Medical Image Analysis 68 (2021) 101902

Contents lists available at ScienceDirect

Medical Image Analysis

journal homepage: www.elsevier.com/locate/media

Brain graph synthesis by dual adversarial domain alignment and target graph prediction from a source graph

Alaa Bessadok[a,b], Mohamed Ali Mahjoub[b], Islem Rekik[a,c,1,*]

[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] School of Science and Engineering, Computing, University of Dundee, UK

ARTICLE INFO

Article history:
Received 24 January 2020
Revised 9 November 2020
Accepted 10 November 2020
Available online 16 November 2020

Keywords:
Brain graph prediction
Domain alignment
Generative adversarial learning
Geometric deep learning
Dual adversarial learning
Adversarial autoencoders

ABSTRACT

Developing predictive intelligence in neuroscience for learning how to generate multimodal medical data from a single modality can improve neurological disorder diagnosis with *minimal data acquisition resources*. Existing deep learning frameworks are mainly tailored for images, which might fail in handling geometric data (e.g., brain graphs). Specifically, predicting a target brain graph from a single source brain graph remains largely unexplored. Solving such problem is generally challenged with *domain fracture* caused by the difference in distribution between source and target domains. Besides, solving the prediction and domain fracture *independently* might not be optimal for both tasks. To address these challenges, we unprecedentedly propose a *Learning-guided Graph Dual Adversarial Domain Alignment (LG-DADA)* framework for predicting a target brain graph from a source brain graph. The proposed LG-DADA is grounded in three fundamental contributions: (1) a source data pre-clustering step using manifold learning to firstly handle source data heterogeneity and secondly circumvent mode collapse in generative adversarial learning, (2) a domain alignment of source domain to the target domain by adversarially learning their latent representations, and (3) a dual adversarial regularization that jointly learns a source embedding of training and testing brain graphs using two discriminators and predict the training target graphs. Results on morphological brain graphs synthesis showed that our method produces better prediction accuracy and visual quality as compared to other graph synthesis methods.

1. Introduction

One major objective of existing machine learning-based methods in medical imaging aims to alleviate the high costs of acquiring multiple medical scans as well as handling medical datasets with incomplete imaging modalities. For instance, a subject might have a magnetic resonance imaging (MRI) scan and lacks a positron emission tomography (PET) scan. However, feeding incomplete multimodal data into a learning-based framework for early disease diagnosis is challenged by missing multimodal medical images, which can provide a more holistic understanding of the underlying mechanisms of the target disease when available. To handle this problem, some of the existing works discarded samples with missing data. However, such techniques led to reducing the performance of the predictive model since it learns from a limited number of observations. Existing methods aiming to solve this problem can be categorized into machine learning based and deep learning based approaches. For instance, in the first category (Huynh et al., 2016) proposed a voxel estimation method that used structured random forest algorithm to predict a CT image from an MRI image. Another learning based work (Jog et al., 2013) proposed to synthesize a T2-weighted MRI data from a T1-weighted data using an ensemble of regression trees.

For the deep learning based category, Bano et al. (2018) designed a fully convolutional network (XmoNet) for cross-modality MR image synthesis. Li et al. (2014) used convolutional neural networks (CNN) to predict positron-emission tomography (PET) image of a specific sample from MRI image. In a follow-up work, Ben-Cohen et al. (2019) combined a fully convolutional network with a conditional Generative Adversarial Network (GAN) to predict PET from CT. A typical GAN (Goodfellow et al., 2014) consists of two neural networks: a *generator* trained to synthesize an output that approximates the real data distribution, and a *discriminator* trained to differentiate between the fake and real images. With its remark-

* Corresponding author at: BASIRA lab (http://basira-lab.com/).
E-mail address: irekik@itu.edu.tr (I. Rekik).
[1] GitHub code: https://github.com/basiralab/LG-DADA

https://doi.org/10.1016/j.media.2020.101902
1361-8415/

able synthesis potential, GAN has been used in a variety of medical imaging applications (Yi et al., 2018) including missing data imputation task. For instance, Olut et al. (2018) used GAN to generate missing Magnetic Resonance Angiography (MRA) from T1- and T2-weighted MRI images. However, most of the existing methods are only applied to Euclidian structured data such as MRI scans and electrocardiogram (ECG) signals (Cho et al., 2018). Hence, they might fail in handling non-Euclidian structured data or 'geometric data' types such as graphs and manifolds (Bronstein et al., 2017).

Recently, the nascent field of *geometric deep learning* (GDL) has cross-pollinated *network neuroscience*, where GDL architectures were trained on brain graphs or connectomes to diagnose neurological diseases. The brain connectome is a graph representation of biological activity across a set of anatomical regions of interest (ROIs) in the brain. For instance, Ktena et al. (2017) used Graph Convolution Network (GCN) (Kipf and Welling, 2016) to learn a similarity metric between two functional brain graphs extracted from resting-state fMRI (rs-fMRI) data for Autism Spectrum Disorder (ASD) diagnosis. Later on, Parisot et al. (2017) proposed to predict the disease state (healthy or affected) of a subject from a partially labeled graph using GCN. Nodes of the brain graph represent functional brain connectivities of a subject extracted from rs-fMRI images and edges represent the similarity between subjects using their brain connectomes and their phenotypic information (age and gender). Another recent work (Arslan et al., 2018) introduced a gender prediction framework (i.e., male or female) based on functional brain graphs extracted from rs-fMRI. Using GCN, they selected the most relevant ROIs for gender classification. While these frameworks presented promising results on brain graphs, they overlooked the problem of *'graph synthesis'*. Especially, predicting a target brain graph from a source graph where each is derived from different metrics (i.e, they have different statistical distributions) remains largely unexplored.

However, to make such a cross-domain prediction there is a need to handle the problem of *domain fracture* resulting in the difference in distribution between the source and target domains. Several works aimed to solve this problem by proposing a GAN-based framework. For instance, Pan et al. (2018b) used cycle-consistent GAN to predict PET images from MRI data. They used a bi-directional domain mapping (i.e., domain alignment) where they first mapped the MRI source domain to the PET target domain and then learned the reverse mapping. The synthesized PET data were used for early Alzheimer's Disease diagnosis. Additionally, Yang et al. (2018) assumes that cycle GAN does not have a constraint between the generated target image and its ground truth target image. So they added a structure-consistency loss to the original cycleGAN and adopted it to predict MRI data from CT data. However, these GAN-based methods focused mainly on synthesizing medical images rather than geometric data while many works demonstrated the ability of GAN in accurately learning from graphs. (Wang et al., 2018) recently introduced GraphGAN, a graph embedding method where graphs were projected into a low-dimensional space. As existing graph representation methods were rooted in generative or discriminative learning frameworks, this work consists of a GAN-based method that combines both classes. Moreover, Liu et al. (2017) demonstrated how GAN can learn topological features of any kind of graph. Specifically, they proposed a graph topology interpolator method to divide a graph into multiple subgraphs. Then, in order to better capture topological features, the subgraphs were fed to GAN.

Several neuroscientists have long suspected that abnormal mental behaviour shown in disordered subjects correlate to specific connectivity features of the brain. For instance, Mahjoub et al. (2018) proposed a brain graph-based representation named multiplex to distinguish between late MCI and AD subjects and detected biomarkers which are morphological connectivities that fingerprint the difference between both stages. More recently, Mhiri et al. (2020) proposed a high-resolution brain graph generation framework and discovered several discriminative functional connectivities in the produced brain graphs. We notice here that such results were not derived from a flat data representation (e.g, MRI), but from a deeper one named brain graph which is a wiring map of the neural connections in a human brain (Fornito et al., 2013). This is explicable because a connectional biomarker of a neurological disorder means that if an ROI is considered as the most affected brain region at a specific stage of that disorder the other ROI connected to it will also be affected. Hence, once equipped with the ability to construct the brain graphs, neuroscientists will be able to discover more biomarkers which help them develop more personalized treatments and advance the surgical interventions. However, to date they have lacked complete medical dataset necessary to fully investigate these hypotheses since the existing ones usually have missing modalities. Thus, to circumvent the need to acquire multiple brain modalities such as functional MRI or diffusion MRI for the purpose of extracting respectively a functional brain graph and a structural brain graph, a brain graph synthesis framework need to be developed. To fill this gap, we unprecedentedly propose a GAN-based framework for predicting target brain graphs from a single source graph. Typically, predicting brain connectomes has several applications in network neuroscience. First, different connectional aspects of the brain (i.e, source and target brain graphs) can provide complementary information of the whole brain which helps boost brain disease diagnosis as in Liu et al. (2016) and Wang et al. (2020). For instance, Liu et al. (2016) proposed a connectomic feature selection strategy to early identify the high-grade glioma (HGG) subjects with survival time over 650. The most reliable features of structural and functional brain graphs respectively derived from diffusion tensor imaging (DTI) and rs-fMRI were selected then fed into an SVM classifier to predict the disease outcome. Another recent paper (Wang et al., 2020) proposed learning-based framework for multi-class ASD classification using multi-modal brain data extracted from rs-fMRI. More specifically, a sparse representation classifier (SRC) (Wright et al., 2008) was leveraged to classify multi-view ASD connectivity features which are extracted from white matter and gray matter data. Second, the brain graph synthesis task helps understand the holistic connectional map of the brain. If the generated graphs are very reliable and biologically sound, one can use them to create integral connectional maps of the brain called connectional brain template (CBT) (Dhifallah et al., 2019). Mainly, it aims to produce a normalized connectional representation of a population using multi-view brain graphs. The generated CBTs from four connectomic datasets were shown to be effective in preserving the connectivity patterns of the population. Such concept was recently leveraged in Goktas et al. (2020) to predict the evolution trajectory of a brain graph where an adversarial conncetome encoder learned the CBT's embeddings which are passed on to a sample selection block for the target prediction task.

