

## Outline

- **Background:** Deep learning models have several challenges on **efficiency, interpretability** and **reliability**, which are major obstacles for widespread applications of ML technology.
- **Goal:** Our team aims to develop innovative machine learning theory, models and algorithms to address these challenges. In particular, tensor methods are exploited with expectation to develop efficient, robust and interpretable approaches, with applications to computer vision, neuroscience and AI for science.

## Robust generalization of adversarial training

**Problem:** Adversarial training is the most effective approach to defend DNN against adversarial attacks. However, there are atypical samples that hurt its robust generalization performance.

(Zhang et al. AAAI'23) **Loss function with reweighting**

**Key techniques**

- Reweight training and adversarial samples.
- Self-supervised technique to avoid overfitting on atypical adversarial samples.
- Memory bank is built as prototypes while the memorization weight is calculated based on the distance to prototypes.

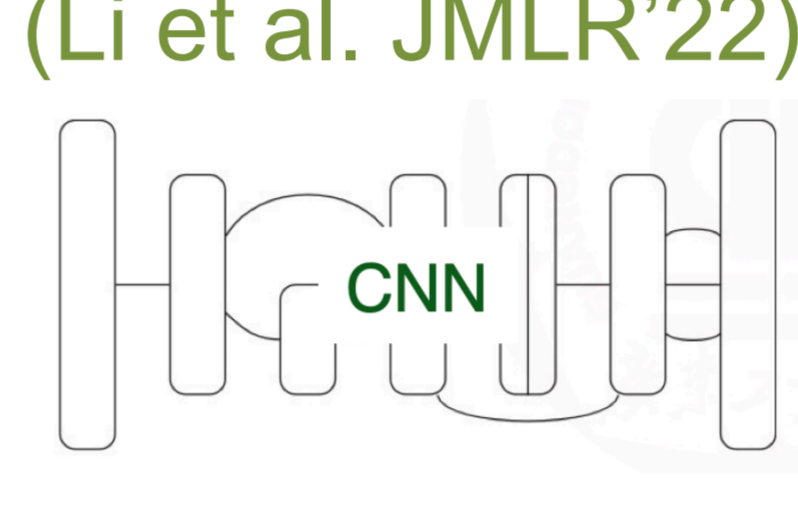
$$\begin{aligned} \ell_{AT}^{MeoW}(f_{\theta}(\hat{\mathbf{x}}_i), y_i) \\ = \alpha \cdot w^+(\mathbf{F}) \cdot \text{CE}(\mathbf{p}(\mathbf{x}; \theta), y) \\ + (1 - \alpha) \cdot w^-(\hat{\mathbf{F}}) \cdot \text{CE}(\mathbf{p}(\hat{\mathbf{x}}; \theta), y). \end{aligned}$$



The proposed technique can enhance AT's robust generalization performance by boosting both robust accuracy and natural accuracy.

## Interpretability of learning models

- Understanding CNN from Volterra convolution perspective (Li et al. JMLR'22)



