

Functional Connectivity Prediction With Deep Learning for Graph Transformation

Negar Etemadyrad^{1b}, Yuyang Gao^{1b}, Qingzhe Li^{1b}, Xiaojie Guo^{1b}, Frank Krueger,
Qixiang Lin^{1b}, Deqiang Qiu^{1b}, and Liang Zhao^{1b}

Abstract—Inferring resting-state functional connectivity (FC) from anatomical brain wiring, known as structural connectivity (SC), is of enormous significance in neuroscience for understanding biological neuronal networks and treating mental diseases. Both SC and FC are networks where the nodes are brain regions, and in SC, the edges are the physical fiber nerves among the nodes, while in FC, the edges are the nodes' coactivation relations. Despite the importance of SC and FC, until very recently, the rapidly growing research body on this topic has generally focused on either linear models or computational models that rely heavily on heuristics and simple assumptions regarding the mapping between FC and SC. However, the relationship between FC and SC is actually highly nonlinear and complex and contains considerable randomness; additional factors, such as the subject's age and health, can also significantly impact the SC-FC relationship and hence cannot be ignored. To address these challenges, here, we develop a novel SC-to-FC generative adversarial network (SF-GAN) framework for mapping SC to FC, along with additional metafeatures based on a newly proposed graph neural network-based generative model that is capable of learning the stochasticity. Specifically, a new graph-based conditional generative adversarial nets model is proposed, where edge convolution layers are leveraged to encode the graph patterns in the SC in the form of a graph representation. New edge deconvolution layers are then utilized to decode the representation back to FC. Additional metafeatures of subjects' profile information are integrated into the graph representation with newly designed sparse-regularized layers that can automatically select features that impact FC. Finally, we have also proposed new *post hoc* explainer of our SF-GAN, which can identify which subgraphs in SC strongly influence which subgraphs in FC by a new multilevel edge-correlation-guided graph clustering problem. The results of experiments conducted to test the new model confirm that it significantly outperforms

existing state-of-the-art methods, with additional interpretability for identifying important metafeatures and subgraphs.

Index Terms—Convolutional neural network, deep learning, functional connectivity (FC), interpretable graph neural networks, metafeatures, structural connectivity (SC), subgraph mining.

NOMENCLATURE

G, G'	Graphs of N nodes representing SC and FC.
A, A'	Weighted adjacency matrices of G and G' .
V	Set of nodes of g .
V_i	i th node of G .
u	Random noise term of distribution $\mathcal{U}$.
M	Vector of K metafeatures.
$\tilde{M}$	Concatenation of N copies of metafeatures.
$\mathcal{F}$	FC conditional generator.
$\mathcal{D}$	FC conditional discriminator.
$F_{j,k}$	Value of the j th node under the k th latent feature.
$E^{l,m}$	Correlation matrix of the l th deconvolutional layer and the m th feature map.
λ_1, λ_2	Hyperparameters to control regularization for the generator.
ψ_m	N2E kernel for the m th feature map (outgoing).
ϕ_m	N2E kernel for the m th feature map (incoming).
$W'_{m,k}$	N2E kernel for the m th feature map and the k th metafeature (outgoing).
$W_{m,k}$	N2E kernel for the m th feature map and the k th metafeature (incoming).
P	Size of each node embedding.
K	Total number of metafeatures.

I. INTRODUCTION

STRUCTURAL connectivity (SC) and functional connectivity (FC) are considered as two forms of brain connectivity networks or “connectomes.” SC represents the synaptic connections among neurons that can be quantified by diffusion magnetic resonance imaging (MRI) techniques such as diffusion tensor imaging, which can map diffusion process of molecules in biological tissues *in vivo* and non-invasively [1]. FC can be inferred by using (resting state) functional MRI (fMRI) to calculate the strength of long-range, temporal correlations (slow fluctuations, <0.1 Hz) [2] of activation signals in various regions of brain [3]. fMRI relies on the blood oxygenation level-dependent (BOLD) responses

Manuscript received 21 August 2021; revised 26 February 2022 and 6 June 2022; accepted 23 July 2022. (Corresponding author: Liang Zhao.)

This work involved human subjects or animals in its research. Approval of all ethical and experimental procedures and protocols was granted by Washington University at St. Louis.

Negar Etemadyrad is with the Department of Electrical and Computer Engineering, George Mason University, Fairfax, VA 22030 USA.

Yuyang Gao and Liang Zhao are with the Department of Computer Science, College of Arts and Sciences, Emory University, Atlanta, GA 30322 USA (e-mail: liang.zhao@emory.edu).

Qingzhe Li and Xiaojie Guo are with the IST Department, George Mason University, Fairfax, VA 22030 USA.

Frank Krueger is with the School of Systems Biology, George Mason University, Fairfax, VA 22030 USA.

Qixiang Lin and Deqiang Qiu are with the Department of Radiology and Imaging Sciences, Emory University, Atlanta, GA 30322 USA.

This article has supplementary material provided by the authors and color versions of one or more figures available at <https://doi.org/10.1109/TNNLS.2022.3197337>.

Digital Object Identifier 10.1109/TNNLS.2022.3197337

2162-237X © 2022 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission.

See <https://www.ieee.org/publications/rights/index.html> for more information.

to activities triggered due to tasks/stimuli and measures brain activity by detecting changes associated with blood oxygenation levels and flow [4]. Resting-state fMRI (rs-fMRI) can also reveal the subjects' baseline BOLD variance. The study of the relationship between SC and FC is one of the most crucial yet open problems neuroscientists currently face. Learning the structural-to-functional mapping will deepen our understanding of how the structural properties of our brain impact its function and could thus suggest potential novel solutions to address neurologic disease and improve well-being. In addition to SC, based on evidence in [5] and [7] and according to [8], demographic and behavioral measures corresponding to the subjects under study strongly influence patterns of brain connectivity and could thus enhance the process of learning the mapping by employing appropriate informative data to delineate FC.

Predicting the mapping from structural to FC is a rapidly developing and very promising domain. Existing works can be categorized into three main approaches. The methods in the first category focus on developing linear diffusion models that strive to fit the linear correlation between the correlation matrices of SC and FC [3], while methods in the second category concentrate on modeling the dynamics of neurons' activity based on their connections, generally by proposing straightforward hypotheses related to the dynamic processes involved [9], [10]. For example, Abdelnour *et al.* [3] assumed that the coactivation patterns reflected by FC are actually derived from the process of signal diffusion in SC. The third category employs graph theory-based knowledge in their prediction, using spectral-based approaches, or linear algebra as discussed in [11] and [12], respectively. Existing works typically focus on the mapping between structural to FC to infer FC from SC samples, ignoring the possible contributions of other factors in addition to SC, which affect the estimation process.

Despite the numerous recent works in this domain, it is still in its nascent stage, with several critical challenges remaining unsolved. These include the following.

- 1) *The Complex and Nonlinear Mapping Between Structural and Functional Connectivities:* Both SC and FC are graph-structured, so there could be dependencies among the nodes inside them due to the complex connections. Moreover, at present, the true mechanisms governing the mapping between the SC and the FC are still largely unknown and cannot be effectively fit based solely on the simple linear transformations between them. This has been demonstrated by several studies where kernel-based [13] and heuristic-based [14] approaches considering nonlinear relationships tend to outperform those based on linearity assumptions. However, existing works still generally make a strong assumption regarding the types of nonlinearity, thus potentially limiting the expressiveness and flexibility of the resulting models.
- 2) *Stochasticity of the Mapping Between Structural and Functional Connectivities:* FC is unlikely to be fully determined by SC and models that ignore the potential randomness in their mapping may well not be

sufficiently robust. It was initially proposed in [14] that during the resting state, intrinsic random fluctuations modeled as white noise (corresponding to channel noise and different sources of biophysical variability) can stimulate firing rates of neurons and propagate through the network, thus determining patterns of spontaneous activity.

- 3) *Insufficient Consideration of Environmental and Intrinsic Features of the Subjects:* Metafeatures, such as the environmental (e.g., substance use) and intrinsic features (e.g., age, demographic, and behavioral measures) for individual subjects, could also impact the FC. However, the existing methods typically fail to take them into account sufficiently [5]. This is partially due to the fact that the SC is graph-structured, while the metafeatures are vector-structured, creating serious technical difficulties when it comes to ingesting and learning such heterogeneous inputs in a seamless way.
 - 4) *Tradeoff Between Model Interpretability and Expressiveness:* The sophisticated, unknown mapping between SC and FC calls for highly expressive models, but highly complex models typically suffer from limited interpretability and are prone to overfitting. It is highly beneficial to establish a model that can not only fit complex patterns between SC and FC but also brings transparency and enables users to distill new knowledge from large amounts of historical data, for example when identifying which metafeatures affect the FC patterns. Moreover, it is also crucial to identify which subgraphs in SC strongly impact which subgraphs in FC. However, the existing graph neural network explainer cannot handle this subgraph-based explanation problem where the input and output graphs have different topologies [6]. To address all the above challenges, in this article, we propose a new generic framework for deep graph transformations designed to learn the relationship between SC and FC. Our proposed SC-to-FC generative adversarial network (SF-GAN) model is a novel graph conditional GAN that encodes SC, after which the encoded latent representation and metafeatures are jointly decoded into FC. The leveraged multiple edge convolution and deconvolution layers ensure high model expressiveness in fitting sophisticated nonlinear patterns, with additional dropout operations to encompass randomness. Various newly proposed sparsity regularization graph deconvolution layers are integrated into the model optimization process in order to select the key metafeatures that impact FC. A new subgraph-based graph neural network explainer has been proposed that automatically identifies which subgraphs in SC impact which subgraphs in FC. Finally, our SF-GAN is validated to scale with at most quadratic computation in terms of the number of nodes of a graph, making it suitable for at least modest scale graphs. The contributions of this work are summarized as follows.
- Propose a New Deep Graph Learning-Based Framework for SC-to-FC Transformations:* We formulate SC-to-FC transformation as a generative process, utilizing a

graph conditional GAN that can model sophisticated transformation process with randomness.