So far, we have identified only a single work on brain graph synthesis (Zhu and Rekik, 2018) which proposed a machine learning-based framework leveraging multi-kernel manifold learning (MKML) technique to predict multiple target brain graphs nested in different domains from a single source brain graph. This landmark work handled the domain shift problem by mapping each pair of source and target brain graphs onto a shared space where their distributions are aligned and domain shift is reduced using canonical correlation analysis (CCA). Although promising, the existing image synthesis frameworks based on deep learning and the CCA-based framework (Zhu and Rekik, 2018) are limited to regarding the domain shift between the source and target data and multimodal image prediction as *independent* tasks. More specifically, they solved both problems in a sequential manner: they first

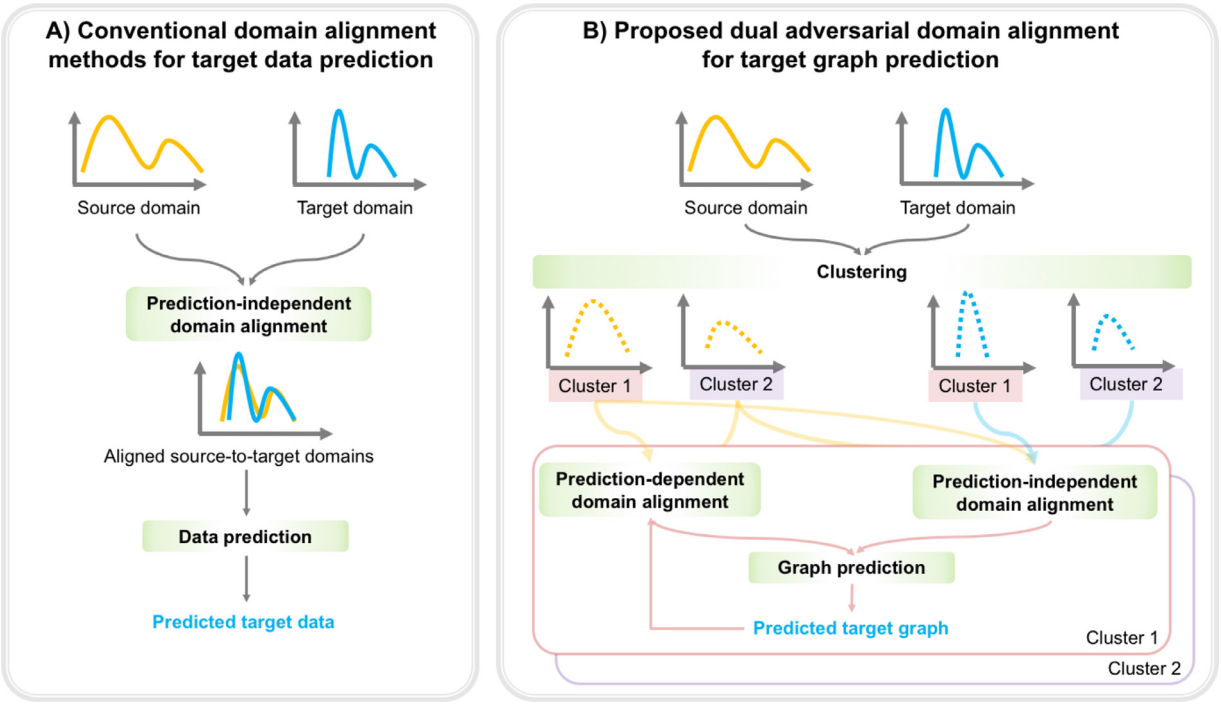

Fig. 1. Conventional domain alignment methods and proposed dual adversarial domain alignment for target graph prediction. (A) In this illustration we explicit the strategy adopted in several medical image synthesis works that is: aligning the source to the target data then predicting the target data. Unfortunately, such strategy is limited to solving the data prediction and domain alignment tasks separately. Moreover, most frameworks based on Adversarial Network (GAN) cannot handle the source data heterogeneity which causes the mode collapse problem. (B) To address these issues, we first propose to cluster the data into homogeneous groups which helps disentangle heterogeneous source data distributions. Next, we propose for each cluster two adversarial domain alignment models: (1) prediction-independent domain alignment where we simply move the source domain to the target domain, (2) prediction-dependent domain alignment where we synergistically integrate the target graph prediction in the domain alignment learning. Hence, with a single unified adversarial framework we enforce each of the domain alignment and the graph prediction tasks to boost the performance of each other while handling the data heterogeneity.

learn how to align the source data to the target data, second they predict the target data using the aligned source-to-target domains (Fig. 1A).

To address these limitations, we propose a *unified* Learning-guided Graph Dual Adversarial Domain Alignment (LG-DADA) architecture, which predicts a target brain graph from a source graph while aligning both domains. Specifically, we leverage the adversarially-regularized generative autoencoder (ARGA) proposed in Pan et al. (2018a) which extended the concept of autoencoder and GAN (Goodfellow et al., 2014) to graphs. ARGA comprises a generator G defined as a GCN (Kipf and Welling, 2016) and a discriminator D_{align} defined as a multilinear perceptron. *However, ARGA was originally designed for a graph embedding task and not for graph prediction and domain alignment.* In this work, we propose to extend it for jointly solving the domain shift problem and graph prediction. Although ARGA is a good starting point for solving our problem, it is an instance of GANs, which generally suffer from the *mode collapse* problem (Goodfellow, 2016). This issue occurs when training the generator G. Ideally, one would learn a *generalizable* generator, which is able to generate diverse target samples covering well the distribution of the target data domain. Still, in practice, the generator ends up producing graphs that approximate a few examples of target graphs (i.e., one mode), thereby identifying a single mode of the real data distribution. To circumvent this problem, we propose to cluster the source with heterogeneous distribution into different homogeneous clusters, where a cluster-specific generator is constrained to generating a specific mode of the target data distribution. Moreover, we aim to simultaneously bridge the distribution shift between source and target graph domains and synthesize the target graph by an alternative bidirectional learning where bridging the domain shift step improves the target graph prediction and vice versa in an iterative progressive manner (Fig. 1B). Fundamentally, our LG-DADA framework has four stages:

(1) *Feature extraction and clustering*. We represent both source and target brain graphs, each encoded in a matrix, by feature vectors. Next, to better learn the inherent statistical distribution of the source data to align with the target, we propose to cluster the source graphs into different homogeneous groups. Then, for each cluster we proceed to the next three steps for target graph prediction.
(2) *Adversarial domain alignment*. We propose to align the source domain to the target domain using training samples. This *prediction-independent* domain alignment is regularized using one discriminator D_{align} that maps the distribution of the source domain to the target domain.
(3) *Dual adversarial regularization*. We propose a *prediction-dependent* domain alignment where we learn a source embedding for training and testing samples by alternating between two discriminators: the first one D_{align} matches the distribution of the embedded source graphs with the distribution of the original source graphs, and the second one D_{pred} enforces the embedded source distribution to match the distribution of the predicted target graphs of training subjects.
(4) *Target brain graph prediction*. To predict the target graph we first learn a connectomic manifold of the source embedding using the training and testing subjects and the aligned source-to-target graph embedding using the training sub-

jects. To predict the target brain graph of a testing subject, we select the closest neighboring source graphs to the testing graph then average their corresponding target graphs.

2. Related works

2.1. Graph synthesis

So far, seminal geometric deep learning works have been successfully leveraged in a few undirected graph synthesis tasks. Some studies employed Recurrent Neural Network (RNN) to sequentially generate subgraphs consisting in a subset of nodes and their connectivities from the whole graph (Su et al., 2019; Liao et al., 2019). In the first work (Su et al., 2019), two RNNs were used to learn the nodes embeddings and another RNN was used to generate the edges linking the embedded nodes of the subgraph. A second work (Liao et al., 2019) assumed that the subgraph generation process of standard recurrent networks does not depend on the overall graph topology. So, they proposed to use a graph neural network along with an attention mechanism to link the generated subgraphs to each other while leveraging the graph topology. On the other hand, a few studies leveraged Graph AutoEncoder (GAE) for molecular graph synthesis. For example, Bresson and Laurent (2019) proposed two successive decoders to generate the molecule structure. The first one generates the formula of the molecule and passes it to the second decoder which creates the structure of the molecule in terms of nodes and their connectivities. More recently, Flam-Shepherd et al. (2020) proposed an autoencoder based on message passing neural network. The graph decoder iteratively used the node's neighbors information called message to generate the graph structure. However, despite their success in handling graph synthesis tasks, these frameworks are only designed to generate graphs from the same domain (e.g, a molecular graph from a molecular formula). Thus, they have limited generalizability to synthesizing target graphs from a totally different source domain, thereby overlooking the potential domain shift between source and target domains.

2.2. Domain alignment

To overcome the limitation of existing graph synthesis works, one can leverage the concept of domain alignment also known as domain adaptation (Wilson and Cook, 2020; Hoffman et al., 2018; Toldo et al., 2020). It mainly refers to the task of predicting labels of target samples, given a labeled source domain (e.g, training set) and an unlabeled target domain (e.g, testing set) where both domains have different statistical distributions (Redko et al., 2020; Wilson and Cook, 2020). This task was proposed to reduce the need for costly labeling the samples in the target domain. Domain alignment has benefited various applications in many practical scenarios, including but not limited to image-to-image translation (Zhu et al., 2017; Choi et al., 2018), and medical image segmentation and reconstruction (Shen and Gao, 2019; Zhou et al., 2019). Most of the existing works performed an adversarial domain alignment where a min-max loss function was used to match the source distribution to the target one. For example, CycleGAN proposed in Zhu et al. (2017) adopted a cycle consistency loss to map two non-overlapping domains such as mapping a horse to a zebra. However, the existing image-to-image translation works (i.e, including CycleGAN) are inefficient in learning the mapping among multiple domains using a single GAN model. Thus, StarGAN added to the input data a label vector representing the target domain to guide the distribution alignment (Choi et al., 2018). From a medical data analysis perspective, Shen and Gao (2019) proposed a self-supervised learning framework performing a multi-channel MRI alignment for the purpose of brain tumor segmentation task. Another adversarial domain alignment framework named Tomo-GAN was proposed in Zhou et al. (2019) to synthesize missing sinograms which refers to different viewing angles of Computed Tomography (CT) and used the resulting views to reconstruct CT images. Critically, the aforementioned frameworks were only designed for image alignment which cannot be generalized to the geometric data types including graphs. Additionally, to the best of our knowledge up to now no existing works have investigated the domain alignment task for brain graph prediction.

3. Methodology

In this section, we detail our joint graph prediction and domain alignment framework. We illustrate in Fig. 2 the four proposed steps: (1) extraction and clustering of source and target brain graphs, (2) alignment of the source to the target domain, (3) dual adversarial regularization of source graph embedding, and (4) prediction of target brain graph. For easy reference, we summarize the major mathematical notations in Table 1.

