

Brain Multigraph Prediction using Topology-Aware Adversarial Graph Neural Network

Alaa Bessadok^{a,b,c}, Mohamed Ali Mahjoub^{b,c}, Islem Rekik^{a,d,*}

^a*BASIRA lab, Faculty of Computer and Informatics, Istanbul Technical University, Istanbul, Turkey*

^b*Higher Institute of Informatics and Communication Technologies, University of Sousse, Tunisia*

^c*National Engineering School of Sousse, University of Sousse, LATIS- Laboratory of Advanced Technology and Intelligent Systems, 4023, Sousse, Tunisia*

^d*School of Science and Engineering, Computing, University of Dundee, UK*

Abstract

Brain graphs (i.e, connectomes) constructed from medical scans such as magnetic resonance imaging (MRI) have become increasingly important tools to characterize the abnormal changes in the human brain. Due to the high acquisition cost and processing time of multimodal MRI, existing deep learning frameworks based on Generative Adversarial Network (GAN) focused on predicting the missing multimodal medical images from a few existing modalities. While brain graphs help better understand how a particular disorder can change the connective facets of the brain, synthesizing a target brain *multigraph* (i.e, multiple brain graphs) from a single source brain graph is strikingly lacking. Additionally, existing graph generation works mainly learn one model for each target domain which limits their scalability in *jointly* predicting multiple target domains. Besides, while they consider the global topological scale of a graph (i.e., graph connectivity structure), they overlook the local topology at the node scale (e.g., how central a node is in the graph). To address these limitations, we introduce topology-aware graph GAN architecture (topoGAN), which jointly predicts multiple brain graphs from a single brain graph while preserving the topological structure of each target graph. Its three key innovations are: (i) designing a novel graph ad-

*Corresponding author; Dr Islem Rekik (irekik@itu.edu.tr), <http://basira-lab.com/>, GitHub code: <https://github.com/basiralab/topoGAN>. Supplementary materials are available in this file.

versarial auto-encoder for predicting multiple brain graphs from a single one, (ii) clustering the encoded source graphs in order to handle the mode collapse issue of GAN and proposing a *cluster-specific decoder*, (iii) introducing a *topological loss* to force the prediction of topologically sound target brain graphs. The experimental results using five target domains demonstrated the outperformance of our method in brain multigraph prediction from a single graph in comparison with baseline approaches.

Keywords: Brain multigraph prediction, Generative adversarial learning, Geometric deep learning, Adversarial autoencoders

1. Introduction

Multimodal neuroimaging data such as magnetic resonance imaging (MRI) and positron emission tomography (PET) provides complementary information for diagnosing neurological disorders. Nevertheless such data is not conventionally acquired for clinical diagnosis. Therefore, predicting modalities from minimal resources becomes a fundamental task in the neuroscience field. Existing deep learning methods aiming to solve this problem can be categorized into *one-target* (i.e, one-to-one) and *multi-target* (i.e, one-to-many) prediction approaches. For instance, in the first category, (Zeng and Zheng, 2019) proposed a framework based on Generative Adversarial Network (GAN) (Goodfellow et al., 2014) to predict Computed Tomography (CT) images from MRI where a cyclic reconstruction loss were introduced to improve the synthesis task. Similarly, (Pan et al., 2019) adopted the cyclic loss proposed in (Zhu et al., 2017) to predict PET from MR images for an early Alzheimer’s Disease identification. While these works predicted the target image using a single source modality, (Li et al., 2019) introduced DiamondGAN, a multi-modal GAN-based framework to predict double inversion recovery (DIR) scan from three source modalities (i.e, Flair, T1 and T2). A potential limitation of such *one-target* prediction frameworks, is that they are incapable of jointly predicting multiple target modalities in a single learning model (**Fig. 1-A**).

To this end, several attempts were made in the literature which are embedded into the second category that is *multi-target* prediction frameworks. For example, (Huang et al., 2019b) designed an autoencoder adversarially regularized by a discriminator to predict three target MR images (i.e, T1-weighted, T2-weighted, and FLAIR) from a single source T1 MRI scan.

While it is the single *multi-target* prediction work we identified in the medical imaging field, many frameworks were designed for computer vision tasks. Recently, (Wu et al., 2019) proposed a GAN-based model where the image synthesis step is first conditioned by a relative attribute vector representing the desired target domain, second it is adversarially regularized using three discriminators. To further improve the quality of the synthetic images, (Cao et al., 2019) exploited the correlation existing across multiple target domains by proposing a Wasserstein GAN-based framework. However, all the aforementioned models belonging to both categories were designed for synthesizing images, which limits their generalization to geometric data types such as graphs and manifolds (Bronstein et al., 2017). In particular, predicting brain graphs (i.e, connectomes), which models the functional, structural, or morphological interactions between brain regions is of paramount importance for charting the brain dysconnectivity patterns (van den Heuvel and Sporns, 2019; Bassett and Sporns, 2017).

A brain graph is an undirected graph conventionally encoded in a symmetric connectivity matrix where each element (i.e, edge connecting two nodes) measures the connectivity strength between pairs of region of interest (ROIs). Leveraging such brain data representation for the purpose of neurological disorder diagnosis can eventually improve prognosis (Fornito et al., 2015). By reason of its importance for understanding normal brain function and disordered brain dysfunction, such brain representation have been used for many purposes such as brain graph integration (Mhiri et al., 2020b; Yang et al., 2020; Bessadok and Rekik, 2018; Gurbuz and Rekik, 2020), disease early detection (Song et al., 2020; Li et al., 2020), developmental trajectories prediction (Ghribi et al., 2019; Goktas et al., 2020; Nebli et al., 2020). These studies demonstrate that leveraging different types of brain connectivities such as functional, structural and morphological ones provide more accurate results compared to neuroimaging since brain graphs represent a comprehensive mapping of neural activities. Thus, we highly need the connectomic data types for early diagnosis of neurological diseases. However, constructing brain graphs is limited by (i) the incompleteness of existing multimodal medical datasets and (ii) the pre-processing pipeline including different steps such as cortical parcellation and the surface registration is time-consuming for a single raw MRI (e.g, T1- and T2-weighted scans) (Li et al., 2013). Hence, these challenges dictate priorities for brain graph prediction. Especially, predicting missing target brain graphs from an existing source graph is highly required for learning the holistic brain mapping in healthy and disor-

der cases. In this regard, several recent studies were proposed for predicting brain graphs (Bessadok et al., 2019b,a, 2020a). However, to predict a target brain multigraph from a single source graph using these frameworks, we need to train the model for each target graph independently. Thus, such one-to-one prediction frameworks have a limited robustness. Consequently, we propose topoGAN, the first geometric deep learning framework aiming to jointly predict multiple brain graphs from a single graph in an end-to-end learning architecture (Fig. 1-B). We root our framework in the recently designed adversarial autoencoder model (Cao et al., 2019), which is a multi-domain translation technique primarily designed for images. Although effective, (Cao et al., 2019) has two major limitations: (i) it fails to operate on graphs as it was primarily designed for Euclidean data, (ii) it overlooks GAN mode collapse, where the generator (i.e., decoder) produces data that mimic a few modes of the target domain. To this end, we first propose to cluster the source graphs into homogeneous groups, which helps disentangle heterogeneous source data distributions. Second, we include the topological measurements (e.g, closeness centrality) into the adversarial learning which aligns the global and local graph topology of the predicted target graphs with that of the ground-truth ones. Fundamentally, we summarize the main contributions of this paper as follows:

1. *Source brain graph embedding clustering.* We learn the source graph embedding using an encoder E defined as a Graph Convolutional Network (GCN) (Kipf and Welling, 2016). Second, we cluster the resulting embeddings of the whole training population with heterogeneous distribution into homogeneous clusters. In that way, we enforce our *multi-target* prediction model to circumvent the mode collapse issue of GAN-based works (Goodfellow et al., 2014; Cao et al., 2019).
2. *Cluster-specific multi-target graph prediction.* Given the clustered source graphs embeddings, we define a set of synergetic generators for each target domain, each representing a *cluster-specific* GCN decoder. Hence, the graph prediction is learned more synergistically using our proposed cluster-specific generators, rather than using a single generator for each target domain. This generative process is regularized using one discriminator, which enforces the generated graphs to match the original target graphs.
3. *Topology-aware adversarial loss function.* In order to preserve both global and local topological properties of the original graphs, we un-

precedently introduce a topological loss function which enforces the generated graphs to retain a centrality score of each nodes in the original target brain graph.

Note that a preliminary version of this work was published in (Bessadok et al., 2020b). This journal version presents the following extensions. (1) We removed the graph reconstruction loss in the adversarial loss of the generator since it had a low affect in the prediction accuracy of the graphs. (2) We carried out more experiments to show the effectiveness of our method compared with the benchmark methods. Specifically, we report the results of jointly predicting multiple brain graphs using six different source graphs (i.e. views). (3) We further compared our framework to its variant architectures including graph attention network (GAT) and graph convolutional network (GCN). (4) We added more topology-focused evaluation metrics such as mean absolute error (i.e. MAE) between the real and predicted graphs, MAE between the real and predicted PageRank centrality, effective size and clustering coefficient of the graphs. We further reported p -value results using two-tailed paired t -test and the Kullback-Leibler divergence between the real and predicted scores. (5) Finally, we added a visual comparison of the real and predicted graphs where we display the residual of the predicted multigraphs.

2. Related work

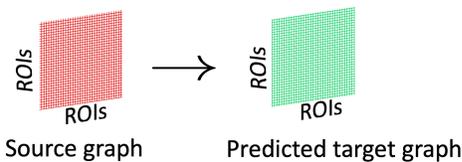
Brain Graph synthesis. A few recent papers have investigated geometric deep learning methods for brain graph prediction (Bessadok et al., 2019b,a; Sserwadda and Rekik, 2020; Zhang et al., 2020a) where in the first two works the synthesis task was partially formalized as a domain adaptation problem. For example, (Bessadok et al., 2019b) symmetrically aligned the training source and target brain graphs and adversarially regularized the training and testing source graphs embeddings using two discriminators. The second geometric deep learning (Hamilton et al., 2017) work, namely HADA (Bessadok et al., 2019a), hierarchically aligned each source graphs to the target graphs of training subjects and optimized the whole framework using a single discriminator. Next, to predict the target brain graph of a representative subject both works averaged the target graphs of the training subjects that share similar local neighborhoods across source and target domains. Although promising, these works are not designed in an *end-to-end* learning fashion (Fig. 1-A). They mainly dichotomize the model into separate parts that do not co-learn which leads to relatively high accumulated

errors across the learning steps. To overcome this limitation, (Sserwadda and Rekik, 2020) and (Zhang et al., 2020a) designed end-to-end GAN-based frameworks for brain graph synthesis. In particular, (Sserwadda and Rekik, 2020) adopted a cycle-consistency loss function to accurately perform the bidirectional mapping between the source and target brain graphs. Additionally, a new topological constraint was proposed to enforce the connectivity strength of the brain regions in the predicted graph to be similar to those of the ground-truth graph. On the other hand, (Zhang et al., 2020a) parallelised multiple GCN models and fused the resulting learned representations to predict a target brain graph from a source one. Furthermore, a structure-preserving loss function was proposed to stabilize the training of both generator and discriminator models. A shared shortcoming of all these brain graph synthesis works lies in their limited scalability for jointly predicting target brain *multigraph* from a single source graph. Regarded as a holistic representation of brain connectivities, a *multigraph* is a set of brain graphs stacked in a tensor where each captures a particular type of interactions between brain regions. Such brain representation plays an important role in modeling the dysfunctions in connectivity patterns existing between brain regions (van den Heuvel and Sporns, 2019). Thus, to predict a target brain *multigraph* using the existing frameworks we should learn one model for each target domain. Such frameworks, however, have a limited robustness in predicting more than one target domain. To solve this issue, our MICCAI 2020 conference paper (Bessadok et al., 2020b) presents the first work that jointly predicts a set of target brain graphs (i.e. target brain *multigraph*) from a single source graph. However, its experiments were restricted to only use a single source view to predict the target brain *multigraph*. Even such experiments demonstrated the superiority of our model over the comparison methods, we aim in this work to further show its prediction performance using five additional source views from both hemispheres to the one we used in the conference paper.

Graph generation. Besides, plenty of efforts have been dedicated to synthesizing different types of graphs and have shown remarkable results in various applications such as road network generation (Belli and Kipf, 2019), scene graph generation (Yang et al., 2018) and biological molecules synthesis (Mitton et al.). Other studies (Su et al., 2019; Liao et al., 2019) proposed to sequentially generate subgraphs consisting in a subset of nodes and their connectivities in order to generate the whole graph. Some other studies (Bresson and Laurent, 2019; Flam-Shepherd et al., 2020) proposed graph au-

toencoder frameworks where the encoded graph structure is decoded using a set of decoders. Recent works adopted a GAN-based solution combined with a reinforcement learning approach where an additional network is designed to further optimize the graph synthesis task (De Cao and Kipf, 2018; You et al., 2018). Despite their ineffectiveness in *multi-target* prediction tasks, such graph synthesis studies fail to preserve the node-wise topological properties of the target domain. Specifically, they only learn the *global graph structure* (i.e., number of nodes and edges weights). However, the brain wiring has both global and local topological properties which makes the ebb and flow of brain activity acts like the fingerprint of a subject (Fornito et al., 2015). In fact, neurological disorders such as Alzheimer’s disease alter the brain cortex and more importantly it atrophies its ROIs in varying degrees. Hence, focusing only on learning the global graph properties and overlooking the *local graph structure* may undervalue the role of specific ROIs in early diagnosing the disease. Therefore, devising a scalable and accurate framework for predicting a target brain multigraph from a single source graph that preserves the global and local topological properties of the original brain graphs is of great interest (Zhang et al., 2020c,b). Such local topology awareness can be defined by learning the node’s influence in the graph measured using path-length based metric which is defined in graph theory as “centrality”. Several network metrics such as betweenness centrality, closeness centrality and PageRank centrality have been introduced in the graph theory literature. For instance, (Huang et al., 2019a) proposed an analysis study for Parkinson’s disease (PD) where many centrality measurements were leveraged to identify the potential biomarkers in the disease progression using functional brain graphs extracted from resting-state functional MRI data (rs-fMRI). In a follow-up work, (Cid et al., 2019) used five centrality metrics to create a graph-based subject profile for a Pulmonary tuberculosis (TB) classification purpose. To the best of our knowledge, up to now no existing works have investigated the learning of node centrality for brain graph prediction (Zhou et al., 2018).