Design a Scenario to Jointly Encode Heterogeneous Inputs of SC and Metafeatures: A large amount of metafeatures are integrated seamlessly with the learned representation of SC, which are jointly used to decode the FC.

Develop New Sparsity-Regularized Graph Deconvolution Layers for Metafeature Selection: To identify the key metafeatures, we propose new graph deconvolution layers that enforce the sparsity of feature weights with various regularization terms. This not only reduces the risk of overfitting but also automatically identifies metafeatures with nonzero weights that impact the FC patterns.

Develop a New Subgraph Graph Neural Network Explainer That Explains the Learned Mapping Patterns Between SC and FC: We formulate into a new multilevel graph clustering problem, which detects the clusters of nodes that have strong correlations in SC, FC, and in-between SC and FC. More importantly, the in-between correlations are innovatively derived from the edge-to-edge (E2E) correlations from the saliency maps of the mapping learned by SF-GAN.

Conduct Extensive Experiments to Validate the Effectiveness and Efficiency of the Proposed Model: Extensive experiments on two real-world resting-state MRI datasets demonstrated that our proposed SF-GAN can indeed predict FC close to ground-truth target graphs and outperforms existing state-of-the-art methods by a large margin.

The remainder of this article is organized as follows. In Section II, we introduce our proposed SF-GAN framework and provide details on the newly developed deconvolution layers, in addition to the new graph neural network explainer for SF-GAN. In Section III, we describe the experiment settings and performance analyses, followed by Section IV in which we apply our novel design to find the best partition of nodes in the SC and FC graphs. We provide a summary of related work in Section V and a conclusion of our work in Section VI.

II. METHOD

In this section, we describe the proposed SF-GAN framework for predicting FC from SC. First, in Section II-A, we introduce the notations in brain network data and machine learning models used in this domain, which is followed by Section II-B that summarizes the overall architecture for the SF-GAN model. We then elucidate the details of the proposed graph transformation model based on newly developed graph deconvolution layers with various regularization settings in Sections II-C and II-D. Finally, in Sections II-F and II-G, we describe our new graph neural network explanation techniques for SF-GAN and present analysis for model complexity, respectively. The schematic for our method is included in Section II-B in Supplementary Material.

A. Problem Formulation

This work focuses on brain networks with graph representation. Specifically, we define SC as $G = (V, A)$, where G serves as an undirected weighted graph. V is the set of N nodes $V = \{v_i | i \in [1, N]\}$ representing the gray matter regions in the brain. The weighted adjacency matrix $A \in \mathbb{R}^{N \times N}$ corresponds to a collection of node pairs where $A_{i,j} \in [0, \infty)$ is the weighted connectivity of the edge between nodes i and j . It reflects the connectivity strength between any two nodes (i, j) based on the proportion of fiber tracts connecting the two regions per unit surface, normalized by the average fiber length measured through dMRI tractography [10]. An FC is defined as $G' = (V, A')$, which has the same node set as the SC G except that the weighted adjacency is denoted as A' , where $A'_{i,j}$ is the temporal correlation between the time series of different brain regions i and j that reveals the “coactivation” relationship between them. Table I describes the various parameters and variables used in this work.

In neuroscience, modeling and understanding how the SC determines and influences the patterns of FC is a critical issue that could be highly beneficial for potential breakthroughs in understanding mental diseases. In this article, the goal is to generate FC of a subject, given the subject’s SC and other metafeatures (e.g., age, health status, and family history of disease).

Formal Definition of the Problem: This problem can be mathematically formulated as a graph transformation problem that learns the mapping

$$\mathcal{F} : \{\mathcal{G}, \mathbb{R}^{1 \times K}, \mathcal{U}\} \longrightarrow \mathcal{G}'$$

from an input graph, namely, the SC $G \in \mathcal{G}$, along with a set of metafeatures $M \in \mathbb{R}^{1 \times K}$, and a random noise term $u \sim \mathcal{U}$, to a target graph, namely, the FC $G' \in \mathcal{G}'$.

The above-formulated research problem involves several technical challenges, including the difficulty of learning the complex and nonlinear relation between the structural and FC; the difficulty of accounting for the stochastic factors that influence the generated FC; the integration of the heterogeneous inputs of SC and metafeatures, with graph structure and vector structure, respectively; and the need to ensure model interpretability in order to indicate important associations between the input and the output. For example, it is important to know which metafeatures have a strong impact on the formation of FC. This is a challenging task that involves teasing out its influence on FC from its joint coupling with SC.

B. SF-GAN Overall Architecture

To achieve SC-to-FC transformation and address all the abovementioned challenges, this article proposes a new model, SF-GAN. SF-GAN is a new graph conditional GAN, which generates an output graph (i.e., FC) based on an input graph and additional input features, as shown in Fig. 1. Going beyond graph prediction methods, which aim to predict deterministic graph output, the advantage of our proposed GAN-based architecture is its ability to account for the randomness of the generated complex graph. It consists of two major components, namely, the conditional FC generator and the conditional FC

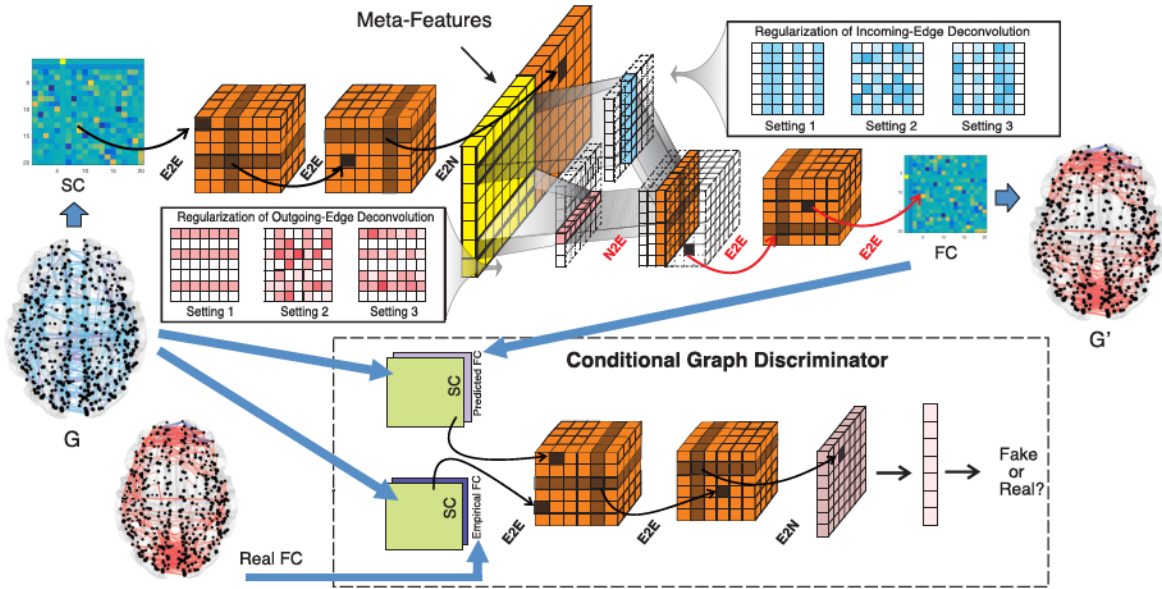


Fig. 1. Proposed SF-GAN framework. The E2E and E2N convolution layers are labeled in black and shown with black arrows, while the deconvolution layers have red labels. Graph skip-nets used in [15] are omitted for more clarity of our model presentation.

discriminator. The former generates an FC as realistically as possible so that it can convince any graph classifier that it is real, while the latter accurately distinguishes the generated graphs as “fake,” compared to the real observed FC. The goals of these two components are adversarial, and hence, the training of our SF-GAN models relies on an iterative process of optimization of each alternately, during which they strengthen each other until no further improvement is found. This process is equivalent to optimize the following general objective function:

$$\begin{aligned} \min_{\mathcal{F}} \max_{\mathcal{D}} \mathcal{L}(\mathcal{F}, \mathcal{D}) = & \mathbb{E}_{G, G'} [\log \mathcal{D}(G'|G)] \\ & + \mathbb{E}_{G, U} [\log(1 - \mathcal{D}(\mathcal{F}(G, M, U)|G))] \\ & + \lambda_1 \mathcal{R}_1(\mathcal{F}) + \lambda_2 \mathcal{R}_2(\mathcal{F}) \end{aligned} \quad (1)$$

where $\mathcal{F}$ and $\mathcal{D}$ denote the FC conditional generator and the FC conditional discriminator, respectively. On the right-hand side, the first term denotes the loss in the discriminator due to classifying real observed FCs, while the second term represents the loss in the generator due to generating the FCs in the conditional generator $\mathcal{F}$. The third and fourth terms introduce additional regularization over the model parameters of the generator. $\mathcal{R}_1(\mathcal{F})$, is the L_1 loss, which enforces sparsity similarity. It is controlled by the hyperparameter λ_1 and is equal to $\mathbb{E}_{A, A', U} [\|A' - \mathcal{F}(G, M, U)\|_1]$. The last term will be presented in Section II-D, after we formally describe the conditional generator in Section II-C where we formulate our proposed deconvolution layers.

C. FC Conditional Generator

The proposed FC conditional generator utilizes a graph encoder–graph decoder framework, where the input of the encoder includes the SC and random noise. The additional metafeatures are input into the information bottleneck, namely, the graph embedding encoded by the aforementioned graph encoder. This addresses the challenges related to heterogeneity

of the two input structures in terms of graphs and vectors. Once past, the information bottleneck, the graph embedding, and the metafeatures are decoded back into the graph domain to form the FC. Graph skip nets are also adopted into the encoding–decoding process, to map the learned latent relations between the input and target graphs but are omitted from Fig. 1 for simplicity. Specifically, the output of the first edge deconvolution layer in the decoder is concatenated with the output of the first edge convolution layer and then input into the second deconvolution layer. A similar technique is applied to the second layers of the encoder and decoder. In the following, we describe the graph convolution and proposed deconvolution layers that constitute the encoder and decoder architecture.

