Learning Similarity-Preserving Representations of Brain Structure-Function Coupling

Yang Li and Gonzalo Mateos
Dept. of Electrical and Computer Engineering, University of Rochester, Rochester, NY, USA

***Abstract*—Advances in graph signal processing for network neuroscience offer a unique pathway to integrate brain structure and function, with the goal of revealing some of the brain's organizing principles at the system level. In this direction, we develop a supervised graph representation learning framework to model the relationship between brain structural connectivity (SC) and functional connectivity (FC) via a graph encoder-decoder system. Specifically, we propose a Siamese network architecture equipped with graph convolutional encoders to learn graph (i.e., subject)-level embeddings that preserve application-dependent similarity measures between brain networks. This way, we effectively increase the number of training samples and bring in the flexibility to incorporate additional prior information via the prescribed target graph-level distance. While information on the brain structure-function coupling is implicitly distilled via reconstruction of brain FC from SC, our model also manages to learn representations that preserve the similarity between input graphs. The superior discriminative power of the learnt representations is demonstrated in downstream tasks including subject classification and visualization. All in all, this work advocates the prospect of leveraging learnt graph-level, similarity-preserving embeddings for brain network analysis, by bringing to bear standard tools of metric data analysis.**

***Index Terms*—Brain connectomics, graph representation learning, Siamese network, graph convolutional network.**

I. INTRODUCTION

The human brain is an immensely complex yet highly efficient network consisting of distributed regions conducting individual tasks, while actively communicating with each other to accomplish collaborative operations [1]. This so-termed integrative nature of the brain motivates well the adoption of graph-centric signal and information processing tools to study the relationship between brain structure (of neural connections) [2] and functional activity [3].

Brain connectomes consist of two major types of networks [4]. Structural connectivity (SC) measures the anatomical tracts of axonal bundles [5], whereas functional connectivity (FC) represents pairwise statistical correlation between (proxies of) neural activities across various brain regions of interest (RoI). Previous studies have revealed that FC correlates with SC at an aggregate level [6], where strong functional connections exist between RoIs with scarce anatomical pathways [7]. These intriguing findings sparked multiple works aiming at predicting FC from SC by exploiting both direct and indirect anatomical connections [6]–[9], and in estimating SC from FC [10], [11]. There is now reasonable consensus that the coupling between brain SC and FC can be of great importance towards understanding the underpinnings of human brain function, and remains a timely research area in (network) neuroscience; see e.g., [12], [13].

In this paper, we explore the coupling between brain SC and FC by learning low-dimensional representations of the SC-to-FC mapping obtained from graph convolutional features [14]–[16]. Specifically, we propose and implement a supervised graph representation learning (GRL) framework to jointly learn: (i) low-dimensional *node* embeddings generated from brain SC networks to reconstruct empirical FC networks; and (ii) (subject-level) *graph* embeddings to represent whole connectomes for downstream classification and visualization tasks. The proposed GRL pipeline incorporates a Siamese network with two identical branches, an architecture that has been proven effective in various pattern recognition tasks [17]–[20], but is otherwise less explored for network neuroscience; see [21], [22] for a recent contribution to metric learning. Each branch can be viewed as a graph encoder-decoder system [23], [24]; see Section II-B for more details. The input of the Siamese network is a pair of SC graphs, and for each graph, the encoder learns low-dimensional node embeddings that preserve application-dependent similarity measures between brain networks. To aggregate information from multiple hops within the network, layered graph convolution operations interleaved with point-wise nonlinearities are used to compute nodal features. This procedure calls for graph convolutional networks (GCNs), which serve as a proper encoder model to output connectome representations that capture indirect brain connections. The learnt node embeddings are then used to reconstruct the corresponding FC of the input SC graph.

