ICASSP 2023 - 2023 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP) | 978-1-7281-6327-7/23/$31.00 ©2023 IEEE | DOI: 10.1109/ICASSP49357.2023.10095707

TENSOR-BASED COMPLEX-VALUED GRAPH NEURAL NETWORK FOR DYNAMIC COUPLING MULTIMODAL BRAIN NETWORKS

Yanwu Yang[1,2]*,Guoqing Cai*[1] *Chenfei Ye*[1,2]*,Yang Xiang*[2*]*, Ting Ma*[1,2*]

[1]Harbin Institute of Technology at Shenzhen, China
[2]Peng Cheng Laboratory, Shenzhen, China

ABSTRACT

The multi-modal neuroimage study has dramatically facilitated disease diagnosis. Tensor-based methods are commonly used to represent multi-modal data as multi-dimensional arrays and usually implement matrix decomposition. These methods can be seen as a linear algebraic way for the lossy compression of an array. However, involved lossy operations might have a negative impact on performance, and overlook underlying important complementary information between modalities. This study proposes a Tensor-based Complex-valued Graph Neural Network (TC-GNN) to model multimodal neuroimages as complex-valued tensor graphs by investigating underlying complementary associations and cross-modality message aggregation. Experiments on two real-world datasets demonstrate our method's consistent improvements and superiority over other baseline models in multi-modal brain disease analysis.

Index Terms— Multimodal, Tensor, Graph neural network, Neuroimage, Gating.

1. INTRODUCTION

Advanced neuroimaging tools, e.g. functional MRI (fMRI) and Diffusion Tensor Imaging (DTI), have become promising candidates for brain disorder diagnosis, such as Alzheimer's disease (AD), and Parkinson's disease (PD) [1, 2]. Functional and structural brain networks, which are derived by mapping corresponding fMRI and DTI neuroimages into a brain parcellation template, play a key role in characterizing the interplay of anatomical and functional brain alterations. Among state-of-the-art deep learning technologies, Graph Neural Networks (GNN) has been widely applied in medical scenario, and have been state-of-art tools to relate graph representations for automatic diagnosis.

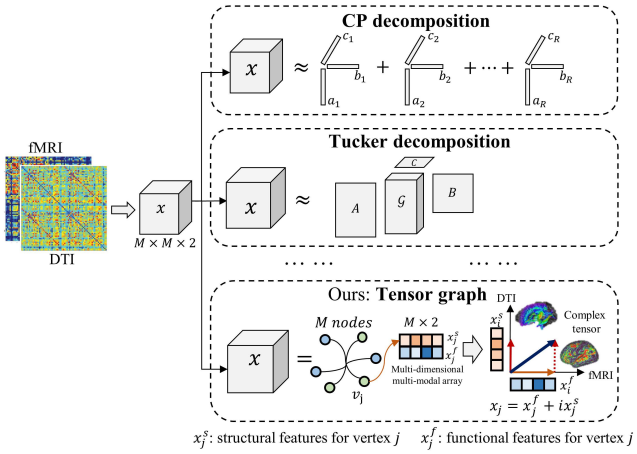

Fig. 1. Comparison of conventional tensor-based approaches (CP and Tucker Decomposition) with proposed complex-valued tensor graph.

Recently, integrating functional and structural networks has become an important topic of study. There is mounting evidence demonstrates that mapping the interplay of multiple modalities is more feasible to encode meaningful and regularized representations [3, 4, 5]. For example, [6] proposes to perform a two-layer convolution on the fMRI and DTI data simultaneously. [4] regularizes convolution on functional connectivity with structural graph Laplacian. The hypergraph strategy provides a way of associating multiple relationships, as well as capturing multimodal dependencies [7, 8].