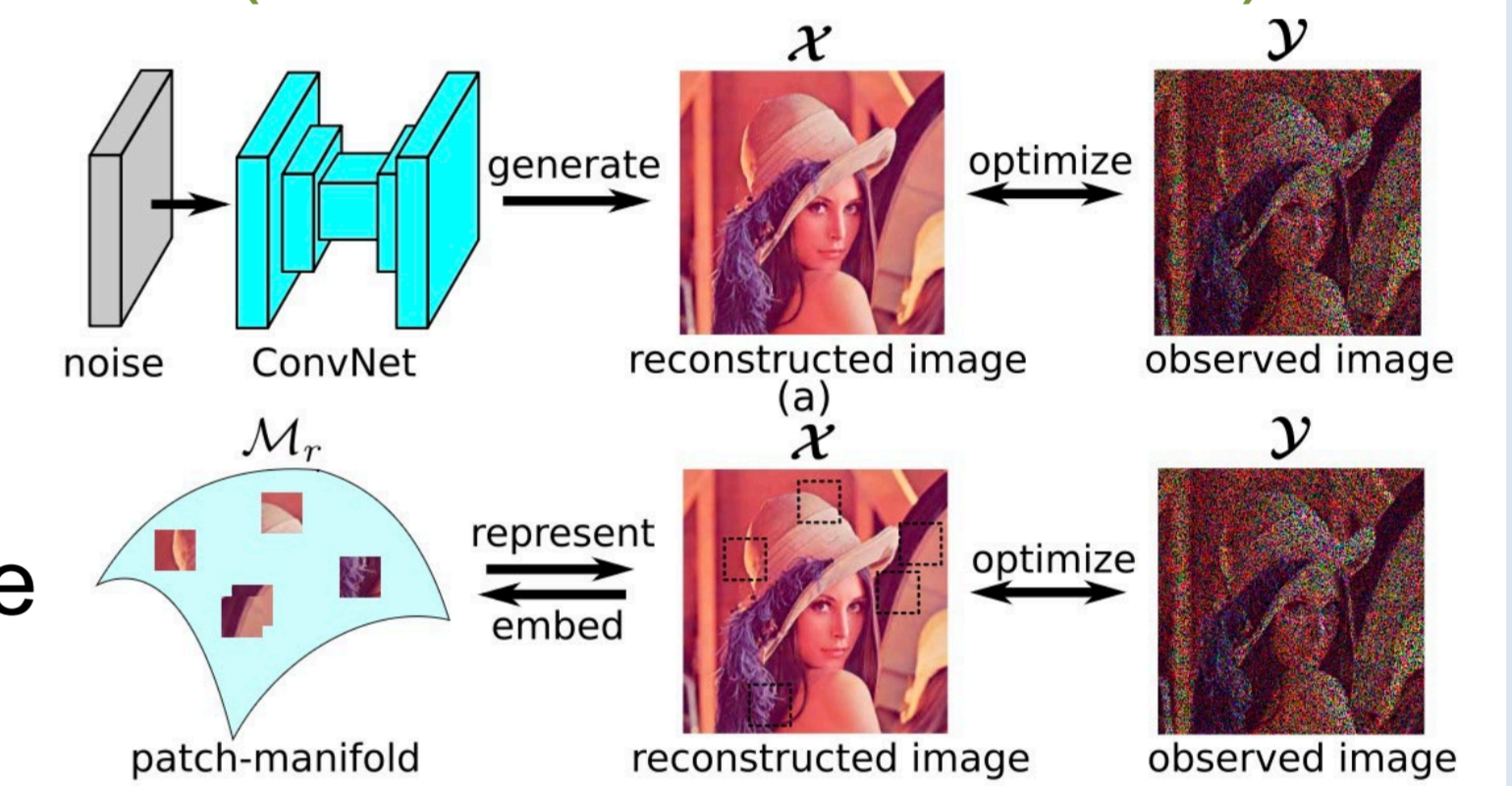
$$\left( \sum_{n=0}^{+\infty} \mathbf{H}_n * \mathbf{x}^n \right) (t) = \sum_{n=0}^{+\infty} \int_{-\infty}^{+\infty} \dots \int_{-\infty}^{+\infty} H_n(\tau_1, \dots, \tau_n) \prod_{i=1}^n (x(t - \tau_i) d\tau_i)$$

NOT n-dimensional convolution

VC net is convenient for the theoretical analysis of CNN and the **robustness analysis** w.r.t. perturbations.

- **Manifold modeling in tensor embedded space:** An interpretable approach as alternative to deep image prior

(Yokota et al. TNNLS'22)