3.1. Brain graph feature extraction and clustering

We aim in the following step to align the source feature vectors to the target vectors using ARGA. However, training such a GAN-based framework might fail in aligning all source samples to all target samples due to the generative mode collapse issue (Arjovsky et al., 2017), which would eventually align training source graphs to a single target graph mode. To handle this problem, we first cluster training and testing source brain graphs $\mathbf{F}_S$ into homogeneous clusters with high inter-subject similarities using multiple kernel manifold learning (MKML) technique (Wang et al., 2017). We choose MKML for its three major advantages. First, it outperformed other clustering methods such as PCA (Jolliffe and Cadima, 2016) and t-SNE (Maaten and Hinton, 2008) when dealing with biological dataset (Wang et al., 2017). Second, by learning multiple kernels that efficiently fit the statistical distribution of the data it provides a better similarity matrix between brain graphs. Third, it uses graph diffusion to overcome the problem of learning weak similarities between samples. Therefore, MKML produces a similarity matrix of size $(n \times n)$ modeling the similarities between pairs of source graphs, with c diagonal blocks representing different clusters. Next, the obtained source similarity matrix is mapped into a lower dimensional space using t-SNE (Maaten and Hinton, 2008) resulting in a latent matrix of size $(n \times c)$. Once we obtain the latent matrix, k-means algorithm is used to cluster the subjects (Fig. 2A). To predict the target graph of a testing subject, we first pre-cluster both testing and training source brain graphs. Next, we identify the cluster c_{tst} including the testing subject and train LG-DADA using the set of training subjects $(n_{c_{tst}} - 1)$ in cluster c_{tst}. Last, we test the trained LG-DADA on the left-out testing sample.

3.2. Domain alignment of source and target domains

To predict a target brain graph from a source graph, we assume that n_s subjects similar to a given subject in the source domain should be also similar in the target domain. So we hypothesize that the target graph of a testing subject should be predicted by averaging the target graphs of the training subjects that share similar local neighborhoods across source and target domains. To do so, we propose to learn a target and a source connectomic manifolds that capture the relationship between subjects in the target and source domains, respectively. We further propose to learn the connectomic manifolds using the embeddings of graphs that result from ARGA model training using the original graphs and their learned similarity matrix. Hence, we can learn the target graph

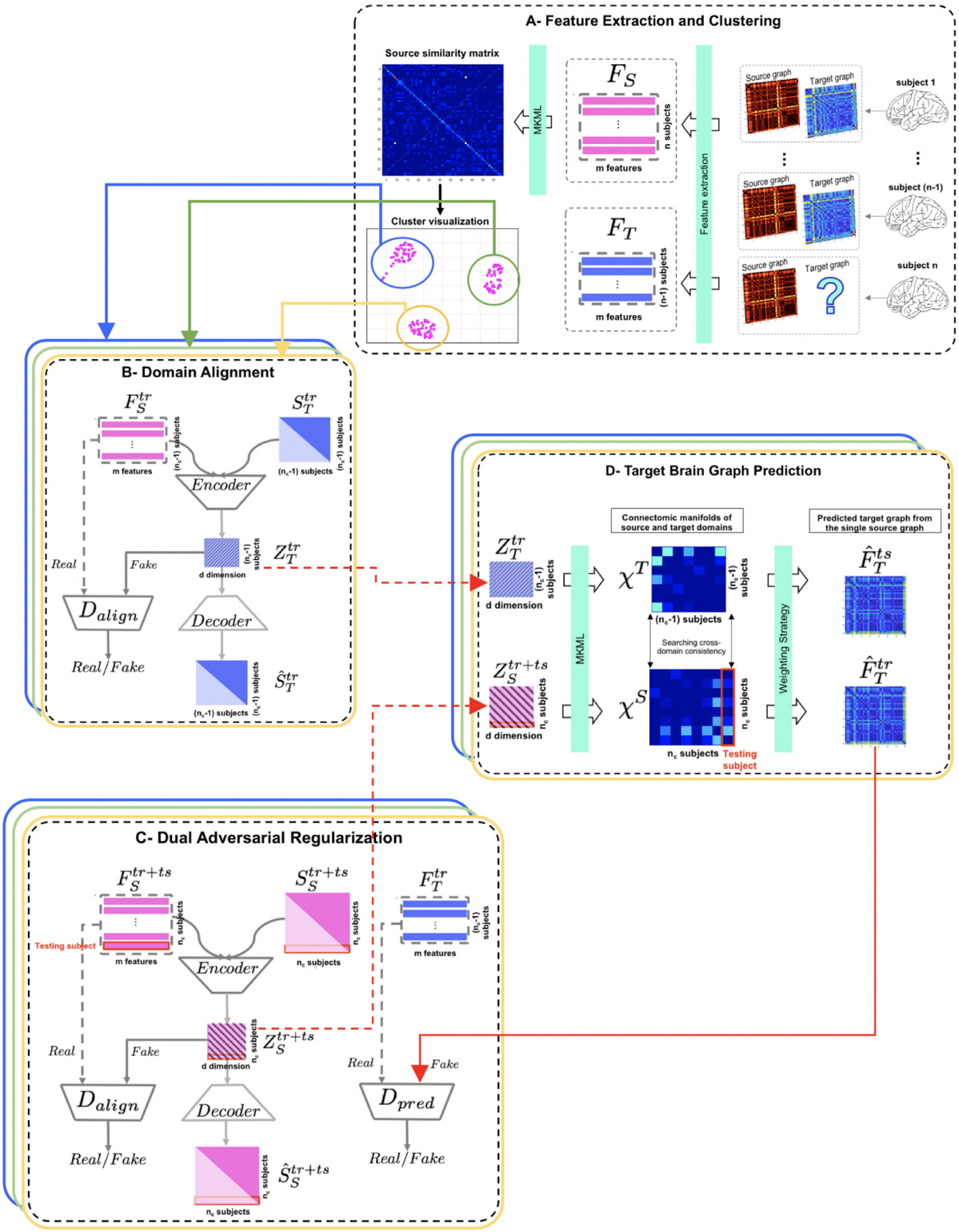

Fig. 2. Proposed framework of Learning-guided Graph Dual Adversarial Domain Alignment (LG-DADA) for target brain graph prediction from a single source graph. **(A) Feature extraction and clustering.** Extraction of feature vectors from source and target brain graphs for each training subject. In this illustration, we have left the n^{th} subject out for testing, hence it only has a source graph. Then, we use multiple kernel manifold learning to cluster the source brain graphs. **(B) Domain alignment.** For each cluster, we train one ARGA model to align the source domain to the target domain in the low-dimensional latent space and regularized by a discriminator D_{align}. **(C) Dual adversarial regularization.** We regularize the source embedding of training and testing subjects by alternating between two discriminators D_{align} and D_{pred}. **(D) Target graph prediction.** It consists of learning a source manifold, that nests the encoded graphs of the training and testing subjects, and a target manifold that nests the encoded graphs of the training subjects. To predict the final testing target graph, we select the training subjects most similar to the testing subject in both manifolds.

Table 1
Major mathematical notations used in this paper.

Mathematical notation	Definition
C	number of clusters
n_c	number of subjects in cluster c
n	number of subjects
m	number of features
d	number of features of the embedded graph
$\mathbf{F}_S^{tr+ts}$	stacked source feature vectors of (n_c) training and testing subjects in $\mathbb{R}^{n_c \times m}$
$\mathbf{F}_k^{tr}$	stacked k source or target feature vectors of $(n_c - 1)$ training subjects in $\mathbb{R}^{(n_c-1)\times m}$
$\hat{\mathbf{F}}_T^{tr}$	predicted target brain graphs of $(n_c - 1)$ training subjects in $\mathbb{R}^{(n_c-1)\times m}$
$\mathcal{S}_S^{tr+ts}$	adjacency matrix between n_c subjects using their source graphs in $\mathbb{R}^{n_c \times n_c}$
$\mathcal{S}_T^{tr}$	adjacency matrix between $(n_c - 1)$ subjects using their target graphs in $\mathbb{R}^{(n_c-1)\times(n_c-1)}$
$\hat{\mathcal{S}}_S^{tr+ts}$	reconstructed similarity matrix between n_c subjects using their source graphs in $\mathbb{R}^{n_c \times n_c}$
$\hat{\mathcal{S}}_T^{tr}$	reconstructed similarity matrix between $(n_c - 1)$ subjects using their target graphs in $\mathbb{R}^{(n_c-1)\times(n_c-1)}$
$\mathbf{Z}_T^{tr}$	embedding of the aligned source-to-target graphs using training subjects of cluster c in $\mathbb{R}^{(n_c-1)\times d}$
$\mathbf{Z}_S^{tr+ts}$	source graph embedding using training and testing subjects in $\mathbb{R}^{n_c \times d}$
χ^T	connectomic manifold quantifying the relationship between the training subjects using their aligned source-to-target embedded graphs in $\mathbb{R}^{(n_c-1)\times(n_c-1)}$
χ^S	connectomic manifold quantifying the relationship between the training and testing subjects using their source embedded graphs in $\mathbb{R}^{(n_c)\times(n_c)}$
$\mathbf{W}^{(l)}$	weight matrix used as filter to learn the graph convolution encoder in $\mathbb{R}^{n_c \times l}$ or $\mathbb{R}^{(n_c-1)\times l}$ where l is the number of neurons in the activated layer
$G_{DA}(\mathbf{F}_S^{tr}, \mathbf{S}_T^{tr})$	generator used for domain alignment taking as input the training source graphs $\mathbf{F}_S^{tr}$ and the adjacency matrix computed using the target graphs $\mathbf{S}_T^{tr}$
$G_S(\mathbf{F}_S^{tr+ts}, \mathbf{S}_S^{tr+ts})$	generator used for source embedding taking as input the source graphs of training and testing subjects $\mathbf{F}_S^{tr+ts}$ and their adjacency matrix computed using the source graphs $\mathbf{S}_S^{tr+ts}$
D_{align}	discriminator used for domain alignment taking as inputs the real data distribution$\mathbf{F}_S^{tr}$ or $\mathbf{F}_S^{tr+ts}$ and the embedded graphs $\mathbf{Z}_T^{tr}$ or $\mathbf{Z}_S^{tr+ts}$
D_{pred}	discriminator used for the dual adversarial regularization of the training and testing source graph embeddings taking as inputs $\hat{\mathbf{F}}_T^{tr}$ and $\mathbf{F}_T^{tr}$

embedding using target graphs and their learned similarity matrix. However, knowing that source and target domains have different statistical distributions there is a need to handle the fracture existing between both domains. Thus, we contribute in learning the source-to-target embeddings using the source graphs of the training subjects and the similarity matrix learned using their target graphs.