A) Conventional one-to-one graph generation framework for brain graph prediction



B) Proposed one-to-many graph generation framework for brain multigraph prediction

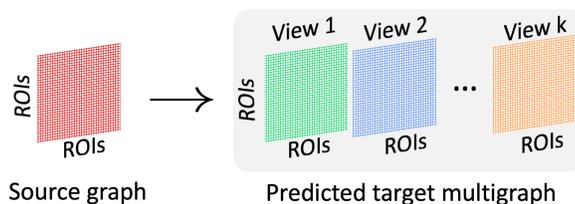


Figure 1: *Conventional brain graph prediction methods and the proposed brain multigraph prediction architecture.* **A)** In this illustration, we show the *one-to-one* strategy adopted in recent brain graph synthesis works aiming to predict a target graph from a source one. Unfortunately, such strategies fail to simultaneously predict multiple target graphs from a single one, which limits their scalability to brain *multigraph* derived from one or more magnetic resonance imaging (MRI) modalities (e.g., T1-weighted or resting-state functional MRI). **B)** To fill this gap, we propose a *one-to-many* learning architecture aiming to predict a target brain multigraph from a source graph.

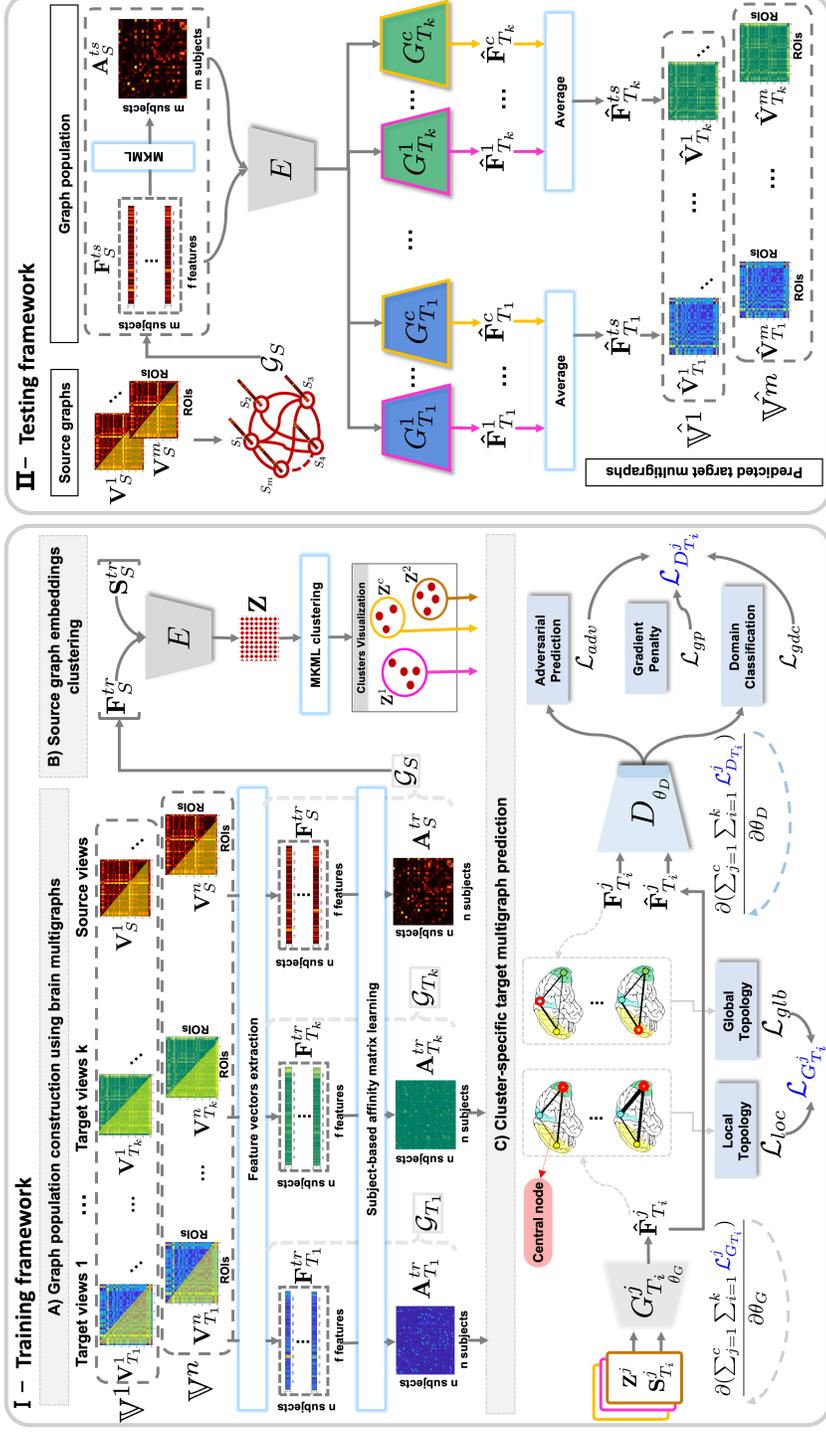


Figure 2: Outline of the proposed topology-aware graph Generative Adversarial Network (*topoGAN*) for predicting a target brain multigraph from a single source graph. **I- Training pipeline.** (A) Graph population construction using brain multigraph. For each training subject, we extract the source and k target feature vectors from their brain graphs. Then, we use multiple kernel manifold learning to learn the affinity between subjects. Thus, we construct a graph population for a specific domain using the resulting extracted features and the learned affinity matrix. (B) Source graph embeddings clustering. We learn the graph embedding of the source population and cluster it into c groups. (C) Cluster-specific target multigraph prediction. We train c cluster-specific generators for each target domain which are adversarially regularized by a discriminator D aiming to enforce the target domain classification. We further regularize each generator $G^j_{T_i}$ by proposing a local and a global topology losses which preserves the local node topology and the global graph connectivity structure, respectively. **II-Testing pipeline.** We construct a graph population capturing the affinity between testing subjects using their brain features. Next, we learn the graph embedding of the source population and decode it using c cluster-specific generators dedicated to each of the k target domains. Then, we predict the testing brain multigraph’s frontal views by averaging the graphs produced by our c generators in a specific target domain.

3. Proposed Method

Problem Definition. A brain graph is defined as $\mathcal{B} = \{\mathcal{B}^r, \mathcal{B}^e, \mathbf{V}\}$ where \mathcal{B}^r is a set of nodes (i.e, ROIs) and \mathcal{B}^e is a set of weighted edges encoding the biological connectivity between nodes. Each training subject s in our dataset is represented by a brain *multigraph* which captures the functional, structural, or morphological connectivities between brain regions. More compactly, it can be written as a tensor \mathbb{V}^s in $\mathbb{R}^{r \times r \times v}$ where its tensor layers denote a set of multi-view brain graphs $\{\mathbf{V}_i^s\}_{i=0}^v$, each view capturing a particular type of interactions between brain regions (e.g, morphological, functional and structural). $\mathbf{V} \in \mathbb{R}^{r \times r}$ denotes the connectivity matrix measuring the pairwise edge weight between nodes using a particular view, where r is the number of ROIs. Given a testing subject s' represented by a single source graph $\mathbf{V}_S^{s'}$, our objective is to predict its missing target brain *multigraph* $\mathbb{V}^{s'}$ in $\mathbb{R}^{r \times r \times k}$ where $k = v - 1$ and its frontal views are denoted by $\{\mathbf{V}_{T_i}^{s'}\}_{i=1}^k$.

Fig. 2 provides an overview of our proposed method and **Table 1** summarizes the major mathematical notations we used in this paper. There are three major steps in the pipeline: 1) extraction of multi-view brain features from source and target graphs and construction of a graph population for each domain, 2) embedding and clustering of the source graphs, and 3) prediction of the target brain multigraph using cluster-specific generators.

3.1. Graph population representation using multi-view brain graphs

To map the source brain graph of a subject to its target brain multigraph, we need to model the relationship between training samples using their connectivity features (e.g, weights). In fact, we hypothesize that samples (i.e, brain graphs) with strong affinity in the source domain will also maintain such a strong affinity in the target domains to some extent. Hence, we propose to learn the source graph embedding of each training subject using an encoder E defined as a Graph Convolutional Network (GCN) (Kipf and Welling, 2016). Thus, we create in this step for each of the source and target domains a graph population encoded in an affinity matrix where nodes denote subjects represented by their brain features and the edges represent the pairwise affinity between subjects. Since each brain graph is encoded in a symmetric matrix, we propose to vectorize the off-diagonal upper-triangular part which helps eliminate redundancy in the brain graphs. Then, we vertically stack the extracted feature vectors of size f for n training subjects which results in a feature matrix \mathbf{F}_v^{tr} in $\mathbb{R}^{n \times f}$ where $v \in \{S, T_1, \dots, T_k\}$. Next, given these

matrices we create a subject-based affinity matrix \mathbf{A}_v^{tr} in $\mathbb{R}^{n \times n}$ by leveraging multi-kernel manifold learning (MKML) algorithm (Wang et al., 2017) which learns the affinity between training feature vectors. We choose MKML for its appealing aspect of learning multiple kernels to efficiently fit the true underlying statistical distribution of the data. Hence, for a specific view v , we define our graph population for a specific domain as $\mathcal{G}_v = \{\mathcal{G}_v^n, \mathcal{G}_v^e, \mathbf{F}_v\}$, \mathcal{G}_v^n denotes a set of nodes (i.e, subjects), \mathcal{G}_v^e a set of weighted edges encoding the affinity between subjects and \mathbf{F}_v denotes a feature matrix. We create a set of graph populations $\{\mathcal{G}_i^{tr}\}_{i=0}^v$ for the source and target domains where each is represented by a set of feature matrices $\{\mathbf{F}_i^{tr}\}_{i=0}^v$ and a set of learned affinity matrices $\{\mathbf{A}_i^{tr}\}_{i=0}^v$ for the training subjects (Fig. 2-A). The resulting graph populations will be used in the two following steps to learn the graph embeddings of the source population (i.e, step B) and to map the source graph to the target multigraph of a training subject (i.e, step C).

3.2. Source brain graphs embedding and clustering

The proposed topoGAN is a graph autoencoder comprising an encoder E and a set of domain-specific decoders (i.e, generators) $\{G_{T_i}\}_{i=1}^k$ (i.e, a decoder for each of the k target domains) regularized by a discriminator D . Considering such an adversarial learning-based framework causes the mode collapse problem where the generator produces very limited number of modes (Goodfellow et al., 2014). In such case, no matter how big is our training set in terms of number of subjects because only few of them contribute to the graph synthesis. Thus, all generated graphs will look similar. We propose to solve this problem by clustering the source brain graphs which naturally have a heterogeneous statistical distribution. Since clustering samples encoded in high-dimensional feature vectors might be a complex task, we first propose to map the source brain graph into a low-dimensional space which helps reduce its dimensionality while preserving its topological structure. Consequently, the mode collapse issue is handled in two consecutive steps:

(1) we first learn the population graph embedding in the source domain which maps each subject-specific features into a lower representative dimensional space. To this end, we use an encoder $E(\mathbf{F}_S^{tr}, \mathbf{A}_S^{tr})$ defined as a GCN with two layers inputting the source feature matrix \mathbf{F}_S^{tr} of the training subjects and the learned subject-based affinity matrix \mathbf{A}_S^{tr} . GCN (Kipf and Welling, 2016) is originally defined using convolutions in the spectral domain which

Table 1: Major mathematical notations used in this paper

Notation	Dimension	Definition
r	\mathbb{N}	number of brain regions (i.e, ROIs)
v	\mathbb{N}	number of brain views (i.e, source and target views)
k	\mathbb{N}	number of target views (i.e, $k = v - 1$)
s	\mathbb{N}	total number of subjects including training and testing ones
n	\mathbb{N}	number of training subjects
m	\mathbb{N}	number of testing subjects
f	\mathbb{N}	number of features extracted from the original brain graph
d	\mathbb{N}	number of features in the embedded graphs
c	\mathbb{N}	number of clusters
e	\mathbb{N}	number of edges in a graph population
e'	\mathbb{N}	number of edges in a brain graph
$\mathcal{B} = \{\mathcal{B}^r, \mathcal{B}^e, \mathbf{V}\}$	—	a brain graph of a specific subject
\mathcal{B}^r	$\mathbb{N}^{1 \times r}$	a set of anatomical brain regions representing nodes in the brain graph
\mathcal{B}^e	$\mathbb{N}^{1 \times e'}$	a set of edges connecting the brain regions representing either functional, structural or morphological connectivities in the brain.
\mathbf{V}	$\mathbb{R}^{r \times r}$	connectivity matrix measuring the pairwise edge weights between nodes (i.e, ROIs)
\mathbf{V}_i	$\mathbb{R}^{r \times r}$	brain graph constructed from a single view i where $i \in \{0, \dots, v\}$
$\hat{\mathbf{V}}_i$	$\mathbb{R}^{n \times n}$	predicted brain graph for a specific target domain i where $i \in \{0, \dots, k\}$
\mathbb{V}_s	$\mathbb{R}^{r \times r \times v}$	a brain multigraph tensor of a training subject s stacking a set of source and target brain graphs
$\hat{\mathbb{V}}_{s'}^T$	$\mathbb{R}^{r \times r \times k}$	predicted target brain multigraph tensor of a testing subject s' stacking a set of target brain graphs
$\mathcal{G}_v = \{\mathcal{G}_v^n, \mathcal{G}_v^e, \mathbf{F}_v\}$	—	graph population representing the similarity between subjects belonging to a population
\mathcal{G}_v^n	$\mathbb{N}^{1 \times s}$	a set of graph population nodes or subjects in a population
\mathcal{G}_v^e	$\mathbb{N}^{1 \times e}$	a set of edges connecting pairs of subjects representing the similarity between them based on their brain graphs
\mathbf{F}_v	$\mathbb{R}^{n \times s}$	feature matrix vertically stacking feature vectors extracted from the view v of s subjects belonging to a specific population
\mathbf{F}_v^{tr}	$\mathbb{R}^{n \times f}$	feature matrix vertically stacking feature vectors extracted from the view v of n training subjects
\mathbf{A}_v^{tr}	$\mathbb{R}^{n \times n}$	subject-based affinity matrix between n training subjects using the brain feature vectors belonging to the view v
\mathbf{Z}	$\mathbb{R}^{n \times d}$	learned source graph embeddings of the training subjects
$(\mathbf{X}_{T_i}^j)^C$	$\mathbb{R}^{n \times r}$	centrality matrix of the real target graphs in the domain T_i and the cluster j of the training subjects computed using the centrality metric C
$(\hat{\mathbf{X}}_{T_i}^j)^C$	$\mathbb{R}^{n \times r}$	centrality matrix of the predicted target graphs in the domain T_i and the cluster j of the training subjects computed using the centrality metric C

is expressed by the following convolution function:

$$f_\phi(\mathbf{F}^{(l)}, \mathbf{A}_S^{tr} | \mathbf{W}^{(l)}) = \phi(\tilde{\mathbf{D}}^{-\frac{1}{2}} \tilde{\mathbf{A}}_S^{tr} \tilde{\mathbf{D}}^{-\frac{1}{2}} \mathbf{F}^{(l)} \mathbf{W}^{(l)}) \quad (1)$$