Whole graph embeddings are also obtained by applying mean pooling mechanism on the node embeddings [25], with the goal of performing (subject-level) graph classification. If the two input graphs come from the same class (here, non-drinkers or heavy drinkers), we expect that they share higher similarity and consequently, their graph embeddings shall be close in the latent space. To learn discriminative representations, embeddings of graphs with the same labels are grouped into clusters while the distances between clusters in latent space are maximized. All in all, we propose a multi-task GRL model with the dual goal of reconstructing brain FC from input SC data, and estimating similarity between each pair of

---

This work was supported in part by NSF Awards under Grants CCF-1750428, CCF-1934962, and ECCS-1809356. *(Corresponding author: Gonzalo Mateos)*. Author emails: yli131@ur.rochester.edu and gmateosb@ece.rochester.edu.

graphs. The graph reconstruction task learns a parsimonious representation of the population-level SC-FC relationship, and the similarity estimation uses subject labels as additional inputs for supervised classification. Accordingly, the model strikes the right balance between deciphering population patterns that shape the SC-FC coupling, and distilling subject-level variability to facilitate graph-level metric learning.

We train and test our model on a population of 412 subjects with 191 non-drinkers and 221 heavy drinkers from the Human Connectome Project[1] (Section III). To evaluate the proposed GRL model, the estimated pairwise similarities are used to assign a class membership for each sample in the test set according to a K-nearest neighbor (KNN) classifier. Satisfactory classification performance is obtained, outperforming relevant baselines [13], [26]. We also conduct embedding visualization by feeding the learnt graph representations to the t-SNE [27] algorithm to project them into a 2-D space, and corroborate that graph embeddings of drinkers and non-drinkers are well separated. By leveraging the Siamese framework, we manage to significantly increase the amount of training data; by using the similarity-preserving mechanism, the learnt graph representations possess better discriminative power thus boosting classification performance.

II. PRELIMINARIES AND PROPOSED MODEL

In this section, we review the background on graph convolutional networks (GCNs) that form the encoder in our model; see also [16], [28] for further details. Then we formally introduce the proposed Siamese GRL framework.

A. Graph convolutional networks

Consider a weighted, undirected graph denoted by $\mathcal{G}(\mathcal{V}, \mathbf{A})$, where $\mathcal{V}$ is a set of N nodes corresponding to brain RoIs, and $\mathbf{A} \in \mathbb{R}_+^{N\times N}$ is the symmetric adjacency matrix with $\mathbf{A}_{ij} = \mathbf{A}_{ji} \geq 0$ representing the structural connection strengths between RoIs i and j. The graph Laplacian matrix is defined as $\mathbf{L} := \mathbf{D} - \mathbf{A}$, where $\mathbf{D}$ is the diagonal degree matrix. The symmetric Laplacian $\mathbf{L}$ can be further decomposed as $\mathbf{L} = \mathbf{U}\mathbf{\Lambda}\mathbf{U}^\top$, where $\mathbf{U} \in \mathbb{R}^{N\times N}$ denotes the set of orthonormal eigenvectors and $\mathbf{\Lambda}$ contains all eigenvalues on its diagonal. The eigenvectors $\mathbf{U}$ serve as a graph Fourier basis [29]. Consider a vertex-valued signal $\mathbf{x} \in \mathbb{R}^N$, where x_i denotes the signal value at RoI i, e.g., nodal attributes or features on the brain FC (or SC) network. With the graph Fourier transform (GFT) of $\mathbf{x}$ given by $\hat{\mathbf{x}} = \mathbf{U}^\top\mathbf{x}$ [29], [30], the graph convolution can be defined as

$$\mathbf{H}\mathbf{x} = (\sum_{i=0}^{K} h_i \mathbf{L}^i)\mathbf{x} = \mathbf{U}(\sum_{i=0}^{K} h_i \mathbf{\Lambda}^i)\mathbf{U}^\top\mathbf{x} = \mathbf{U}\hat{\mathbf{H}}\hat{\mathbf{x}}, \quad (1)$$

where $\mathbf{H} = \sum_{i=0}^{K} h_i \mathbf{L}^i$ is the graph filter with coefficients $\mathbf{h} := [h_0, \ldots, h_K]^\top$ and frequency response $\hat{\mathbf{H}} = \sum_{i=0}^{K} h_i \mathbf{\Lambda}^i$. Graph convolutional networks are then neural network architectures with stacked layers of graph convolutional filters (1) and pointwise nonlinear activation functions [31].

