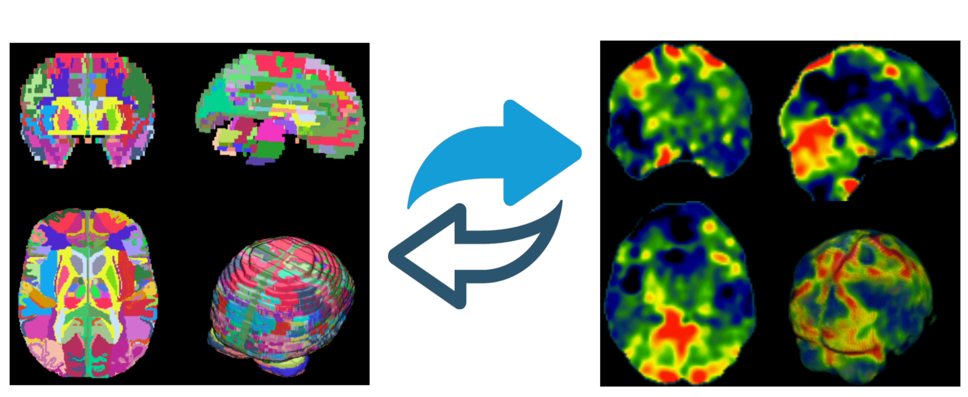
Snapshot Selection

- **Local maximums:** $\mathbf{S}_i^* = \left\{ \arg_{\mathbf{S}_i} \; \phi_i \;\middle|\; \underbrace{\frac{\partial \phi_i}{\partial \mathbf{S}_i}}_{\text{Onsets}} = 0 \text{ and } \frac{\partial^2 \phi_i}{\partial \mathbf{S}_i \mathbf{S}_i} > 0 \right\}$ — Smoothed design matrix (ϕ_i); Onsets



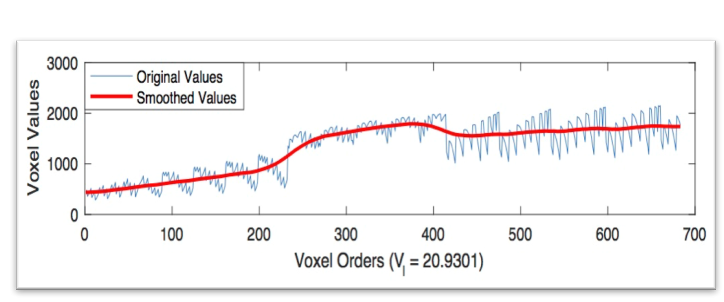
Snapshot

Feature Extraction



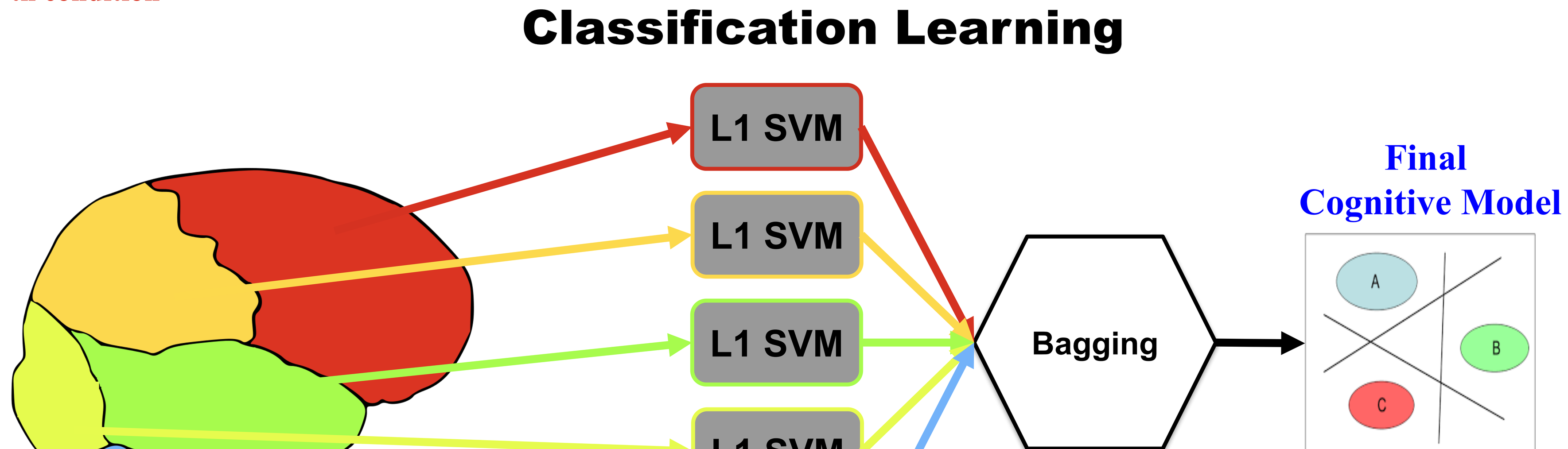
Anatomical Atlas — **Functional Activities**