ϕ denotes the *ReLU* and *linear* activation functions we used in the first and second layers, respectively. We define \mathbf{F} in the first layer as the source feature matrix \mathbf{F}_S^{tr} while we define it in the second layer as the resulting embeddings learned from the first layer l . $\mathbf{W}^{(l)}$ is a filter used to learn the convolution in the GCN in each layer l . $\tilde{\mathbf{D}}_{ii} = \sum_j \tilde{\mathbf{A}}_S(ij)$ is a diagonal matrix and $\tilde{\mathbf{A}}_S^{tr} = \mathbf{A}_S^{tr} + \mathbf{I}$ with \mathbf{I} being an identity matrix used for regularization. Ultimately, we define the layers of our GCN encoder as follows:

$$\mathbf{Z}^{(1)} = f_{ReLU}(\mathbf{F}_S^{tr}, \mathbf{A}_S^{tr} | \mathbf{W}^{(0)}); \mathbf{Z}^{(2)} = f_{linear}(\mathbf{Z}^{(1)}, \mathbf{A}_S^{tr} | \mathbf{W}^{(1)}) \quad (2)$$

(2) We cluster the resulting source embeddings \mathbf{Z} into homogeneous groups which helps disentangle the heterogeneous distribution thereby reducing the generator’s risk to match a few unimodal samples of the target domain. To do so, we leverage MKML since it outperformed PCA (Jolliffe and Cadima, 2016) and t-SNE (Maaten and Hinton, 2008) clustering methods when dealing with biological dataset (Wang et al., 2017). More importantly, it is widely used for brain graph analysis tasks and showed promising results in neurological disorder diagnosis such as Autism spectrum disorder (ASD) (Soussia and Rekik, 2018; Bessadok et al., 2019b; Mhiri and Rekik, 2020). Specifically, it first produces a pairwise affinity matrix measuring the similarity between training subjects using their source graph embeddings \mathbf{Z} and having c diagonal blocks denoting the clusters. Next, the obtained source affinity matrix is projected into a lower dimension using t-SNE (Maaten and Hinton, 2008) which results in a latent matrix in $\mathbb{R}^{n \times c}$. Last, k-means algorithm (Jain, 2010) is used to cluster the subjects into c clusters based on the resulting learned latent matrix (Fig. 2-B).

3.3. Cluster-specific multi-target graph prediction

For each target view k , we design a set of *cluster-specific generators* $\{G_{T_i}^j\}_{i=1, j=1}^{k,c}$, each learning to match the distribution of a cluster j belonging to a target domain T_i (Fig. 2-C). Our goal is to enforce each generator to learn from all examples in the cluster c thereby avoiding the mode collapse issue of GAN-based models. We define our generators as GCN decoders with similar architecture to the encoder (Eq. (2)). Specifically, a generator $G_{T_i}^j$ designed to predict the target graphs of the domain T_i and the cluster j

takes as input the learned source embeddings \mathbf{Z}^j and the subject-based affinity matrix $\mathbf{A}_{T_i}^j$. Specifically, we learn the affinity between feature vectors of the target domain T_i belonging to the same cluster j . In other words, decoding the source embeddings with the affinity matrix learned using the target graphs in T_i enforces the generated graphs to approximate the real target domain structure of a specific cluster j . We further propose to optimize the target graph prediction using a discriminator D which is a GCN with three layers aiming to enforce the generated target graph to approximate the ground-truth target distribution of a specific target domain. To achieve this, we propose three loss functions which optimize the discriminator D .

Adversarial loss. We introduce an adversarial loss function differently from the vanilla GAN (Goodfellow et al., 2014) where we compute the Wasserstein distance among all domains in order to measure the realness of the generated graphs. This distance has been widely used in the GAN literature as it stabilizes the training process of the model thereby making it less sensitive to hyperparameter regularization (Gulrajani et al., 2017). Thus, we formulate it as follows:

$$\mathcal{L}_{adv}^j = -\mathbb{E}_{\mathbf{F}' \sim \mathbb{P}_{\mathbf{F}'_S}} [D(\mathbf{F}')] + \frac{1}{k} \sum_{i=1}^k \mathbb{E}_{\mathbf{F}'' \sim \mathbb{P}_{\hat{\mathbf{F}}_{T_i}^j}} [D(\mathbf{F}'')] \quad (3)$$

where $\mathbb{P}_{\mathbf{F}'_S}$ is the real source graph distribution of the cluster j , and the distribution $\mathbb{P}_{\hat{\mathbf{F}}_{T_i}^j}$ is the generated distribution by $G_{T_i}^j$ in the target domain T_i .

Graph domain classification loss. We recall that the goal of training the cluster-specific generators $\{G_{T_i}^j\}_{i=1, j=1}^{k,c}$ is to produce graphs for the cluster j which are properly classified by the discriminator to the specific target domain T_i . Thus, we define a binary classifier D_C on top of our discriminator D which classifies the synthetic graphs $\hat{\mathbf{F}}_{T_i}^j$ as 0 and the real target graphs $\mathbf{F}_{T_i}^j$ as 1. In detail, the former is defined as:

$$\mathcal{L}_{gdc}^j = \sum_{i=1}^k \mathbb{E}_{\mathbf{F}'' \sim \mathbb{P}_{\hat{\mathbf{F}}_{T_i}^j} \cup \mathbb{P}_{\mathbf{F}_{T_i}^j}} [\ell_{MSE}(D_C(\mathbf{F}''), y(\mathbf{F}''))] \quad (4)$$

ℓ_{MSE} is the mean squared loss. $D_C(\mathbf{F}'')$ and $y(\mathbf{F}'')$ denote the predicted and ground-truth labels corresponding to the graph \mathbf{F}'' , respectively. More specifically, we compute the ℓ_{MSE} for the predicted and real graphs separately

(i.e., $\hat{\mathbf{F}}_{T_i}^j, \mathbf{F}_{T_i}^j$) then we sum both values which we mathematically denote by the following notation for simplicity $\mathbb{E}_{\mathbf{F}'' \sim \mathbb{P}_{\hat{\mathbf{F}}_{T_i}^j} \cup \mathbb{P}_{\mathbf{F}_{T_i}^j}}$.

Gradient penalty loss. To improve the training stability of our model, we adopt the gradient penalty loss used in (Cao et al., 2019) which is formulated as follows:

$$\mathcal{L}_{gp}^j = (\max\{0, \mathbb{E}_{\tilde{\mathbf{F}} \sim \mathbb{P}_{\tilde{\mathbf{F}}_k^j}} \|\nabla D(\tilde{\mathbf{F}})\| - \sigma\})^2 \quad (5)$$

$\tilde{\mathbf{F}}$ is sampled between the source graph distribution $\mathbb{P}_{\mathbf{F}_S^j}$ and the predicted target graph distribution $\mathbb{P}_{\tilde{\mathbf{F}}_k^j}$ where $\tilde{\mathbf{F}}_k^j$ is a matrix stacking vertically the generated target graphs for all k domains in the cluster j . Particularly, $\tilde{\mathbf{F}} \leftarrow \alpha \mathbf{F}_S^j + (1 - \alpha) \tilde{\mathbf{F}}_k^j$ where $\alpha \sim U[0, 1]$ and U is a uniform distribution. As suggested in (Cao et al., 2019), we set the hyper-parameter σ to k .

Ultimately, the cost function of the discriminator which helps the generators of each cluster produce brain graphs each associated with a specific target domain is formulated as follows:

$$\mathcal{L}_D = \sum_{j=1}^c (\mathcal{L}_{adv}^j + \lambda_{gp} \cdot \mathcal{L}_{gp}^j + \lambda_{gdc} \cdot \mathcal{L}_{gdc}^j), \quad (6)$$

λ_{gdc} and λ_{gp} are hyper-parameters to be tuned. Hence, by maximizing the discriminator loss function defined above **Eq. (6)** the generators are optimally trained to produce graphs that belong to a specific target domain.

Moreover, brain graphs have rich topological properties including their percolation threshold, hubness and modularity (Bassett and Sporns, 2017). Such unique properties should be preserved when synthesizing the target brain graphs (Liu et al., 2017; Joyce et al., 2010). Although regarded as an efficient graph embedding model, graph autoencoder is limited to only learning the *global graph structure* such as number of nodes and edges in the graph while the *local graph structure* should be also learned since it reflects the node importance in the graph. To this aim, we unprecedentedly introduce a topological loss function to guide the training process of the generators and constrains each of them to preserve the local nodes properties while learning the global graph structure (**Fig. 2-C**). Specifically, we adopt three centrality metrics to compute a score for each ROI in the brain graph. We choose three centrality metrics widely used in graph theory (Borgatti and Everett, 2006): closeness centrality, betweenness centrality and eigenvector centrality.

The closeness centrality CC quantifying the closeness of a node to all other nodes (Freeman, 1977) is defined as follows:

$$CC(r^a) = \frac{r - 1}{\sum_{r^a \neq r^b} p_{r^a r^b}} \quad (7)$$

r denotes the number of nodes (i.e, ROIs) and $p_{r^a r^b}$ is the length of the shortest path between nodes r^a and r^b .

The betweenness centrality BC measuring the number of shortest paths which pass across a node (Beauchamp, 1965) can be defined as:

$$BC(r^a) = \frac{2}{(r - 1)(r - 2)} \times \sum_{r^a \neq r^b \neq r^c} \frac{P_{(r^c, r^b)}(r^a)}{P_{(r^c, r^b)}} \quad (8)$$

$P_{(r^c, r^b)}(r^a)$ denotes the number of shortest paths between two nodes r^c and r^b that pass through (r^a) .

The eigenvector centrality EC capturing the centralities of a node’s neighbors (Bonacich, 2007) is defined as follows:

$$EC(r^a) = \mathbf{x}^a = \frac{1}{\lambda} \sum_{b=1}^r \mathbf{V}_{ab} \mathbf{x}^b \quad (9)$$

\mathbf{V}_{ab} represents the b neighbor of the node a , \mathbf{x}^a and \mathbf{x}^b are the eigenvectors resulting from the eigen decomposition of the adjacency matrix \mathbf{V} and λ is a positive proportionality factor.

Next, we compute $\ell_{MAE}(\cdot)^c$ the absolute difference between the real and predicted centrality scores of each node in the target graph which represents our *local topology loss*. In particular, given a centrality metric \mathcal{C} where $\mathcal{C} \in \{CC, BC, EC\}$, a cluster j and a target domain T_i , we define $(\mathbf{X}_{T_i}^j)^c$ and $(\hat{\mathbf{X}}_{T_i}^j)^c$ both in $\mathbb{R}^{n \times r}$ as the centralities for the real brain graphs $\mathbf{V}_{T_i}^j$ and the generated ones $\hat{\mathbf{V}}_{T_i}^j$ reconstructed from the feature matrices $\mathbf{F}_{T_i}^j$ and $\hat{\mathbf{F}}_{T_i}^j$, respectively. To preserve the relationship between brain regions in terms of number of edges and their weights we compute the absolute difference between the real and predicted feature matrices \mathbf{F}_{T_i} and $\hat{\mathbf{F}}_{T_i}$, which mainly represents our *global topology loss* function. Ultimately, to take advantage of both local and global losses we propose to fuse them into a single function.

Topological loss. This is one of the key contributions for our proposed architecture which regularizes the cluster-specific generators (Fig. 2-C). It is computed as follows:

$$\mathcal{L}_{top}^j(\mathcal{C}) = \underbrace{\left(\sum_{i=1}^k \ell_{MAE}(\mathbf{X}_{T_i}^j, \hat{\mathbf{X}}_{T_i}^j)^{\mathcal{C}} \right)}_{\text{local topology loss}} + \underbrace{\sum_{i=1}^k \ell_{MAE}(\mathbf{F}_{T_i}^j, \hat{\mathbf{F}}_{T_i}^j)}_{\text{global topology loss}} \quad (10)$$

Information maximization loss. Since our k target domains are correlated we integrate the information maximization loss term to force the *cluster-specific generators* $\{G_{T_i}^j\}_{i=1, j=1}^{k, c}$ to correlate the predicted graphs with a specific target domain T_i . As in (Cao et al., 2019), we define it using this formula:

$$\mathcal{L}_{inf}^j = \sum_{i=1}^k \ell_{BCE}(y = 1, D_C(\hat{\mathbf{F}}_{T_i}^j)) \quad (11)$$

ℓ_{BCE} is the binary cross entropy. Given the above definitions of the topological and information maximization losses, we introduce the overall *topology-aware adversarial loss function* of each generator as:

$$\mathcal{L}_G = \sum_{j=1}^c \left(-\frac{1}{k} \cdot \sum_{i=1}^k \mathbb{E}_{\mathbf{F}'' \sim \hat{\mathbf{F}}_{T_i}^j} [D(\mathbf{F}'')] \right) + \lambda_{top} \cdot \mathcal{L}_{top}^j(\mathcal{C}) + \lambda_{inf} \cdot \mathcal{L}_{inf}^j \quad (12)$$

where λ_{top} and λ_{inf} are hyper-parameters that control the relative importance of topological loss and information maximization losses, respectively. As illustrated in (Fig. 2-II), given a testing source graph we predict each view of the target brain multigraph by averaging the graphs of a specific target domain T_i produced by the *cluster-specific generators*.

Table 2: Average prediction results of a target brain multigraph from three source views derived from the left hemisphere using different evaluation metrics.