Adoption of first-order ($K = 1$) graph convolutional filters was advocated for the GCN model in [14], working with the degree-normalized Laplacian $\mathbf{D}^{-1/2}\mathbf{L}\mathbf{D}^{-1/2}$ and a suitable reparametrization of the filter coefficients so that (1) simplifies to ($\mathbf{I}_N$ stands for the $N \times N$ identity matrix)

$$\mathbf{H}\mathbf{x} = \theta(\mathbf{I}_N + \mathbf{D}^{-1/2}\mathbf{A}\mathbf{D}^{-1/2})\mathbf{x}. \quad (2)$$

This motivates a simple per-layer filtering update implemented to refine the nodal embeddings, namely

$$\mathbf{X}^{(\ell)} = \text{ReLU}\left(\tilde{\mathbf{A}}\mathbf{X}^{(\ell-1)}\mathbf{\Theta}^{(\ell)}\right), \quad (3)$$

where $\tilde{\mathbf{A}} := \mathbf{I}_N + \mathbf{D}^{-1/2}\mathbf{A}\mathbf{D}^{-1/2}$, $\mathbf{X}^{(\ell)} \in \mathbb{R}^{N\times d_\ell}$ are the nodal representations at layer ℓ, and $\mathbf{\Theta}^{(\ell)} \in \mathbb{R}^{d_{\ell-1}\times d_\ell}$ stores the learnable parameters of d_ℓ filters acting on $d_{\ell-1}$ input features. For the encoder of the proposed Siamese GRL pipeline, we leverage GCN layers in (3) due to their simplicity and satisfactory performance [14], [32], [33]. By stacking multiple layers, the GCN-based encoder is capable to integrate multi-hop information, making it suitable to capture both direct and indirect interactions over brain SC networks [13], [26].

B. Problem statement and Siamese model architecture

Given brain SC networks, the twofold goal is to build and train a model to: (i) reconstruct FC networks using learnt node embeddings; and (ii) learn whole graph representations that approximately preserve some pre-defined similarity measure between pairs of input graphs. The learnt representations then contain information of both population patterns while they also capture some degree of subject-level variability. To this end, we propose a supervised Siamese network with two sister graph encoder-decoder pipelines as shown in Figure 1.

Taking as input a brain SC network whose N nodes represent brain RoIs, the encoder generates a lower-dimensional representation for each node. Among various node embedding models, neighbor aggregation methods such as GCN are permutation-invariant and inductive [17]. As a result, a multi-layer GCN is used in the encoder to generate latent variables that consolidates both nodal attributes such as the known intrinsic properties of each brain region, and the network topology information like the connection strengths among regions in SC networks. In the Siamese framework, the two sister GCN encoders share the same learnable weights and for each encoder, the input is a subject's SC network $\mathbf{A}$ and the corresponding feature matrix $\mathbf{X} \in \mathbb{R}^{N\times T}$, where each row represents a nodal attribute of length T. Following the re-normalization idea in [14], the normalized adjacency matrix is $\tilde{\mathbf{A}} := \hat{\mathbf{D}}^{-1/2}\hat{\mathbf{A}}\hat{\mathbf{D}}^{-1/2}$, where $\hat{\mathbf{A}} := \mathbf{I}_N + \mathbf{A}$ and $\hat{\mathbf{D}} = \text{diag}(\hat{\mathbf{A}}\mathbf{1})$ is the degree matrix of $\hat{\mathbf{A}}$. Identity matrix is added to the input graph $\mathbf{A}$ to introduce self loops to assure that the attribute on the node itself also contributes to the new node embedding during graph convolution. Inspired by [17], [32], we concatenate node embeddings learnt in intermediate GCN layers $\ell = 1, \ldots, L-1$, thus forming $\mathbf{X}_C \in \mathbb{R}^{N\times\sum_\ell d_\ell}$.