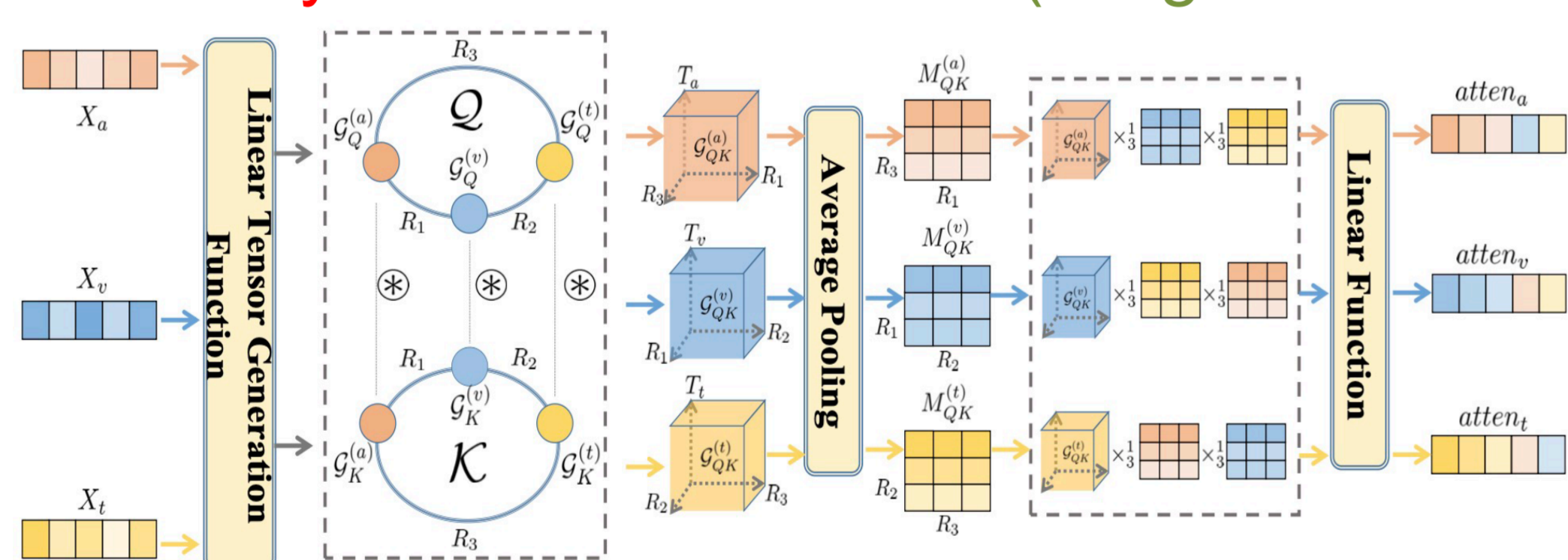


- Fast algorithm for tensor completion, **100x speedup** while maintaining high accuracy (Yokota et al. CVPR'22)

## Highly expressive model with tensor network

**Problem:** How to boost expressive power without significantly increasing the model size?

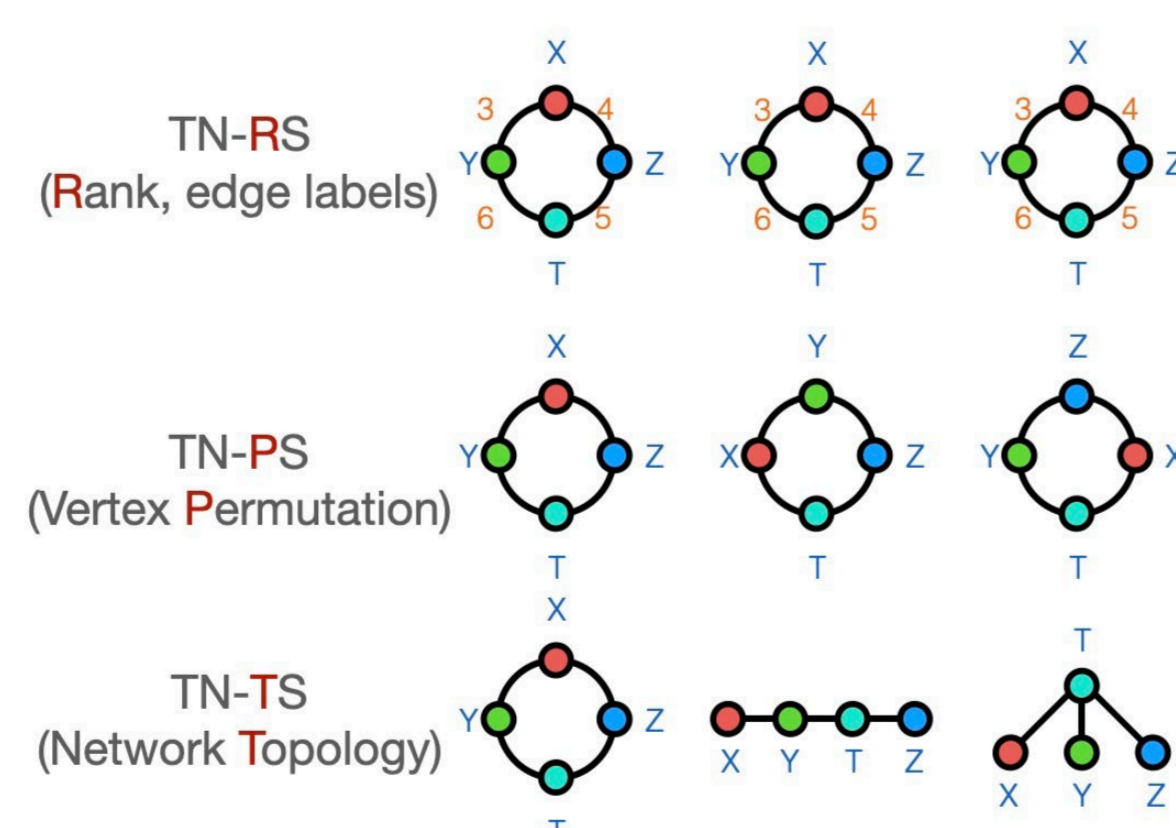
**Multimodal attention** (Tang et al. IJCAI'22)