View 1		MAE	MAE(CC)	MAE(BC)	MAE(EC)	MAE(PC)	MAE(EFF)	MAE(CIst)
Adapted MWGAN	GAT	0.1781	0.2988	0.0134	0.022	0.0035	7.4249	0.429
	GCN	0.192	0.2178	0.0087	0.013	0.0033	6.4845	0.2879
Adapted MWGAN (clustering)	GAT	0.2003	0.1823	0.007	0.0118	0.0071	5.6466	0.2305
	GCN	0.201	0.0461	0.0015	0.0055	0.0049	1.5199	0.0495
topoGAN+CC	GAT	0.1887	0.2331	0.0096	0.0201	0.0041	6.4751	0.3002
	GCN	0.1827	0.0581	0.0019	0.0059	0.0042	1.8975	0.0628
topoGAN+BC	GAT	0.1818	0.1133	0.0041	0.0102	0.005	3.6689	0.1319
	GCN	0.2043	0.0552	0.0019	0.0064	0.0045	1.8496	0.061
topoGAN+EC	GAT	0.2152	0.1678	0.0063	0.0098	0.0055	5.3983	0.2137
	GCN	0.1903	0.0262	0.0007	0.0041	0.005	0.7618	0.0242
MultiGraphGAN+CC	GCN	0.1927	0.1103	0.004	0.0089	0.0055	3.6406	0.1301
	GCN	0.192	0.0567	0.0018	0.0055	0.0039	1.905	0.063
MultiGraphGAN+BC	GCN	0.1889	0.0374	0.0012	0.0048	0.0045	1.2506	0.0403
	GCN	0.1889	0.0374	0.0012	0.0048	0.0045	1.2506	0.0403
View 2	GAT	0.2317	0.293	0.0131	0.0253	0.0059	7.4492	0.4206
	GCN	0.2152	0.2665	0.0115	0.02	0.0048	7.0436	0.3604
Adapted MWGAN	GAT	0.2243	0.1804	0.0069	0.012	0.0045	5.6719	0.231
	GCN	0.2252	0.2257	0.0093	0.0193	0.0052	6.3599	0.2881
topoGAN+CC	GAT	0.2247	0.2581	0.011	0.0212	0.0042	6.7694	0.3355
	GCN	0.2266	0.2322	0.0096	0.0162	0.0057	6.6855	0.3096
topoGAN+BC	GAT	0.2302	0.186	0.0072	0.0116	0.0054	5.7792	0.2381
	GCN	0.2231	0.1744	0.0067	0.0117	0.0051	5.4989	0.2211
topoGAN+EC	GAT	0.2282	0.1414	0.0052	0.0107	0.0058	4.4982	0.1694
	GCN	0.2239	0.2468	0.0103	0.0148	0.0056	7.1952	0.3464
MultiGraphGAN+CC	GCN	0.2228	0.1961	0.0077	0.0145	0.0051	5.9207	0.25
	GCN	0.224	0.2318	0.0095	0.0184	0.0058	6.6384	0.3067
MultiGraphGAN+BC	GCN	0.2218	0.1777	0.0068	0.0126	0.0052	5.513	0.2233
	GCN	0.2218	0.1777	0.0068	0.0126	0.0052	5.513	0.2233
View 3	GAT	0.1483	0.2404	0.01	0.0179	0.0034	6.7071	0.3171
	GCN	0.1156	0.1806	0.0069	0.011	0.0028	5.635	0.2296
Adapted MWGAN	GAT	0.1556	0.1732	0.0066	0.012	0.0044	5.4467	0.2183
	GCN	0.1206	0.0548	0.0018	0.0052	0.0032	1.8435	0.0608
topoGAN+CC	GAT	0.1935	0.1784	0.0068	0.0123	0.0063	5.566	0.2257
	GCN	0.1087	0.0626	0.0021	0.0065	0.0028	2.0644	0.0687
topoGAN+BC	GAT	0.1234	0.1303	0.0047	0.0098	0.0033	4.187	0.1548
	GCN	0.1324	0.0594	0.002	0.0063	0.0028	1.9572	0.0648
topoGAN+EC	GAT	0.1221	0.1445	0.0053	0.0102	0.0033	4.5529	0.1726
	GCN	0.1323	0.0545	0.0018	0.0052	0.0037	1.8202	0.0601
MultiGraphGAN+CC	GCN	0.1336	0.0769	0.0026	0.0078	0.0029	2.502	0.085
	GCN	0.1195	0.0686	0.0023	0.0062	0.0027	2.2984	0.0772
MultiGraphGAN+BC	GCN	0.1545	0.0799	0.0027	0.0082	0.0031	2.5771	0.0877
	GCN	0.1545	0.0799	0.0027	0.0082	0.0031	2.5771	0.0877

View 1: maximum principal curvature (LH). View 2: average curvature (LH). View 3: mean sulcal depth (LH). MAE: mean absolute error. CC: close-

Table 3: Average prediction results of a target brain multigraph from three source views derived from the right hemisphere using different evaluation metrics.

View 4		MAE	MAE(CC)	MAE(BC)	MAE(EC)	MAE(PC)	MAE(EFF)	MAE(CIst)
Adapted MWGAN	GAT	0.2096	0.3242	0.015	0.0245	0.0076	7.8095	0.5078
	GCN	0.1844	0.2193	0.0088	0.014	0.0033	6.6041	0.2949
Adapted MWGAN (clustering)	GAT	0.1977	0.1878	0.0073	0.0132	0.0074	5.6507	0.2333
	GCN	0.1886	0.0456	0.0015	0.0049	0.0042	1.5334	0.05
topoGAN+CC	GAT	0.1978	0.175	0.0067	0.0117	0.0063	5.4843	0.2201
	GCN	0.1878	0.0558	0.0019	0.0069	0.0047	1.7895	0.0591
topoGAN+BC	GAT	0.1906	0.1448	0.0053	0.0103	0.004	4.5976	0.1747
	GCN	0.1758	0.0672	0.0022	0.0076	0.0044	2.1657	0.0725
topoGAN+EC	GAT	0.231	0.1535	0.0057	0.0104	0.0054	4.8375	0.1864
	GCN	0.1875	0.0569	0.0018	0.0072	0.0046	1.7222	0.0566
MultiGraphGAN+CC	GAT	0.1816	0.04621	0.0015	0.0046	0.0043	1.5496	0.0506
	GCN	0.2008	0.0845	0.0029	0.0082	0.0048	2.7354	0.0939
MultiGraphGAN+BC	GAT	0.195	0.0435	0.0014	0.0054	0.0046	1.4378	0.05
	GCN	0.195	0.0435	0.0014	0.0054	0.0046	1.4378	0.05
View 5		MAE	MAE(CC)	MAE(BC)	MAE(EC)	MAE(PC)	MAE(EFF)	MAE(CIst)
Adapted MWGAN	GAT	0.2267	0.3022	0.0136	0.0245	0.0058	7.5284	0.4402
	GCN	0.2175	0.2583	0.0109	0.0176	0.0042	7.1924	0.3595
Adapted MWGAN (clustering)	GAT	0.2256	0.1475	0.0055	0.0133	0.0029	4.5758	0.1747
	GCN	0.2198	0.169	0.0064	0.0122	0.0042	5.3084	0.2118
topoGAN+CC	GAT	0.2223	0.1542	0.0058	0.0101	0.0033	4.887	0.1884
	GCN	0.2215	0.1691	0.0065	0.0126	0.0041	5.23	0.2083
topoGAN+BC	GAT	0.2317	0.1982	0.0078	0.016	0.0035	5.8761	0.2491
	GCN	0.2232	0.1993	0.0079	0.0155	0.0051	5.9041	0.2511
topoGAN+EC	GAT	0.2218	0.1673	0.0063	0.0083	0.0035	5.3628	0.2118
	GCN	0.2224	0.2179	0.0088	0.0168	0.0047	6.3244	0.2816
MultiGraphGAN+CC	GAT	0.2206	0.1748	0.0067	0.0117	0.0045	5.4439	0.2193
	GCN	0.2243	0.2385	0.0099	0.0199	0.0043	6.7307	0.3176
MultiGraphGAN+BC	GAT	0.2189	0.2044	0.0081	0.0161	0.0049	5.8971	0.2537
	GCN	0.2189	0.2044	0.0081	0.0161	0.0049	5.8971	0.2537
View 6		MAE	MAE(CC)	MAE(BC)	MAE(EC)	MAE(PC)	MAE(EFF)	MAE(CIst)
Adapted MWGAN	GAT	0.1409	0.2466	0.0103	0.0176	0.0032	6.9177	0.3326
	GCN	0.1096	0.3107	0.0141	0.0167	0.0046	7.5565	0.4564
Adapted MWGAN (clustering)	GAT	0.1288	0.2056	0.0082	0.0144	0.0057	6.1356	0.2652
	GCN	0.1121	0.1351	0.0049	0.0105	0.0051	4.349	0.1627
topoGAN+CC	GAT	0.1919	0.1854	0.0072	0.0138	0.0049	5.6401	0.2332
	GCN	0.115	0.1004	0.0035	0.0084	0.0039	3.2629	0.1149
topoGAN+BC	GAT	0.0805	0.1597	0.006	0.0097	0.0068	5.151	0.2007
	GCN	0.1028	0.0722	0.0025	0.0081	0.0037	2.2983	0.0776
topoGAN+EC	GAT	0.1008	0.1617	0.0061	0.0127	0.0048	5.0954	0.1999
	GCN	0.0772	0.086	0.003	0.0073	0.0045	2.8036	0.0966
MultiGraphGAN+CC	GAT	0.1825	0.1793	0.0069	0.0144	0.0054	5.5751	0.2273
	GCN	0.1587	0.1671	0.0063	0.0109	0.0052	5.1932	0.2061
MultiGraphGAN+BC	GAT	0.0767	0.1178	0.0042	0.0101	0.0052	3.7543	0.1357
	GCN	0.0767	0.1178	0.0042	0.0101	0.0052	3.7543	0.1357

View 4: maximum principal curvature (RH). View 5: average curvature (RH). View 6: mean sulcal depth (RH). MAE: mean absolute error. CC: close-

4. Experiments and Discussion

4.1. Multi-view brain graph dataset

For evaluation, we use the Autism Brain Imaging Data Exchange (ABIDE¹) including 310 subjects (i.e, structural T1-w MRI). We reconstruct both cortical hemispheres by FreeSurfer (Fischl, 2012) and then parcellate each cortical hemisphere into 35 cortical regions using Desikan-Killiany Atlas. For each hemisphere, we extract three morphological brain graphs (i.e, MBG), which means each subject in our dataset is represented by six MBGs. We use the following cortical measurements to extract a pair of graphs for left and right hemispheres: maximum principal curvature (i.e, view 1 and view 4), average curvature (i.e, view 2 and view 5) and mean sulcal depth (i.e, view 3 and view 6). Recently introduced in (Mahjoub et al., 2018), morphological brain networks (MBNs) quantify the dissimilarity in morphology between pairs of ROIs in the cortex. Essentially, given a particular cortical attribute, (e.g, mean sulcal depth), we compute its average across all vertices in a given brain ROI. Next, we define the edge weight between two ROIs as the absolute difference between their average cortical attributes. MBNs have been widely used for healthy (Dhifallah et al., 2020; Nebli and Rekik, 2019; Mhiri and Rekik, 2020) and disordered brain connectivity analysis (Khelifa and Rekik, 2019; Banka and Rekik, 2019; Soussia and Rekik, 2018).

4.2. Model architecture and parameter setting

As shown in (Fig. 2-I), we use $(c \times k + 2)$ trainable neural networks: an encoder E , a discriminator D and $(c \times k)$ cluster-specific generators $\{G_{T_i}^j\}_{i=1, j=1}^{k,c}$. We construct our encoder with a hidden layer comprising 32 neurons and an embedding layer with 16 neurons. Conversely, we define all generators with two layers each comprising 16 and 32 neurons. The discriminator comprises three layers each has 32, 16 and 1 neurons, respectively. We add to its last layer a softmax activation function representing our domain classifier. We train our model on 1000 iterations using batch size 70, a learning rate of 0.0001, $\beta_1 = 0.5$ and $\beta_2 = 0.999$ for Adam optimizer. We use a grid search to set our hyper-parameters $\lambda_{gdc} = 1$, $\lambda_{gp} = 0.1$, $\lambda_{top} = 0.1$, $\lambda_{rec} = 0.01$ and $\lambda_{inf} = 1$. Specifically, we use the parameter setting that produced the best performance for each method independently. We also train our discriminator

¹<http://fcon.1000.projects.nitrc.org/indi/abide/>

five times and the generators one time in order to have better models. By doing so, we are improving their learning performances. For MKML parameters (Wang et al., 2017), we fix the number of kernels to 10. We vary the number of clusters c between 2 and 4 then we choose the one which gave the best performance $c = 2$.

4.3. Evaluation and comparison methods

We train our model using two different training strategies: (i) a random split where 90% of the dataset represents the training set and 10% of it is used for testing, and (ii) a 3-fold cross-validation where we train the model on two folds and test it on the left fold. To quantitatively evaluate the topoGAN’s performance, we use the mean absolute error (MAE) (Lin et al., 1990) computed between ground-truth target graphs and the generated ones and the MAE between the ground-truth centrality scores and the predicted ones. To comprehensively evaluate our model, we further incorporate an additional centrality evaluation including PageRank centrality (PC), effective size (Eff) and the clustering coefficient (Clst). In particular, PageRank centrality introduced in (Brin, 1998) assumes that a central node is likely to receive more edges from other nodes. The effective size metric introduced in (Burt, 2009) assume that the efficiency of a node is related to the number of non-redundant links in a graph and the clustering coefficient (Saramäki et al., 2007) measures the degree to which nodes in a graph tend to cluster together. Lastly, we consider the average of the resulting MAEs computed for each of the k target domains as the final measures to evaluate our framework. The lower is the MAE the better is the performance of our framework in predicting the target multigraph. Moreover, we compute the Kullback–Leibler divergence (Kullback, 1959) between the real and predicted topology scores (i.e, CC, BC, EC, EFF and Clst). Then, we average the KL-divergence results of all k domains and report the results for each source brain graph. Originated in probability theory and information theory, KL-divergence measures the information lost when the predicted distribution approximates the original distribution. Thus, a lower value indicates that the predicted distribution is almost identical to the real one.