A straightforward and natural way of investigating multimodal data is to model multimodal data as multi-dimensional arrays, which are known as tensors. Tensors have become powerful tools for the discovery of underlying hidden complex data structures in brain data as well as relating common patterns from independent components [9, 10]. Most existing tensor-based approaches utilize tensor decomposition, e.g. Canonical Polyadic (CP) and Tucker Decomposition (TD), to identify representative features along with improving learning efficiency [11, 12]. However, these approaches are proposed based on matrix approximation and factoriza-

* corresponding authors.
This study is supported by grants from the National Natural Science Foundation of China (62106113, 6227608, and 62106115), the Innovation Team and Talents Cultivation Program of National Administration of Traditional Chinese Medicine (NO:ZYYCXTD-C-202004), Basic Research Foundation of Shenzhen Science and Technology Stable Support Program (GXWD20201230155427003-20200822115709001), The National Key Research and Development Program of China (2021YFC2501202), and the Major Key Project of PCL.

tion, where involved lossy operations might decrease the performance. Moreover, these methods potentially ignore underlying cross-modality associations and fail to capture important intrinsic complementary information among modalities for multi-modal learning.

In this regard, we propose to investigate multi-modal data as complex-valued tensors. The idea behind this is to construct a composite tensor for learning latent multimodal representations and generating high-order generalizations of matrices to capture the consistency among modalities. As is shown in Figure 1, apart from conventional tensor-based approaches that decompose tensors into vectors or arrays, we introduce the complex-valued neural network (CVNN) [13] to encode tensor representations. We explicitly model multi-modal data into a complex-valued tensor graph, which allows for richer representations by expressing the neuron's output with multiple indices.

Moreover, in this study, we propose a Tensor-based Complex-valued Graph Neural Network (TC-GNN) for modeling with cross-embedding and cross-aggregation. In addition, existing complex-valued neural networks fail to consider neuronal synchrony, where the gating of information processing and dynamic binding of representations are ignored. In particular, we propose a gated complex-valued graph convolution layer for attentively modeling the interplay between modalities. Experiments on two real-world datasets demonstrate that our proposed method is a superior tool for multimodal brain disorder classification, which outperforms other baselines in prediction performance.

2. METHOD

2.1. Complex-valued Tensor Graph Construction

Functional and structural networks $\mathcal{X}_f, \mathcal{X}_s \in \mathbb{R}^{m\times m}$ are symmetric matrices derived by a template with m regions. A brain network is formulated as a graph $G = (\mathcal{V}, \mathcal{E})$, where the sets $\mathcal{V}$ and $\mathcal{E}$ denote the nodes and links.

Node. As is shown in Figure 1, the functional and structural features for vertex v_i are constructed into a two-dimensional array $x_i \in \mathbb{R}^{2\times m}$, where the node features for each modality are built by the i-th column or row in the matrices $\mathcal{X}$. The array is further formulated into a complex-valued tensor, where the functional and structural networks are represented as the real and imaginary parts.

Edge. Apart from conventional graph data, brain networks have unclear graph structures. A well-known strategy is by dynamic graph convolution [14], where the edges are dynamically embedded by $f_{\mathcal{E}} : \{\mathcal{X}_f, \mathcal{X}_s\} \to \mathcal{E}$.

2.2. Proposed TC-GNN

Figure 2 demonstrates the proposed TC-GNN, with dynamic complex-valued edge embedding (DCEE), complex-valued node embedding (CNE), and gated complex-valued graph convolution (GCGC) layers.

2.2.1. Dynamic Complex-valued Node Embedding.

The multi-modal node features are aggregated from the connection features to the centroid node with a complex-valued multi-layer perception embedding. The output features are obtained by a weighted sum. In detail, given a vertex v_j, the corresponding features h_j are obtained from its connections $\mathcal{N}_j = \{x_{j,1}, x_{j,2}, ..., x_{j,m}\}$:

$$\begin{aligned} h_j &= f_n(\mathcal{N}_j) \\ &= \sum_{k\in\mathcal{N}(j)} \mathrm{Re}(w_k)\,\mathrm{Re}(x_{j,k}) - \sum_{k\in\mathcal{N}(j)} \mathrm{Im}(w_k)\,\mathrm{Im}(x_{j,k}) \\ &\quad + i\sum_{k\in\mathcal{N}(j)} \mathrm{Re}(w_k)\,\mathrm{Im}(x_{j,k}) + i\sum_{k\in\mathcal{N}(j)} \mathrm{Im}(w_k)\,\mathrm{Re}(x_{j,k}) \end{aligned} \tag{1}$$

where Re and Im represent the real part and imaginary part respectively, and w are parameterized by the network.