We aim to use a GCN as an encoder to learn the latent representation (i.e., embedding) of the brain graphs (Fig. 2B). Basically, our generator $G(\mathbf{F}, \mathbf{S})$ takes as input the training source graphs $\mathbf{F}_S^{tr}$ and their adjacency matrix $\mathbf{S}_T^{tr}$ that encodes the similarities between training subjects using their target graphs. We further propose to learn the graph adjacency matrix using MKML algorithm (Wang et al., 2017). Instead of using one predefined kernel (e.g., Euclidian distance), this algorithm adopted multiple kernels to learn the similarity between data points with high dimensionality and heterogeneous distribution. Such data might not be well represented with ordinary distance metric (e.g., Euclidean distance) (Wang et al., 2017) that might fail to capture heterogeneous data distribution. Our GCN encoder is constructed with two layers defined as follows:

$$\mathbf{Z}^{(1)} = f_{ReLU}(\mathbf{F}, \mathbf{S}|\mathbf{W}^{(0)}); \tag{1}$$

$$\mathbf{Z}^{(2)} = f_{linear}(\mathbf{Z}^{(1)}, \mathbf{S}|\mathbf{W}^{(1)}), \tag{2}$$

where $\mathbf{Z}^{(1)}$ and $\mathbf{Z}^{(2)}$ are the results of computing the first and the second layers, respectively. $\mathbf{Z}^{(2)}$ represents the desired alignment of source graphs to the target graphs. $\mathbf{W}^{(l)}$ is a weight matrix used as a filter to learn the convolution of the layers l in the GCN. Rectified Linear Unit, $ReLU(.)$ is the activation function of the first layer and a linear function is used for the second layer. As in GCN (Kipf and Welling, 2016), the graph convolution function $f_{(.)}$ is defined as follows:

$$f_\phi(\mathbf{F}^{(l)}, \mathbf{S}|\mathbf{W}^{(l)}) = \phi(\widetilde{\mathbf{D}}^{-\frac{1}{2}}\widetilde{\mathbf{S}}\widetilde{\mathbf{D}}^{-\frac{1}{2}}\mathbf{F}^{(l)}\mathbf{W}^{(l)}). \tag{3}$$

ϕ is the activation function for a specific layer (l), $\widetilde{\mathbf{S}} = \mathbf{S} + \mathbf{I}$ where $\mathbf{I}$ is the identity matrix, and $\widetilde{\mathbf{D}}_{ii} = \sum_j \widetilde{\mathbf{S}}_{ij}$ is a diagonal matrix. As suggested in Pan et al. (2018a) we propose to decode the embedded graph $\mathbf{Z}$ by reconstructing the target adjacency matrix $\hat{\mathbf{S}}$. Specifically, we measure it by computing the sigmoid function of the dot product of embedded graphs $\mathbf{z}_j$ of the subject (i.e., node) j and the transposed embedded graphs $\mathbf{z}_i^\top$ of the subject i:

$$Dec(\hat{\mathbf{S}}|\mathbf{Z}) = \frac{1}{1 + e^{-(\mathbf{z}_i^\top \cdot \mathbf{z}_j)}}. \tag{4}$$

The goal of training this first autoencoder $ARGA_T^{tr}$ is to minimize its reconstruction error, so that its overall optimization energy is written as follows:

$$\mathcal{L} = \mathbf{E}_{\mathbf{G}(\mathbf{F},\mathbf{S})}[\log Dec(\hat{\mathbf{S}}|\mathbf{Z})]. \tag{5}$$

The key idea in this step is conditioning the embedded graphs by the prior distribution of the source graphs. This is modeled by an adversarial regularizer D_{align} that enforces the latent representation to match the prior distribution of the source domain. This discriminator is a multi-layer perceptron, considered as a binary classifier, tries to minimize the error in discriminating between real data distribution (source graphs) and fake one generated from our encoder. Hence, we formulate the cost function of the *prediction-independent* alignment of source and target domains as follows:

$$\min_{G_{DA}} \max_{D_{align}} \mathbf{E}_{p_{(real)}}[\log D_{align}(\mathbf{F}_S^{tr})] + \mathbf{E}_{p_{(fake)}}[\log(1 - D_{align}(\mathbf{Z}_T^{tr}))], \tag{6}$$

where $\mathbf{E}$ is the cross-entropy cost, $\mathbf{Z}_T^{tr}$ is the embedding of the aligned source-to-target graphs generated from our graph encoder $G_{DA}(\mathbf{F}_S^{tr}, \mathbf{S}_T^{tr})$ and D_{align} is the binary classifier with maximum log likelihood objective.

3.3. Dual adversarial regularization for source graph latent representation

To predict the target graphs of a testing subject, we propose to select its nearest training neighbors in the *embedded* source domain (Fig. 2C). Then, we average their corresponding *aligned embedded* target graphs which represent the predicted target graphs. To do so, we learn a second ARGA for source graph latent representation of training and testing subjects $ARGA_S^{tr+ts}$. Specifically,

we stack the source feature vector of a testing subject below those of the training samples used in the first $ARGA_T^{tr}$, then we generate their latent representation using the generator $G_S(\mathbf{F}_S^{tr+ts}, \mathbf{S}_S^{tr+ts})$. Although regarded as an efficient source graph embedding model, $ARGA_S^{tr+ts}$ suffers from one major limitation: it does not align the source domain to the target domain as $ARGA_T^{tr}$. It basically solves the domain shift and brain graph prediction problems separately. To iron out such weakness, we propose a *prediction-dependent* domain alignment where we simultaneously integrate the target graph prediction and the domain alignment tasks into a single adversarial learning model, where each tasks boosts the performance of the other task in a progressive manner. Specifically, we aim to regularize the source embedding of training and testing subject by alternating two discriminators: *at each epoch, we activate just one discriminator*:

(1) the discriminator D_{align} matches the distribution of the source embedding $\mathbf{Z}_S^{tr+ts}$ generated from our encoder $G_S(\mathbf{F}_S^{tr+ts}, \mathbf{S}_S^{tr+ts})$ with the training and testing source graphs (Fig. 2C). And the resulting embedded source domain $\mathbf{Z}_S^{tr+ts}$ is then fed to the block of target graph prediction illustrated in (Fig. 2D) for target graphs prediction of the testing subject. Next, we predict the target graphs corresponding to each source graph in the training set using nested leave-one-out cross-validation. We fed the predicted target brain graphs to a second discriminator D_{pred} that will regularize, in the next epoch, the source graph embedding of the training and testing subjects.

(2) the discriminator D_{pred} regularizes $ARGA_S^{tr+ts}$ by enforcing the embedded source distribution to match the distribution of the predicted training target graphs. Specifically, it distinguishes between the real target graph of the training samples $\mathbf{F}_T^{tr}$ and their predicted target graphs $\hat{\mathbf{F}}_T^{tr}$. So we define the energy function of $ARGA_S^{tr+ts}$ when training both discriminator as follows:

$$\mathcal{L} = \begin{cases} \min_{G_S} \mathbf{E}_{p_{(fake)}}[\log(1 - D_{align}(\mathbf{Z}_S^{tr+ts}))] \\ \max_{D_{align}} \mathbf{E}_{p_{(real)}}[\log D_{align}(\mathbf{F}_S^{tr+ts})] \\ \quad + \mathbf{E}_{p_{(fake)}}[\log(1 - D_{align}(\mathbf{Z}_S^{tr+ts}))] \\ \max_{D_{pred}} \mathbf{E}_{p_{(real)}}[\log D_{pred}(\mathbf{F}_T^{tr})] \\ \quad + \mathbf{E}_{p_{(fake)}}[\log(1 - D_{pred}(\hat{\mathbf{F}}_T^{tr}))]). \end{cases} \tag{7}$$

Accordingly, the predicted target graphs boost the alignment of source domain to the target domain and simultaneously the aligned source graphs improve the prediction of the target graph. Since both tasks rely on each other, we are jointly solving both problems within a unified adversarial framework.

3.4. Target graphs prediction for the testing subject

To predict the target graph of the testing subject belonging to the cluster c, we assume that when a training subject is similar to the testing subject in the source domain, they are also similar in the target domain. So we aim in this step to find the most similar training subjects to the testing subject using their learned source embeddings. We further propose to enforce a local consistency in source and target neighborhoods for the selected training source neighbors. Mainly, we select source training graphs that have a large overlap in nearest neighbors across the embedded source and target domains. A selected training source sample becomes more reliable with larger local neighborhood overlaps across both domains. To do so, we learn a connectomic manifold (MKML) (Wang et al., 2017) χ^S that quantifies the relationship between training and testing subjects using their aligned source and target embeddings, and learn a second connectomic manifol χ^T using only the training aligned source-to-target embeddings (Fig. 2–D).

Next, we first identify in the source domain the top K-closest training subjects to the testing subject by sorting the learned manifold χ^S. Second, we find for each of the K selected training samples its nb nearest neighbors in both manifolds χ^S and χ^T. Third, we store the nb samples in $\mathbf{L}_S$ and $\mathbf{L}_T$ lists in order to assign a weighted similarity score $w(k)$ for each training subject k. Last, we compute their overlap using the following formula:

$$\mathbf{w}(k) = \exp\left(\left(\frac{\mathbf{L}_S \cap \mathbf{L}_T}{nb}\right) \times \chi S(ts, k)\right). \tag{8}$$

Finally, we average the target graphs of the selected nb neighbors with the highest w scores in order to estimate the target graph of the testing subject. We detail the steps of our proposed LG-DADA framework in the Algorithm 1.

4. Results and discussion

4.1. Connectomic dataset

A set of 150 structural T1-w MRI data for 75 ASD and 75 NC subjects extracted from Autism Brain Imaging Data Exchange (ABIDE[2]) public dataset was used. Both hemispheres of each subject were reconstructed using FreeSurfer (Fischl, 2012). Then using Desikan–Killiany Atlas we parcellated the hemispheres into 35 anatomical regions. Each subject has three morphological brain graphs (MBG) using the following cortical measurements in each hemisphere: mean sulcal depth, maximum principal curvature and average curvature. Each MBG is encoded in a symmetric matrix that quantifies the morphological similarity between pairs of ROIs. Morphological brain networks have been recently introduced to investigate brain morphology on a *connectional level* in health (Dhifallah et al., 2019; Nebli and Rekik, 2020) and disease (Alzheimer's Disease Neuroimaging Initiative et al., 2018; Banka and Rekik, 2019; Lisowska et al., 2017; Mahjoub et al., 2018; Soussia and Rekik, 2018). Mainly, we used the morphological brain graphs as a proof-of-concept for predicting a target graph from a source graph where both source and target domains have different statistical distributions.