- Multimodal attention block by tensor ring representation
- Flexible and linearly scalable to the number of modality

**Tensor network structure search** (Li et al. ICML'22)

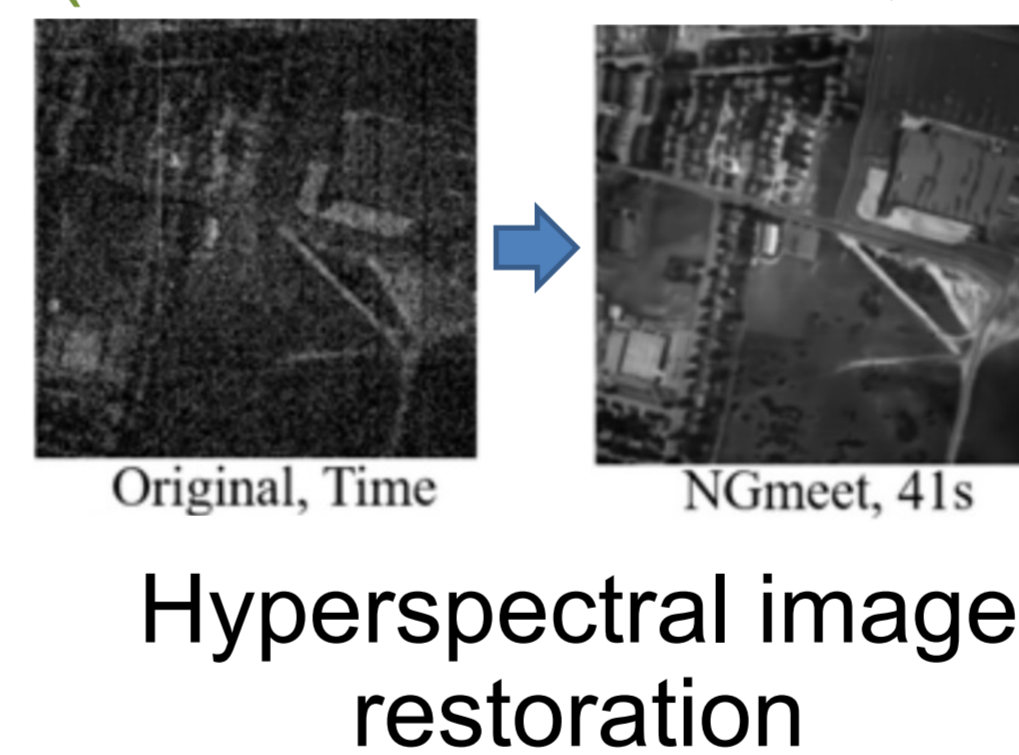
- Find highly expressive TN with small model size
- Efficient learning algorithm by randomly sampling