2.2.2. Dynamic Complex-valued Edge Embedding.

Usually, the edge connections should be conditioned on the node field. We use a transformation matrix $\mathbf{W}$ to estimate the dynamic adjacency matrix $\mathbf{L}$ from the connectivity x by:

$$\mathbf{L} = \sigma(\hat{\mathbf{L}}) \tag{2}$$

$$\begin{aligned} \hat{\mathbf{L}}_{j,k} =& \mathrm{Re}(w_{j,k})\,\mathrm{Re}(x_{j,k}) - \mathrm{Im}(w_{j,k})\,\mathrm{Im}(x_{j,k}) \\ &+ i\,\mathrm{Im}(w_{j,k})\,\mathrm{Re}(x_{j,k}) + i\,\mathrm{Im}(w_{j,k})\,\mathrm{Re}(x_{j,k}) \end{aligned} \tag{3}$$

where σ is a sigmoid function, and $w_{j,k}$ denotes the element of the matrix $\mathbf{W}$ in the j-th row and k-th column.

2.2.3. Tensor-based complex-valued graph convolution.

The complex-valued graph convolution layer is proposed by replacing the real kernel with a complex-valued kernel. Given a complex-valued adjacency matrix $\mathbf{L} = \mathrm{Re}(\mathbf{L}) + i\,\mathrm{Im}(\mathbf{L})$ and an input feature $\mathbf{H} = \mathrm{Re}(\mathbf{H}) + i\,\mathrm{Im}(\mathbf{H})$, the complex-valued graph convolution is obtained by using a combination of real-valued convolutions:

$$\begin{aligned} \mathbf{L}\mathbf{H}^{l-1}\mathbf{W} =& (\mathrm{Re}(\mathbf{L}) + i\,\mathrm{Im}(\mathbf{L}))\cdot(\mathrm{Re}(\mathbf{H}^{l-1}) \\ &+ i\,\mathrm{Re}(\mathbf{H}^{l-1}))\cdot(\mathrm{Re}(\mathbf{W}) + i\,\mathrm{Im}(\mathbf{W})) \end{aligned} \tag{4}$$

2.2.4. Others

Complex-valued Activation Function. To avoid violating the Cauchy-Riemann equations, the $\mathbb{C}ReLU$ function is implemented as the activation for mapping heterogeneous regions of interest within the real and imaginary parts respectively: $\mathbb{C}ReLU(\mathbf{x}) = ReLU(\mathrm{Re}(\mathbf{x})) + iReLU(\mathrm{Im}(\mathbf{x}))$