Since SC is a weighted adjacency matrix, the inputs are actually edges rather than nodes. This means that instead of the graph convolution layers normally used for the convolution of node attributes, here, we require edge convolution layers. Specifically, as shown in Fig. 1, we first leverage the E2E convolution layers proposed by Kawahara *et al.* [16], which can aggregate the information for all the adjacent edges and hence learn high-order neighborhood information. Moreover, to further aggregate the result into higher level graph representation, the edge-to-node (E2N) convolution layers have been leveraged to generate graph embedding as shown in Fig. 1 and elaborated in Fig. 2, denoted as $F \in \mathbb{R}^{N \times P}$, where N is the number of nodes and P is the size of each node embedding. We can now describe the proposed graph deconvolution layers as follows.

We proceed by focusing initially on the node-to-edge deconvolution layer, which deconvolves the graph embedding F as well as the metafeatures to a pairwise node relation matrix, and for simplicity, we disregard the bias. Formally, let $E^{0,m} \in \mathbb{R}^{N \times N}$ represent the pairwise node relation matrix for the m th feature map, where $E_{i,j}^{0,m}$ denotes the correlation between nodes i and j . Mathematically, we add four terms to represent

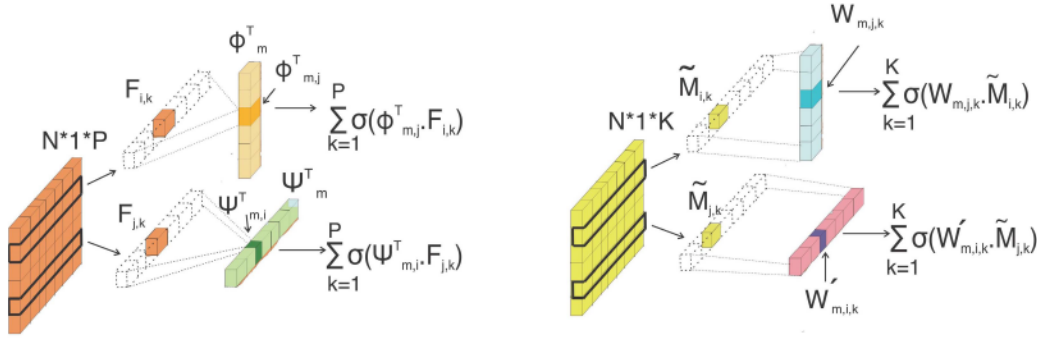


Fig. 2. Node-to-edge deconvolution layer, deconvolves the graph embedding F as well as the metafeatures. The four terms in (2) are demonstrated in this figure.

pairwise node relation between nodes i and j as follows and demonstrate these in Fig. 2:

$$E_{i,j}^{0,m} = \sum_{k=1}^P \sigma(\phi_{m,j}^T \cdot F_{i,k}) + \sum_{k=1}^K \sigma(W_{m,j,k} \cdot \tilde{M}_{i,k}) + \sum_{k=1}^P \sigma(\psi_{m,i}^T \cdot F_{j,k}) + \sum_{k=1}^K \sigma(W'_{m,i,k} \cdot \tilde{M}_{j,k})$$

$$s.t., W_m \in \mathcal{C}, W'_m \in \mathcal{C}' \quad (2)$$

where $\tilde{M} \in \mathbb{R}^{N \times K}$ is the concatenation of N copies of metafeature vectors $M \in \mathbb{R}^{1 \times K}$, $\tilde{M}_{j,k}$ corresponds to the j th node and k th metafeature, $F_{j,k}$ denotes the value of the j th node under the k th latent feature, $\phi_m \in \mathbb{R}^{1 \times N}$ is the node-to-edge deconvolution kernel for “incoming” edges for m th feature map, $\sigma(\cdot)$ is an activation function such as ReLU [17] or sigmoid function, $\psi_m \in \mathbb{R}^{N \times 1}$ is the node-to-edge deconvolution kernel for “outgoing” edges for the m th feature map, $W_m \in \mathbb{R}^{N \times K}$ is the metafeatures’ node-to-edge deconvolution kernel for “incoming” edges for m th feature map, $W_{m,k} \in \mathbb{R}^{N \times 1}$ is a column vector of the k th metafeature, and $W'_m \in \mathbb{R}^{K \times N}$ is the metafeatures’ node-to-edge deconvolution kernel for “outgoing” edges for the m th feature map. Hence, $W'_{m,k} \in \mathbb{R}^{1 \times N}$ is a row vector of the k th metafeature and $\mathcal{C}$ and $\mathcal{C}'$ are all possible desired sparsity patterns of the weights W_m and W'_m , respectively. Some implementations of these are introduced in Section II-D. We can now leverage the E2E deconvolution, which is a reversed process denoted as follows:

$$E_{i,j}^{l,m} = \sum_{n=1}^d \left(\sigma \left(\alpha_j^{l,m} \sum_{k=1}^N E_{i,k}^{l-1,n} \right) + \sigma \left(\beta_j^{l,m} \sum_{k=1}^N E_{k,j}^{l-1,n} \right) \right) \quad (3)$$

where $\alpha^{l,m} \in \mathbb{R}^N$ and $\beta^{l,m} \in \mathbb{R}^N$ are the kernels for the “incoming” and “outgoing” edge deconvolution, respectively; and $E_{i,j}^{l,m}$ is the correlation between nodes i and j in the node correlation matrix of the l th layer and the m th feature map, in the last layer before the final output. Moreover, d refers to the number of relation modes extracted by the previous layer.

D. Regularized Deconvolution for Metafeatures

Based on the mathematical representations of the node-to-edge deconvolution layer provided by (2), we can now describe

the last term, $\mathcal{R}_2(\mathcal{F})$, in (1). The constraints over W_m and W'_m in Section II-C are enforced during the model parameter optimization process via the regularization term $\mathcal{R}_2(\mathcal{F})$. By designing appropriate regularization terms, different forms of the desired constraints can be achieved. In this article, we explore several of the regularized patterns desired for solving the problem of SC-to-FC transformation, introduced as follows.

1) Setting 1 (Identical Sparsity Pattern of Metafeatures):

This regularization requires that among all the metafeatures, only a small subset are important factors that significantly influence the patterns of FC. Each metafeature influences different nodes identically. This notion can be denoted as

$$\forall m, k, i \neq j : W_{m,k,i} \equiv W_{m,k,j}, \text{ and } |\text{supp}(W_m)| \ll N \times K$$

(namely W_m is sparse)

where $\text{supp}(x)$ means the support function of a vector whose outputs are the set of nonzero elements of the input vector x . To achieve the above sparse pattern, in (1), we define $\mathcal{R}_2(\mathcal{F}) = \sum_m \|W_m\|_1 (\forall m, k, i \neq j : W_{m,k,i} \equiv W_{m,k,j})$.

2) Setting 2 (Independent Sparsity Pattern of Metafeatures):

This regularization requires that among all the metafeatures, only a small subset are important factors that significantly influence the patterns of FC. Each metafeature influences different nodes differently. This notion can be denoted as $\text{supp}(W_m) \ll N \times K$ (namely W_m is sparse). This can be easily achieved by once again enforcing the same ℓ_1 norm as that used in Setting 1, but without the equality constraint.

3) Setting 3 (Group Sparsity Pattern of Metafeatures):

Settings 1 and 2 strive to address extreme patterns, where the former assumes that the metafeatures must influence all the nodes the same, while the latter insists that the influence on all the nodes is totally independent. In contrast, Setting 3 tries to take an intermediate case between the above two settings by enforcing the constraint that if a metafeature is (un)important, then it should be (un)important to all the nodes, and the specific strength of the importance can vary across different nodes. This type of regularization aims to achieve a compelling tradeoff between interpretability (depending on whether each metafeature is useful for generating FC) and expressiveness (preserving the model parameter flexibility to handle node heterogeneity due to the different inherent functions of different brain regions). More concretely, the desired pattern can be

expressed in the following:

$$\begin{aligned} \forall m, k, i \neq j : \\ \text{supp}(W_{m,k,i}) \equiv \text{supp}(W_{m,k,j}), \text{ and } |\text{supp}(W_m)| \ll N \times K \\ \text{(namely } W_m \text{ is sparse).} \end{aligned}$$

The above equation calls for a pattern that ensures that for the same feature, W_m will have the same sparsity pattern across all the nodes. This effect can be achieved by enforcing a group sparsity term such as an $\ell_{2,1}$ norm on the graph deconvolution kernel, namely, $\mathcal{R}_2(\mathcal{F}) = \sum_m \|W_m\|_{2,1} = \sum_m \sum_k \|W_{m,k}\|_2$.

E. FC Conditional Discriminator

The utilized SF-GAN conditional graph discriminator is designed to accept two graphs (one pair) as the input and decide whether the two are related or not. Specifically, the input pair is formed by concatenating SC with either predicted or real FC. The input here contains graph edge information and hence is passed through edge convolution layers to generate node representations. These embeddings are then summed up and fed to a softmax layer to generate the final output. This architecture is also shown in Fig. 1.

F. Subgraph Mapping Discovery Between SC and FC

Beyond merely doing SC-to-FC prediction based on our SF-GAN framework, it is also interesting and important to figure out which subgraphs in the SC majorly influence which subgraphs in the functional one. For example, via sophisticated experimental design, neuroscientists investigated and found that strong FC weights are present between regions with no direct SCs [18]. In [19], the researchers combined FC with DTI data prove that FC reflects, to a large degree, the underlying SC. However, the research domain is still looking forward to an automatic way to quickly probe the data and find candidate pairs of structural and functional subgraphs that have strong correlation.