4.2. Parameter setting

We construct our encoder with two layers, the hidden and the embedding layers, each comprising 32 neurons. We construct both discriminators D_{align} and D_{pred} with two layers composed of 64 and 32 neurons, respectively. As in ARGA (Pan et al., 2018a), we fix both learning rates of the encoder and discriminators to 0.001. For MKML parameters (Wang et al., 2017), we set the number of kernels to 10 and we fix the number of clusters to $c = 3$. We note that we evaluated our framework when varying the number of clusters from $c = 2$ to $c = 4$. The best performance was achieved for $c = 3$ and the improvement was negligible when increasing the number of clusters to $c = 4$. For target graph prediction, we set the number of source neighbors to 5. The same parameters are used for all comparison methods.

4.3. Comparison methods

We compared our LG-DADA method with two state-of-the-art methods that do not use a pre-clustering step:

- **ADA:** We use for this method the original ARGA (Pan et al., 2018a) for adversarial domain alignment (ADA), where we move the source domain to the target one. For the regularization we use only one discriminator D_{align} as in Pan et al. (2018a) and for the input target adjacency matrix estimation we use the MKML algorithm (Wang et al., 2017).

[2] http://fcon_1000.projects.nitrc.org/indi/abide/.

Algorithm 1 Learning-guided graph dual adversarial domain alignment for target graph prediction.

INPUTS: Set of source brain graphs: $\{\mathbf{F}_S^1, \ldots, \mathbf{F}_S^i, \ldots \mathbf{F}_S^n\}$;Set of target brain graphs: $\{\mathbf{F}_T^1, \ldots, \mathbf{F}_T^i, \ldots \mathbf{F}_T^n\}$;$C$: number of clusters;$T$: number of epoch;$K$: number of steps for iterating both discriminators;CV : number of folds in cross-validation strategy
Cluster the source brain graphs $\mathbf{F}_S$ by learning a similarity matrix using MKML algorithm:
for c=1,...,C **do**
  Use leave-one-out cross-validation to generate training and testing subject from source and target graphs of the cluster c.
  Learn adjacency matrices $\mathbf{S}$ using MKML algorithm:(1) $\mathbf{S}_T^{tr}$ for target graph of training samples $\mathbf{F}_T^{tr}$(2) $\mathbf{S}_S^{tr+ts}$ for source graph of training and testing samples $\mathbf{F}_S^{tr+ts}$
  Domain alignment and target graph embedding for training subjects using $ARGA_T^{tr}$:$\mathbf{Z}^{(1)} = f_{ReLU}(\mathbf{F}_S^{tr}, \mathbf{S}_T^{tr} \mid \mathbf{W}^{(0)})$;$\mathbf{Z}_T^{tr} = f_{linear}(\mathbf{Z}^{(1)}, \mathbf{S}_T^{tr} \mid \mathbf{W}^{(1)})$
  for iterator = 1,2,3, ..., T **do**
    for k = 1,2, ..., K **do**
      Sample m entities { $\mathbf{z}^{(1)}$,...,$\mathbf{z}^{(m)}$} from latent matrix $\mathbf{Z}_T^{tr}$
      Sample m entities { $\mathbf{f}^{(1)}$,...,$\mathbf{f}^{(m)}$} from the prior distribution $\mathbf{F}_S^{tr}$
      Update the discriminator D_{align} with its stochastic gradient: $\nabla \frac{1}{m} \sum_{i=1}^{m} [\log D_{align}(\mathbf{f^i}) + \log(1 - D_{align}(\mathbf{z^i}))]$
    end for
    Update the graph autoencoder with its stochastic gradient by Eq.(5).
  end for
  Dual adversarial regularization and source graph latent representation of training and testing subjects $ARGA_S^{tr+ts}$:$\mathbf{Z}^{(1)} = f_{ReLU}(\mathbf{F}_S^{tr+ts}, \mathbf{S}_S^{tr+ts} | \mathbf{W}^{(0)})$;$\mathbf{Z}_S^{tr+ts} = f_{linear}(\mathbf{Z}^{(1)}, \mathbf{S}_S^{tr+ts} | \mathbf{W}^{(1)})$
  for iterator = 1,2,3, ..., T **do**
    if iterator = odd **then** Activate D_{align}
      for k = 1,2, ..., K **do**
        Sample m entities { $\mathbf{z}^{(1)}$,...,$\mathbf{z}^{(m)}$} from latent matrix $\mathbf{Z}_S^{tr+ts}$
        Sample m entities { $\mathbf{f}^{(1)}$,...,$\mathbf{f}^{(m)}$} from the prior distribution $\mathbf{F}_S^{tr+ts}$
        Update the discriminator D_{align} with its stochastic gradient: $\nabla \frac{1}{m} \sum_{i=1}^{m} [\log D_{align}(\mathbf{f^i}) + \log(1 - D_{align}(\mathbf{z^i}))]$
      end for
      Update the graph autoencoder with its stochastic gradient by Eq.(5).
      Target graph prediction for the training subjects:(1) Exclude testing subject from $\mathbf{Z}_S^{tr+ts}$ to get $\mathbf{Z}_S^{tr}$(2) Use leave-one-out cross-validation to generate new testing subject from $\mathbf{Z}_S^{tr}$ and also from $\mathbf{Z}_T^{tr}$
      for i = 1, ..., **CV** **do**
        Generate the new nested source and target graph embeddings: (1) stack the source graph embeddings of training and the new testing subject: $\bar{\mathbf{Z}}_S^{tr+ts}$ (2) exclude the testing subject from the new target graph embedding: $\bar{\mathbf{Z}}_T^{tr}$
        Learn a connectomic manifold using MKML algorithm for each of the new generated source and target graph embeddings: $\bar{\chi}_S^{tr+ts}$ and, $\bar{\chi}_T^{tr}$
        Predict the target graph $\hat{\mathbf{F}}_T^i$ of the testing subject i using the weighting strategy defined by Eq.(8).
      end for
      Stack vertically all the predicted target graphs of the training subjects: $\hat{\mathbf{F}}_T^{tr}$.
    else Activate D_{pred}
      for k = 1,2, ..., K **do**
        Sample m entities { $\hat{\mathbf{F}}^{(1)}$,...,$\hat{\mathbf{F}}^{(m)}$} from the predicted target graph $\hat{\mathbf{F}}_T^{tr}$
        Sample m entities { $\mathbf{f}_T^{(1)}$,...,$\mathbf{f}_T^{(m)}$} from the prior distribution of the real target graph $\mathbf{F}_T^{tr}$
        Update the discriminator D_{pred} with its stochastic gradient: $\nabla \frac{1}{m} \sum_{i=1}^{m} [\log D_{pred}(\mathbf{f}_T^i) + \log(1 - D_{pred}(\hat{\mathbf{F}}^i))]$
      end for
      Update the graph autoencoder with its stochastic gradient by Eq.(5).
    end if
  end for
  Predict the missing target graph of the testing subject:(1) Learn a connectomic manifold using MKML algorithm for each of the outputs of step (6) and (15), $\mathbf{Z}_T^{tr}$ and $\mathbf{Z}_S^{tr+ts}$ respectively, to get χ^T and χ^S(2) Predict the target graph of the testing subject using the weighting strategy defined by Eq.(8).
  OUTPUT: Return the predicted brain graph of the testing subject $\mathbf{F}_T^{ts}$.
end for

- **CCA-DA:** We compared LG-DADA with the proposed network prediction method in Zhu and Rekik (2018), where both source and target domains are mapped into a new share space using CCA algorithm (Hardoon et al., 2004). Then, a cross-domain weighting strategy is adopted to predict the target brain graph of the testing subjects.

4.4. Evaluation using cross-validation

Following the graph pre-clustering step, we evaluate our framework using leave-one-out cross-validation (LOO-CV) on each cluster and report the average prediction results. We also use LOO-CV to evaluate the benchmark methods (Pan et al., 2018a; Zhu and Rekik, 2018).