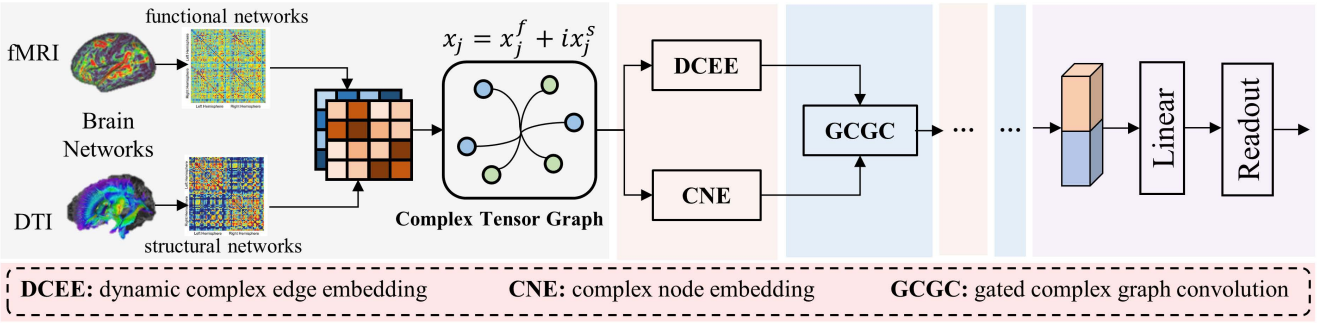

Fig. 2. The framework of the proposed gated complex-valued graph convolution network. Complex-valued tensor graphs derived from functional and structural networks are further fed into dynamic complex-valued edge embedding (DCEE) layers and complex-valued node embedding (CNE) layers for dynamic edge and node embedding. The gated complex-valued graph convolution (GCGC) layer is applied to aggregate complex-valued tensors attentively.

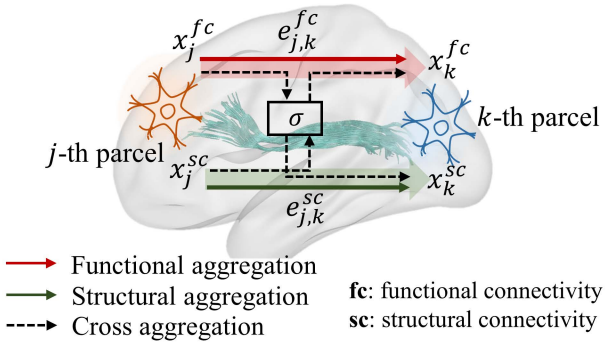

Fig. 3. The demonstration of the gating mechanism, which attentively aggregates cross-axis information.

Readout Function. Given a complex-valued feature vector, the complex-valued numbers are readout into real numbers by: $Readout(\mathbf{x}) = \sqrt{\mathrm{Re}(\mathbf{x})^2 + \mathrm{Im}(\mathbf{x})^2}$

2.3. Gated Complex-valued Graph Convolution

The gating mechanism has become a common tool in a neural network that facilitates the learning of longer-term relationships and protects the cell state from undesired updates [15]. Intuitively, the gates determine the amount of a signal to pass. In this study, it is implemented by an attention mechanism as an element-wise product and is applied to the graph message passing, resulting in the scaling of the intra-axis and inter-axis aggregation.

$$\begin{aligned} g_{r\to i} &= \sigma(\mathbf{W}_{11}\cdot \mathrm{Re}(\mathbf{z}) + \mathbf{W}_{12}\cdot \hat{\mathbf{z}} + \mathbf{b_1}) \\ g_{i\to r} &= \sigma(\mathbf{W}_{21}\cdot \mathrm{Im}(\mathbf{z}) + \mathbf{W}_{22}\cdot \hat{\mathbf{z}} + \mathbf{b_2}) \end{aligned} \tag{5}$$

where $\mathbf{W}$ and $\mathbf{b}$ are learnable parameters. $g_{r\to i}$ and $g_{i\to r}$ denote gates for the aggregation from real to imaginary and from imaginary to real respectively. And σ is the sigmoid activation function. Notably, $\mathbf{z}$ represents the transition state representations, where each element $\mathbf{z}(j)$ determines the j-th node by: $\mathbf{z}(j) = ||\{Avg_{k\in\mathcal{N}(j)}(x_k), Max_{k\in\mathcal{N}(j)}(x_k)\}$. While for $\hat{\mathbf{z}}$, it is obtained by $\sqrt{\mathrm{Re}(x_k)^2 + \mathrm{Im}(x_k)^2}$.