To accomplish the above goal, we propose a new *post hoc* explanation technique for explaining our SC-to-FC prediction results. To discover the underlying mapping patterns between SC and FC, we define the problem as subgraph mapping patterns discovery between SC and FC, which is to find mappings from subgraphs in SC to subgraphs in FC, such that the nodes are well connected within each subgraph (i.e., intraconnectivity) while also having tight intergraph connections between subgraphs in each pair (i.e., interconnectivity). Mathematically, define $g_i \subseteq G$ and $g'_i \subseteq G'$ as a pair of subgraphs from SC G and FC G' , respectively, the objective of finding C such pairs is given as follows:

$$\begin{aligned} \min_{\{(g_n, g'_n) | g_n \subseteq G, g'_n \subseteq G'\}} \sum_{n=1}^C f(g_n, g'_n, A, A', Q) \\ = \sum_{n=1}^C (f_1(g_n, A) + f_1(g'_n, A') + f_2(g_n, g'_n, Q)) \quad (4) \end{aligned}$$

where C is the number of subgraphs after the partition; the function $f_1(g_n, A)$ is to measure the intraconnectivity of a subgraph g_n of a given graph with adjacency matrix A , while

the function $h(g_n, g'_n, Q)$ is to measure the interconnectivity between the subgraphs g_n and g'_n of each pair, where $Q \in \mathbb{R}^{N \times N \times N \times N}$ denotes the correlation between edges in the first graph and the edges in the second graph such that $Q_{i,j,x,y}$ denotes the correlation between the edge (i, j) in the second graph (i.e., FC) and edge (x, y) in the first graph (i.e., SC). The problem now becomes finding the best sets of subgraphs in the SC graph and FC graph as well as defining the measurements for intraconnectivity and interconnectivity. It is not hard to measure the intraconnectivity of the set of nodes within one graph via the edge weights, and however, measuring the interconnectivity $Q \in \mathbb{R}^{N \times N \times N \times N}$ between SC and FC is challenging. Also, to the best of our knowledge, this is still an open problem.

Here, we propose to take advantage of the SF-GAN model that was trained to learn and uncover the mapping from the SC graph to the corresponding FC graph. Specifically, we leverage the gradient information [22] via computing the partial derivatives of each output edge $A'_{i,j}$ (i.e., each edge in FC graph) with respect to each input edge $A_{x,y}$ (i.e., each edge in the SC graph), as previously introduced in Section II-A as follows:

$$Q_{i,j,x,y} = \frac{\partial A'_{i,j}}{\partial A_{x,y}} \quad (5)$$

where $A'_{i,j}$ corresponds to the edge between nodes i and j in the predicted FC graph and $A_{x,y}$ corresponds to edge between nodes x and y in the input SC graph. Clearly, computing these derivatives, for each specific edge in the output, generates an $N \times N$ matrix, resulting in an N^4 tensor for all edges in the predicted FC.

Although this 4-D tensor q will give us the correspondence between each pair of edges in SC and FC, it is still very obscure to be used directly for measuring the interconnectivity between a set of nodes in SC and a set of nodes in FC. Thus, we propose to perform a compression operation by global average pooling (GAP), which reduces the 4-D tensor q to a 2-D matrix $\hat{q}$. In using GAP, we performed the compression over two directions, one from input and one from the output, such that the outcome of the operation will yield the average correspondence between one node from SC and one node from FC. Specifically, given a node i from the SC graph and a node y from the FC graph, the node-level correspondence weight between nodes i and y can be estimated by taking the average values over all the possible paths that connect these two nodes as follows:

$$\hat{Q}_{i,y} = \frac{1}{N^2} \sum \sum q_{i,j,x,y}. \quad (6)$$

Finally, with the information of A , A' , and $\hat{Q} \in \mathbb{R}^{N \times N}$, we can define a supergraph $\hat{G}$, which is a combination among them such that

$$\hat{A} = \begin{bmatrix} A & \hat{Q} \\ \hat{Q}^T & A' \end{bmatrix}. \quad (7)$$

Here, $\hat{G} = (\hat{V} = V \cup V', \hat{A} = A \cup A' \cup \hat{Q})$, where $\hat{A} \in \mathbb{R}^{2N \times 2N}$, and hence, the problem in (4) can be transferred into

the following:

$$\begin{aligned}
& \min_{\{(g_n, g'_n) | g_n \subseteq G, g'_n \subseteq G'\}} \sum_{n=1}^C (f_1(g_n, A) + f_1(g'_n, A') \\
& \quad + f_2(g_n, g'_n, \hat{Q})) \\
& = \min_{\{(g_n, g'_n) | g_n \subseteq G, g'_n \subseteq G'\}} \sum_{i=1}^C f(g_n, g'_n, A, A', \hat{Q}) \\
& \Rightarrow \min_{\{(g_n \cup g'_n) \subseteq \hat{G}\}} \sum_{i=1}^C f_1(g_n \cup g'_n, \hat{A}) \text{ [According to (7)]}.
\end{aligned} \tag{8}$$

Therefore, $f_1(\cdot)$, as mentioned above, is a score function measuring the connectivity of subgraph $g_n \cup g'_n$ and can be a classical community detection objective such as InfoMap [21], [23]. Here, the goal of InfoMap for optimizing (8) is to find the optimal partition that minimizes the description length over all possible network partitions. Accordingly, the network partition that gives the shortest description length [24] and compresses the data the most also best captures the community structure with respect to the dynamics on the network [21]. Following the existing works, we applied the fast stochastic and recursive search algorithm [20] to minimize (8) and find the best descriptive subgraph mapping.

G. Time Complexity Analysis

Here, we compute the overall time complexity of SF-GAN as the sum of the complexity for the conditional generator and the complexity for the conditional discriminator. The total complexity can be presented as $\mathcal{O}(t*((L_1 + L_2 + L_3)*N^2 M_0^2 + N^2 M_0 + N^2 M_0 R))$, where N is the number of nodes, M_0 is the number of feature maps at each convolution layer, L_1 is the number of E2E convolution layers in the encoder in generator, L_2 is the number of E2E deconvolution layers in the decoder in generator, L_3 is the number of E2E convolution layers in the discriminator, R is the length of the last (fully connected) layer, and t is the number of training samples. Our SF-GAN framework is able to provide a scalable [i.e., $\mathcal{O}(N^2)$] algorithm for mapping SC to FC. This is valuable if we need to use other brain parcellations with a larger number of brain regions, compared to the existing graph neural network generative models [25], [26], which can only scale to small graphs (up to $N = 20$) and often have $\mathcal{O}(N^3)$ or even $\mathcal{O}(N^4)$ computational complexity.

III. EXPERIMENTS

This section describes the experiment settings and performance analyses. To evaluate the performance of the proposed model, extensive experiments were performed using a 64-bit machine, with a 40-GB memory, a 4-Core Intel¹ CPU, and an Nvidia¹ RTX-2080 Ti GPU. The deep architecture was fully implemented by tensorflow 1.13.1. The details of the experiments, including the datasets, comparison methods, and evaluation metrics, are described in turn below. MATLAB and the python scripts written to obtain the results are publicly available at <https://github.com/netemady/SF-GAN->. Details on

data splitting and hyperparameter tuning are included in Appendix-A (see the Supplementary Material), and a discussion on model selection is provided in Section II-A in the Supplementary Material, both available in the same repository.

A. Datasets

The SC and FC datasets used in this study were extracted from the Human Connectome Project (HCP), specifically, the 1200 Subjects Release, February 2017 [27]. HCP aimed to study and share MRI data from 1200 young adult (ages 22–35) subjects, along with their behavioral features. The MRI data and metafeatures used in this study can be downloaded from the HCP website (<https://db.humanconnectome.org/>).

1) *Data Sources*: The (structural, diffusion, and functional) MRI data were preprocessed using the HCP pipeline [8]. For the diffusion MRI, this was followed by the Bayesian Estimation of Diffusion Parameters Obtained using Sampling Techniques, modeling crossing X fibers (BEDPOSTX) algorithm in the FMRIB Software Library [28] (FSL), which models white matter fiber orientations and crossing fibers for probabilistic tractography. The resting-state blood-oxygen-level-dependent functional MRI (r-fMRI) time series data were acquired from a total of 823 participants, in four runs of approximately 15 min for each participant, including two runs on two different days (Days 1 and 2). These measurements were collected with the subject supine and still, with eyes open, to track physiological changes in the brain (i.e., changes in blood flow and oxygen levels) that occur in the resting state, when an explicit task is not being performed [29], [30].

The behavioral and demographic data HCP collected for each participant were utilized as additional input for our model. Relevantly, a public data dictionary was released that included 357 features related to each subject in the set of which 161 were selected as the input to our model. A list of the input features is provided in the Supplementary Material. These features were divided into five categories: subject information (e.g., age), health and family history (e.g., height, BMI, and family history of disease), psychiatry and life function (e.g., anxiety and attention problems score), sensory (odor identification, hearing and pain tests, and so on), and substance use (e.g., test result for methamphetamine and cocaine). When training our proposed SF-GAN model, our primary goal was to identify which of these five categories had the greatest impact on improving the prediction compared to the use of SC alone. Once identified, the most influential categories could then be used to ensure a more efficient feature collection process for the FC prediction of additional subjects, with a few key factors delivering a similar performance to that achieved using all 161 metafeatures.

2) *Extracting SC and FC*: To construct the SC matrix for each subject, we ran Probtrackx in FSL with 68 regions of interest (ROIs) obtained from the Desikan–Killiany atlas [31]. For the remaining parameter setting in Probtrackx, we followed the recommendations of the tutorial [32] provided by HCP. Finally, the resulting SC matrices were normalized by dividing the respective row sum from each nonzero value.

Three steps were followed to extract the FC from the r-fMRI time series data, for each day: 1) concatenate the time series

¹Registered trademark.

for the two runs together; 2) for each of the 68 ROIs defined by the Desikan–Killiany atlas, average all the time series to create a single ROI time series; and 3) obtain the FCs by either computing the pairwise ROI time series' Pearson correlations using FSLNets [33] with the full correlation option, thus generating Dataset 1; or following similar three steps as mentioned for Dataset 1, except that we concatenate the time series for the two runs performed in Day 2 together, thus generating Dataset 2.

B. Comparison Methods

The following FC prediction methods were included in the performance comparison. The SC was normalized before using it as input for all the techniques.

