LEHIGH UNIVERSITY

Tensor Learning in Brain Network Analysis

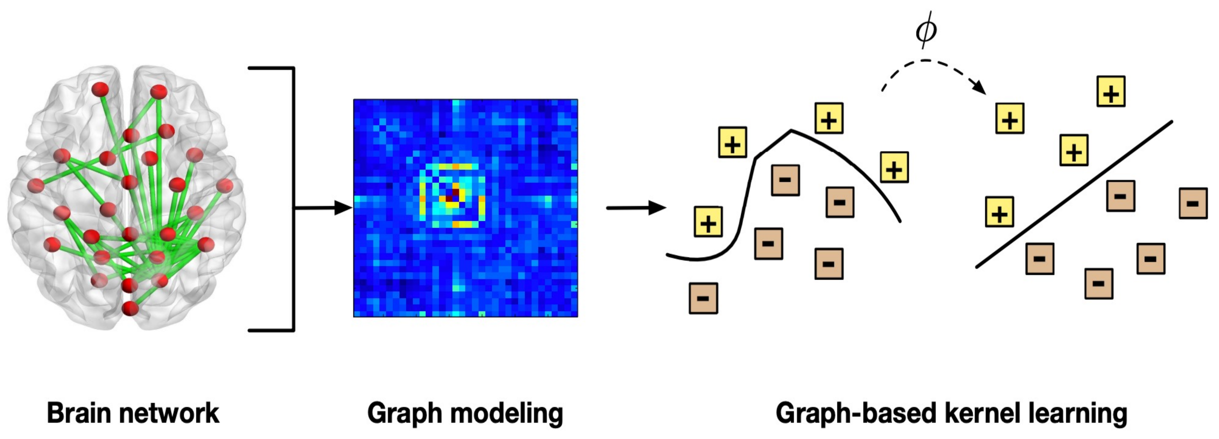
Jun Yu

Homepage: junfish.github.io

E-mail: juy220@lehigh.edu

Structure-Preserving Graph Kernel for Brain Network Classification[1]

Brain network analysis is of great importance in clinical diagnosis and treatments. we present a novel graph-based kernel learning approach for brain network classification. Specifically, we demonstrate how to exploit the natural graph structure of brain networks to encode prior knowledge in the kernel using the tensor product operator. For each brain network, we first propose to apply sparse matrix factorization with a symmetric constraint to extract tensor product based approximation. We then use them to derive a structure-persevering symmetric graph kernel to be fed into the support vector machine (SVM). The framework is shown as follows:



As mentioned above, the brain network is encoded as a symmetric matrix. We firstly try to learn a sparse structure $\widetilde{\boldsymbol{X}}$, as is to be expected, by tensor product:

$$\min_{\mathbf{a}_r} \|\mathbf{X} - \sum_{r=1}^{R} \mathbf{a}_r \otimes \mathbf{a}_r\|_F^2 + \lambda \sum_{r=1}^{R} \|\mathbf{a}_r\|_1,$$