Moreover, in this study, we discard the spectral graph convolution and apply the spatial aggregation to simplify. To summarize, the GCGC layer can be formulated with eq (4) and (5) as:

$$\mathbf{H}^l = \sigma(\mathbf{L} \star \mathbf{H}^{l-1}\mathbf{W}) \tag{6}$$

$$\begin{aligned} \mathbf{L}\star\mathbf{H}^{l-1} &= \mathrm{Re}(\mathbf{L})\,\mathrm{Re}(\mathbf{H}^{l-1}) - \mathrm{Im}(\mathbf{L})\,\mathrm{Im}(\mathbf{H}^{l-1}) \\ &+ g_{r\to i}\odot i\,\mathrm{Im}(\mathbf{L})\,\mathrm{Re}(\mathbf{H}^{l-1}) + g_{i\to r}\odot i\,\mathrm{Re}(\mathbf{L})\,\mathrm{Im}(\mathbf{H}^{l-1}) \end{aligned} \tag{7}$$

3. EXPERIMENTS AND RESULTS

3.1. Dataset

Two datasets cover subjects with fMRI and DTI images are included for evaluation of the classification performance:

ADNI Dataset[1]: 114 subjects from the ADNI database are used for the prediction of the early stage of AD. 51 healthy controls (HC) and 63 mild cognitive impairment (MCI) diagnosed at baseline were included.

XHCMU Dataset [17]: The subjects in this dataset were recruited from the Xuanwu Hospital of Capital Medical University, including 70 HCs and 85 Parkinson's Disease (PD).

The fMRI and DTI images were pre-processed by reference to the standard fmriprep pipeline and MRtrix3. The mean time series of each region is extracted and calculated by the Pearson Correlation to construct the functional network. The number of streamlines representing the connection strength between each pair of regions is used to construct the structural network. The Schaefer-100 atlas is applied for mapping brain regions, where $m = 100$.

[1] http://www.adni-info.org/

Table 1. Classification results on two tasks in terms of accuracy (Acc %), sensitivity (Sen %), specificity (Spe %), and area under the curve (AUC %). The standard deviation cross folds are listed.

	Type	ADNI				XHCMU			
		Acc	Sen	Spe	AUC	Acc	Sen	Spe	AUC
BrainNetCNN [16]	CNN	74.4±9.6	72.8±16.1	77.7±14.0	76.7±11.5	75.6±9.7	77.6±16.5	87.9±9.3	77.6±10.7
TD&MLP	Tensor, NN	74.2±13.4	86.1±13.0	77.1±16.1	79.4±11.6	76.3±9.1	73.6±9.3	89.8±10.4	79.9±6.0
CP&MLP	Tensor, NN	82.6±10.2	93.1±12.4	81.4±14.9	81.2±15.1	77.8±5.3	78.9±12.5	**91.0±14.6**	80.5±8.3
M-GCN [4]	GNN	78.3±11.0	84.3±11.9	77.3±14.7	70.5±14.5	81.3±4.6	80.0±7.6	86.1±9.9	75.4±10.2
HGNN [7]	GNN	82.0±7.6	79.0±8.1	**90.8±14.2**	79.6±10.6	77.7±9.8	75.3±7.5	89.3±9.8	80.3±12.2
DHGNN [8]	GNN	84.8±5.9	83.6±11.0	86.6±11.7	86.6±7.9	78.1±6.8	**80.5±10.2**	83.7±12.3	81.2±4.4
Ours	Tensor, GNN	**88.7±6.5**	**94.1±7.3**	86.8±8.4	**90.6±7.0**	**85.0±5.0**	80.0±6.7	90.4±4.9	**86.6±6.9**

3.2. Implementation Details

For better comparison of the small data sets, 10-fold cross-validation is applied for evaluation. The learning rate is set as $3e-4$, and the weight decay is $5e-5$. All the models are trained for 600 epochs and would be early stopped when the loss has not been decreased for 50 epochs. We trained the models on one NVIDIA 2080-Ti GPU. Within the network, a complex-valued graph convolution group is comprised of a DCEE layer and a CNE layer, followed by a complex-valued leaky ReLU activation and a complex-valued dropout layer. Finally, a three-layer multi-layer perception is applied for classification. In this study, the number of complex-valued graph convolution groups and hidden size is decided within a grid search of $(1, 4)$ and $\{16, 32, 48, 64\}$ respectively.