---

[1] https://www.humanconnectome.org/

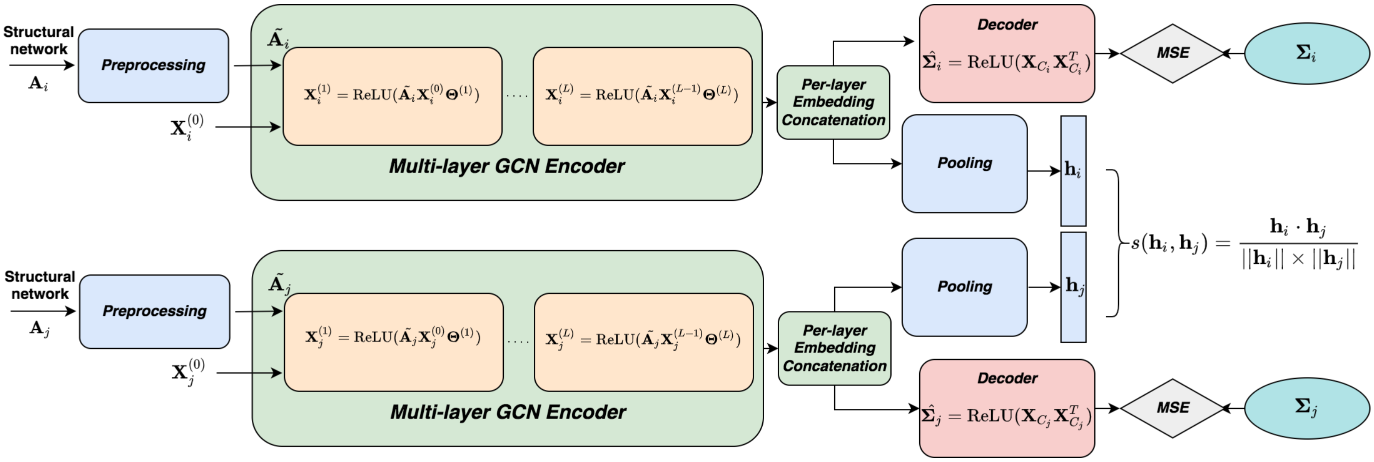


Fig. 1: The Siamese GRL architecture. The inputs are a pair of SC networks $\mathbf{A}_i, \mathbf{A}_j$ and optional nodal attributes. Graph convolution/information propagation occurs within the GCN encoder. The sister GCN encoders share the same learnable weights. Rows in $\mathbf{X}^{(L)}$ are low-dimensional node embeddings at the final layer of the GCN encoder. An outer-product decoder reconstructs the FC network $\hat{\mathbf{\Sigma}}$ and thus implicitly models the SC-to-FC mapping. Given graph (i.e., subject)-level representations $\mathbf{h}_i, \mathbf{h}_j$ obtained via pooling, their similarity is computed using the cosine similarity.

For the FC reconstruction objective of the GRL architecture in Figure 1, node embeddings $\mathbf{X}_C$ go through an outer product decoder

$$\hat{\mathbf{\Sigma}} = \mathrm{ReLU}(\mathbf{X}_C \mathbf{X}_C^\top) \tag{4}$$

to generate a predicted FC adjacency matrix $\hat{\mathbf{\Sigma}}$. The mean squared error (MSE) between the reconstructed graph $\hat{\mathbf{\Sigma}}$ and the empirical FC $\mathbf{\Sigma}$ of each input subject, averaged over a training set and denoted as $\mathcal{L}_{\mathbf{MSE}}(\hat{\mathbf{\Sigma}}, \mathbf{\Sigma})$, is used as reconstruction loss for training the regression branch of the model. In the current setup described in Section III-A, FC networks contain few negative edges with much smaller magnitude compared with the vast amount of positive edges. To avoid such data imbalance problem, we remove all negative FC connections and restrict FC weights to the range $[0, 1]$ as in other studies on FC [34], [35].

Given the node embeddings $\mathbf{X}_C$, we further obtain whole graph embeddings by taking the row-wise average of all the node embeddings. This representation $\mathbf{h}_i$ can be viewed as a summary (in latent space) of the SC-FC coupling of subject i. Node-wise average pooling is a simple yet effective procedure which has been used in many studies; see e.g. [15]. For the input graph pairs $\{\mathbf{A}_i, \mathbf{A}_j\}$ fed to the Siamese network, the learnt representations $\mathbf{h}_i, \mathbf{h}_j$ shall exhibit high similarity (quantified by the cosine similarity $s(\mathbf{h}_i, \mathbf{h}_j)$) if the input graphs share the same labels $l_i = l_j$ (i.e both are non-drinkers, or drinkers). On the other hand, when the input graphs have different labels, we expect to train the model to output embeddings that are comparatively less similar. To this end, with N_s pairs of graphs with the same labels and N_d pairs with different labels, we implement the pairwise similarity global loss function in [22], [36]

$$\mathcal{L}_{\mathbf{SIM}} = (\sigma^{2+} + \sigma^{2-}) + w \times \max(0, m - (\mu^+ - \mu^-)) \tag{5}$$