- 1) *Deep Graph Spectral Evolution Network (GSEN)* [34]: The graph topology evolution problem was modeled by the composition of newly developed generalized graph kernels. GSEN was designed to represent many sophisticated phenomena that require the involvement of sequential or simultaneous graph kernels.
- 2) *Connectome Embedding (CE)* [35]: This technique creates embeddings of input SC to a meaningful low-dimensional vector space. It will then employ the embeddings to construct predictive deep models of functional and SC.
- 3) *CE-Aligned* [36]: Built upon CE, this method employs CEs to improve structural to FC mapping in individuals with a novel embedding alignment approach.
- 4) *Graph Diffusion Model* [3]: A linear network of brain dynamics based on graph diffusion was derived and compared against nonlinear approaches. A diffusion process was modeled based on the underlying structural network to obtain a deterministic solution for FC and the global diffusivity parameter was estimated for each dataset by finding the critical (optimum) value over all subjects using curve fitting to fit an exponential function on all data points.
- 5) *Eigen Decomposition Model* [38]: An eigen relationship between structural and FC networks was established via Laplacian spectra based on the FC and SC found to share eigenvectors whose eigenvalues were exponentially related.
- 6) *Weighted Sum of Matrix Powers* [39]: Using two functional imaging modalities, namely, fMRI and MEG, the mapping between SC and stationary FC was described by a mathematical function represented as a weighted sum of matrix powers. The weights were obtained by computing the normalized sum of squares for each point.
- 7) *Spontaneous Neural Activity Model* [14]: Focusing on the microscopic level to describe the synaptic connections between neurons, a general mathematical model of neural dynamics was derived. This mathematical equation describing the variation in the firing rate in the neurons could then be applied to identify the link between dominant patterns of spontaneous activity and the underlying network architecture.
- 8) *Baseline* [15]: By replacing the ReLU activation in the last layer of generator with tanh, it became possible to

make negative FC predictions as well as positive ones. The hyperparameters were tuned separately for each input dataset.

The performance of our proposed SF-GAN method under the three regularized patterns was compared with the results obtained using the above five methods and reported in Section III-D.

C. Evaluation Metrics

In order to evaluate how effectively SF-GAN and other techniques can reveal the translation rules between the predicted and empirical graphs, the Pearson correlation coefficient, r , was computed for each pair. This metric, which represents the linear correlation between estimated and ground-truth data, is computed between the upper triangular values of the two connectivity matrices and averaged over all pairs in the validation set. Before employing any of the comparison methods or our proposed SF-GAN, the SC matrix must be normalized in a preprocessing step.

The distance measurement between the generated and real graphs was calculated using mean squared error (MSE) and mean absolute error (MAE) metrics for all eight methods over the two datasets. Higher r and lower MSE and MAE show better performance.

D. Performance

1) *Model Accuracy Analysis*: This section examines the performance of the proposed and comparison methods for Datasets 1 and 2. The results for the Pearson correlation, MSE, and MAE are presented in Table I; the best performance for each dataset among 11 different techniques tested is indicated in bold print.

As the data shown in Table I, SF-GAN consistently outperformed the four nondeep learning comparison methods in terms of the Pearson correlation, with a higher discrepancy between the average Pearson correlation of the nondeep learning and SF-GAN techniques (77.14%) being observed for Dataset 2 compared to Dataset 1 (73.33%). As a general observation, for nondeep learning, and all GAN techniques, the Pearson correlation increased moving from Dataset 1 to Dataset 2, by at least 9.09%. On a similar comparison, we can compute an improvement in Pearson correlation equal to 32.45% for Dataset 1, from state-of-the-art deep learning methods (GSEN, CE, and CE-aligned) to SF-GAN techniques. This rise equals 41.32% for Dataset 2, when comparing SF-GAN to state-of-the-art deep learning methods.

The impact of integrating metafeatures to the model is also revealed by the data in Table I, with the metadata significantly improving the Pearson correlation for both datasets; with an improvement of 5.41% for Setting 1 and 10.25% for Settings 2 and 3 (Dataset 1), and 6.67% and 4.55% for Dataset 2. This establishes the value of incorporating metafeatures using node embedding and regularization for the node layer.

We also computed the MSE as another well-recognized graph metric to evaluate the predictions generated by SF-GAN and the comparison methods. The superiority of the deep

TABLE I
PEARSON CORRELATION, MSE, AND MAE RESULTS FOR THE TWO DATASETS

Methods	Dataset 1			Dataset 2		
	$\uparrow r$	$\downarrow$ MSE	$\downarrow$ MAE	$\uparrow r$	$\downarrow$ MSE	$\downarrow$ MAE
Graph diffusion	-	$> 10^3$	2.55	0.07	139.43	0.98
Eigen decompose	0.10	$> 10^3$	$> 10^3$	0.11	$> 10^3$	$> 10^3$
Weighted sum	0.11	145.01	1.01	0.12	$> 10^3$	1.55
Neural activity	0.09	$> 10^3$	$> 10^3$	-	$> 10^3$	7.39
GSEN	0.28	68.19	0.53	0.30	69.12	0.53
CE	0.23	71.14	0.53	0.23	70.18	0.54
CE-aligned	0.25	75.17	0.56	0.24	77.14	0.56
Baseline	0.35	70.17	0.54	0.42	62.33	0.49
SF-GAN Setting 1	0.37	69.28	0.51	0.45	55.06	0.47
SF-GAN Setting 2	0.39	67.51	0.52	0.44	57.02	0.48
SF-GAN Setting 3	0.39	68.06	0.51	0.44	57.05	0.48

TABLE II
EFFICIENCY RESULTS FOR SF-GAN AND FIVE COMPARISON METHODS

Methods	Training time		Testing time	
	mean	std	mean	std
Graph diffusion	8.28	0.97	1.18	0.13
Eigen decomposition	8.65	0.60	1.25	0.06
Weighted sum	8.73	0.86	1.17	0.08
Neural activity	2.58	0.09	2.58	0.08
Baseline	86.25	1.45	1.98	0.09
SF-GAN Setting 1	79.19	1.86	1.97	0.10
SF-GAN Setting 2	86.95	1.56	2.17	0.13
SF-GAN Setting 3	93.04	0.96	2.27	0.10

learning techniques (both SF-GAN and comparison methods) is again demonstrated by this metric, in addition to the findings for the Pearson correlation. Graph diffusion, eigen decomposition, and spontaneous neural activity all performed markedly worse than the others for both datasets, with the poorest value of greater than 10^3 being obtained for nondeep learning methods in several cases. The GAN and the rest of deep learning-based techniques exceed all the comparison methods for both datasets, with a high margin in all instances.

To investigate the predictions of all the techniques more closely, we also used a third metric, MAE (Table I). The GAN techniques for both datasets outperform all other methods, with the best performance for SF-GAN Settings 1 and 3, and SF-GAN Setting 1, for the first and second datasets, respectively.

2) *Model Efficiency Analysis*: The training and test times for the proposed technique and the nondeep learning comparison methods for Dataset 2 are presented in Table II. The training time is reported for 30 epochs for all the GAN methods. The results for the other dataset are not presented here as they followed a similar pattern. The shortest training time is for spontaneous neural activity.

The deep learning-based methods tend to consume more time to train, largely due to their more sophisticated architecture. This is still at most 1.57 min, making the proposed model very practical for real-world applications. The test times for the baseline and SF-GAN are within the same range as the other methods; the times shown are for 100 samples, making the test time per sample less than 0.03 s for all models, which is negligible.

3) *Model Scalability Analysis*: Our proposed method's scalability was evaluated in terms of the number of edges,

metafeatures, and samples for both training and test sets. The number of edges in the SC (and FC) matrices can be gradually increased by relaxing the threshold required for each pair of ROIs to be considered connected. The results on 20 repeated experiments for each setting and both datasets follow the pattern in Figs. 3 and 4. The data show that adding more edges to the graphs by lowering the threshold creates curves that follow a generally constant trend. As the number of metafeatures increases, the training and test time are not affected to a great extent, as shown in Figs. 5 and 6. These results are consistent with the results of the time complexity analysis reported in Section II-G.

Scaling the total number of samples used for training and test with a constant ratio reveals a linear relationship between the time it takes to train/test the model and the number of samples in each set (Figs. 7 and 8). This relationship is also shown in Section II-G. We provide additional discussions of results as the Supplementary Material at github repository addressed in Section III.

E. Model Feature Selection

In this section, we describe the feature selection process utilized by SF-GAN. As explained in Section II, the proposed method integrates and incorporates metafeatures to generate the final prediction. This is possible because the W and W' kernels are leveraged to form the node-to-edge deconvolution layer in the conditional generator architecture shown in Fig. 1. The constraints enforced by the model parameter optimization determine the contribution of each individual metafeature.

To study the influence of the metafeatures, we compared the Pearson correlation r , with comparison methods, for the two datasets and three settings, in Table I. The metadata clearly improved r for both datasets, compared to the baseline. We, therefore, performed further analysis on Datasets 1 and 2, to find out more about how these features contributed in improving the performance.

Statistical Analysis: Here, the goal was to determine and report which category(ies) among the five options presented in Section III-A1 has the greatest impact on improving the model's prediction capability. Computing the mean value for the weights W in the trained model, we obtain a vector with dimension equal to the total number of metafeatures, which represents the mean weight corresponding to each individual feature. We can now calculate the average weight for each

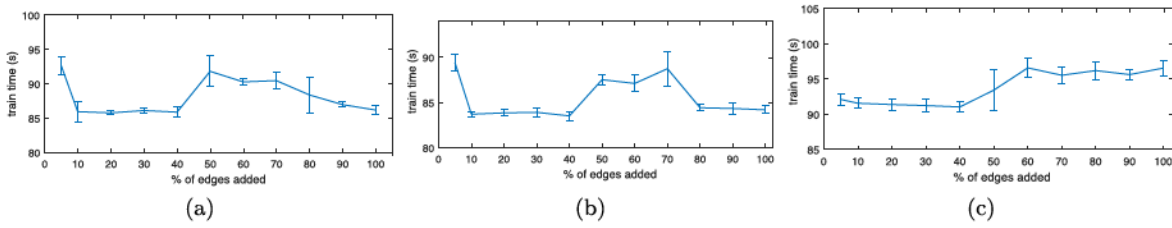


Fig. 3. Training time versus percentage of edges. (a) Setting 1. (b) Setting 2. (c) Setting 3.