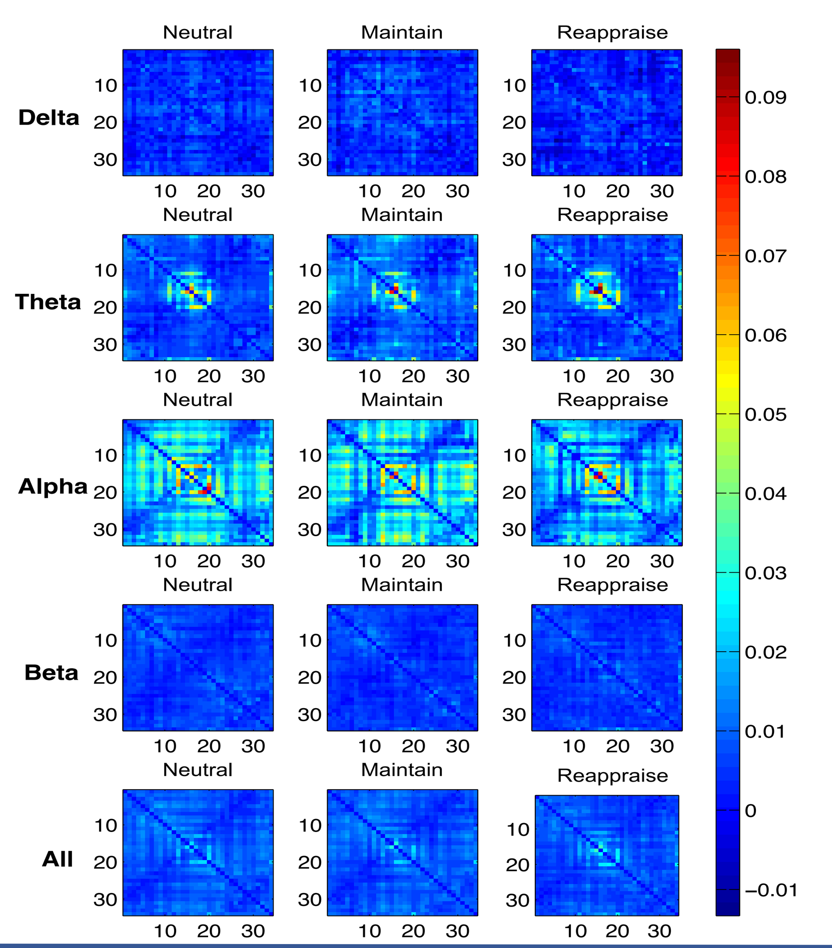
where R is the predefined rank of matrix $\widetilde{\boldsymbol{X}}$. Then, we use a graph-based feature mapping function $\phi(\cdot)$ to transform the graph data into a higher dimensional feature space (Hilbert space). Furthermore, we prove that the matrix factorization can be mapped into the outer product feature space:

$$\phi : \sum_{r=1}^{R} \mathbf{x}_r \otimes \mathbf{x}_r \to \sum_{r=1}^{R} \phi(\mathbf{x}_r) \otimes \phi(\mathbf{x}_r)$$

As usual, the kernel function can be used to simplify this mapping:

$$\begin{aligned} \kappa(\mathbf{X}, \mathbf{Y}) &= \kappa\Big(\sum_{r=1}^{R} \mathbf{x}_r \otimes \mathbf{x}_r, \sum_{r=1}^{R} \mathbf{y}_r \otimes \mathbf{y}_r\Big) \\ &= \Big\langle \sum_{r=1}^{R} \phi(\mathbf{x}_r) \otimes \phi(\mathbf{x}_r), \sum_{r=1}^{R} \phi(\mathbf{y}_r) \otimes \phi(\mathbf{y}_r) \Big\rangle \\ &= \sum_{p=1}^{R}\sum_{q=1}^{R} \kappa(\mathbf{x}_p, \mathbf{y}_q)\kappa(\mathbf{x}_p, \mathbf{y}_q). \end{aligned}$$

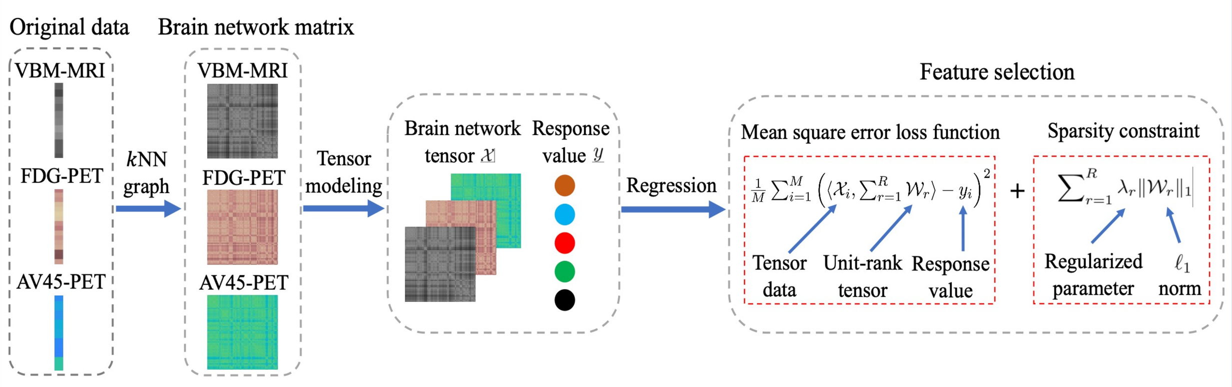
The average EEG-connectome during neutral, maintain, and reappraise and the classification accuracy by competing methods in five different frequency bands are shown as follows:



Category	Method	Frequency Band				
		Delta	Theta	Alpha	Beta	All
Traditional	Edge	42.42	54.55	51.52	51.52	45.45
	CC	54.55	54.55	42.42	51.52	42.42
	CPL	48.48	42.42	45.45	48.48	39.39
	gSpan	39.39	51.52	39.39	54.55	48.48
	DuSK–2D	51.52	63.64	51.51	51.52	54.55
	DuSK–3D	57.58	57.58	57.58	54.55	48.48
	DuSK–4D	54.55	54.55	51.52	54.55	57.58
Deep Learning	CNN–2D	51.11	43.71	43.07	42.54	41.48
	CNN–3D	46.67	45.93	41.48	57.04	44.44
	GCN	41.31	48.08	41.01	40.61	37.37
Ours	$\text{SSGK}_{\text{w/o sparse}}$	57.58	66.67	63.64	54.55	57.58
	SSGK	**63.64**	**69.70**	**72.73**	**60.61**	**57.58**

Tensor-based Multi-Modality Feature Selection for Alzheimer's Diagnosis[2]

The assessment of Alzheimer's Disease (AD) and Mild Cognitive Impairment (MCI) associated with brain changes remains a challenging task. We propose a novel tensor-based multi-modality feature selection and regression model for the diagnosis and biomarker identification of AD and MCI from normal controls. We present the practical advantages of our method for the analysis of ADNI data using three imaging modalities (VBM-MRI, FDG-PET and AV45-PET) with clinical parameters of disease severity and cognitive scores. The framework of our method is shown as follows:



To exploit the high-dimensional structure and correlation in the tensor representation, we employ the following sparse and low-rank tensor regression model:

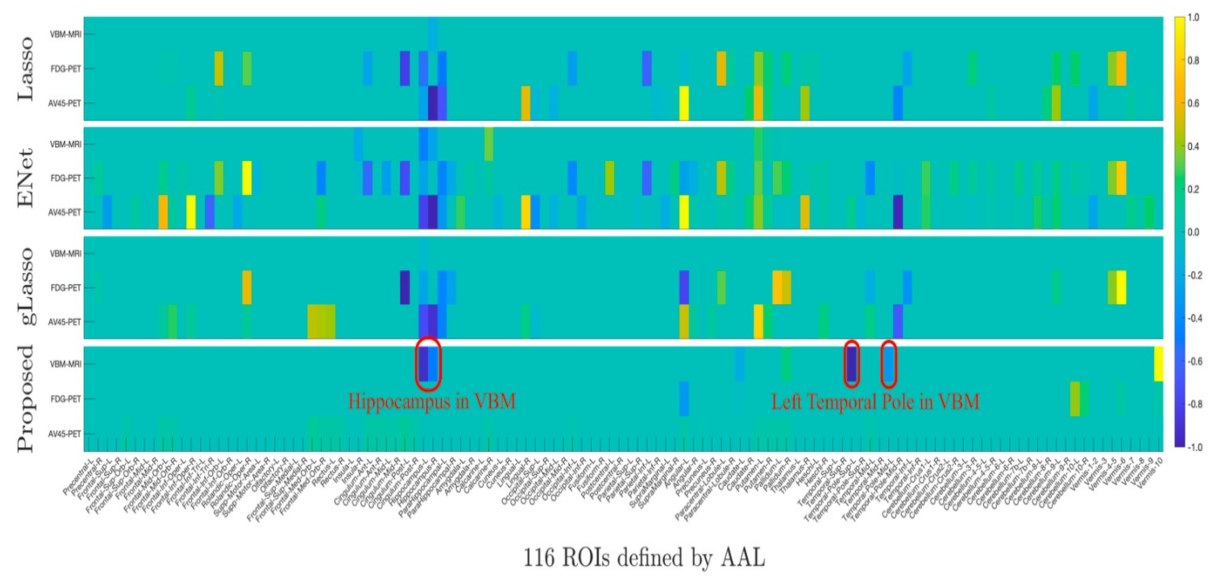
$$\min_{\mathcal{W}_r} \frac{1}{N}\sum_{i=1}^{N}\Big(\langle \mathcal{X}_i, \sum_{r=1}^{R} \mathcal{W}_r\rangle - y_i\Big)^2 + \sum_{r=1}^{R} \lambda_r \|\mathcal{W}_r\|_1, \quad \text{s.t.} \quad \text{CP-rank}(\mathcal{W}_r) \le 1$$