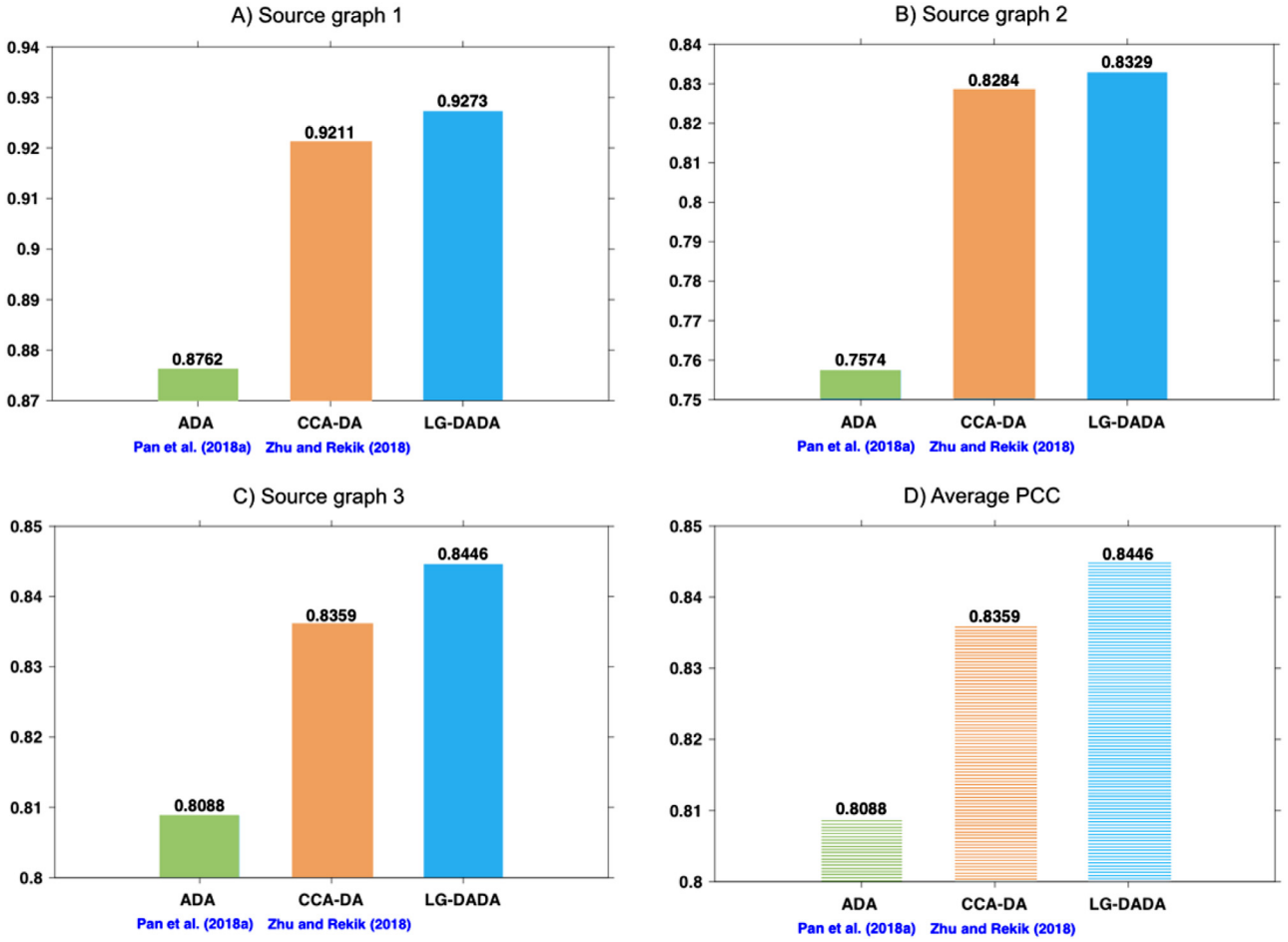

Fig. 3. Comparison of LG-DADA against state-of-the-art methods using pearson correlation coefficient (PCC). Source graph 1: the maximum principal curvature. Source graph 2: mean sulcal depth. Source graph 3: average curvature. ADA: the adversarial domain alignment using ARGA (Pan et al., 2018a) with a learned similarity matrix computed using MKML algorithm. CCA-DA: the multi-target prediction method proposed in Zhu and Rekik (2018) where CCA is used for domain alignment. LG-DADA: the proposed dual adversarial regularization for source-to-target domain alignment using MKML for adjacency matrix estimation and using weighting strategy for training samples selection.

4.5. Evaluation metrics

To evaluate our framework, we used Pearson correlation coefficient (PCC) (Egghe and Leydesdorff, 2009) and mean absolute error (MAE) (Lin et al., 1990). PCC measures the strength of the relationship between $\mathbf{F}_T^i$ the ground truth graph of a representative subject i and the predicted target graph $\hat{\mathbf{F}}_T^i$. We propose first to compute the PCC for each representative subject in the dataset. Then, we consider the average of all resulting PCCs as the final measure to evaluate our framework. On the other hand, the MAE represents the prediction error measured using the average absolute difference between the ground truth graphs and their corresponding predicted graphs. Thus, a small MAE means a better performance. Similar to PCC, we consider the average of MAEs across all subjects as the final measure to evaluate our framework.

4.6. Results and benchmarking

In this work, we proposed LG-DADA, a geometric deep learning framework which jointly performs a domain alignment of source and target brain graphs and predict target brain graph from a single source brain graph. First, following a pre-clustering step to circumvent the issue of generative mode collapse where a very limited number of modes are generated, LG-DADA learns for each cluster how to move the source domain to a target domain by embedding source brain graphs of training samples and their learned adjacency matrices using their target brain graphs. Second, it performs a source graph embedding of training and testing subjects in order to search the most similar training samples to the testing subject in the source domain. Third, it learns a connectomic manifold for each of these resulting latent representations of the source and target domains in order to search for a local shared neighborhood in both domains. Once the shared neighbors across domains are identified, the target graph of the testing subject is predicted by averaging the target graphs of these selected neighbors.

We conducted three different experiments, each taking one of the three morphological brain graphs as a source graph and the two remaining ones as target graphs to predict. We display in Fig. 3 the average PCC for each experiment (e.g., using source graph 1 to predict target graphs 1 and 2). Our method achieved the best prediction performance across all experiments when varying the source and target graphs, demonstrating its adaptiveness to diverse combinations of source and target domains. This also shows that LG-DADA can handle different profiles of domain shifts. Specifically, our method produced the highest PCC result when considering the maximum principal curvature as the source domain (Fig. 3A). As illustrated in (Fig. 3D), LG-DADA has a better overall average PCC across different experiments than both ADA (Pan et al., 2018a) and CCA-DA methods (Zhu and Rekik, 2018). We note that despite the significant outperformance of CCA-based domain alignment (Zhu and Rekik, 2018) in comparison with other baseline techniques, our LG-DADA *consistently* outperformed this

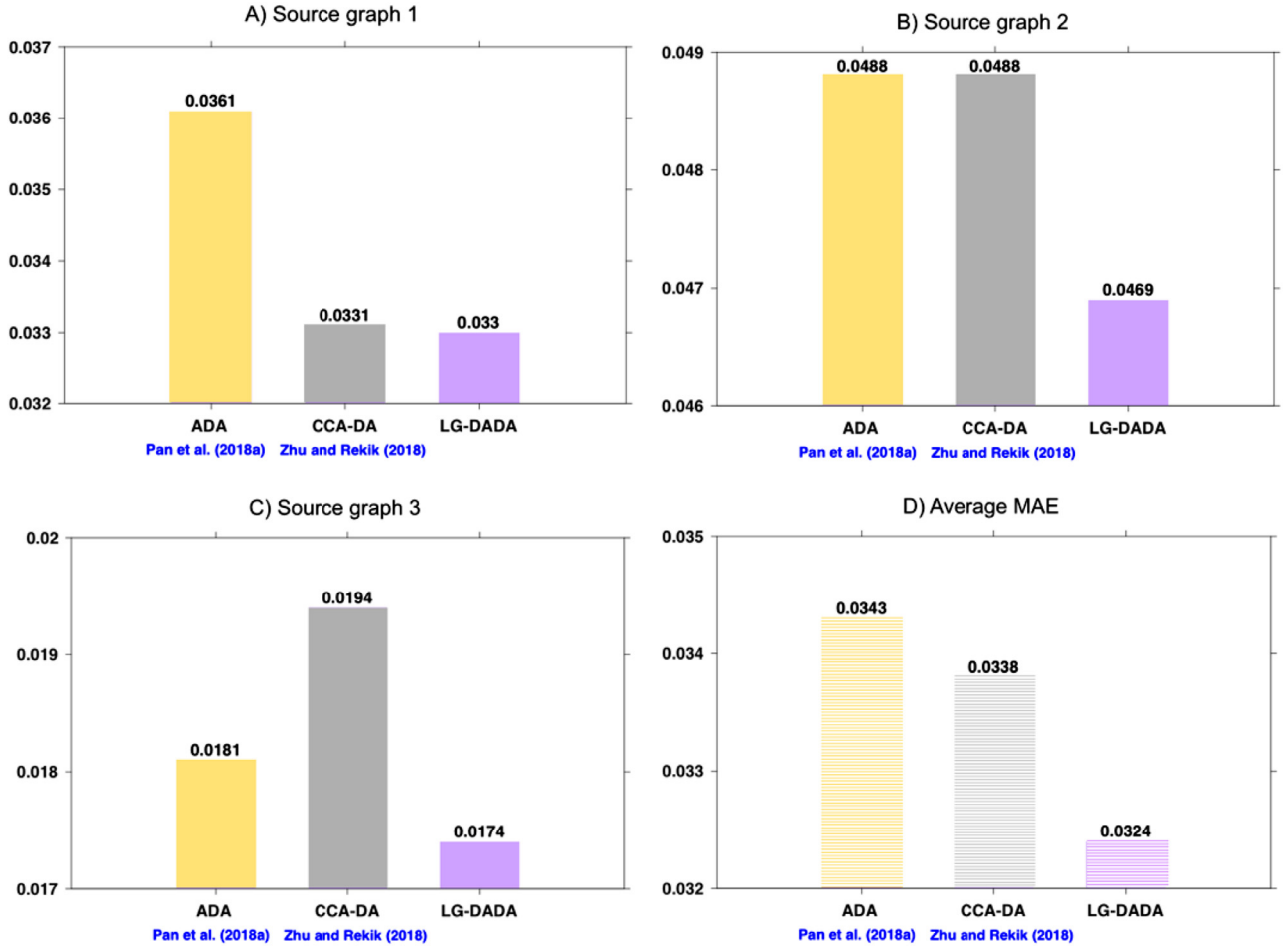

Fig. 4. Comparison of LG-DADA against state-of-the-art methods using mean absolute error (MAE). Source graph 1: the maximum principal curvature. Source graph 2: mean sulcal depth. Source graph 3: average curvature. ADA: the adversarial domain alignment using ARGA (Pan et al., 2018a) with a learned similarity matrix computed using MKML algorithm. CCA-DA: the multi-target prediction method proposed in Zhu and Rekik (2018) where CCA is used for domain alignment. LG-DADA: the proposed dual adversarial regularization for source-to-target domain alignment using MKML for adjacency matrix estimation and using weighting strategy for training samples selection.

method across all experiments using both evaluation metrics. This higher performance can be explained by the fact that our LG-DADA can solve the data fracture issue and the multi-view graph prediction task simultaneously while CCA-DA framework solved each of these problems independently. This points to the advantage of the synergy existing between the domain alignment and the graph prediction tasks modeled by our dual regularization step that efficiently learns the source graph embedding $\mathbf{Z}_S^{tr+ts}$ while performing an alignment of the source graphs to the predicted training target graphs. Moreover, CCA performs a domain alignment of source to target domain without learning their optimized inherent representations for the target prediction tasks (i.e., supervised by the target prediction task). Hence, leveraging ARGA to learn the latent representation of homogeneous brain graphs allows a better alignment of the source and target domains.

Fig. 4 A–C show the lowest MAE produced by LG-DADA when considering different source graphs. In particular, it outperformed the ADA method that trained an ARGA using a single discriminator to regularize both domain alignment and source graph embedding. Such good results demonstrate the advantage of our proposed dual adversarial regularization for source graph embedding in addition to the pre-clustering step for disentangling heterogeneous distributions and avoid GAN mode collapse (Goodfellow et al., 2014; Goodfellow, 2016; Arjovsky et al., 2017). This can be explained by the fact that the clustering step allows to transform the heterogeneous data into homogeneous distributions each presents a single mode of the data. Hence a GAN model can easily trained on homogeneous data. Fig. 5 displays the residual network (i.e., the absolute difference between the predicted and ground truth networks) for a randomly selected subject. Notably, LG-DADA produced a reduced residual error when predicting target graph 1 (derived from the maximum principle curvature) from the source graph (derived from sulcal depth) in comparison with benchmark methods. As for predicting target graph 2 (derived from the average curvature), LG-DADA outperformed ADA, however it produced a slightly higher residual than CCA-DA. We included this example in our results to show that while the *average* results of LG-DADA were remarkable and outperform comparison methods, our method might lag behind for a few particular subjects.