which minimises the mean similarity between embeddings of different classes $\mu^- = \sum s(\mathbf{h}_i, \mathbf{h}_j)/N_d, l_i \neq l_j$ and maximises the mean similarity between embeddings of the same classes $\mu^+ = \sum s(\mathbf{h}_i, \mathbf{h}_j)/N_s, l_i = l_j$. At the same time, the variance of similarities within and between classes are also minimized. The margin between the means of matching and non-matching pairs is denoted by m, and w tunes the balance between the mean and variance terms. The overall loss function is

$$\mathcal{L} = \mathcal{L}_{\mathbf{SIM}} + \lambda \times \mathcal{L}_{\mathbf{MSE}}(\hat{\mathbf{\Sigma}}, \mathbf{\Sigma}), \tag{6}$$

where hyperparameter λ controls the trade-off between FC reconstruction and similarity estimation. By training the model end-to-end, we aim to strike the right balance between FC reconstruction and similarity estimation and achieve satisfactory performance on both objectives.

III. NUMERICAL TEST

In this section, the performance of the proposed Siamese architecture is evaluated on a real-world neuroimaging dataset, and compared with our previous work on the same dataset.

A. Neuroimaging data

We adopt a neuroimaging dataset with $P = 412$ subjects from the Human Connectome Project (HCP) [37], [38]. The cohort is partitioned into two classes: 191 non-drinkers and 221 heavy drinkers according to information available on the lifetime maximum number of drinks had in a single day. The threshold is set to 21. The Desikan-Killiany atlas is used to specify brain RoIs [39]. Hence, $\mathcal{V}$ in both FC and SC networks correspond to $N = 68$ cortical surface regions, with 34 nodes in each hemisphere. The SC network $\mathbf{A}$ of each subject is extracted from the dMRI and structural MRI data [40], [41]. Brain functional activities are given by the resting-state BOLD

TABLE I: Subject classification performance

Model	Accuracy	F score
Siamese model	$\mathbf{0.6843 \pm 0.016}$	$\mathbf{0.7391 \pm 0.016}$
End-to-end model in [13]	0.6610 ± 0.043	0.6962 ± 0.030

signals measured using fMRI. Brain FC networks $\mathbf{\Sigma}$ are then constructed so that edge weights are the Pearson correlation coefficient between the BOLD signals at the incident RoIs. Currently we use one-hot encoding to define graph signals as $\mathbf{X} = \mathbf{I}_{68}$. Additional information regarding specific brain anatomical features, e.g. volumes, can be integrated into the nodal features to bring in add-on advantages, which we leave for future work.

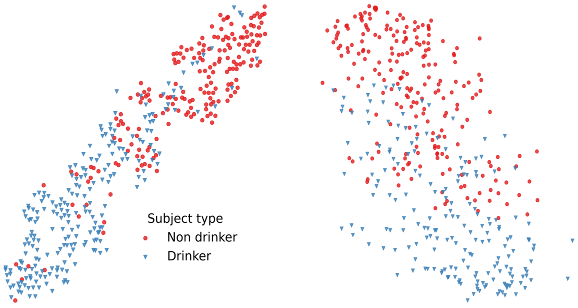
B. Implementation

The proposed model is implemented in TensorFlow. We opt to choose the same GCN encoder architecture as in [13] which contains 3 layers with dimensions of $32 \times 16 \times 8$. Xavier initialization [42] was used to initialize the weight coefficients $\mathbf{\Theta}$. λ is set to be 0.2 in (6) via grid search, and $m = w = 1$ in (5). We carry out a 5-fold cross validation where each time the whole dataset is partitioned randomly into 60% training, 20% validation and 20% testing set. From graphs in the training set, all the possible pairs are generated and fed to the Siamese model to learn representations that reconstruct FC from input SC and simultaneously maximize/minimize the similarity between input graphs if they have same/different labels. To avoid overfitting, early stopping is applied to monitor the performance on validation sets, and to stop the training once the performance is not improving in the last 10 training epochs. The Adam [43] optimizer is used with learning rate 0.001.