We compare our topoGAN framework with two baseline methods and three ablated versions of our model where we use three different centrality metrics. Essentially, each method is implemented with two graph neural network architectures (i.e, GCN (Kipf and Welling, 2016) and graph attention network (GAT) (Veličković et al., 2017)):

1. **Adapted MWGAN:** we use the same architecture proposed in (Cao et al., 2019) that we adapted to graph data types where we neither include the clustering step nor our proposed topological loss term (Eq. (10)).
2. **Adapted MWGAN (clustering):** it is a variant of the first method where we add the MKML clustering of the source graph embeddings (Wang et al., 2017).
3. **topoGAN+CC:** it is an ablated version of our framework which adopts a clustering strategy and a topological loss function that includes the closeness centrality (Freeman, 1977) in the topological loss function (i.e, $\mathcal{C} = CC$ in Eq. (10)).
4. **topoGAN+BC:** in this method we incorporate the betweenness centrality (Beauchamp, 1965) in the topological loss function (i.e, $\mathcal{C} = BC$ in Eq. (10)).
5. **topoGAN+EC:** in this method we include the eigenvector centrality (Bonacich, 2007) in the topological loss function (i.e, $\mathcal{C} = EC$ in Eq. (10)).

4.4. Results and benchmarking

We conducted six different experiments, each taking one of the six MBGs as a source graph and the five remaining ones as views stacked in the target brain multigraph to predict. Clearly, supplementary **Table ??**, **Table ??** and **Table ??** show that our topoGAN consistently and significantly (p -value < 0.05 using two-tailed paired t -test) outperformed benchmark methods in predicting brain multigraph from a single source graph.

4.4.1. Impact of the clustering and global topology loss function

To evaluate the effectiveness of clustering the source graph embeddings and learning the global topological structures of the target graphs, we report in **Table ??** and **Table ??** the average MAE (i.e, first column) computed between the real and predicted graphs of five target domains. These results show that three variants of topoGAN outperformed two baseline methods. Although it produced a slightly higher MAE using one source graph (i.e, view 5), our framework achieved the lowest MAE results using five source graphs (i.e, views 1,2,3,4 and 6). Specifically, methods adopting the CC and BC both ranked first best in six experiments while the method adopting EC ranked second best using the MAE metric. Notably, these results show that our topoGAN using the proposed topological loss function significantly outperforms the baseline methods in preserving the global topology of the original

target graphs. We also show that our framework using the *global topology loss* term is better than GCN-based methods which simply leverage graph convolution to learn the graph structure. Additionally, **Fig. 3** displays the source graph, the ground-truth target graphs and the predicted ones by topoGAN using three centrality measures and the two baseline methods (Adapted MWGAN and Adapted MWGAN (clustering)) for a representative subject. We display below each predicted graph its residual graph, representing the absolute difference between the ground-truth and predicted target graph. We observe that the residual was noticeably reduced by our method using both CC and BC measures. When predicting the target graph 2 (i.e, mean sulcal depth (LH)), Adapted MWGAN is marginally better than three variants of topoGAN. This shows that while the average results of topoGAN outperform benchmark methods, our framework might lag behind for a few particular subjects. However, we argue that although our framework works relatively poorly in predicting the target graph 2, it still achieved better results than the Adapted MWGAN (clustering) in predicting four target graphs for the same representative subject (**Fig. 3**). This demonstrates the advantage of our *cluster-specific generators* in eventually avoiding the mode collapse problem thereby boosting the performance of topoGAN in the target brain multigraph prediction.

4.4.2. Impact of the local topology loss and the reconstruction loss evaluation

To further evaluate the effectiveness of learning the local topological structures of the original target graphs, we compute the MAE between the ground-truth and the predicted centrality scores, effective size and clustering coefficient. Results reported in **Table ??** and **Table ??** show that our method consistently achieved the best topology-preserving predictions compared with the baseline methods using the six graph topology evaluation measures. Notably, four out of six experiments (i.e, views 1,2,3 and 6) results highlight the importance of using BC and EC which both gave the lowest MAE using different evaluation metrics. This is explicable since considering the node neighborhoods (i.e., EC) and the frequency of being on the shortest path between nodes in the graph (i.e., BC) have much impact on identifying the most influential node rather than focusing on the average shortest path existing between two nodes. As for predicting five target graphs from the source view 4 (derived from the maximum principal curvature of the right hemisphere), topoGAN outperformed both baseline methods (i.e, Adapted MWGAN and Adapted MWGAN (clustering)) in terms of MAE computed

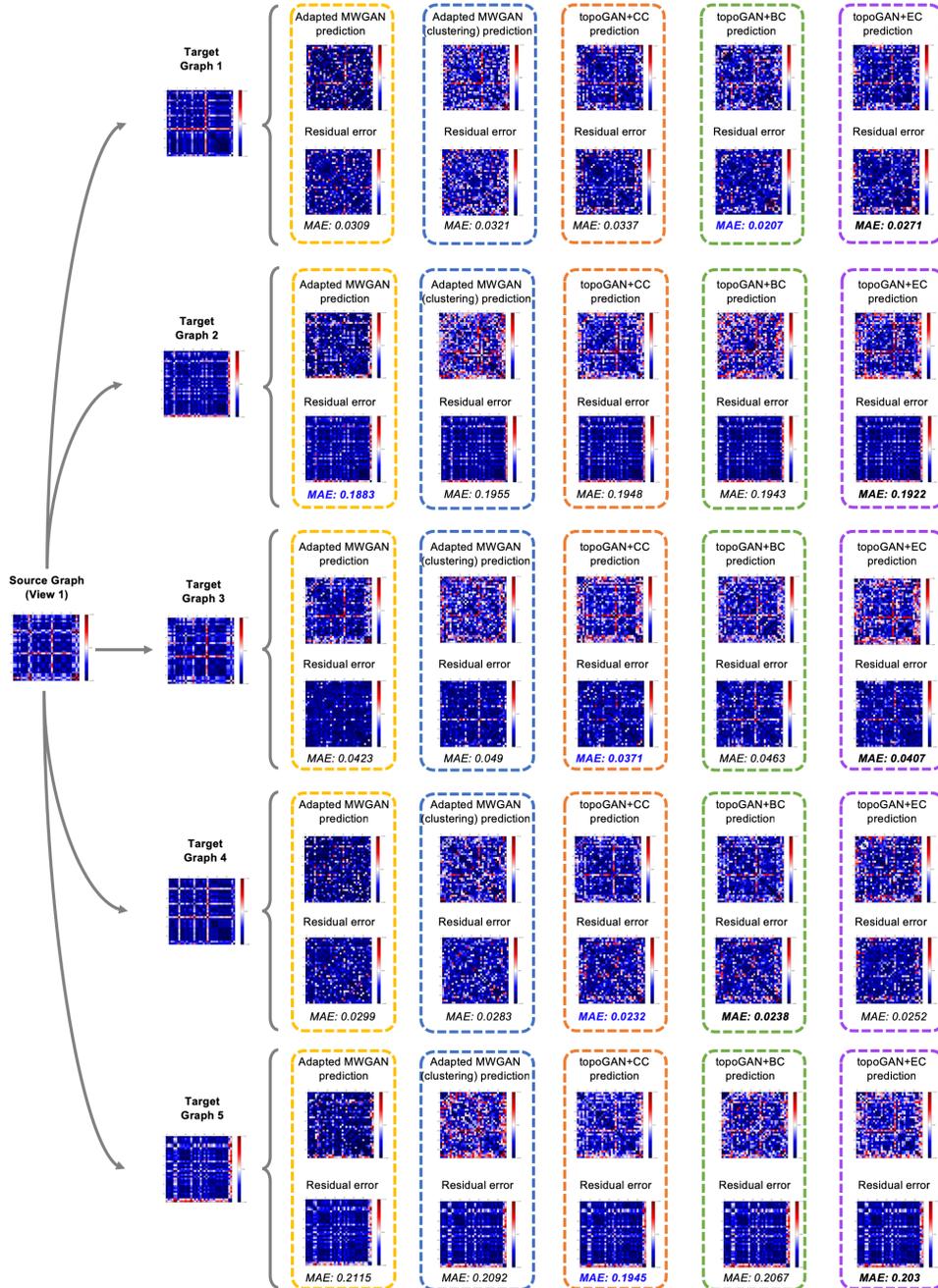


Figure 3: Visual comparison between the ground-truth and the predicted five views stacked in the target brain multigraph of a representative testing subject using different methods: Adapted MWGAN Cao et al. (2019), Adapted MWGAN (clustering), and topoGAN using different centrality metrics. We display the residual error computed using mean absolute error (MAE) metric between the original brain graph and the predicted graph. Source brain graph is derived from maximum principal curvature (LH) (i.e. view1). Target brain graphs are derived from the average curvature (LH) (Target graph 1), mean sulcal depth (LH) (Target graph 2), maximum principal curvature (RH) (Target graph 3), average curvature (RH) (Target graph 4) and mean sulcal depth (RH) (Target graph 5).

using the whole graph, however it produced a slightly higher MAE results computed using the topological measurements (i.e, CC, BC, EC, EFF and Clst). We can see similar results for the view 5 (average curvature of the right hemisphere). We included this example in our results to show that while the average results of topoGAN were remarkable and outperformed comparison methods, our method might lag behind for a few particular source views. Still the Adapted MWGAN (clustering) which is the clustering-based baseline method achieved better results than the Adapted MWGAN method. This clearly demonstrates the advantage of our *cluster-specific generators* for avoiding the mode collapse problem. We can conclusively confirm that our proposed *local topology loss* term optimally improves the learning of the local graph structure when predicting the target graphs. On the other hand, the results reported in **Table ??** and **Table ??** show that the fact of removing the graph reconstruction loss term introduced in (Bessadok et al., 2020b) from the adversarial loss remarkably improves the learning of our model. Mainly, five out of six experiments (i.e, views 1,2,3,5 and 6) results proved that the graph reconstruction loss does not impact the accuracy of predicting the target multigraph. This is explicable because our aim is to make our prediction similar to the real target multigraph so there is no need to enforce the model to preserve the topology of the source domain in the predicted graphs. This further demonstrates that our proposed framework outperformed the existing state-of-the-art method named MultiGraphGAN (Bessadok et al., 2020b). Interestingly, our model achieved better results than comparison methods using two different training strategies: random split and cross-validation. Specifically, we report in supplementary **Table ??** results for three GCN-based comparison methods and our GCN-based topoGAN learned using eigenvector centrality in our loss function. More results can be found in the supplementary material.

4.4.3. Impact of the graph representation learning using GCN

To further evaluate the effectiveness of GCN in learning the brain representation, we implemented a GAT-based version with 8 heads for each benchmark method. Results reported in **Table ??** and **Table ??** demonstrate that our GCN-based framework outperformed the GAT version for six experiments. This can be explained by the fact that assigning larger weights to the most important nodes in brain graphs is not effective since each ROI have a distinct role in the graph thereby equally learning the nodes' representations is highly needed. Originally, the graph convolution operation

updates the local deep features by aggregating the features in a local neighborhood of a particular sample (i.e, node) (Zhang et al., 2019). Thus, our GCN version of topoGAN combined with the topological constraint perfectly models the larger contextual region within the brain which is important for learning the correct anatomical structures in connectomes. We observe in the **Table ??** and **Table ??** that the average KL-divergence over five target views was noticeably reduced by our method. Still, our framework produces higher KL-distance using the source views 4 and 5, which is similar to the MAE results reported in **Table ??**. Essentially, views 1, 2, 3 and 6 show that betweenness and eigenvector centralities present better choices to boost the topological property preservation in jointly predicting multiple target graphs. Such good results demonstrate the advantage of our GCN-based framework in learning the source brain graph representation and preserving the structure of the original target graphs which is in line with recent GCN-based brain analysis works (Banka and Rekik, 2019; Bessadok et al., 2019b).

4.4.4. Comparison using cross-validation

For a thorough comparison, we also compare our topoGAN with two baseline methods (i.e., Adapted MWGAN and Adapted MWGAN (clustering)) and MultiGraphGAN method using 3-fold cross-validation strategy. Since GCN-based variants along with the inclusion of EC in our loss function outperformed the GAT-based ones and GCN-based methods that include CC and BC during the learning process, we only include the experimental results of the GCN-based methods. Supplementary **Table 4** reports the average results over three folds on the same dataset. Our topoGAN achieved better performance than comparison methods when predicting five target graphs from the source views 1, 4 and 5 in terms of MAE computed using the topological measurements (i.e, CC, BC, EFF and Clst), however it produced a slightly higher MAE results computed using the EC and PC evaluation measurements. This is in line with previous results reported in **Table ??** and **Table ??** for the topoGAN+EC method. On the other hand, our topoGAN ranked second best when predicting five target graphs using the source views 2 and 3. This is because the alignment of a source domain to multiple target domains is very challenging, even a one-to-one domain alignment has been recently shown to be a hard task when performed on brain graphs (Wang et al., 2020; Bessadok et al., 2020a). Thus, integrating a domain alignment module in our topoGAN is our future avenue which will better learn the adaptation of the source distribution to multiple target distributions. Still,

these results reported in the supplementary **Table 4** indicate the superiority of our topoGAN over three experiments (i.e, view 1, 4 and 5) which demonstrate the advantage of our proposed cluster-specific generators and our topological loss function compared to its ablated versions as well as the state-of-the-art method.

4.5. Discussion and future recommendations

In this paper, we introduced the first geometric deep learning framework designed for jointly predicting multiple brain graphs from a single brain graph. Our model (i) is a novel graph adversarial auto-encoder that includes an encoder and a set of decoders (i.e., generators) all regularized by a single discriminator, (ii) clusters the source embeddings and defines a set of *cluster-specific decoder* to overcome the mode collapse issue of GAN, (iii) includes a *topological loss* to preserve both global and local topological properties of the original graphs. For the first time, we take the connectomics field one step further into predicting missing brain graphs from existing minimal resources (i.e., a single brain graph). Our topoGAN achieved significantly better performances than the ablated versions and the state-of-the-art method –namely MultiGraphGAN (Bessadok et al., 2020b). Moreover, it consistently outperformed all comparison methods not only when evaluated using random split (i.e., 90/10%) but also when using cross-validation strategy. Although the proposed target brain multigraph prediction framework is generic and can also be applied to any type of brain graph (e.g, functional or structural), there are still some issues that can affect its performance. One major issue is that the more clusters we introduce, the more easily the model can overfit the training set. This might be circumvented by monitoring the training and testing loss functions (Rice et al., 2020).