We leverage BrainNetCNN [16] as a baseline, which is a state-of-the-art CNN method for brain network learning. Tensor-based approaches including Tucker (TD) and CP decomposition are implemented. The decomposed features are fed into a multi-layer perception for classification. The rank is decided with a grid search. Moreover, several well-established graph models are leveraged for comparison, including M-GCN [4], HGNN [7], and DHGNN [8]. For all the experiments, we evaluated the performance by diagnosis accuracy (Acc), sensitivity (Sen), specificity (Spe), and Area Under the Curve (AUC).

3.3. Multi-modal Graph Classification Performance

The classification results are listed in Table 1, where the best results are shown in bold, and the second best are with underline. We can see that CP decomposition outperforms BrainNetCNN, while TD performs worse on the ADNI dataset. The tensor-based methods are simple and efficient ways of learning multimodal data. Compared with these approaches, the graph-based models can generally improve the classification performances on both two datasets. Moreover, our approach outperforms the state-of-the-art graph models, where 3.9%/3.7% improvements in accuracy and 4%/5.4% improvements in AUC are achieved in the ADNI/XHCMU datasets. This can be attributed to both complex-valued tensor and graph networks in modeling multiple modalities.

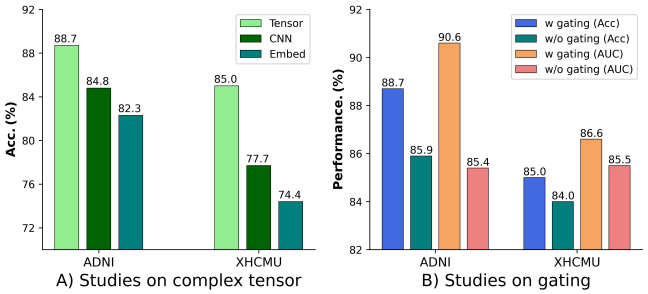

Fig. 4. Ablation results in terms of the complex-valued tensor and the gating mechanism in multi-modal modeling.

3.4. Ablation study

Moreover, we perform ablation studies on the complex-valued tensor and gating mechanism in graphs. We compared the proposed complex-valued tensor with convolution and embedding for evaluating the power of modeling multimodal data. As is shown in Figure 4 A), our complex-valued tensor approach significantly outperforms other forms of learning in both ADNI and XHCMU datasets. This indicates that the complex-valued tensor plays a significant role in multimodal learning. Moreover, Figure 4 B) demonstrates the comparison results on the gating mechanism. With the gating mechanism, our proposed TC-GNN achieves consistent improvements by 2.8% and 1% on two datasets respectively.

4. CONCLUSION

In this paper, we propose a novel paradigm, called a complex-valued tensor graph, by introducing complex-valued tensors into graphs for coupling multimodal complementary representations. The experimental results on two real-world datasets demonstrate that our proposed model achieves promising results in terms of brain disorder classification. Consistent improvements are achieved on both two tasks compared with the current state-of-the-art baselines, indicating that our proposed TC-GNN is feasible to enrich representations of multiple modalities and provides a powerful way of multimodal learning.

5. REFERENCES

[1] Juliette Valenchon and Mark Coates, "Multiple-graph recurrent graph convolutional neural network architectures for predicting disease outcomes," in *ICASSP 2019-2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. IEEE, 2019, pp. 3157–3161.

[2] Jiayue Cai, Aiping Liu, Taomian Mi, Saurabh Garg, Wade Trappe, Martin J McKeown, and Z Jane Wang, "Dynamic graph theoretical analysis of functional connectivity in parkinson's disease: The importance of fiedler value," *IEEE journal of biomedical and health informatics*, vol. 23, no. 4, pp. 1720–1729, 2018.

[3] Makoto Fukushima, Richard F Betzel, Ye He, Martijn P van den Heuvel, Xi-Nian Zuo, and Olaf Sporns, "Structure–function relationships during segregated and integrated network states of human brain functional connectivity," *Brain Structure and Function*, vol. 223, no. 3, pp. 1091–1106, 2018.