4.7. Limitations and recommendations for future work

Although our method produced the best results in predicting different target brain graphs from a source graph, it has a few limitations. *First*, we predicted each target graph independently. We aim in the future to jointly predict all target graphs from the source graph. Specifically, we will improve the domain alignment step by leveraging a recent GAN-based method (Pei et al., 2018) that is able to align different data distributions using multiple adversarial regularizers. However, this method only works

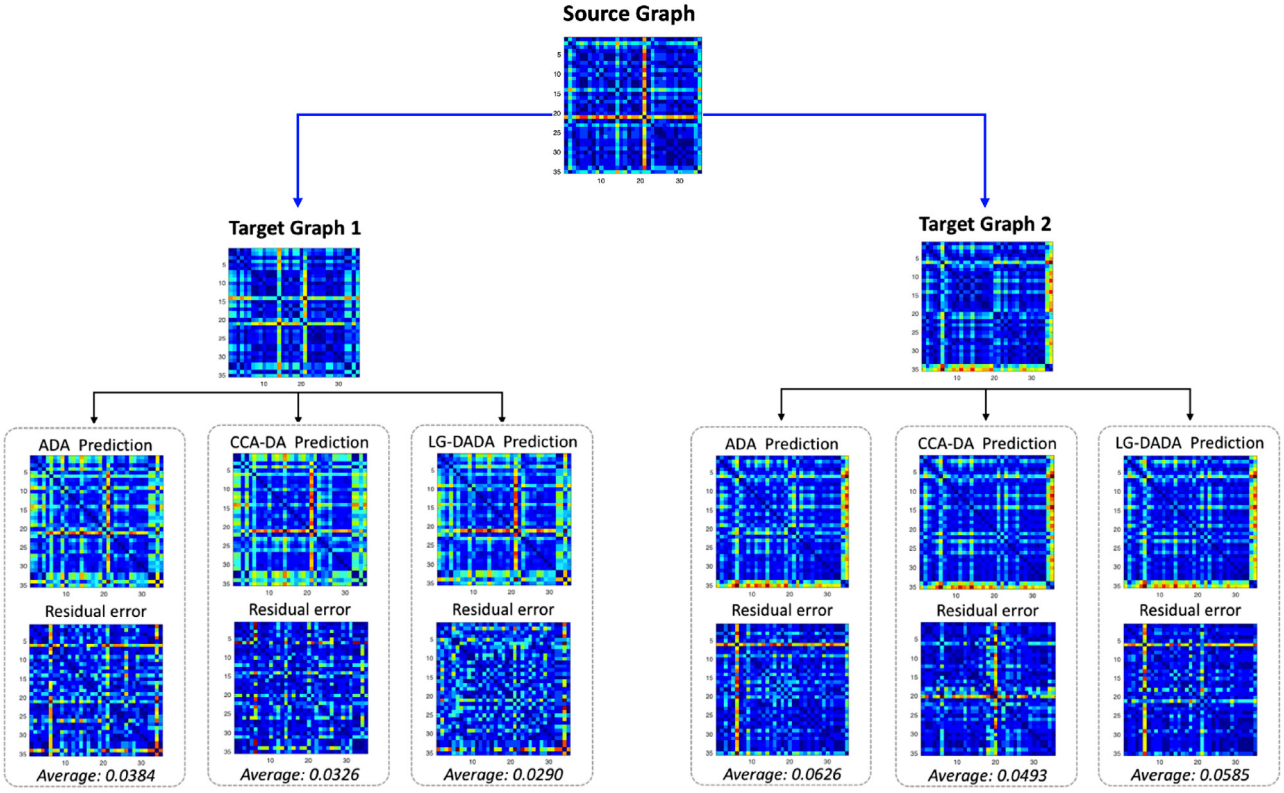

Fig. 5. Visual comparison between the original and the predicted target graph of a representative testing subject using three different methods: ADA (Pan et al., 2018a), CCA-DA (Zhu and Rekik, 2018) and LG-DADA. We display the residual error computed using mean absolute error metric between the original brain graph and the predicted graph. Source brain graph is mean sulcal depth. Target brain graphs are the maximum principal curvature (Target graph 1) and average curvature (Target graph 2).

on images so we aim to adopt it to geometric data. *Second*, in our LG-DADA framework, we used the standard autoencoder. However, as demonstrated in Pan et al. (2018a), we can use the variational autoencoder (Pu et al., 2016) which added some improvement to autoencoder model. Particularly, one can use variational GCN which outperformed the standard GCN in very recent study (Tiao et al., 2019). *Third*, we can increase the number of neurons of the encoder and the dual discriminators, as recommended by Pan et al. (2018a). To effectively define the number of layers and the number of neurones in each layer we can leverage a framework (Alvarez and Salzmann, 2016) that can automatically estimate these parameters and thus improve the computation cost of training our LG-DADA framework. Moreover, this study was limited to only evaluating our framework on morphological brain graphs. Although our evaluation connectomic dataset is heterogeneous, it would be interesting to test our model on multimodal graphs such as functional (Mhiri and Rekik, 2020) and structural (Wen et al., 2017) brain graphs which are highly heterogeneous. Notably, our framework produced the lowest error and the highest PCC when predicting three target brain graphs from different source graphs, which demonstrates the potential utility of our predictive framework in disease diagnosis tasks. In this paper, we primarily focused on brain graph synthesis which is a challenging task, however, we aim in our future work to leverage the predicted target graphs in early disease identification such as classifying mild cognitive impairment subjects and Alzheimer's disease subjects.. Ultimately, our target brain graph prediction framework from a source graph provides a new and exciting venue for better predicting missing brain graphs while preserving the real connectivity patterns existing between brain regions. Thus, we will investigate in our future work the discriminative power of the predicted brain graphs in the diagnosis of different neurological disorders such as Alzheimer's disease (Soussia and Rekik, 2019; 2018).

5. Conclusion

In this work, we introduced LG-DADA, a geometric deep learning framework for target brain graph prediction from a single source graph. Our key contribution consists in designing: (1) a *domain alignment* of source domain to the target domain by learning their latent representations, and (2) a *dual adversarial regularization* that synergistically learns a source embedding of training and testing brain graphs using two discriminators and predicts the training target graphs. We evaluated our framework on morphological brain graphs of healthy and disordered subjects. In our future work, we will first leverage cycleGAN, a recent GAN-based algorithm (Zhu et al., 2017), to improve domain alignment by moving the source to the target and translating back the target to the source domain while leveraging our proposed dual discriminators. This algorithm has been widely used for medical image synthesis tasks (Hiasa et al., 2018). We aim to adopt it for graph synthesis task. Second, we will learn a joint alignment of all source and target graphs that will maximize the correlation between these domains (i.e., source and target) and, third we will predict all target graphs simultaneously. It would be also interesting to include other brain graphs such as functional or structural (Liu et al., 2016) and to compare our model to existing domain adaptation methods (Hoffman et al., 2017; Tzeng et al., 2017).

Declaration of Competing Interest

Authors declare that they have no conflict of interest.

CRediT authorship contribution statement

Alaa Bessadok: Methodology, Software, Formal analysis, Validation, Visualization, Writing - original draft. **Mohamed Ali Mahjoub:** Supervision, Writing - original draft. **Islem Rekik:** Conceptualization, Supervision, Methodology, Resources, Writing - review & editing, Funding acquisition.

Acknowledgements

This project has been funded by the 2232 International Fellowship for Outstanding Researchers Program of TUBITAK (Project No:118C288, http://basira-lab.com/reprime/). However, all scientific contributions made in this project are owned and approved solely by the authors.