A second major issue is considering a single discriminator to regularize the encoder and our set of *clustering-specific* generators while recent works demonstrate that multiple discriminators improve the learning process even using small datasets (Durugkar et al., 2016; Neyshabur et al., 2017). Hence, we will improve the cluster-specific multi-target graph prediction step (**Fig. 2-C**) by leveraging a recent multi-objective optimization framework (Albuquerque et al., 2019) that ensures an accurate data generation using multiple adversarial regularizers. However, this method was originally designed for image generation so we aim to adopt it to geometric data. To further improve the GCN learning on brain graphs, we aim to use variational GCN (Tiao et al.) in combination with a recent adversarial graph embedding

Table 4: Prediction results for three source views derived from the left hemisphere using KL-divergence evaluation metric.

View 1		CC	BC	EC	PC	EFF	Clst
Adapted MWGAN	GAT	0.0023	0.088	0.0203	0.0124	0.0213	0.0025
	GCN	0.0019	0.0461	0.0063	0.0093	0.0083	0.0003
Adapted MWGAN (clustering)	GAT	0.0022	0.0501	0.0061	0.0702	0.0104	0.0003
	GCN	0.0009	0.0226	0.001	0.027	0.0038	1.6E-05
topoGAN+CC	GAT	0.0035	0.1064	0.0144	0.0194	0.0222	0.0021
	GCN	0.001	0.0209	0.0011	0.0206	0.0031	1.6E-05
topoGAN+BC	GAT	0.0024	0.0369	0.0039	0.0309	0.0072	0.0001
	GCN	0.0011	0.027	0.0014	0.0223	0.004	2.1E-05
topoGAN+EC	GAT	0.0017	0.0374	0.0034	0.0383	0.0056	0.0001
	GCN	0.0006	0.0215	0.0006	0.0282	0.0025	0,0000072
MultiGraphGAN+CC	GCN	0.00167	0.0351	0.0026	0.0394	0.0055	0,000071
MultiGraphGAN+BC	GCN	0.0007	0.0178	0.0009	0.0155	0.0017	0,0000086
MultiGraphGAN+EC	GCN	0.0007	0.0211	0.0008	0.0229	0.0033	0,000013
View 2		CC	BC	EC	PC	EFF	Clst
Adapted MWGAN	GAT	0.0038	0.1718	0.0235	0.0415	0.0314	0.0045
	GCN	0.0034	0.1336	0.0179	0.0283	0.026	0.0014
Adapted MWGAN (clustering)	GAT	0.002	0.0366	0.0051	0.0241	0.0067	0.0002
	GCN	0.0041	0.1092	0.0135	0.0374	0.0237	0.0009
topoGAN+CC	GAT	0.0039	0.1646	0.0196	0.021	0.0398	0.0033
	GCN	0.003	0.068	0.0128	0.0415	0.0132	0.0006
topoGAN+BC	GAT	0.0022	0.0481	0.0055	0.031	0.0084	0.0003
	GCN	0.0018	0.0324	0.0045	0.0317	0.0061	0.0002
topoGAN+EC	GAT	0.0023	0.0501	0.0042	0.0371	0.0086	0.0001
	GCN	0.0018	0.0371	0.0079	0.0352	0.0078	0.0009
MultiGraphGAN+CC	GCN	0.0029	0.0557	0.0074	0.0312	0.0116	0.0005
MultiGraphGAN+BC	GCN	0.0033	0.0873	0.0122	0.0394	0.0181	0.0008
MultiGraphGAN+EC	GCN	0.0024	0.0441	0.0059	0.0358	0.0083	0.0002
View 3		CC	BC	EC	PC	EFF	Clst
Adapted MWGAN	GAT	0.0032	0.0893	0.0134	0.0109	0.0189	0.0012
	GCN	0.0017	0.0332	0.0045	0.0052	0.0063	0.0002
Adapted MWGAN (clustering)	GAT	0.0022	0.0362	0.0052	0.021	0.0077	0.0003
	GCN	0.0007	0.0147	0.0009	0.0088	0.0024	0,000014
topoGAN+CC	GAT	0.0026	0.0491	0.0063	0.0431	0.0082	0.0002
	GCN	0.0012	0.0282	0.0015	0.0064	0.0064	0,000047
topoGAN+BC	GAT	0.0024	0.0508	0.0045	0.0123	0.0086	0.0001
	GCN	0.0012	0.0265	0.0015	0.0065	0.0057	0,000035
topoGAN+EC	GAT	0.0024	0.0409	0.0047	0.0105	0.0092	0.0003
	GCN	0.0008	0.012	0.0009	0.0113	0.0016	0,0000098
MultiGraphGAN+CC	GCN	0.0015	0.0247	0.0021	0.0068	0.0054	0,000048
MultiGraphGAN+BC	GCN	0.0017	0.0266	0.0013	0.0061	0.0037	0,000024
MultiGraphGAN+EC	GCN	0.0019	0.0465	0.0028	0.0079	0.0103	0,000097

View 1: maximum principal curvature (LH). View 2: average curvature (LH). View 3: mean sulcal depth (LH). MAE: mean absolute error. CC: closeness centrality. BC: betweenness centrality. EC: eigenvector centrality. PC: PageRank centrality. EFF: effective size. Clst: clustering coefficient. We highlight in red and blue colors the lowest KL-distance resulting from a particular evaluation metric for the GCN and GAT versions of topoGAN, respectively.

Table 5: Prediction results for three source views derived from the right hemisphere using KL-divergence evaluation metric.

View 4		CC	BC	EC	PC	EFF	Clst
Adapted MWGAN	GAT	0.0028	0.1416	0.0268	0.0661	0.0223	0.0062
	GCN	0.0021	0.0552	0.0075	0.0102	0.0095	0.0003
Adapted MWGAN (clustering)	GAT	0.0029	0.0773	0.0085	0.0766	0.0165	0.0005
	GCN	0.0008	0.0167	0.001	0.0188	0.002	0,000009
topoGAN+CC	GAT	0.0023	0.061	0.0052	0.0455	0.0115	0.0004
	GCN	0.0014	0.034	0.0018	0.0235	0.0055	0,00003
topoGAN+BC	GAT	0.0025	0.0417	0.0049	0.0178	0.0067	0.0002
	GCN	0.0015	0.0349	0.0019	0.0198	0.005	0,000022
topoGAN+EC	GAT	0.0021	0.0421	0.0042	0.034	0.0075	0.0002
	GCN	0.0015	0.0418	0.0018	0.0236	0.006	0,000026
MultiGraphGAN+CC	GCN	0.0006	0.0117	0.0007	0.0185	0.0014	0,000008
MultiGraphGAN+BC	GCN	0.0019	0.0461	0.0029	0.0267	0.0088	0,000085
MultiGraphGAN+EC	GCN	0.0008	0.0153	0.001	0.0255	0.0029	0,000015
View 5		CC	BC	EC	PC	EFF	Clst
Adapted MWGAN	GAT	0.0034	0.1783	0.023	0.0426	0.0337	0.0045
	GCN	0.0024	0.0688	0.0116	0.0195	0.0123	0.0015
Adapted MWGAN (clustering)	GAT	0.0032	0.0543	0.0066	0.0074	0.0116	0.0003
	GCN	0.0025	0.0386	0.0056	0.0198	0.006	0.0003
topoGAN+CC	GAT	0.0021	0.0374	0.0042	0.0154	0.0087	0.0002
	GCN	0.0026	0.052	0.0065	0.018	0.0116	0.0005
topoGAN+BC	GAT	0.0036	0.0738	0.0105	0.014	0.0156	0.0006
	GCN	0.0035	0.0725	0.01	0.0332	0.0157	0.0005
topoGAN+EC	GAT	0.0013	0.0405	0.0026	0.0111	0.0065	0.0001
	GCN	0.0035	0.0845	0.0114	0.0253	0.0171	0.001
MultiGraphGAN+CC	GCN	0.0022	0.0431	0.0052	0.0244	0.0069	0.0003
MultiGraphGAN+BC	GCN	0.0038	0.1026	0.015	0.0232	0.0185	0.0012
MultiGraphGAN+EC	GCN	0.0036	0.0746	0.0099	0.0295	0.0175	0.0008
View 6		CC	BC	EC	PC	EFF	Clst
Adapted MWGAN	GAT	0.0027	0.0883	0.0117	0.0099	0.018	0.0011
	GCN	0.0014	0.0791	0.0095	0.026	0.0169	0.002
Adapted MWGAN (clustering)	GAT	0.003	0.0708	0.0098	0.0414	0.0122	0.0006
	GCN	0.0018	0.0257	0.004	0.031	0.0045	0.0002
topoGAN+CC	GAT	0.0026	0.0386	0.0068	0.0314	0.006	0.0003
	GCN	0.0017	0.037	0.0026	0.0202	0.0065	0.0001
topoGAN+BC	GAT	0.0017	0.0416	0.0034	0.0581	0.0076	0.0002
	GCN	0.0019	0.0295	0.0028	0.0189	0.0065	0.0001
topoGAN+EC	GAT	0.0026	0.0481	0.0054	0.0304	0.0096	0.0003
	GCN	0.0015	0.0254	0.0022	0.027	0.0052	0.0001
MultiGraphGAN+CC	GCN	0.0027	0.0491	0.0066	0.0426	0.0077	0.0003
MultiGraphGAN+BC	GCN	0.0018	0.029	0.0056	0.0342	0.0075	0.0003
MultiGraphGAN+EC	GCN	0.0026	0.0548	0.0043	0.0351	0.0124	0.0002

View 4: maximum principal curvature (RH). View 5: average curvature (RH). View 6: mean sulcal depth (RH). MAE: mean absolute error. CC: closeness centrality. BC: betweenness centrality. EC: eigenvector centrality. PC: PageRank centrality. EFF: effective size. Clst: clustering coefficient. We highlight in red and blue colors the lowest KL-distance resulting from a particular evaluation metric for the GCN and GAT versions of topoGAN, respectively.

technique (Pan et al., 2018). We hypothesize that such architecture improvement will boost the source graph embedding. As another research direction, we will evaluate our framework using a larger dataset including functional and structural brain graphs. Since in this work we only focused on graph synthesis task we aim in the future to work on early disease identification. Essentially, we will combine both graph prediction and disease classification tasks in a single and unified geometric deep learning framework trained in an end-to-end manner which will be used for diagnosing different neurological disorders. Ultimately, this model can serve as a stepping stone to develop more holistic brain prediction models such as predicting spatiotemporal trajectory (Ghribi et al., 2019; Vohryzek et al., 2020) and super-resolution brain graphs (Cengiz and Rekik, 2019; Mhiri et al., 2020a).

5. Conclusion

Very few models exist for predicting a single target brain graph from a source graph (i.e, one-to-one prediction task). This is a recently emerging field with high-level meaningful implications in neurological disorder diagnosis. We presented in this paper the first geometric deep learning framework, namely topoGAN, for jointly predicting multiple brain views represented by a *target multigraph* from a single source graph both derived from MRI (i.e, one-to-many prediction task). Our architecture has two compelling strengths: (i) clustering the learned source graphs embeddings then training a set of *cluster-specific generators* which synergistically predict the target brain graphs, (ii) introducing a *topological loss function* using a centrality measure which enforces the generators to preserve both local and global topologies of the original target graphs. Our proposed brain multigraph prediction framework can be further tailored to predict the evolution of target brain multigraph over time from a single source brain graph (Ghribi et al., 2019; Vohryzek et al., 2020). Such framework harbors a powerful tool to examine how alterations in the connectome may lead to a progressive brain dysfunction in neurological disorder. Eventually, we envision to evaluate our framework on larger connectomic datasets and cover a diverse range of brain graphs such structural and functional networks (Wen et al., 2017; Mhiri and Rekik, 2020).

6. Acknowledgements

This work was funded by generous grants from the European H2020 Marie Skłodowska-Curie action (grant no. 101003403, <http://basira-lab.com/>

[normnets/](#)) to I.R. and the Scientific and Technological Research Council of Turkey to I.R. under the TUBITAK 2232 Fellowship for Outstanding Researchers (no. 118C288, <http://basira-lab.com/reprime/>). However, all scientific contributions made in this project are owned and approved solely by the authors. A.B. is supported by the same the TUBITAK 2232 Fellowship for Outstanding Researchers.

Table 6: P-value results using two-tailed paired t-test for predicting a target brain multigraph using two source views derived from the left hemisphere.

View 1		topoGAN+CC	topoGAN+BC	topoGAN+EC
Adapted MWGAN	View 1	1.1661×10^{-10}	1.7835×10^{-5}	4.8519×10^{-10}
	View 2	0.0354	0.0911	0.1138
	View 3	0.1259	0.0321	0.0516
	View 4	9.8979×10^{-7}	0.0404	0.0717
	View 5	0.0369	0.0694	0.0808
Adapted MWGAN (clustering)	View 1	0.016	0.0924	0.1008
	View 2	0.0643	0.0679	0.0281
	View 3	0.0146	0.0007	0.0014
	View 4	7.6946×10^{-7}	0.0045	0.0553
	View 5	0.0014	0.1216	0.0454
MultiGraphGAN+CC	View 1	0.0001	0.0012	0.0023
	View 2	2.3×10^{-20}	0.0001	0.0041
	View 3	0.0299	2.7×10^{-7}	0.0018
	View 4	0.0403	0.0011	2.02×10^{-8}
	View 5	0.0003	2.7×10^{-7}	2.5×10^{-12}
MultiGraphGAN+BC	View 1	0.0789	0.0649	0.0306
	View 2	0.0976	0.0245	0.0002
	View 3	0.086	0.0015	0.0002
	View 4	0.0599	0.0115	0.0211
	View 5	0.0805	0.0549	0.0102
MultiGraphGAN+EC	View 1	0.0679	0.1076	0.0059
	View 2	1.4×10^{-7}	0.1027	0.0723
	View 3	0.0177	0.0917	0.0264
	View 4	8.5×10^{-11}	0.0054	0.0609
	View 5	0.0161	0.0548	0.086
View 2		topoGAN+CC	topoGAN+BC	topoGAN+EC
Adapted MWGAN	View 1	1.7883×10^{-15}	8.9525×10^{-12}	0.0008
	View 2	2.4751×10^{-15}	7.0026×10^{-10}	0.0007
	View 3	4.6296×10^{-20}	9.9692×10^{-20}	9.9618×10^{-11}
	View 4	4.6198×10^{-7}	0.0603	0.0567
	View 5	1.0612×10^{-12}	0.0001	0.0003
Adapted MWGAN (clustering)	View 1	0.2857	0.1867	0.0294
	View 2	0.0433	0.2578	0.1832
	View 3	0.0599	0.0553	0.4429
	View 4	0.4283	0.0005	0.0292
	View 5	0.1386	0.0361	0.2063
MultiGraphGAN+CC	View 1	0.1786	0.3806	0.0246
	View 2	0.0361	0.1927	0.263
	View 3	0.4158	0.4336	0.1761
	View 4	0.14	0.0002	0.0003
	View 5	1.01×10^{-5}	0.2049	0.1333
MultiGraphGAN+BC	View 1	0.0129	0.1223	0.3667
	View 2	7.8×10^{-7}	0.0359	0.2009
	View 3	0.0099	0.0372	0.2489
	View 4	5.8×10^{-24}	6.98×10^{-7}	1.8×10^{-10}
	View 5	0.0184	0.3352	0.5021
MultiGraphGAN+EC	View 1	0.1926	0.3907	0.0498
	View 2	0.3544	0.3772	0.0046
	View 3	0.01	0.031	0.2065
	View 4	3.3×10^{-6}	0.0407	0.0343
	View 5	0.0014	0.3154	0.2411

Adapted MWGAN: the graph-based architecture of the method introduced in (Cao et al., 2019). Adapted MWGAN (clustering): a variant of the adapted method (Cao et al., 2019) with a clustering step. topoGAN+CC, topoGAN+BC, topoGAN+EC: the proposed method that includes closeness, betweenness and eigenvector. MultiGraphGAN+CC, MultiGraphGAN+BC, MultiGraphGAN+EC: the state-of-the-art method (Bessadok et al., 2020b) that includes the reconstruction loss and the closeness, betweenness and eigenvector centralities. View 1: maximum principal curvature (LH). View 2: average curvature (LH). We highlight in bold the p -value < 0.05 using two-tailed paired t -test between row-wise and column-wise methods.