C. Results

During testing, the learnt representations of all the graphs in the training set and validation set (with known labels) are combined to establish a selection pool in the vector space. Each graph in the test set goes through the encoder to generate its low-dimensional representation. We locate its position in the vector space, and classify it by a majority vote of its closest 5 neighbors measured by the cosine similarity between their embeddings. Repeat this process for each test set of the 5 folds, and the overall classification performance across all test sets is given in Table I.

Table I shows that the proposed Siamese framework outperforms our previous work where we proposed a supervised graph representation learning model that reconstruct FC from SC and conduct subject classification via a logistic regression classifier [13]. This indicates that by building the Siamese network and training it with graph pairs significantly increase the amount of training data, leading to the superior performance in subject classification. In addition, the training objective of similarity estimation is more directed to the distinguishing of subjects' labels in the vector space, and by utilizing the global loss function from [36], the similarity metric is optimally estimated for graph representations of subjects with the same labels, and with different labels. Since in [13] we already reveal that the SC-FC coupling, captured by the regression branch in Figure 1, helps achieve better classification performance, here we only compare with the reported results in [13] to reflect the advantage of the proposed Siamese network plus the metric learning procedure of similarity estimation.



(a) Proposed (b) Model in [13]

Fig. 2: Visualization of the learnt graph representations via t-SNE algorithm. Clear separation is observed between red circles of non-drinkers and blue triangles of drinkers, indicating the effectiveness of the proposed Siamese network to capture subject uniqueness. Circles within each group are closer to each other on the left, indicating the advantage of the similarity estimation proposed in this work.

Visualizing the embeddings on a two-dimensional space is a popular way to assess the representation learning performance. To this end, we select the model trained using the data from one of the 5 folds that offers the highest classification accuracy on the test set, and put all the 412 graphs through the GCN encoder to generate 412 embeddings. Via the t-SNE algorithm [27], we plot these embeddings on the 2-D plane shown in Figure 2a. It is obvious that a nice separation exists between the drinkers represented by the blue triangles, and the non-drinkers of the red circles. Compared with [13] in Figure 2b, the representations in Figure 2a are closer within each group, which is the expected result from the similarity learning. Such visualization further verifies the discriminative power of the proposed Siamese network trained with similarity estimation between graph representations.

IV. CONCLUSION

In this work, we propose a Siamese network framework that consists of two sister graph encoder-decoder pipelines. Besides reconstructing FC networks from SC which captures the SC-FC coupling, the Siamese architecture also entails a metric-learning mechanism where we manage to learn graph representations that are more similar if the input graphs have the same labels (i.e. drinking habits), and are less similar if they have different labels. Such modification in the GRL model architecture significantly increases the amount of training data, and the learnt similarity-preserving representations offer more discriminative power to distinguish drinkers from non-drinkers. Experiments on a large cohort of subjects from HCP reveal superior performance on subject classification than our previous work, suggesting the prospect of leveraging graph-level, similarity-preserving embeddings to measure SC-FC coupling for brain network analysis. Future work shall be devoted to further exploit the power of the Siamese network. For example, the margin that measures the difference between the means of matching and non-matching pairs can be designed to reflect more intrinsic differences between drinkers and non-drinkers, or the differences within the FC networks between input subjects.

V. ACKNOWLEDGMENT

The authors would like to thank Zhengwu Zhang for the collaboration on neuroimaging data and the advises from biomedical and bio-statistical perspective.

REFERENCES

[1] M. Van Den Heuvel and H. Pol, "Exploring the brain network: a review on resting-state fMRI functional connectivity," *Eur. Neuropsychopharmacol.*, vol. 20, no. 8, pp. 519–534, 2010.

[2] A. Fornito, A. Zalesky, and E. Bullmore, *Fundamentals of Brain Network Analysis*. Academic Press, 2016.

[3] J. D. Power, A. L. Cohen, S. M. Nelson *et al.*, "Functional network organization of the human brain," *Neuron*, vol. 72, no. 4, pp. 665–678, 2011.

[4] O. Sporns, *Networks of the Brain*. MIT Press, 2010.

[5] E. Bullmore and O. Sporns, "Complex brain networks: Graph theoretical analysis of structural and functional systems," *Nat. Rev. Neurosci.*, vol. 10, no. 3, p. 186, 2009.

[6] C. Honey, O. Sporns, L. Cammoun, X. Gigandet, J.-P. Thiran, R. Meuli, and P. Hagmann, "Predicting human resting-state functional connectivity from structural connectivity," *Proc. Natl. Acad. Sci. U.S.A.*, vol. 106, no. 6, pp. 2035–2040, 2009.