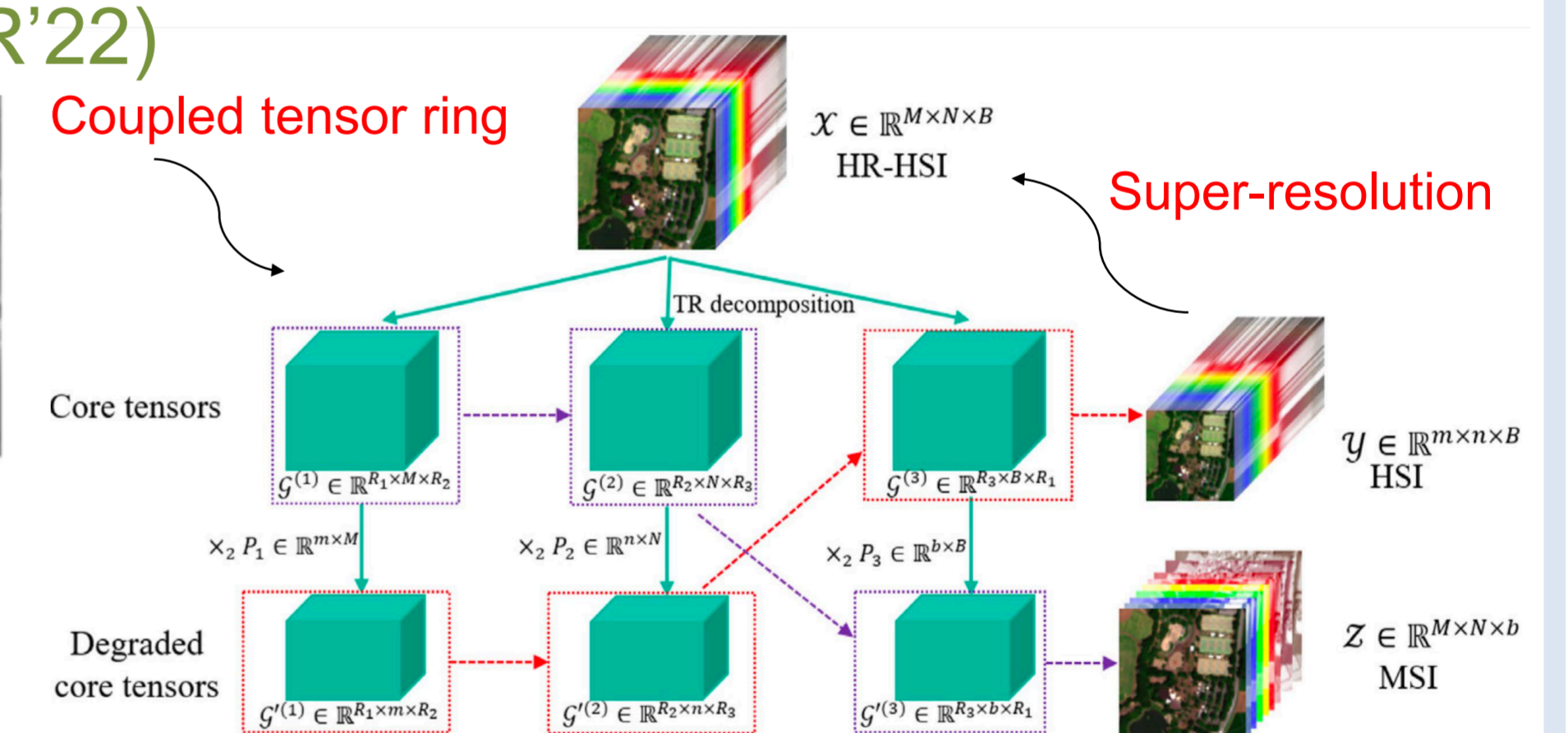
TN can be leveraged to boost expressive power of DL models with an efficient model size. TN structure search is underexplored.