Table 7: P-value results using two-tailed paired t-test for predicting a target brain multigraph using six source views.

View 3		topoGAN+CC	topoGAN+BC	topoGAN+EC
Adapted MWGAN	View 1	0.0252	0.1635	0.0961
	View 2	4.2923×10^{-13}	0.0265	0.005
	View 3	0.1796	0.0043	0.0611
	View 4	0.024	0.0006	0.0059
	View 5	0.0583	0.0547	0.0479
Adapted MWGAN (clustering)	View 1	0.0016	0.179	0.0437
	View 2	6.1182×10^{-7}	0.0858	0.0527
	View 3	0.0217	0.091	0.0292
	View 4	0.0204	0.0068	0.0317
	View 5	0.0202	0.0595	0.023
MultiGraphGAN+CC	View 1	0.0001	0.0012	0.0023
	View 2	2.3×10^{-20}	0.0001	0.0041
	View 3	0.0299	2.7×10^{-7}	0.0018
	View 4	0.0403	0.0011	2.02×10^{-8}
	View 5	0.0003	2.7×10^{-7}	2.5×10^{-12}
MultiGraphGAN+BC	View 1	0.0789	0.0649	0.0306
	View 2	0.0976	0.0245	0.0002
	View 3	0.086	0.0015	0.0002
	View 4	0.0599	0.0115	0.0211
	View 5	0.0805	0.0549	0.0102
MultiGraphGAN+EC	View 1	0.0679	0.1076	0.0059
	View 2	1.4×10^{-7}	0.1027	0.0723
	View 3	0.0177	0.0917	0.0264
	View 4	8.5×10^{-11}	0.0054	0.0609
	View 5	0.0161	0.0548	0.086
View 4		topoGAN+CC	topoGAN+BC	topoGAN+EC
Adapted MWGAN	View 1	0.0437	0.0762	0.0008
	View 2	0.2806	0.2209	0.0847
	View 3	0.1332	0.1404	8.6744×10^{-7}
	View 4	0.1761	0.0564	0.0316
	View 5	0.166	0.0047	0.1258
Adapted MWGAN (clustering)	View 1	0.1384	1.2571×10^{-16}	0.1709
	View 2	0.0002	3.1156×10^{-11}	0.373
	View 3	0.0462	6.7905×10^{-7}	0.0099
	View 4	0.0269	6.8363×10^{-11}	4.0492×10^{-37}
	View 5	0.1468	2.1769×10^{-6}	0.2629
MultiGraphGAN+CC	View 1	0.0835	2.1×10^{-14}	0.1468
	View 2	0.0793	0.0015	0.0321
	View 3	0.1515	0.0023	0.0044
	View 4	0.1276	0.0076	0.0007
	View 5	0.0231	3.1×10^{-10}	0.0276
MultiGraphGAN+BC	View 1	0.0122	1.96×10^{-6}	0.0105
	View 2	0.1649	0.1468	0.0384
	View 3	0.0287	3.97×10^{-13}	0.0211
	View 4	1.3×10^{-5}	0.0389	0.1389
	View 5	0.0006	1.3×10^{-16}	5.65×10^{-8}
MultiGraphGAN+EC	View 1	0.1068	0.0314	0.1585
	View 2	0.0181	0.027	0.2697
	View 3	0.0994	0.0048	0.0004
	View 4	0.0001	0.0464	0.1533
	View 5	0.1148	0.0444	0.1643

Adapted MWGAN: the graph-based architecture of the method introduced in (Cao et al., 2019). Adapted MWGAN (clustering): a variant of the adapted method (Cao et al., 2019) with a clustering step. topoGAN+CC, topoGAN+BC, topoGAN+EC: the proposed method that includes closeness, betweenness and eigenvector. MultiGraphGAN+CC, MultiGraphGAN+BC, MultiGraphGAN+EC: the state-of-the-art method (Bessadok et al., 2020b) that includes the reconstruction loss and the closeness, betweenness and eigenvector centralities. View 3: mean sulcal depth (LH). View 4: maximum principal curvature (RH). We highlight in bold the p -value < 0.05 using two-tailed paired t -test between row-wise and column-wise methods.

Table 8: P-value results using two-tailed paired t-test for predicting a target brain multigraph using six source views.

View 5		topoGAN+CC	topoGAN+BC	topoGAN+EC
Adapted MWGAN	View 1	0.03	1.9211×10^{-15}	0.0002
	View 2	0.0967	2.9925×10^{-6}	0.2356
	View 3	7.8964×10^{-6}	2.9087×10^{-11}	0.0006
	View 4	1.3136×10^{-7}	9.1108×10^{-9}	4.8735×10^{-10}
	View 5	0.029	0.0002	0.0037
Adapted MWGAN (clustering)	View 1	0.0997	1.3734×10^{-9}	0.0489
	View 2	0.046	0.0002	0.2289
	View 3	0.1573	0.0319	0.2678
	View 4	0.1849	0.2604	0.108
	View 5	0.1197	0.012	0.0808
MultiGraphGAN+CC	View 1	0.2674	0.0074	0.1941
	View 2	0.3214	0.0206	0.4486
	View 3	0.39	0.1531	0.0641
	View 4	0.1074	0.0933	0.0707
	View 5	0.099	0.0514	0.095
MultiGraphGAN+BC	View 1	0.3093	0.034	0.1881
	View 2	0.1106	0.013	0.1803
	View 3	0.2799	0.208	0.0787
	View 4	0.1296	0.0765	0.1991
	View 5	0.0954	0.3384	0.1493
MultiGraphGAN+EC	View 1	0.3618	0.0036	0.1621
	View 2	0.2133	0.27	0.1087
	View 3	0.0444	0.0326	0.1248
	View 4	0.0509	0.0295	0.0323
	View 5	0.2233	0.0551	0.1527
View 6		topoGAN+CC	topoGAN+BC	topoGAN+EC
Adapted MWGAN	View 1	0.047	0.0035	0.0396
	View 2	0.1528	0.0824	0.0838
	View 3	0.0129	0.0019	0.0044
	View 4	0.1271	0.0515	0.0626
	View 5	0.0904	0.0647	0.1407
Adapted MWGAN (clustering)	View 1	0.034	0.0399	0.0522
	View 2	0.034	0.0335	0.0019
	View 3	0.0152	0.0295	0.0916
	View 4	0.0134	0.0592	0.075
	View 5	0.0908	0.0732	0.0014
MultiGraphGAN+CC	View 1	0.0631	0.0466	0.0562
	View 2	0.0805	0.0273	0.015
	View 3	0.0866	0.0304	0.0751
	View 4	0.0584	0.0619	0.0694
	View 5	0.028	0.0217	0.0526
MultiGraphGAN+BC	View 1	0.0113	0.0601	8.95E-06
	View 2	0.0006	0.1182	0.0014
	View 3	0.1567	0.0928	0.0771
	View 4	0.0308	0.0644	0.0073
	View 5	0.0011	0.0286	0.0423
MultiGraphGAN+EC	View 1	0.0005	0.0063	0.0325
	View 2	0.0009	0.1067	0.0039
	View 3	0.0109	0.0402	0.0224
	View 4	0.0225	0.043	0.0247
	View 5	0.0072	0.0057	0.0174

Adapted MWGAN: the graph-based architecture of the method introduced in (Cao et al., 2019). Adapted MWGAN (clustering): a variant of the adapted method (Cao et al., 2019) with a clustering step. topoGAN+CC, topoGAN+BC, topoGAN+EC: the proposed method that includes closeness, betweenness and eigenvector. MultiGraphGAN+CC, MultiGraphGAN+BC, MultiGraphGAN+EC: the state-of-the-art method (Bessadok et al., 2020b) that includes the reconstruction loss and the closeness, betweenness and eigenvector centralities. View 5: average curvature (RH). View 6: mean sulcal depth (RH). We highlight in bold the p -value < 0.05 using two-tailed paired t -test between row-wise and column-wise methods.

Table 9: Prediction results for six source views derived from the left and right hemispheres using 3-fold cross-validation strategy.

View 1	MAE	MAE(CC)	MAE(BC)	MAE(EC)	MAE(PC)	MAE(EFF)	MAE(Clst)
Adapted MWGAN	0.19319	0.27894	0.01212	0.01743	0.00375	7.55474	0.40396
Adapted MWGAN (clustering)	0.20226	0.12416	0.00449	0.00936	0.00517	4.00669	0.14722
MultiGraphGAN	0.20063	0.13705	0.00502	0.00973	0.00523	4.39254	0.16524
topoGAN	0.19342	0.09916	0.00348	0.00783	0.0046	3.26974	0.11495
View 2	MAE	MAE(CC)	MAE(BC)	MAE(EC)	MAE(PC)	MAE(EFF)	MAE(Clst)
Adapted MWGAN	0.20901	0.3032	0.01362	0.01755	0.00552	7.63256	0.44805
Adapted MWGAN (clustering)	0.21127	0.16951	0.00648	0.01323	0.00462	5.1742	0.20651
MultiGraphGAN	0.21245	0.18279	0.00708	0.01242	0.0062	5.58634	0.2296
topoGAN	0.21138	0.18863	0.00733	0.01236	0.00545	5.78864	0.24022
View 3	MAE	MAE(CC)	MAE(BC)	MAE(EC)	MAE(PC)	MAE(EFF)	MAE(Clst)
Adapted MWGAN	0.12093	0.21652	0.00871	0.01469	0.0037	6.41082	0.28415
Adapted MWGAN (clustering)	0.15036	0.14922	0.00556	0.01116	0.00385	4.72162	0.18153
MultiGraphGAN	0.14884	0.12242	0.00442	0.00936	0.00405	3.96439	0.14465
topoGAN	0.13851	0.13969	0.00515	0.01044	0.00358	4.42449	0.16662
View 4	MAE	MAE(CC)	MAE(BC)	MAE(EC)	MAE(PC)	MAE(EFF)	MAE(Clst)
Adapted MWGAN	0.18549	0.27491	0.01193	0.01761	0.00435	7.32622	0.3872
Adapted MWGAN (clustering)	0.19355	0.10838	0.00385	0.00853	0.00471	3.52877	0.12668
MultiGraphGAN	0.19864	0.12713	0.0046	0.0092	0.00474	4.11026	0.15115
topoGAN	0.19029	0.09762	0.00342	0.00891	0.00478	3.2047	0.1125
View 5	MAE	MAE(CC)	MAE(BC)	MAE(EC)	MAE(PC)	MAE(EFF)	MAE(Clst)
Adapted MWGAN	0.20877	0.29294	0.01303	0.02137	0.00506	7.52582	0.42447
Adapted MWGAN (clustering)	0.21394	0.18547	0.00721	0.0131	0.00565	5.66924	0.23389
MultiGraphGAN	0.21372	0.22346	0.00916	0.01538	0.00536	6.45255	0.30052
topoGAN	0.21143	0.1744	0.00669	0.01261	0.00521	5.4011	0.21785
View 6	MAE	MAE(CC)	MAE(BC)	MAE(EC)	MAE(PC)	MAE(EFF)	MAE(Clst)
Adapted MWGAN	0.09501	0.32472	0.01514	0.02613	0.00543	7.4837	0.48978
Adapted MWGAN (clustering)	0.05621	0.16894	0.00643	0.01149	0.00436	5.30278	0.21254
MultiGraphGAN	0.06627	0.18414	0.0071	0.01143	0.00431	5.71455	0.2348
topoGAN	0.10415	0.17302	0.00664	0.01254	0.00456	5.32186	0.2146

View 1: maximum principal curvature (LH). View 2: average curvature (LH). View 3: mean sulcal depth (LH). View 4: maximum principal curvature (RH). View 5: average curvature (RH). View 6: mean sulcal depth (RH). MAE: mean absolute error. CC: closeness centrality. BC: betweenness centrality. EC: eigenvector centrality. PC: PageRank centrality. EFF: effective size. Clst: clustering coefficient. We highlight in blue the lowest MAE resulting from a particular evaluation metric for the GCN-based comparison methods and our topoGAN method learned using EC in our loss function.