[7] J. Goñi, M. P. van den Heuvel, A. Avena-Koenigsberger, N. V. de Mendizabal, R. F. Betzel, A. Griffa, P. Hagmann, B. Corominas-Murtra, J.-P. Thiran, and O. Sporns, "Resting-brain functional connectivity predicted by analytic measures of network communication," *Proc. Natl. Acad. Sci. U.S.A.*, vol. 111, no. 2, pp. 833–838, 2014.

[8] C. Stam, E. Van Straaten, E. Van Dellen, P. Tewarie *et al.*, "The relation between structural and functional connectivity patterns in complex brain networks," *Intl. J. Psychophysiol.*, vol. 103, pp. 149–160, 2016.

[9] Y. Li, R. Shafipour, G. Mateos, and Z. Zhang, "Mapping brain structural connectivities to functional networks via graph encoder-decoder with interpretable latent embeddings," *IEEE Global Conf. Signal and Info. Process. (GlobalSIP)*, pp. 1–5, 2019.

[10] Y. Li and G. Mateos, "Identifying structural brain networks from functional connectivity: A network deconvolution approach," *IEEE Intl. Conf. Acoust., Speech and Signal Process. (ICASSP)*, pp. 1135–1139, 2019.

[11] F. Abdelnour, H. U. Voss, and A. Raj, "Network diffusion accurately models the relationship between structural and functional brain connectivity networks," *Neuroimage*, vol. 90, pp. 335–347, 2014.

[12] T. Sarwar, Y. Tian, B. T. Yeo, K. Ramamohanarao, and A. Zalesky, "Structure-function coupling in the human connectome: A machine learning approach," *NeuroImage*, vol. 226, p. 117609, 2021.

[13] Y. Li, G. Mateos, and Z. Zhang, "Learning to model the relationship between brain structural and functional connectomes," *arXiv preprint arXiv:2112.09906*, 2021.

[14] T. N. Kipf and M. Welling, "Semi-supervised classification with graph convolutional networks," *Proc. ICLR*, 2017.

[15] D. Duvenaud, D. Maclaurin *et al.*, "Convolutional networks on graphs for learning molecular fingerprints," *Adv. Neural. Inf. Process. Syst.*, pp. 2224–2232, 2015.

[16] F. Gama, A. G. Marques, G. Leus, and A. Ribeiro, "Convolutional neural network architectures for signals supported on graphs," *IEEE Trans. Signal Process.*, vol. 67, no. 4, pp. 1034–1049, 2018.

[17] Y. Bai, H. Ding, Y. Qiao, A. Marinovic, K. Gu, T. Chen *et al.*, "Unsupervised inductive graph-level representation learning via graph-graph proximity," *Intl. Joint Conf. Artificial Intelligence*, pp. 1988–1994, 2019.

[18] X. Dong and J. Shen, "Triplet loss in siamese network for object tracking," in *Comput. Vis. ECCV*, 2018, pp. 459–474.

[19] G. Koch, R. Zemel, R. Salakhutdinov *et al.*, "Siamese neural networks for one-shot image recognition," in *ICML Deep Learning Workshop*, vol. 2. Lille, 2015, p. 0.

[20] D. Chicco, "Siamese neural networks: An overview," *Artificial Neural Netw.*, pp. 73–94, 2021.

[21] S. I. Ktena, S. Parisot, E. Ferrante, M. Rajchl, M. Lee, B. Glocker, and D. Rueckert, "Distance metric learning using graph convolutional networks: Application to functional brain networks," in *Med. Image. Comput. Comput. Assist. Interv.* Springer, 2017, pp. 469–477.

[22] ——, "Metric learning with spectral graph convolutions on brain connectivity networks," *NeuroImage*, vol. 169, pp. 431–442, 2018.

[23] T. N. Kipf and M. Welling, "Variational graph auto-encoders," *NIPS Workshop on Bayesian Deep Learning*, 2016.

[24] I. Chami, S. Abu-El-Haija, B. Perozzi, C. Ré, and K. Murphy, "Machine learning on graphs: A model and comprehensive taxonomy," *arXiv preprint arXiv:2005.03675*, 2020.