where $\mathcal{W}_r = \mathbf{w}_r^{(1)} \otimes \cdots \otimes \mathbf{w}_r^{(N)}$ is a unit-rank tensor defined upon the CP rank. Here we adopt the fast Stagewise Unit-Rank tensor Factorization (SURF) algorithm to solve the optimization problem above.

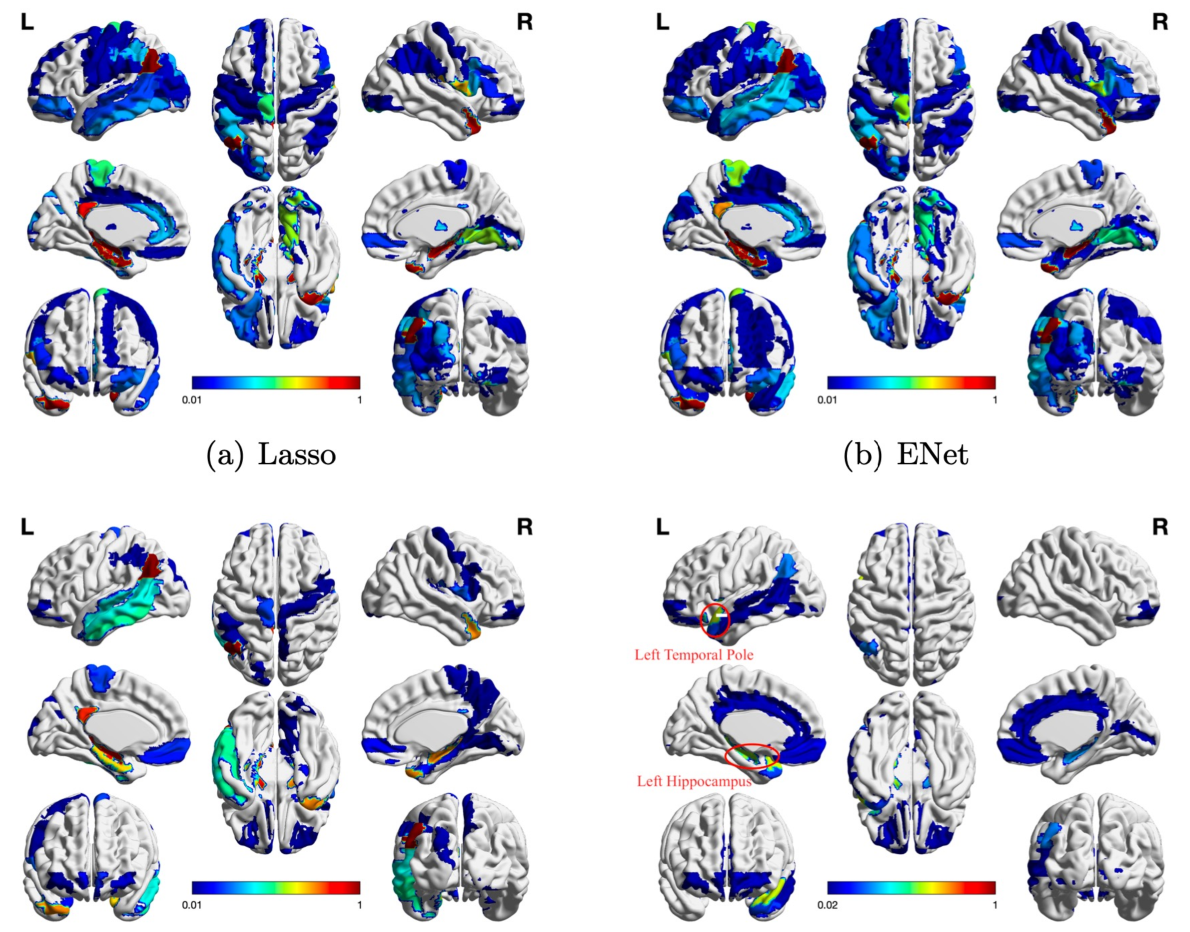
Performance comparison over different feature sizes on ADNI dataset is shown below:

Features	Scores	Metrics	PCA+LR	Lasso	ENet	gLasso	Proposed
116 × 3	DSS	RMSE↓	$.39 \pm .02$	$.37 \pm .01$	$.36 \pm .02$	$.33 \pm .01$	$\mathbf{.31 \pm .01}$
		Sparsity↑	$.00 \pm .00$	85.12 ± 4.14	77.41 ± 7.54	$88.39 \pm .56$	$\mathbf{97.16 \pm .43}$
	ADAS-13	RMSE↓	$.20 \pm .02$	$.23 \pm .01$	$.23 \pm .01$	$.18 \pm .01$	$\mathbf{.17 \pm .00}$
		Sparsity↑	$.00 \pm .00$	85.17 ± 2.76	82.82 ± 2.52	$91.21 \pm .88$	$\mathbf{98.56 \pm .41}$
	MMSE	RMSE↓	$.27 \pm .01$	$.26 \pm .20$	$.26 \pm .02$	$.24 \pm .02$	$\mathbf{.22 \pm .02}$
		Sparsity↑	$.00 \pm .00$	87.70 ± 6.05	85.23 ± 5.46	91.95 ± 1.06	$\mathbf{96.67 \pm .16}$
116 × 116	DSS	RMSE↓	$.36 \pm .03$	$.38 \pm .01$	$.38 \pm .01$	$.31 \pm .02$	$\mathbf{.29 \pm .01}$
		Sparsity↑	$.00 \pm .00$	$98.34 \pm .55$	97.39 ± 1.74	$97.69 \pm .47$	$\mathbf{99.75 \pm .04}$
	ADAS-13	RMSE↓	$.23 \pm .01$	$.23 \pm .01$	$.23 \pm .01$	$.18 \pm .01$	$\mathbf{.16 \pm .01}$
		Sparsity↑	$.00 \pm .00$	$98.77 \pm .09$	$97.99 \pm .80$	$98.86 \pm .30$	$\mathbf{99.89 \pm .02}$
	MMSE	RMSE↓	$.29 \pm .02$	$.27 \pm .01$	$.27 \pm .01$	$.22 \pm .02$	$\mathbf{.20 \pm .02}$
		Sparsity↑	$.00 \pm .00$	$99.10 \pm .24$	$98.38 \pm .86$	$98.84 \pm .26$	$\mathbf{99.64 \pm .03}$
116 × 116 × 3	DSS	RMSE↓	$.34 \pm .02$	$.37 \pm 0.02$	$.36 \pm .02$	$.32 \pm .01$	$\mathbf{.28 \pm .01}$
		Sparsity↑	$.00 \pm .00$	$99.38 \pm .18$	$98.42 \pm .76$	$98.49 \pm .25$	$\mathbf{99.97 \pm .01}$
	ADAS-13	RMSE↓	$.19 \pm .02$	$.23 \pm .01$	$.22 \pm .01$	$.18 \pm .01$	$\mathbf{.14 \pm .01}$
		Sparsity↑	$.00 \pm .00$	$99.55 \pm .06$	$99.35 \pm .17$	$98.85 \pm .25$	$\mathbf{99.97 \pm .00}$
	MMSE	RMSE↓	$.22 \pm .02$	$.26 \pm .02$	$.26 \pm .02$	$.21 \pm .02$	$\mathbf{.18 \pm .02}$
		Sparsity↑	$.00 \pm .00$	$99.77 \pm .05$	$99.62 \pm .12$	$98.70 \pm .44$	$\mathbf{99.97 \pm .01}$

To verify that our proposed method can learn a better sparse structure, we visualize the coefficient weights in terms of each imaging modality:



We also use BrainNet Viewer to visualize the brain structure and highlight the regions that the proposed method used to make the predictions against other compared methods:



(a) Lasso (b) ENet (c) gLasso (d) Proposed

Tensor-based Multivariate Regression for Alzheimer's Diagnosis[3]

Based on previous study, we are further proposing multivariate tensor regression. Specifically, we introduce K related losses to regress jointly as follows:

$$\min \sum_{k=1}^{K} \mathcal{L}_k\left(\langle \mathcal{W}^k, \mathcal{X}\rangle, y_k\right) + \lambda\Omega(\mathcal{W}),$$

where $\mathcal{L}_k$ indicates specific loss function for respective response variate, and $\mathcal{W} = [\mathcal{W}^1, \cdots, \mathcal{W}^K] \in \mathbb{R}^{D_1 \times \cdots \times D_M \times K}$ denotes the final coefficients tensor.

If we introduce $\mathbf{e}_k = [\underbrace{0, \cdots, 0}_{k-1}, 1, 0, \cdots, 0]^{\mathrm{T}} \in \mathbb{R}^K$ as the response indicator vector and the Mean Square Error (MSE) as the loss function, the final optimization problem of sparse multivariate tensor regression is as follows:

$$\min_{\mathcal{W}_r} \frac{1}{N}\sum_{i=1}^{N}\sum_{k=1}^{K}\Big(\langle \mathcal{X}_i \otimes \mathbf{e}_k, \sum_{r=1}^{R} \mathcal{W}_r\rangle - y_{i,k}\Big)^2 + \sum_{r=1}^{R} \lambda_r \|\mathcal{W}_r\|_1$$

$$\text{s.t.} \quad \text{CP-rank}(\mathcal{W}_r) \le 1.$$

The preliminary results compared with single variate regression are shown as follows:

Features	Scores	Metrics	Single Variate	Multi-Variate
116 × 3	DSS	RMSE↓	0.336 ± 0.016	$\mathbf{0.327 \pm 0.009}$
	ADAS-13	RMSE↓	$\mathbf{0.141 \pm 0.032}$	0.145 ± 0.019
	MMSE	RMSE↓	0.156 ± 0.019	$\mathbf{0.142 \pm 0.011}$
	Total	RMSE↓	$0.397 \pm .0251$	$\mathbf{0.385 \pm 0.010}$
116 × 116	DSS	RMSE↓	$\mathbf{0.328 \pm 0.010}$	$\mathbf{0.328 \pm .010}$
	ADAS-13	RMSE↓	0.150 ± 0.021	$\mathbf{0.148 \pm 0.031}$
	MMSE	RMSE↓	0.154 ± 0.021	$\mathbf{0.153 \pm 0.016}$
	Total	RMSE↓	0.392 ± 0.019	$\mathbf{0.391 \pm 0.021}$
116 × 116 × 3	DSS	RMSE↓	0.318 ± 0.021	$\mathbf{0.317 \pm 0.010}$
	ADAS-13	RMSE↓	$\mathbf{0.149 \pm 0.031}$	0.151 ± 0.013
	MMSE	RMSE↓	0.155 ± 0.020	$\mathbf{0.142 \pm 0.011}$
	Total	RMSE↓	0.384 ± 0.029	$\mathbf{0.379 \pm 0.010}$

References

[1] Jun Yu, Zhaoming Kong, Aditya Kendre, Hao Peng, Carl Yang, Lichao Sun, Alex Leow, and Lifang He. "Structure-Preserving Graph Kernel for Brain Network Classification." Accepted in IEEE 19th Inter-national Symposium on Biomedical Imaging (ISBI), 2022.

[2] Jun Yu, Zhaoming Kong, Liang Zhan, Li Shen, Lifang He. "Tensor-based Multi-Modality Feature Selection for Alzheimer's Disease Diagnosis." Submitted to the 25th International Conference on Medical Image Computing and Computer-Assisted Intervention (MICCAI). Springer, 2022.

[3] Jun Yu, Li Shen, Lifang He. "Tensor-based Multivariate Regression for Alzheimer's Disease Diagnosis." In progress.