References

- Albuquerque, I., Monteiro, J., Doan, T., Considine, B., Falk, T., Mitliagkas, I., 2019. Multi-objective training of generative adversarial networks with multiple discriminators. arXiv preprint arXiv:1901.08680 .
- Banka, A., Rekik, I., 2019. Adversarial connectome embedding for mild cognitive impairment identification using cortical morphological networks. *International Workshop on Connectomics in Neuroimaging* , 74–82.
- Bassett, D.S., Sporns, O., 2017. Network neuroscience. *Nature neuroscience* 20, 353–364.
- Beauchamp, M.A., 1965. An improved index of centrality. *Behavioral science* 10, 161–163.
- Belli, D., Kipf, T., 2019. Image-conditioned graph generation for road network extraction. arXiv preprint arXiv:1910.14388 .
- Bessadok, A., Mahjoub, M.A., Rekik, I., 2019a. Hierarchical adversarial connectomic domain alignment for target brain graph prediction and classification from a source graph. *International Workshop on Predictive Intelligence In Medicine* , 105–114.
- Bessadok, A., Mahjoub, M.A., Rekik, I., 2019b. Symmetric dual adversarial connectomic domain alignment for predicting isomorphic brain graph from a baseline graph. *International Conference on Medical Image Computing and Computer-Assisted Intervention* , 465–474.
- Bessadok, A., Mahjoub, M.A., Rekik, I., 2020a. Brain graph synthesis by dual adversarial domain alignment and target graph prediction from a source graph. *Medical Image Analysis* , 101902.
- Bessadok, A., Mahjoub, M.A., Rekik, I., 2020b. Topology-aware generative adversarial network for joint prediction of multiple brain graphs from a single brain graph, in: *International Conference on Medical Image Computing and Computer-Assisted Intervention*, Springer. pp. 551–561.
- Bessadok, A., Rekik, I., 2018. Intact connectional morphometricity learning using multi-view morphological brain networks with application to autism spectrum disorder, in: *International Workshop on Connectomics in Neuroimaging*, Springer. pp. 38–46.

- Bonacich, P., 2007. Some unique properties of eigenvector centrality. *Social networks* 29, 555–564.
- Borgatti, S.P., Everett, M.G., 2006. A graph-theoretic perspective on centrality. *Social networks* 28, 466–484.
- Bresson, X., Laurent, T., 2019. A two-step graph convolutional decoder for molecule generation. *arXiv preprint arXiv:1906.03412* .
- Brin, S., 1998. The PageRank citation ranking: bringing order to the web. *Proceedings of ASIS, 1998* 98, 161–172.
- Bronstein, M.M., Bruna, J., LeCun, Y., Szlam, A., Vandergheynst, P., 2017. Geometric deep learning: going beyond euclidean data. *IEEE Signal Processing Magazine* 34, 18–42.
- Burt, R.S., 2009. *Structural holes: The social structure of competition*. Harvard university press.
- Cao, J., Mo, L., Zhang, Y., Jia, K., Shen, C., Tan, M., 2019. Multi-marginal Wasserstein GAN. *Advances in Neural Information Processing Systems* , 1774–1784.
- Cengiz, K., Rekik, I., 2019. Predicting high-resolution brain networks using hierarchically embedded and aligned multi-resolution neighborhoods. *International Workshop on PRedictive Intelligence In MEdicine* , 115–124.
- Cid, Y.D., Jimenez-del Toro, O., Poletti, P.A., Müller, H., 2019. A graph model of the lungs with morphology-based structure for tuberculosis type classification. *International Conference on Information Processing in Medical Imaging* , 372–383.
- De Cao, N., Kipf, T., 2018. Molgan: An implicit generative model for small molecular graphs. *arXiv preprint arXiv:1805.11973* .
- Dhifallah, S., Rekik, I., Initiative, A.D.N., 2020. Estimation of connectional brain templates using selective multi-view network normalization. *Medical Image Analysis* 59, 101567.
- Durugkar, I., Gemp, I., Mahadevan, S., 2016. Generative multi-adversarial networks. *arXiv preprint arXiv:1611.01673* .

- Fischl, B., 2012. Freesurfer. *Neuroimage* 62, 774–781.
- Flam-Shepherd, D., Wu, T., Aspuru-Guzik, A., 2020. Graph deconvolutional generation. *arXiv preprint arXiv:2002.07087* .
- Fornito, A., Zalesky, A., Breakspear, M., 2015. The connectomics of brain disorders. *Nature Reviews Neuroscience* 16, 159–172.
- Freeman, L.C., 1977. A set of measures of centrality based on betweenness. *Sociometry* , 35–41.
- Ghribi, O., Li, G., Lin, W., Shen, D., Rezik, I., 2019. Progressive infant brain connectivity evolution prediction from neonatal MRI using bidirectionally supervised sample selection. *International Workshop on Predictive Intelligence In Medicine* , 63–72.
- Goktas, A.S., Bessadok, A., Rezik, I., 2020. Residual embedding similarity-based network selection for predicting brain network evolution trajectory from a single observation. *arXiv preprint arXiv:2009.11110* .
- Goodfellow, I., et al., 2014. Generative adversarial nets. *Advances in neural information processing systems* , 2672–2680.
- Gulrajani, I., Ahmed, F., Arjovsky, M., Dumoulin, V., Courville, A.C., 2017. Improved training of wasserstein gans. *Advances in neural information processing systems* , 5767–5777.
- Gurbuz, M.B., Rezik, I., 2020. Deep graph normalizer: A geometric deep learning approach for estimating connectional brain templates, in: *International Conference on Medical Image Computing and Computer-Assisted Intervention*, Springer. pp. 155–165.
- Hamilton, W.L., Ying, R., Leskovec, J., 2017. Representation learning on graphs: Methods and applications. *arXiv preprint arXiv:1709.05584* .
- van den Heuvel, M.P., Sporns, O., 2019. A cross-disorder connectome landscape of brain dysconnectivity. *Nature reviews neuroscience* 20, 435–446.
- Huang, L.C., Wu, P.A., Lin, S.Z., Pang, C.Y., Chen, S.Y., 2019a. Graph theory and network topological metrics may be the potential biomarker in Parkinson’s disease. *Journal of Clinical Neuroscience* 68, 235–242.

- Huang, P., et al., 2019b. CoCa-GAN: Common-feature-learning-based context-aware generative adversarial network for glioma grading. *International Conference on Medical Image Computing and Computer-Assisted Intervention* , 155–163.
- Jain, A.K., 2010. Data clustering: 50 years beyond k-means. *Pattern recognition letters* 31, 651–666.
- Jolliffe, I.T., Cadima, J., 2016. Principal component analysis: a review and recent developments. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences* 374, 20150202.
- Joyce, K.E., Laurienti, P.J., Burdette, J.H., Hayasaka, S., 2010. A new measure of centrality for brain networks. *PloS one* 5, e12200.
- Khelifa, O.B., Rekik, I., 2019. Graph morphology-based genetic algorithm for classifying late dementia states. *International Workshop on Connectomics in Neuroimaging* , 21–31.
- Kipf, T.N., Welling, M., 2016. Semi-supervised classification with graph convolutional networks. *arXiv preprint arXiv:1609.02907* .
- Kullback, S., 1959. *Information theory and statistics* wiley. New York .
- Li, H., et al., 2019. Diamondgan: Unified multi-modal generative adversarial networks for MRI sequences synthesis. *International Conference on Medical Image Computing and Computer-Assisted Intervention* , 795–803.
- Li, W., Andreasen, N.C., Nopoulos, P., Magnotta, V.A., 2013. Automated parcellation of the brain surface generated from magnetic resonance images. *Frontiers in neuroinformatics* 7, 23.
- Li, X., Zhou, Y., Dvornek, N.C., Zhang, M., Zhuang, J., Ventola, P., Duncan, J.S., 2020. Pooling regularized graph neural network for fmri biomarker analysis, in: *International Conference on Medical Image Computing and Computer-Assisted Intervention*, Springer. pp. 625–635.
- Liao, R., et al., 2019. Efficient graph generation with graph recurrent attention networks. *Advances in Neural Information Processing Systems* , 4257–4267.

- Lin, J.H., Sellke, T.M., Coyle, E.J., 1990. Adaptive stack filtering under the mean absolute error criterion. *IEEE transactions on acoustics, speech, and signal processing* 38, 938–954.
- Liu, J., et al., 2017. Complex brain network analysis and its applications to brain disorders: a survey. *Complexity* 2017.
- Maaten, L.v.d., Hinton, G., 2008. Visualizing data using t-SNE. *Journal of machine learning research* 9, 2579–2605.
- Mahjoub, I., Mahjoub, M.A., Rekik, I., 2018. Brain multiplexes reveal morphological connectional biomarkers fingerprinting late brain dementia states. *Scientific reports* 8, 4103.
- Mhiri, I., Khalifa, A.B., Mahjoub, M.A., Rekik, I., 2020a. Brain graph super-resolution for boosting neurological disorder diagnosis using unsupervised multi-topology connectional brain template learning. *Medical Image Analysis* , 101768.
- Mhiri, I., Mahjoub, M.A., Rekik, I., 2020b. Supervised multi-topology network cross-diffusion for population-driven brain network atlas estimation, in: *International Conference on Medical Image Computing and Computer-Assisted Intervention*, Springer. pp. 166–176.
- Mhiri, I., Rekik, I., 2020. Joint functional brain network atlas estimation and feature selection for neurological disorder diagnosis with application to autism. *Medical image analysis* 60, 101596.
- Mitton, J., Senn, H.M., Wynne, K., Murray-Smith, R., . A graph vae and graph transformer approach to generating molecular graphs .
- Nebli, A., Kaplan, U.A., Rekik, I., 2020. Deep evographnet architecture for time-dependent brain graph data synthesis from a single timepoint, in: *International Workshop on Predictive Intelligence In Medicine*, Springer. pp. 144–155.
- Nebli, A., Rekik, I., 2019. Gender differences in cortical morphological networks. *Brain Imaging and Behavior* , 1–9.
- Neyshabur, B., Bhojanapalli, S., Chakrabarti, A., 2017. Stabilizing GAN training with multiple random projections. *arXiv preprint arXiv:1705.07831* .

- Pan, S., Hu, R., Long, G., Jiang, J., Yao, L., Zhang, C., 2018. Adversarially regularized graph autoencoder for graph embedding. arXiv preprint arXiv:1802.04407 .
- Pan, Y., Liu, M., Lian, C., Xia, Y., Shen, D., 2019. Disease-image specific generative adversarial network for brain disease diagnosis with incomplete multi-modal neuroimages. International Conference on Medical Image Computing and Computer-Assisted Intervention , 137–145.
- Rice, L., Wong, E., Kolter, J.Z., 2020. Overfitting in adversarially robust deep learning. arXiv preprint arXiv:2002.11569 .
- Saramäki, J., Kivelä, M., Onnela, J.P., Kaski, K., Kertesz, J., 2007. Generalizations of the clustering coefficient to weighted complex networks. Physical Review E 75, 027105.
- Song, X., Frangi, A., Xiao, X., Cao, J., Wang, T., Lei, B., 2020. Integrating similarity awareness and adaptive calibration in graph convolution network to predict disease, in: International Conference on Medical Image Computing and Computer-Assisted Intervention, Springer. pp. 124–133.
- Soussia, M., Rekik, I., 2018. Unsupervised manifold learning using high-order morphological brain networks derived from T1-w MRI for autism diagnosis. Frontiers in Neuroinformatics 12.
- Sserwadda, A., Rekik, I., 2020. Topology-guided cyclic brain connectivity generation using geometric deep learning. Journal of Neuroscience Methods , 108988.
- Su, S.Y., Hajimirsadeghi, H., Mori, G., 2019. Graph generation with Variational Recurrent Neural Network. arXiv preprint arXiv:1910.01743 .
- Tiao, L., Elinas, P., Nguyen, H., Bonilla, E.V., . Variational Graph Convolutional Networks .
- Veličković, P., Cucurull, G., Casanova, A., Romero, A., Lio, P., Bengio, Y., 2017. Graph attention networks. arXiv preprint arXiv:1710.10903 .
- Vohryzek, J., et al., 2020. Dynamic spatiotemporal patterns of brain connectivity reorganize across development. Network neuroscience 4, 115–133.

- Wang, B., Ramazzotti, D., De Sano, L., Zhu, J., Pierson, E., Batzoglu, S., 2017. SIMLR: a tool for large-scale single-cell analysis by multi-kernel learning. *bioRxiv* , 118901.
- Wang, J., Chen, J., Lin, J., Sigal, L., de Silva, C.W., 2020. Discriminative feature alignment: Improving transferability of unsupervised domain adaptation by gaussian-guided latent alignment. *arXiv preprint arXiv:2006.12770* .
- Wen, H., et al., 2017. Disrupted topological organization of structural networks revealed by probabilistic diffusion tractography in tourette syndrome children. *Human Brain Mapping* 38, 3988–4008.
- Wu, P.W., Lin, Y.J., Chang, C.H., Chang, E.Y., Liao, S.W., 2019. Relgan: Multi-domain image-to-image translation via relative attributes. *Proceedings of the IEEE International Conference on Computer Vision* , 5914–5922.
- Yang, J., Lu, J., Lee, S., Batra, D., Parikh, D., 2018. Graph r-cnn for scene graph generation. *Proceedings of the European conference on computer vision (ECCV)* , 670–685.
- Yang, J., Zhu, Q., Zhang, R., Huang, J., Zhang, D., 2020. Unified brain network with functional and structural data, in: *International Conference on Medical Image Computing and Computer-Assisted Intervention*, Springer. pp. 114–123.
- You, J., Liu, B., Ying, Z., Pande, V., Leskovec, J., 2018. Graph convolutional policy network for goal-directed molecular graph generation. *Advances in neural information processing systems* , 6410–6421.
- Zeng, G., Zheng, G., 2019. Hybrid generative adversarial networks for deep MR to CT synthesis using unpaired data. *International Conference on Medical Image Computing and Computer-Assisted Intervention* , 759–767.
- Zhang, L., Wang, L., Zhu, D., 2020a. Recovering brain structural connectivity from functional connectivity via multi-gcn based generative adversarial network, in: *International Conference on Medical Image Computing and Computer-Assisted Intervention*, Springer. pp. 53–61.

- Zhang, L., Wang, M., Liu, M., Zhang, D., 2020b. A survey on deep learning for neuroimaging-based brain disorder analysis. arXiv preprint arXiv:2005.04573 .
- Zhang, S., Tong, H., Xu, J., Maciejewski, R., 2019. Graph convolutional networks: a comprehensive review. Computational Social Networks 6, 11.
- Zhang, Z., Cui, P., Zhu, W., 2020c. Deep learning on graphs: A survey. IEEE Transactions on Knowledge and Data Engineering .
- Zhou, J., et al., 2018. Graph neural networks: A review of methods and applications. arXiv preprint arXiv:1812.08434 .
- Zhu, J.Y., Park, T., Isola, P., Efros, A.A., 2017. Unpaired image-to-image translation using cycle-consistent adversarial networks. Proceedings of the IEEE international conference on computer vision , 2223–2232.