1. **Normalization:** $\mathbf{T}_i$: $\widehat{\beta}_i \in \mathbb{R}^m \rightarrow \beta_i \in \mathbb{R}^n$, **where** $\mathbf{T}_i = \arg\min(NMI(\widehat{\beta}_i, \mathbf{Ref}))$ (Original Space → Standard Space)
2. **Segmentation:** $\mathbf{\Theta}_{(j,\ell)} = \{\theta_j^k \mid \theta_j^k \in \mathbf{\Theta}_j \text{ and } k \in \mathbf{A}_\ell\}$ ($\mathbf{\Theta}_{(j,\ell)}$: voxels for ℓ-th region j-th condition; $\mathbf{\Theta}_j$: voxels for j-th condition; $\mathbf{A}_\ell$: voxels for ℓ-th region)
3. **Smoothing:** $\forall \ell = L1 \dots L2 \rightarrow \mathbf{X}_{(j,\ell)} = \mathbf{\Theta}_{(j,\ell)} * \mathbf{V}_\ell$, $\mathbf{X}_j = \{\mathbf{X}_{(j,L1)}, \dots, \mathbf{X}_{(j,\ell)}, \dots \mathbf{X}_{(j,L2)}\}$ ($\mathbf{X}_j$: Extracted features for j-th condition; $\mathbf{V}_\ell$: Gaussian kernel)



Smoothed voxels belonged to a unique anatomical region

Classification Learning



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

DATASETS

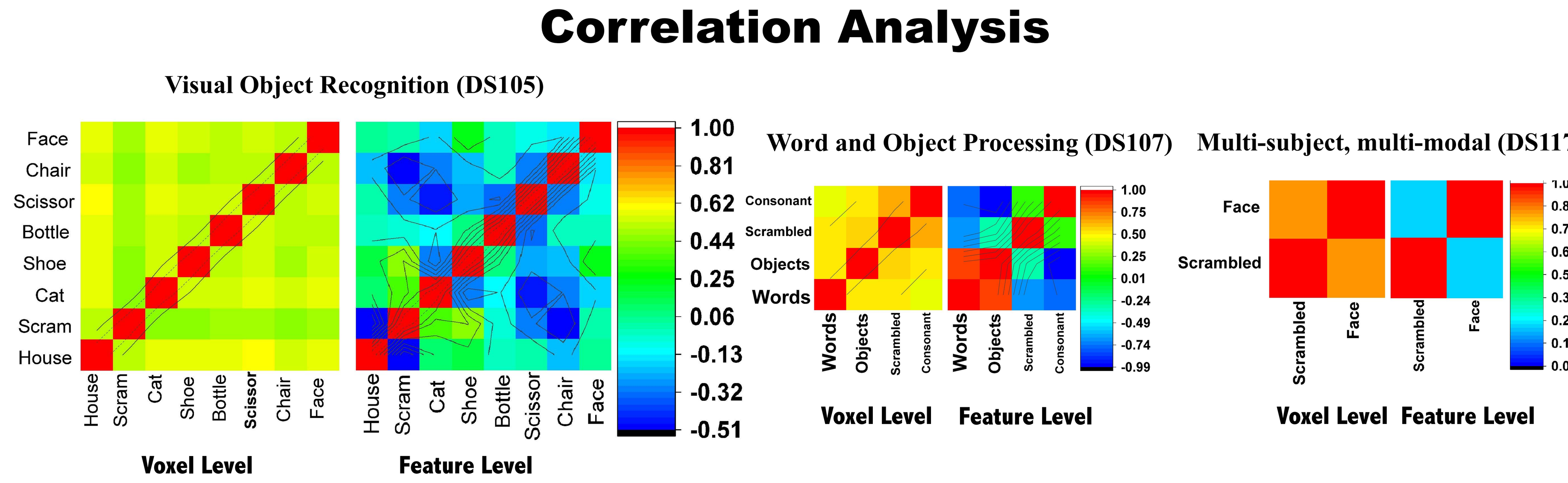
- **This paper utilizes three data sets, shared by www.openfmri.org, for running empirical studies:**

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.