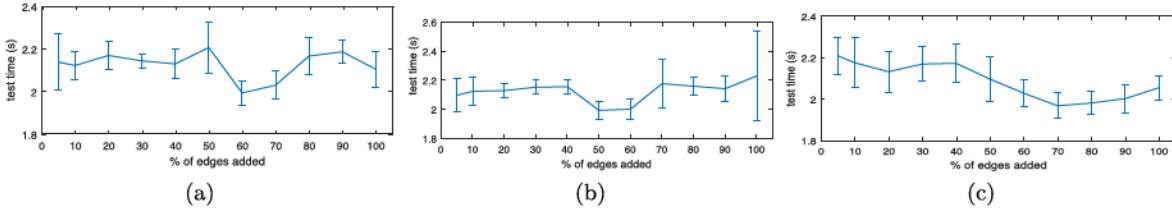


Fig. 4. Testing time versus percentage of edges. (a) Settings 1. (b) Setting 2. (c) Setting 3.

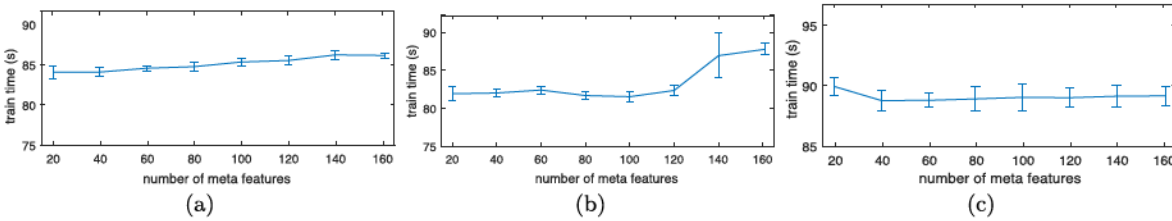


Fig. 5. Training time versus number of metafeatures. (a) Setting 1. (b) Setting 2. (c) Setting 3.

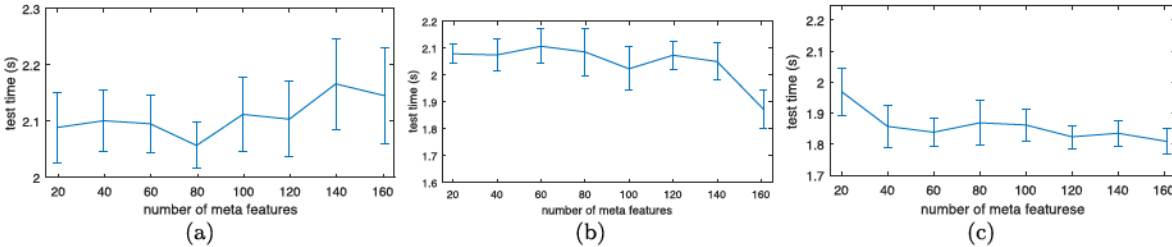


Fig. 6. Testing time versus number of metafeatures. (a) Setting 1. (b) Setting 2. (c) Setting 3.

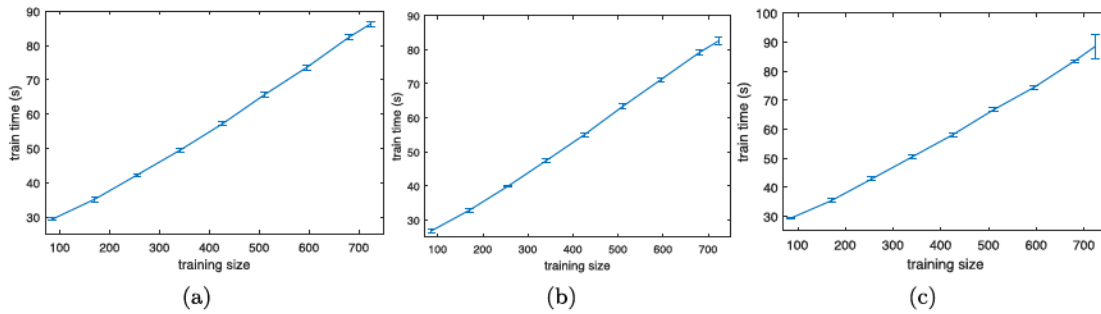


Fig. 7. Training time versus size of training set. (a) Setting 1. (b) Setting 2. (c) Setting 3.

category. The results show that the “subject information” category exhibits the highest mean value overall. Continuing the analysis, a one-way ANOVA with five groups identifies that the results are significant at the 5% significance level. This is revealed by computing $F_{\text{critical}} = F(x, y) = F(4, 156)$ (for x equal to between groups degree of freedom and y equal to within groups degree of freedom), and $F_{\text{anova}} = 4.83$. As it can be seen, $F_{\text{critical}} < F_{\text{anova}}$, with $p\text{-value} = 0.001$. A “*post hoc*” analysis of the group means using the least significant difference (LSD) test showed that the “subject information”

category was statistically significant compared to the other categories. These results are plotted as a bar graph in Fig. 9. As the list of metafeatures provided in the Supplementary Material demonstrates, age is included in this category, which has already been shown to be influential by previous studies reported in the literature [7].

F. Qualitative Analysis of the Visualized Connectivities

During the experimental process, we observed numerous interesting predictions made by our models on the information

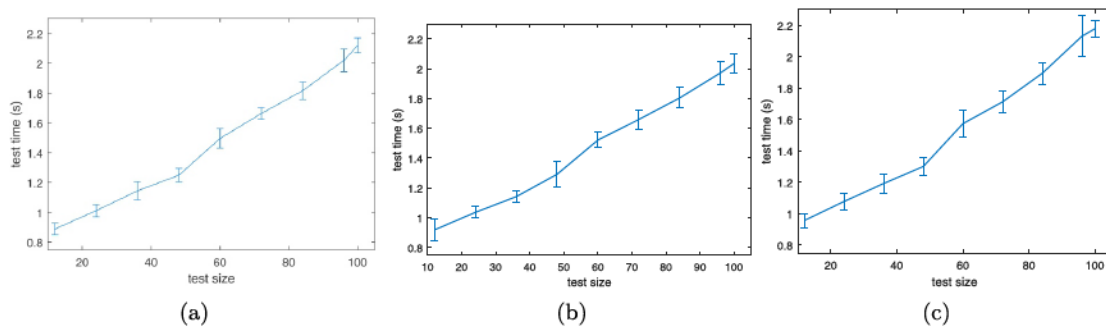


Fig. 8. Testing time versus size of test set. (a) Setting 1. (b) Setting 2. (c) Setting 3.

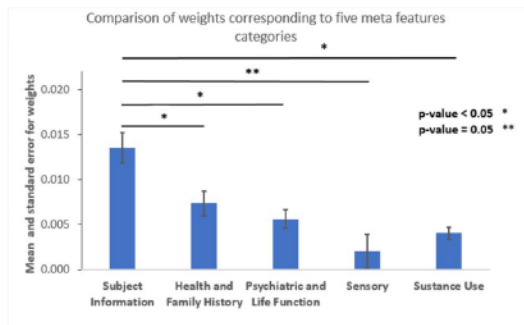


Fig. 9. Bar graph comparing the weights of each of the five categories of metafeatures.

contained in Datasets 1 and 2. Due to space limitations, we are only able to present an example here. Fig. 10 shows the case study for subjects 21. Each figure includes a single subject's SC (the left-hand column), empirical FC (the middle column), and predicted FC (the right-hand column). The predicted FC using each subject's SC is very close to the same subject's empirical FC, even though they possess very similar adjacency matrices for SC. This is because the SC is defined by the human brain's anatomical neural network, which has relatively few individual differences among human beings [34]. The case studies for more subjects are shown in the Supplementary Material.

The SC and FC between brain regions are visualized in Fig. 11. The nodes represent regions in the brain in terms of the coordinates utilized in the HCP database, and the graphs are plotted using the BrainNet Viewer toolbox. The corresponding ranges for SC and FC are shown by the various colorbars. As we can see, the true and predicted values for FC include both positive and negative edges.

IV. SUBGRAPH DISCOVERY AND *Post Hoc* ANALYSIS OF RESULTS

As proposed in Section II-F, we can apply our novel design to find the best partition of nodes in the SC and FC graphs. These partitions would have tight connections within, in addition to between subgraphs (i.e., intraconnectivity and interconnectivity). To do that, we optimized the objective defined in 8, to find the best partition M , with minimum description length, or $L(M)$. We used the adjacency matrices A and A' , in SC and FC, plus the compressed gradient using GAP, defined by 5, to compute all terms in 8. The result of partitions for Subject 1, Dataset 1, is shown in Fig. 12.

In each subfigure, the left image shows the partition in the SC, and the right image corresponds to partition in FC graphs. Each specific color is used in the two images, to show the subgraphs that are mapped from input to output. We have included additional images showing these patterns for more subjects, in the Supplementary Material.

Here, we report and summarize our findings, for 500 subjects, as follows.

- 1) In more than 70% of the mappings discovered, the subgraphs had exactly the same nodes that exist in SC and FC. For the remaining mappings, between 50% and 80% of nodes were the same.
- 2) In more than 70% of the mappings, nodes in the right hemisphere in SC determine nodes in the right hemisphere of FC. This value was about 78% for the left hemisphere.
- 3) We also investigated the number of edges in each subgraph of SC, for all mappings, and all subjects, that contributed to predicting one or more edges in FC. In more than 92% of input subgraphs, several (more than 5) edges existed. This can show that, rather than a one-to-one mapping, there is a cluster-to-cluster correspondence from SC to FC.
- 4) In addition, for about 89% of mappings, we found several significant structural edges being responsible for fewer of the functional edges.