[4] Niharika Shimona Dsouza, Mary Beth Nebel, Deana Crocetti, Joshua Robinson, Stewart Mostofsky, and Archana Venkataraman, "M-gcn: A multimodal graph convolutional network to integrate functional and structural connectomics data to predict multidimensional phenotypic characterizations," in *Medical Imaging with Deep Learning*. PMLR, 2021, pp. 119–130.

[5] Selen Atasoy, Isaac Donnelly, and Joel Pearson, "Human brain networks function in connectome-specific harmonic waves," *Nature communications*, vol. 7, no. 1, pp. 1–10, 2016.

[6] Yan Wang, Yanwu Yang, Xin Guo, Chenfei Ye, Na Gao, Yuan Fang, and Heather T Ma, "A novel multimodal mri analysis for alzheimer's disease based on convolutional neural network," in *2018 40th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*. IEEE, 2018, pp. 754–757.

[7] Yifan Feng, Haoxuan You, Zizhao Zhang, Rongrong Ji, and Yue Gao, "Hypergraph neural networks," in *Proceedings of the AAAI conference on artificial intelligence*, 2019, vol. 33, pp. 3558–3565.

[8] Jianwen Jiang, Yuxuan Wei, Yifan Feng, Jingxuan Cao, and Yue Gao, "Dynamic hypergraph neural networks.," in *IJCAI*, 2019, pp. 2635–2641.

[9] Fabien Lotte, Laurent Bougrain, Andrzej Cichocki, Maureen Clerc, Marco Congedo, Alain Rakotomamonjy, and Florian Yger, "A review of classification algorithms for eeg-based brain–computer interfaces: a 10 year update," *Journal of neural engineering*, vol. 15, no. 3, pp. 031005, 2018.

[10] Bilal Ahmad, Liana Khamidullina, Alexey A Korobkov, Alla Manina, Jens Haueisen, and Martin Haardt, "Joint model order estimation for multiple tensors with a coupled mode and applications to the joint decomposition of eeg, meg magnetometer, and gradiometer tensors," in *ICASSP 2022-2022 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. IEEE, 2022, pp. 1186–1190.

[11] Justin Dauwels, K Srinivasan, Reddy M Ramasubba, and Andrzej Cichocki, "Multi-channel eeg compression based on matrix and tensor decompositions," in *2011 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. IEEE, 2011, pp. 629–632.

[12] Bhaskar Sen and Keshab K Parhi, "Extraction of common task signals and spatial maps from group fmri using a parafac-based tensor decomposition technique," in *2017 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. IEEE, 2017, pp. 1113–1117.

[13] Trabelsi Chiheb, O Bilaniuk, D Serdyuk, et al., "Deep complex networks," in *International Conference on Learning Representations. https://openreview.net/forum*, 2017.

[14] Kanhao Zhao, Boris Duka, Hua Xie, Desmond J Oathes, Vince Calhoun, and Yu Zhang, "A dynamic graph convolutional neural network framework reveals new insights into connectome dysfunctions in adhd," *Neuroimage*, vol. 246, pp. 118774, 2022.

[15] Moritz Wolter and Angela Yao, "Complex gated recurrent neural networks," *Advances in neural information processing systems*, vol. 31, 2018.

[16] Jeremy Kawahara, Colin J Brown, Steven P Miller, Brian G Booth, Vann Chau, Ruth E Grunau, Jill G Zwicker, and Ghassan Hamarneh, "Brainnetcnn: Convolutional neural networks for brain networks; towards predicting neurodevelopment," *NeuroImage*, vol. 146, pp. 1038–1049, 2017.

[17] Yanwu Yang, Chenfei Ye, Junyan Sun, Li Liang, Haiyan Lv, Linlin Gao, Jiliang Fang, Ting Ma, and Tao Wu, "Alteration of brain structural connectivity in progression of parkinson's disease: a connectome-wide network analysis," *NeuroImage: Clinical*, vol. 31, pp. 102715, 2021.