EXPERIMENTAL RESULTS

Correlation Analysis



Classification Analysis

Table 1: Accuracy of binary predictors

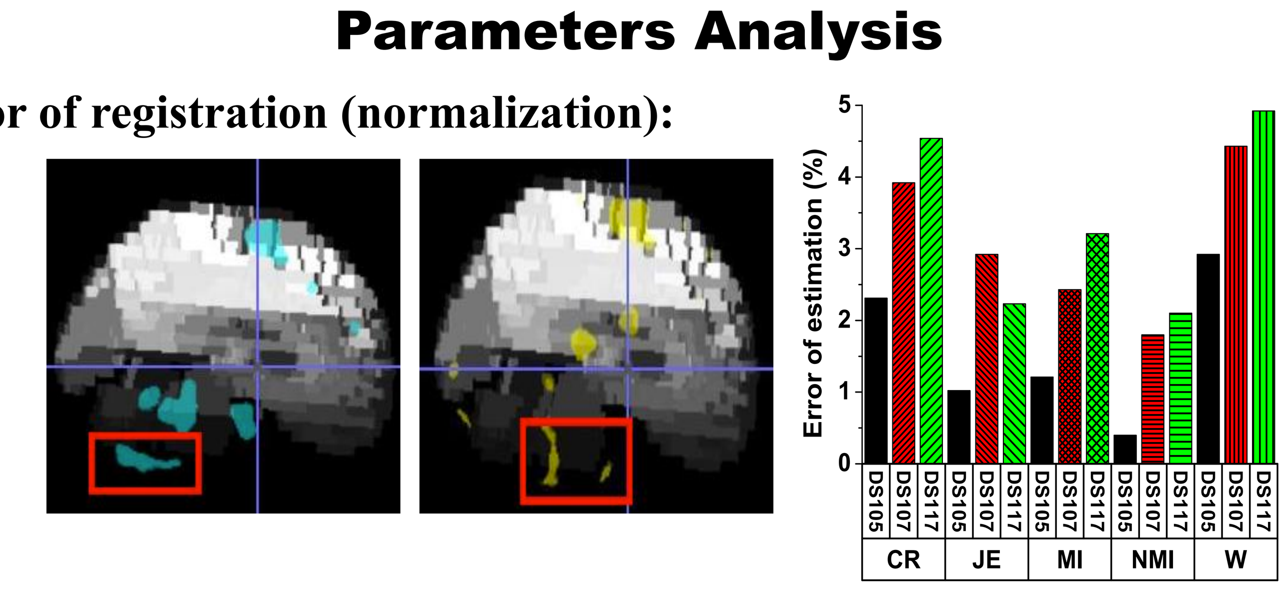
Data Sets	SVM	Graph Net	Elastic Net	L1-Reg. SVM	Osher et al.	Proposed method
DS105: Objects vs. Scrambles	71.65±0.97	81.27±0.59	83.06±0.36	85.29±0.49	90.82±1.23	**94.32±0.16**
DS107: Words vs. Others	82.89±1.02	78.03±0.87	88.62±0.52	86.14±0.91	90.21±0.83	**92.04±0.09**
DS107: Consonants vs. Others	67.84±0.82	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

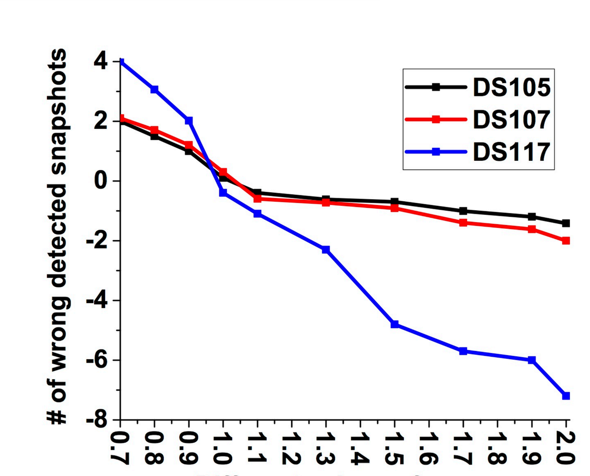
Data Sets	SVM	Graph Net	Elastic Net	L1-Reg. SVM	Osher et al.	Proposed method
DS105: Objects vs. Scrambles	68.37±1.01	70.32±0.92	82.22±0.42	80.91±0.21	88.54±0.71	**93.25±0.92**
DS107: Words vs. Others	80.76±0.91	77.91±1.03	86.35±0.39	84.23±0.57	87.61±0.62	**91.86±0.17**
DS107: Consonants vs. Others	63.84±1.45	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

- **The error of registration (normalization):**



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



CONCLUSION

There is a wide range of challenges in the Multivariate Pattern (MVP) techniques, i.e. decreasing noise and sparsity, defining effective regions of interest (ROIs), visualizing results, and the cost of brain studies. In overcoming these challenges, 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 clearly show the superiority of our proposed method in comparison with state-of-the-art methods. In addition, the time complexity of the proposed method is naturally lower than the previous methods because it employs a snapshot of brain image for each stimulus rather than using the whole of time series. In future, we plan to apply the proposed method to different brain tasks such as risk, emotion and etc.