In addition to studying subjects individually, we computed the average interconnectivity and intraconnectivity values for all subjects in our dataset. Here, we also report the observations we had for subgraphs discovered by applying these average connectivity values to 8. In summary, we identified nine subgraphs among the brain regions considering both SC and FC. Subgraph 1 included 28 and 30 nodes from SC and FC, respectively, accounting for a large percentage of the total 68 brain regions. This suggests that the association between SC and FC for this subgraph is largely global and less modular. There was also a significant overlap of nodes from SC and FC being grouped to the same subgraph. For example, in subgraph 1 out of 28 and 30 brain regions from the SC and FC graphs, respectively, 26 nodes are common brain regions. In subgraph 6, right entorhinal cortex, right parahippocampal gyrus, and right fusiform gyrus were found to form subgraph in SC and influence ten brain regions in FC across bilateral hemispheres. Entorhinal cortex, parahippocampal gyrus, and fusiform gyrus are implicated in memory and human face

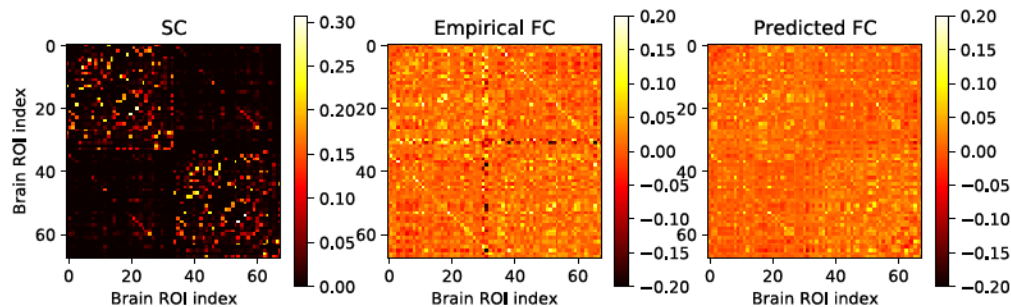


Fig. 10. Case study: heat-maps for subject 21.

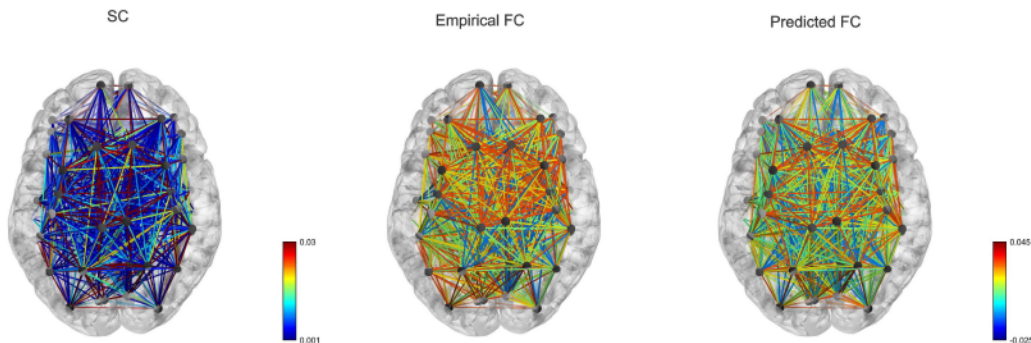


Fig. 11. Case study: brain connectivity for subject 21.

recognition. Interestingly, we also found subgraphs that only involve SC nodes, including subgraphs 4 and 9.

V. RELATED WORK

A. Deep Learning for Graph Transformation

Graph neural networks have been widely studied with the general focus on graph recurrent networks or graph convolutional networks. The former, which is based on recursive neural networks and was initially proposed in [40], has been an attractive topic. Bacciu *et al.* [41] proposed a novel approach based on a node ordering procedure and two RNNs, each of which generated one endpoint of an edge in the predicted graph. The latter, which is based on the generalization of the idea of convolutions from grids to graph data, utilizes a series of building blocks to form more complex graph neural network models, including generative models [42]. Generative models can be used for graph topology prediction [43], through design of an encoder-decoder framework. Along the same track, Guo *et al.* [44] established a new framework based on disentangled representation learning for deep generative models.

B. Feature Selection

Feature selection, initially proposed as a data preprocessing approach, has successfully solved critical issues, such as the curse of dimensionality, overfitting, high memory, and computational complexity in data analytics-related fields [45]. For the purposes of the current study, it is convenient to describe traditional feature selection algorithms for conventional data in terms of four groups. The first group assigns value to features based on how well they preserve similarity in the data. By defining some utility function over the features and

selecting the top k features that maximize their individual utility, this approach has proved useful in both supervised and unsupervised problems and has low computational complexity, although most methods in this category cannot handle feature redundancy. The second group is based on a conditional likelihood maximization framework. Although these methods maximize feature relevance and minimize feature redundancy, they can only be applied to discrete variables. The third group consists of sparse regularization terms that eliminate some features through a typical learning algorithm. This method has high interpretability and good performance, but at a high computational cost. Finally, the fourth group consists of methods that rely on various statistical measures to assess feature relevance. These methods do not address feature redundancy issues.

A newly proposed feature selection method that applies the Group Lasso penalty to inputs of neural networks is proposed by Zhang *et al.* [46], where they claim that their method represents an improvement over the primary Group Lasso, eliminates the need for additional weights, and promises to solve the nondifferentiability problem at the origin.

C. SC-to-FC Prediction

Inferring functional interactions between different brain regions from patterns of anatomical connectivity has been the focus of many studies over the last few decades. The literature in this domain can be categorized into three types: linear models based on a diffusion pattern, heuristic models based on modeling the dynamics of neurons' activity, and graph theory-based models. Examples of works that fall into the first category by incorporating diffusion process in their model include: Suárez *et al.* [48] argued that the correspondence between SC and FC in most existing models is imperfect and

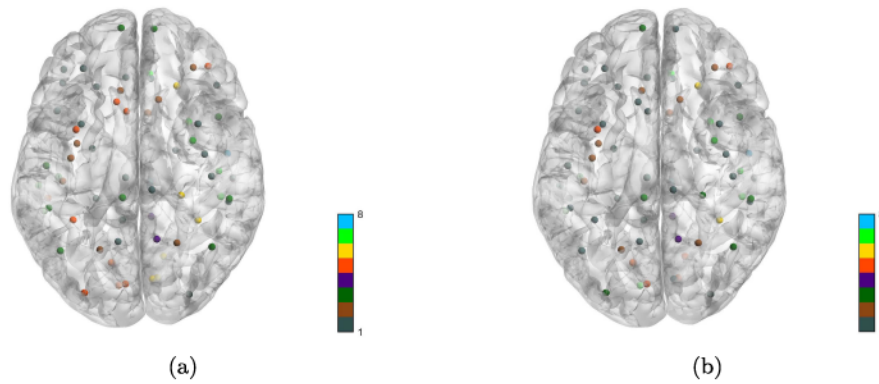


Fig. 12. Subgraph mapping for subject 1-Dataset 1. (a) Patterns of SC. (b) Patterns of FC.

suggest the use of statistical, communication, and biophysical models to assign weights to connectivities, incorporate one or more signaling mechanisms (i.e., routing or diffusion), or generate stimulation-induced activity. These models are regionally heterogeneous and integrated with cellular and molecular metadata, and however, the complexity for such model is not discussed or compared with other findings. Tarun *et al.* [49] proposed an integrated diffusion/fMRI approach based on the analysis of the default mode network (DMN) and the brain regions reported to be more involved during the resting state.

The second category of techniques includes those that make assumptions about the dynamics of neuronal activity. The comprehensive comparison of seven computational-based models on the prediction of FC from SC presented in [10] concluded that all these models can be simplified to a single-core stationary linear process based on the simultaneous autoregressive (SAR) model and discussed the limited predicted power of such models in their current form with respect to empirical FC. The work reported in [9] focused on the origins of FC, proposing a new modeling method for the spiking attractor network of the brain. This approach represents a heuristic model very effectively as it makes assumptions regarding the existing mechanisms at the level of each single brain area based on spiking neurons and realistic synapses. You *et al.* [50] proposed a regularized multiple regression approach in estimating the interactions between SC and FC for task-evoked fMRI data but did not apply their method to rsfMRI measurements. Proposed a regularized multiple-regression approach that adapts to non-Gaussian data. Finally, the third type of models incorporates graph theory-based knowledge in their prediction. Becker *et al.* [11] proposed a spectral graph theory-based approach that applies a nonlinear mapping of SC to resting-state FC by computing a weighted combination of the structural matrix and then changing the coordinates to align the eigenmodes of the target and input graphs. Using linear algebra on one side and empirical and synthetic data on the other, Tewarie *et al.* [12] concluded that functional networks can be described by all possible walks in the structural network, and this is equivalent to using the eigenmodes of the structural network.

An extensive search of the existing literature in this domain revealed that joint consideration of the nonlinearity, complexity, and stochasticity of the SC-FC link remains largely

unexplored. In addition to this challenge, assessing the influence of various features of the subjects and proposing designs that provide different dynamic patterns corresponding to different brain regions within an integrated framework can potentially enhance the generalizability of the model considerably.

VI. CONCLUSION

In this article, we have presented SF-GAN, a novel graph-based conditional generative-adversarial-nets technique to infer resting-state FC from anatomical structural measurements. Our proposed design is capable of capturing highly nonlinear, complex, stochastic relationships between the two connectivities and innovatively take the subjects' profiles into account by integrating factors such as age, health, and family history of disease into our model. To make this possible, a new framework is proposed, namely, the conditional generator, which is designed to encode the SC into a graph representation, integrate the result with the metadata, and decode the combination back into the graph domain to represent FC. New sparse-regularized edge deconvolution layers are proposed, which can automatically select the key features needed for constructing FC. In addition, *post hoc* analysis of our SF-GAN model can identify which subgraphs in SC strongly influence which subgraphs in FC by solving a new multilevel edge-correlation-guided graph clustering problem. As a current limitation of this work, there are additional metafeatures provided by HCP, which are not considered here. Also, investigating task MRI data is of interest besides resting-state datasets. We plan to move on to explore these topics in the subsequent work in the near future.

REFERENCES

- [1] L.-A. Harsan *et al.*, "In vivo diffusion tensor magnetic resonance imaging and fiber tracking of the mouse brain," *NMR Biomed.*, vol. 23, no. 7, pp. 884–896, Aug. 2010.
- [2] H. Finger *et al.*, "Modeling of large-scale functional brain networks based on structural connectivity from DTI: Comparison with EEG derived phase coupling networks and evaluation of alternative methods along the modeling path," *PLOS Comput. Biol.*, vol. 12, no. 8, Aug. 2016, Art. no. e1005025.
- [3] 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, Apr. 2014.
- [4] S. A. Huettel, A. W. Song, and G. McCarthy, *Functional Magnetic Resonance Imaging*, vol. 1. Sunderland, MA, USA: Sinauer Associates, 2004.