## Applications to hyperspectral image and EEG

(He et al. TPAMI'22, PR'22)



Hyperspectral image restoration

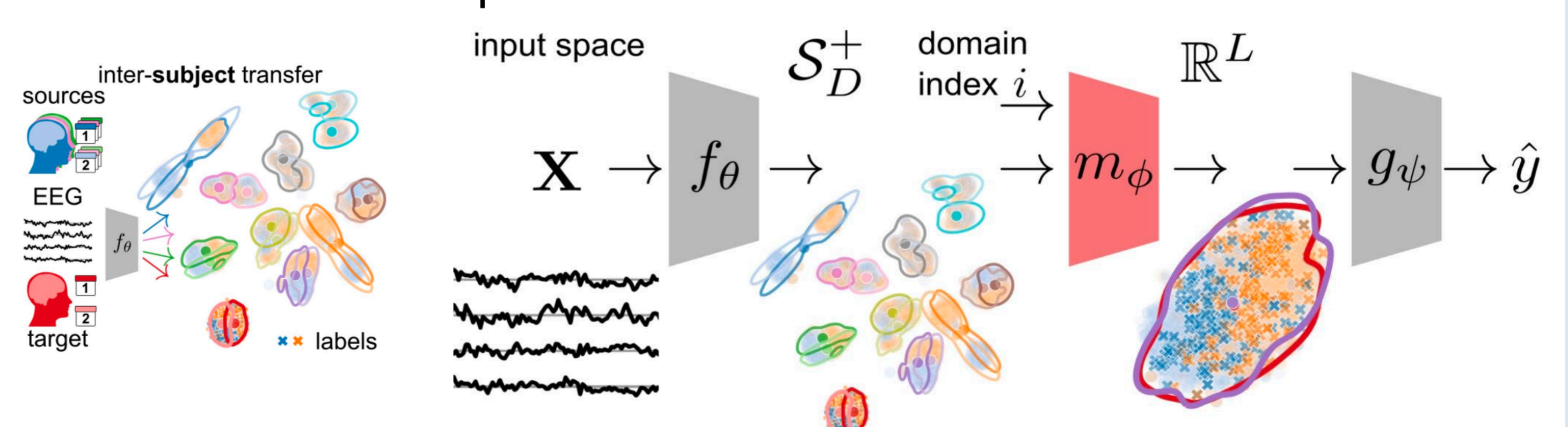


Tensor methods can achieve better results than deep learning and training dataset is not required.

(Kobler et al. NeurIPS'22)

**Problem:** poor generalization across sessions and subjects

**Solution:** domain-specific momentum batch normalization



Obtain the state-of-the-art results on inter-subject transfer learning

## Related Publications

- J. Zhang, Y. Hong, and Q. Zhao, "Memorization weights for instance reweighting in adversarial training," in AAAI 2023.
- T. Li, G. Zhou, Y. Qiu, and Q. Zhao, "Understanding convolutional neural networks from theoretical perspective via Volterra convolution," JMLR 2022.
- T. Yokota, H. Hontani, Q. Zhao, and A. Cichocki, "Manifold modeling in embedded space: An interpretable alternative to deep image prior," IEEE TNNLS 2022.
- R. Yamamoto, H. Hontani, A. Imakura, and T. Yokota, "Fast algorithm for low-rank tensor completion in delay-embedded space," in CVPR 2022.
- J. Tang, K. Li, M. Hou, X. Jin, W. Kong, Y. Ding, and Q. Zhao, "MMT: Multi-way multi-modal transformer for multimodal learning," in IJCAI 2022.
- C. Li, J. Zeng, Z. Tao, and Q. Zhao, "Permutation search of tensor network structures via local sampling," in ICML 2022.
- W. He, Q. Yao, C. Li, N. Yokoya, Q. Zhao, H. Zhang, and L. Zhang, "Non-local meets global: An iterative paradigm for hyperspectral image restoration," IEEE TPAMI 2022.
- R. Kobler, J. Hirayama, Q. Zhao, and M. Kawanabe, "SPD domain-specific batch normalization to crack interpretable unsupervised domain adaptation in EEG," in NeurIPS 2022.