[25] M. Cheung, J. Shi, L. Jiang, O. Wright, and J. M. Moura, "Pooling in graph convolutional neural networks," in *Asilomar Conf. on Signals, Systems, and Computers*, 2019, pp. 462–466.

[26] Y. Li, R. Shafipour, G. Mateos, and Z. Zhang, "Supervised graph representation learning for modeling the relationship between structural and functional brain connectivity," *IEEE Intl. Conf. Acoust., Speech and Signal Process. (ICASSP)*, pp. 9065–9069, 2020.

[27] L. Maaten and G. Hinton, "Visualizing data using t-SNE," *J. Mach. Learn. Res.*, vol. 9, no. Nov, pp. 2579–2605, 2008.

[28] Z. Wu, S. Pan, F. Chen, G. Long, C. Zhang, and P. S. Yu, "A comprehensive survey on graph neural networks," *arXiv preprint arXiv:1901.00596*, 2019.

[29] A. Ortega, P. Frossard *et al.*, "Graph signal processing: Overview, challenges, and applications," *Proc. IEEE*, vol. 106, no. 5, pp. 808–828, 2018.

[30] W. Huang, T. A. Bolton, J. D. Medaglia *et al.*, "A graph signal processing perspective on functional brain imaging," *Proc. IEEE*, vol. 106, no. 5, 2018.

[31] F. Gama, E. Isufi, G. Leus, and A. Ribeiro, "Graphs, convolutions, and neural networks: From graph filters to graph neural networks," *IEEE Signal Process. Mag.*, vol. 37, no. 6, pp. 128–138, 2020.

[32] K. Xu, W. Hu, J. Leskovec, and S. Jegelka, "How powerful are graph neural networks?" in *Intl. Conf. on Learning Representations (ICLR)*, 2019.

[33] J. You, Z. Ying, and J. Leskovec, "Design space for graph neural networks," in *Adv. Neural. Inf. Process. Syst. (NeurIPS)*, vol. 33, 2020.

[34] J. D. Power, D. A. Fair, B. L. Schlaggar, and S. E. Petersen, "The development of human functional brain networks," *Neuron*, vol. 67, no. 5, pp. 735–748, 2010.

[35] M. Rubinov and O. Sporns, "Complex network measures of brain connectivity: uses and interpretations," *Neuroimage*, vol. 52, no. 3, pp. 1059–1069, 2010.

[36] V. Kumar BG, G. Carneiro, and I. Reid, "Learning local image descriptors with deep siamese and triplet convolutional networks by minimising global loss functions," in *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit.*, 2016, pp. 5385–5394.

[37] D. C. Van Essen, S. M. Smith, D. M. Barch, T. E. Behrens, E. Yacoub, K. Ugurbil, W.-M. H. Consortium *et al.*, "The Wu-Minn human connectome project: an overview," *Neuroimage*, vol. 80, pp. 62–79, 2013.

[38] M. F. Glasser, S. M. Smith, D. S. Marcus, J. L. Andersson, E. J. Auerbach, T. E. Behrens, T. S. Coalson, M. P. Harms, M. Jenkinson, S. Moeller *et al.*, "The human connectome project's neuroimaging approach," *Nat. Neurosci.*, vol. 19, no. 9, pp. 1175–1187, 2016.

[39] R. S. Desikan, F. Ségonne, B. Fischl, B. T. Quinn *et al.*, "An automated labeling system for subdividing the human cerebral cortex on MRI scans into gyral based regions of interest," *Neuroimage*, vol. 31, no. 3, pp. 968–980, 2006.

[40] Z. Zhang, G. I. Allen, H. Zhu, and D. Dunson, "Tensor network factorizations: Relationships between brain structural connectomes and traits," *Neuroimage*, vol. 197, pp. 330–343, 2019.

[41] Z. Zhang, M. Descoteaux, J. Zhang, G. Girard, M. Chamberland, D. Dunson, A. Srivastava, and H. Zhu, "Mapping population-based structural connectomes," *NeuroImage*, vol. 172, pp. 130–145, 2018.

[42] X. Glorot and Y. Bengio, "Understanding the difficulty of training deep feed forward neural networks," *Intl. Conf. on Artificial Intelligence and Statistics*, pp. 249–256, 2010.

[43] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," *ICLR*, 2015.