- [5] S. Kann, S. Zhang, P. Manza, H.-C. Leung, and C.-S. R. Li, "Hemispheric lateralization of resting-state functional connectivity of the anterior insula: Association with age, gender, and a novelty-seeking trait," *Brain Connect.*, vol. 6, no. 9, pp. 724–734, 2016.
- [6] H. Yuan, H. Yu, S. Gui, and S. Ji, "Explainability in graph neural networks: A taxonomic survey," 2020, *arXiv:2012.15445*.
- [7] B. B. Biswal *et al.*, "Toward discovery science of human brain function," *Proc. Nat. Acad. Sci. USA*, vol. 107, no. 10, pp. 4734–4739, 2010.
- [8] H. C. P. WU-Minn. (2017). *1200 Subjects Data Release Reference Manual*. [Online]. Available: <https://www.humanconnectome.org>
- [9] G. Deco and V. K. Jirsa, "Ongoing cortical activity at rest: Criticality, multistability, and ghost attractors," *J. Neurosci.*, vol. 32, no. 10, pp. 3366–3375, 2012.
- [10] A. Messé, D. Rudrauf, A. Giron, and G. Marrelec, "Predicting functional connectivity from structural connectivity via computational models using MRI: An extensive comparison study," *NeuroImage*, vol. 111, pp. 65–75, May 2015.
- [11] C. O. Becker *et al.*, "Spectral mapping of brain functional connectivity from diffusion imaging," *Sci. Rep.*, vol. 8, no. 1, pp. 1–15, Dec. 2018.
- [12] P. Tewarie *et al.*, "Mapping functional brain networks from the structural connectome: Relating the series expansion and eigenmode approaches," *NeuroImage*, vol. 216, Aug. 2020, Art. no. 116805.
- [13] S. G. Surampudi, S. Naik, R. B. Surampudi, V. K. Jirsa, A. Sharma, and D. Roy, "Multiple kernel learning model for relating structural and functional connectivity in the brain," *Sci. Rep.*, vol. 8, no. 1, pp. 1–14, Dec. 2018.
- [14] R. F. Galán, "On how network architecture determines the dominant patterns of spontaneous neural activity," *PLoS ONE*, vol. 3, no. 5, p. e2148, May 2008.
- [15] X. Guo, L. Wu, and L. Zhao, "Deep graph translation," *IEEE Trans. Neural Netw. Learn. Syst.*, early access, Mar. 17, 2022, doi: 10.1109/TNNLS.2022.3144670.
- [16] J. Kawahara *et al.*, "BrainNetCNN: Convolutional neural networks for brain networks; towards predicting neurodevelopment," *NeuroImage*, vol. 146, pp. 1038–1049, Feb. 2017.
- [17] P. Veličković, G. Cucurull, A. Casanova, A. Romero, P. Liò, and Y. Bengio, "Graph attention networks," 2017, *arXiv:1710.10903*.
- [18] C. J. Honey *et al.*, "Predicting human resting-state functional connectivity from structural connectivity," *Proc. Nat. Acad. Sci. USA*, vol. 106, no. 6, pp. 2035–2040, Feb. 2009.
- [19] M. D. Greicius, K. Supekar, V. Menon, and R. F. Dougherty, "Resting-state functional connectivity reflects structural connectivity in the default mode network," *Cerebral Cortex*, vol. 19, no. 1, pp. 72–78, 2009.
- [20] M. Rosvall and C. T. Bergstrom, "Maps of random walks on complex networks reveal community structure," *Proc. Nat. Acad. Sci. USA*, vol. 105, no. 2, pp. 1118–1123, 2008.
- [21] M. De Domenico, A. Lancichinetti, A. Arenas, and M. Rosvall, "Identifying modular flows on multilayer networks reveals highly overlapping organization in interconnected systems," *Phys. Rev. X*, vol. 5, no. 1, Mar. 2015, Art. no. 011027.
- [22] K. Simonyan, A. Vedaldi, and A. Zisserman, "Deep inside convolutional networks: Visualising image classification models and saliency maps," 2013, *arXiv:1312.6034*.
- [23] M. Rosvall, D. Axelsson, and C. T. Bergstrom, "The map equation," *Eur. Phys. J. Special Topics*, vol. 178, no. 1, pp. 13–23, Nov. 2009.
- [24] P. D. Grünwald, I. J. Myung, and M. A. Pitt, *Advances in Minimum Description Length: Theory and Applications*. Cambridge, MA, USA: MIT Press, 2005.
- [25] M. Simonovsky and N. Komodakis, "GraphVAE: Towards generation of small graphs using variational autoencoders," in *Proc. Int. Conf. Artif. Neural Netw.* Cham, Switzerland: Springer, 2018, pp. 412–422.
- [26] B. Samanta *et al.*, "NEVAE: A deep generative model for molecular graphs," *J. Mach. Learn. Res.*, vol. 21, no. 114, pp. 1–33, Apr. 2020.
- [27] D. C. Van Essen *et al.*, "The WU-Minn human connectome project: An overview," *NeuroImage*, vol. 80, pp. 62–79, Oct. 2013.
- [28] M. Jenkinson, C. F. Beckmann, T. E. Behrens, M. W. Woolrich, and S. M. Smith, "FSL," *NeuroImage*, vol. 62, no. 2, pp. 782–790, 2012.
- [29] R. L. Buckner, F. M. Krienen, and B. T. T. Yeo, "Opportunities and limitations of intrinsic functional connectivity MRI," *Nature Neurosci.*, vol. 16, no. 7, pp. 832–837, Jul. 2013, doi: 10.1038/nn.3423.
- [30] B. B. Biswal, "Resting state fMRI: A personal history," *NeuroImage*, vol. 62, no. 2, pp. 938–944, Aug. 2012.
- [31] R. S. Desikan *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, Jul. 2006.
- [32] *HCP Tutorial*. Accessed: Aug. 1, 2020. [Online]. Available: <https://wustl.app.box.com/s/wna2cu94pqgt8zskg687mj8z-lmfj1pq7>
- [33] *FSL-Scripts*. Accessed: Aug. 1, 2020. [Online]. Available: https://github.com/aheckel/fsl-scripts/blob/master/rsc/scripts/bin/FSLNets/nets_netmats.m
- [34] N. Etemadyrad, Q. Li, and L. Zhao, "Deep graph spectral evolution networks for graph topological evolution," in *Proc. AAAI Conf. Artif. Intell.*, vol. 35, no. 8, 2021, pp. 1–9.
- [35] G. Rosenthal *et al.*, "Mapping higher-order relations between brain structure and function with embedded vector representations of connectomes," *Nature Commun.*, vol. 9, no. 1, pp. 1–12, Dec. 2018.
- [36] G. Levakov, J. Faskowitz, G. Avidan, and O. Sporns, "Mapping individual differences across brain network structure to function and behavior with connectome embedding," *NeuroImage*, vol. 242, Nov. 2021, Art. no. 118469.
- [37] S. M. Smith *et al.*, "Functional connectomics from resting-state fMRI," *Trends Cognit. Sci.*, vol. 17, no. 12, pp. 666–682, 2013.
- [38] F. Abdelnour, M. Dayan, O. Devinsky, T. Thesen, and A. Raj, "Functional brain connectivity is predictable from anatomic network's Laplacian Eigen-structure," *NeuroImage*, vol. 172, pp. 728–739, 2018.
- [39] J. Meier *et al.*, "A mapping between structural and functional brain networks," *Brain Connectivity*, vol. 6, no. 4, pp. 298–311, May 2016.
- [40] M. Gori, G. Monfardini, and F. Scarselli, "A new model for learning in graph domains," in *Proc. IEEE Int. Joint Conf. Neural Netw.*, Jul. 2005, pp. 729–734.
- [41] D. Bacciu, A. Micheli, and M. Podda, "Edge-based sequential graph generation with recurrent neural networks," *Neurocomputing*, vol. 416, pp. 177–189, Nov. 2020.
- [42] Z. Wu, S. Pan, F. Chen, G. Long, C. Zhang, and P. S. Yu, "A comprehensive survey on graph neural networks," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 32, no. 1, pp. 4–24, Jan. 2020.
- [43] M. Sun and P. Li, "Graph to graph: A topology aware approach for graph structures learning and generation," in *Proc. 22nd Int. Conf. Artif. Intell. Statist.*, 2019, pp. 2946–2955.
- [44] X. Guo, L. Zhao, Z. Qin, L. Wu, A. Shehu, and Y. Ye, "Interpretable deep graph generation with node-edge co-disentanglement," in *Proc. 26th ACM SIGKDD Int. Conf. Knowl. Discovery Data Mining*, Aug. 2020.
- [45] J. Li *et al.*, "Feature selection: A data perspective," *ACM Comput. Surv.*, vol. 50, no. 6, pp. 1–45, 2017.
- [46] H. Zhang, J. Wang, Z. Sun, J. M. Zurada, and N. R. Pal, "Feature selection for neural networks using group lasso regularization," *IEEE Trans. Knowl. Data Eng.*, vol. 32, no. 4, pp. 659–673, Apr. 2020.
- [47] R. F. Betzel *et al.*, "Multi-scale community organization of the human structural connectome and its relationship with resting-state functional connectivity," *Netw. Sci.*, vol. 1, no. 3, pp. 353–373, Dec. 2013.
- [48] L. E. Suárez, R. D. Markello, R. F. Betzel, and B. Misic, "Linking structure and function in macroscale brain networks," *Trends Cognit. Sci.*, vol. 24, no. 4, pp. 302–315, Apr. 2020.
- [49] A. Tarun, H. Behjat, T. Bolton, D. Abramian, and D. Van De Ville, "Structural mediation of human brain activity revealed by white-matter interpolation of fMRI," *NeuroImage*, vol. 213, Jun. 2020, Art. no. 116718.
- [50] H. You, S. Lin, and A. Singh, "Learning interpretable models for coupled networks under domain constraints," in *Proc. AAAI Conf. Artif. Intell.*, vol. 35, 2021, pp. 1–10.