References

Alvarez, J.M., Salzmann, M., 2016. Learning the number of neurons in deep networks. In: Advances in Neural Information Processing Systems, pp. 2270–2278.
Alzheimer's Disease Neuroimaging Initiative, Lisowska, A., Rekik, I., 2018. Joint pairing and structured mapping of convolutional brain morphological multiplexes for early dementia diagnosis. Brain Connect. 9, 22–36.
Arjovsky, M., Chintala, S., Bottou, L., 2017. Wasserstein generative adversarial networks. In: International Conference on Machine Learning, pp. 214–223.
Arslan, S., Ktena, S. I., Glocker, B., Rueckert, D., Graph saliency maps through spectral convolutional networks: application to sex classification with brain connectivity. arXiv preprint arXiv:1806.01764.
Banka, A., Rekik, I., 2019. Adversarial connectome embedding for mild cognitive impairment identification using cortical morphological networks. In: International Workshop on Connectomics in Neuroimaging, pp. 74–82.
Bano, S., Asad, M., Fetit, A.E., Rekik, I., 2018. XmoNet: a fully convolutional network for cross-modality MR image inference. In: International Workshop on PRedictive Intelligence In MEdicine, pp. 129–137.
Ben-Cohen, A., Klang, E., Raskin, S.P., Soffer, S., Ben-Haim, S., Konen, E., Amitai, M.M., Greenspan, H., 2019. Cross-modality synthesis from CT to PET using FCN and GAN networks for improved automated lesion detection. Eng. Appl. Artif. Intell. 78, 186–194.
Bresson, X., Laurent, T., A two-step graph convolutional decoder for molecule generation. arXiv preprint arXiv:1906.03412.
Bronstein, M.M., Bruna, J., LeCun, Y., Szlam, A., Vandergheynst, P., 2017. Geometric deep learning: going beyond euclidean data. IEEE Signal Process. Mag. 34 (4), 18–42.
Cho, J., Kim, Y., Lee, M., 2018. Prediction to atrial fibrillation using deep convolutional neural networks. In: International Workshop on PRedictive Intelligence In MEdicine, pp. 164–171.
Choi, Y., Choi, M., Kim, M., Ha, J.-W., Kim, S., Choo, J., 2018. Stargan: unified generative adversarial networks for multi-domain image-to-image translation. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 8789–8797.
Dhifallah, S., Rekik, I., Initiative, A.D.N., et al., 2019. Clustering-based multi-view network fusion for estimating brain network atlases of healthy and disordered populations. J. Neurosci. Methods 311, 426–435.
Egghe, L., Leydesdorff, L., 2009. The relation between pearson's correlation coefficient r and salton's cosine measure. J. Am. Soc. Inf. Sci. Technol. 60 (5), 1027–1036.
Fischl, B., 2012. Freesurfer. Neuroimage 62 (2), 774–781.
Flam-Shepherd, D., Wu, T., Aspuru-Guzik, A., Graph deconvolutional generation. arXiv preprint arXiv:2002.07087.
Fornito, A., Zalesky, A., Breakspear, M., 2013. Graph analysis of the human connectome: promise, progress, and pitfalls. Neuroimage 80, 426–444.
Goktas, A.S., Bessadok, A., Rekik, I., 2020. Residual embedding similarity-based network selection for predicting brain network evolution trajectory from a single observation. International Workshop on PRedictive Intelligence In MEdicine.
Goodfellow, I., NIPS 2016 tutorial: generative adversarial networks. arXiv preprint arXiv:1701.00160.
Goodfellow, I., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., Courville, A., Bengio, Y., 2014. Generative adversarial nets. In: Advances in neural information processing systems, pp. 2672–2680.
Hardoon, D.R., Szedmak, S., Shawe-Taylor, J., 2004. Canonical correlation analysis: an overview with application to learning methods. Neural Comput. 16 (12), 2639–2664.
Hiasa, Y., Otake, Y., Takao, M., Matsuoka, T., Takashima, K., Carass, A., Prince, J.L., Sugano, N., Sato, Y., 2018. Cross-modality image synthesis from unpaired data using CycleGAN. In: International Workshop on Simulation and Synthesis in Medical Imaging, pp. 31–41.
Hoffman, J., Tzeng, E., Park, T., Zhu, J.-Y., Isola, P., Saenko, K., Efros, A., Darrell, T., 2018. Cycada: cycle-consistent adversarial domain adaptation. In: International Conference on Machine Learning, pp. 1989–1998.
Hoffman, J., Tzeng, E., Park, T., Zhu, J.-Y., Isola, P., Saenko, K., Efros, A. A., Darrell, T., Cycada: cycle-consistent adversarial domain adaptation. arXiv preprint arXiv:1711.03213.
Huynh, T., Gao, Y., Kang, J., Wang, L., Zhang, P., Lian, J., Shen, D., 2016. Estimating CT image from MRI data using structured random forest and auto-context model. IEEE Trans. Med. Imaging 35 (1), 174–183.
Jog, A., Roy, S., Carass, A., Prince, J.L., 2013. Magnetic resonance image synthesis through patch regression. In: 2013 IEEE 10th International Symposium on Biomedical Imaging, pp. 350–353.
Jolliffe, I.T., Cadima, J., 2016. Principal component analysis: a review and recent developments. Philos. Trans. R. Soc. A 374 (2065), 20150202.
Kipf, T. N., Welling, M., Semi-supervised classification with graph convolutional networks. arXiv preprint arXiv:1609.02907.
Ktena, S.I., Parisot, S., Ferrante, E., Rajchl, M., Lee, M., Glocker, B., Rueckert, D., 2017. Distance metric learning using graph convolutional networks: application to functional brain networks. In: International Conference on Medical Image Computing and Computer-Assisted Intervention, pp. 469–477.
Li, R., Zhang, W., Suk, H.-I., Wang, L., Li, J., Shen, D., Ji, S., 2014. Deep learning based imaging data completion for improved brain disease diagnosis. In: International Conference on Medical Image Computing and Computer-Assisted Intervention, pp. 305–312.
Liao, R., Li, Y., Song, Y., Wang, S., Hamilton, W., Duvenaud, D.K., Urtasun, R., Zemel, R., 2019. Efficient graph generation with graph recurrent attention networks. In: Advances in Neural Information Processing Systems, pp. 4257–4267.
Lin, J.-H., Sellke, T.M., Coyle, E.J., 1990. Adaptive stack filtering under the mean absolute error criterion. IEEE Trans. Acoust. Speech Signal Process. 38 (6), 938–954.
Lisowska, A., Rekik, I., Initiative, A.D.N., et al., 2017. Pairing-based ensemble classifier learning using convolutional brain multiplexes and multi-view brain networks for early dementia diagnosis. In: International Workshop on Connectomics in Neuroimaging, pp. 42–50.
Liu, L., Zhang, H., Rekik, I., Chen, X., Wang, Q., Shen, D., 2016. Outcome prediction for patient with high-grade gliomas from brain functional and structural networks. In: International Conference on Medical Image Computing and Computer-Assisted Intervention, pp. 26–34.
Liu, W., Chen, P.-Y., Cooper, H., Oh, M. H., Yeung, S., Suzumura, T., Can GAN learn topological features of a graph? arXiv preprint arXiv:1707.06197.
Maaten, L.v.d., Hinton, G., 2008. Visualizing data using t-SNE. J. Mach. Learn. Res. 9 (Nov), 2579–2605.
Mahjoub, I., Mahjoub, M.A., Rekik, I., 2018. Brain multiplexes reveal morphological connectional biomarkers fingerprinting late brain dementia states. Sci. Rep. 8 (1), 4103.
Mhiri, I., Khalifa, A.B., Mahjoub, M.A., Rekik, I., 2020. Brain graph super-resolution for boosting neurological disorder diagnosis using unsupervised multi-topology connectional brain template learning. Med. Image Anal. 65, 101768.
Mhiri, I., Rekik, I., 2020. Joint functional brain network atlas estimation and feature selection for neurological disorder diagnosis with application to Autism. Med. Image Anal. 60, 101596.
Nebli, A., Rekik, I., 2020. Gender differences in cortical morphological networks. Brain Imaging Behav. 14 (5), 1831–1839.
Olut, S., Sahin, Y. H., Demir, U., Unal, G., Generative adversarial training for MRA image synthesis using multi-contrast MRI. arXiv preprint arXiv:1804.04366.
Pan, S., Hu, R., Long, G., Jiang, J., Yao, L., Zhang, C., Adversarially regularized graph autoencoder. arXiv preprint arXiv:1802.04407.
Pan, Y., Liu, M., Lian, C., Zhou, T., Xia, Y., Shen, D., 2018b. Synthesizing missing PET from MRI with cycle-consistent generative adversarial networks for Alzheimer's disease diagnosis. In: International Conference on Medical Image Computing and Computer-Assisted Intervention, pp. 455–463.
Parisot, S., Ktena, S.I., Ferrante, E., Lee, M., Moreno, R.G., Glocker, B., Rueckert, D., 2017. Spectral graph convolutions for population-based disease prediction. In: International Conference on Medical Image Computing and Computer-Assisted Intervention, pp. 177–185.
Pei, Z., Cao, Z., Long, M., Wang, J., 2018. Multi-adversarial domain adaptation. In: Thirty-Second AAAI Conference on Artificial Intelligence.
Pu, Y., Gan, Z., Henao, R., Yuan, X., Li, C., Stevens, A., Carin, L., 2016. Variational autoencoder for deep learning of images, labels and captions. In: Advances in neural information processing systems, pp. 2352–2360.
Redko, I., Morvant, E., Habrard, A., Sebban, M., Bennani, Y., A survey on domain adaptation theory: learning bounds and theoretical guarantees. arXiv preprint arXiv:2004.11829.
Shen, Y., Gao, M., 2019. Brain tumor segmentation on MRI with missing modalities. In: International Conference on Information Processing in Medical Imaging, pp. 417–428.
Soussia, M., Rekik, I., A review on image-and network-based brain data analysis techniques for Alzheimer's disease diagnosis reveals a gap in developing predictive methods for prognosis. arXiv preprint arXiv:1808.01951.
Soussia, M., Rekik, I., 2018. Unsupervised manifold learning using high-order morphological brain networks derived from T1-w MRI for autism diagnosis. Front. Neuroinformatics 12, 70.
Soussia, M., Rekik, I., 2019. 7 years of developing seed techniques for Alzheimer's Disease diagnosis using brain image and connectivity data largely bypassed prediction for prognosis. In: International Workshop on PRedictive Intelligence In MEdicine, pp. 81–93.
Su, S.-Y., Hajimirsadeghi, H., Mori, G., 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.
Toldo, M., Maracani, A., Michieli, U., Zanuttigh, P., Unsupervised domain adaptation in semantic segmentation: a review. arXiv preprint arXiv:2005.10876.

Tzeng, E., Hoffman, J., Saenko, K., Darrell, T., 2017. Adversarial discriminative domain adaptation. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 7167–7176.

Wang, B., Ramazzotti, D., De Sano, L., Zhu, J., Pierson, E., Batzoglou, S., 2017. SIMLR: a tool for large-scale single-cell analysis by multi-kernel learning. bioRxiv, 118901.

Wang, H., Wang, J., Wang, J., Zhao, M., Zhang, W., Zhang, F., Xie, X., Guo, M., 2018. GraphGAN: graph representation learning with generative adversarial nets. In: Thirty-Second AAAI Conference on Artificial Intelligence.

Wang, J., Zhang, L., Wang, Q., Chen, L., Shi, J., Chen, X., Li, Z., Shen, D., 2020. Multi-class ASD classification based on functional connectivity and functional correlation tensor via multi-source domain adaptation and multi-view sparse representation. IEEE Trans. Med. Imaging.

Wen, H., Liu, Y., Rekik, I., Wang, S., Zhang, J., Zhang, Y., Peng, Y., He, H., 2017. Disrupted topological organization of structural networks revealed by probabilistic diffusion tractography in tourette syndrome children. Hum. Brain Mapp. 38 (8), 3988–4008.

Wilson, G., Cook, D.J., 2020. A survey of unsupervised deep domain adaptation. ACM Trans. Intell. Syst. Technol. (TIST) 11 (5), 1–46.

Wright, J., Yang, A.Y., Ganesh, A., Sastry, S.S., Ma, Y., 2008. Robust face recognition via sparse representation. IEEE Trans. Pattern Anal. Mach. Intell. 31 (2), 210–227.

Yang, H., Sun, J., Carass, A., Zhao, C., Lee, J., Xu, Z., Prince, J., 2018. Unpaired brain MR-to-CT synthesis using a structure-constrained CycleGAN. In: Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support. Springer, Cham, pp. 174–182.

Yi, X., Walia, E., Babyn, P., Generative adversarial network in medical imaging: a review. arXiv preprint arXiv:1809.07294.

Zhou, B., Lin, X., Eck, B., 2019. Limited angle tomography reconstruction: synthetic reconstruction via unsupervised sinogram adaptation. In: International Conference on Information Processing in Medical Imaging, pp. 141–152.

Zhu, J.-Y., Park, T., Isola, P., Efros, A. A., 2017. Unpaired image-to-image translation using cycle-consistent adversarial networks. arXiv preprint.

Zhu, M., Rekik, I., 2018. Multi-view brain network prediction from a source view using sample selection via CCA-based multi-kernel connectomic manifold learning. In: International Workshop on PRedictive Intelligence In MEdicine, pp. 94–102.