A Structure-Guided Effective and Temporal-Lag Connectivity Network for Revealing Brain Disorder Mechanisms

Zhengwang Xia[®], Tao Zhou[®], *Member, IEEE*, Saqib Mamoon[®], Amani Alfakih, and Jianfeng Lu[®], *Member, IEEE*

Abstract—Brain network provides important insights for the diagnosis of many brain disorders, and how to effectively model the brain structure has become one of the core issues in the domain of brain imaging analysis. Recently, various computational methods have been proposed to estimate the causal relationship (i.e., effective connectivity) between brain regions. Compared with traditional correlation-based methods, effective connectivity can provide the direction of information flow, which may provide additional information for the diagnosis of brain diseases. However, existing methods either ignore the fact that there is a temporal-lag in the information transmission across brain regions, or simply set the temporal-lag value between all brain regions to a fixed value. To overcome these issues, we design an effective temporal-lag neural network (termed ETLN) to simultaneously infer the causal relationships and the temporal-lag values between brain regions, which can be trained in an end-to-end manner. In addition, we also introduce three mechanisms to better guide the modeling of brain networks. The evaluation results on the Alzheimer's Disease Neuroimaging Initiative (ADNI) database demonstrate the effectiveness of the proposed method.

Index Terms—Brain network, causal inference, temporallag estimation, brain disorder identification.

I. INTRODUCTION

F UNCTIONAL brain networks have been widely used to understand the principles of brain organization [1], [2] and explore the sensitive biomarkers of neuropsychiatric diseases [3]. Brain networks depict the complex patterns of interactions between brain regions and provide a powerful tool for detecting several brain disorders such as autism spectrum disorder [4] and Alzheimer's disease [5].

Manuscript received 15 September 2022; revised 1 March 2023; accepted 4 April 2023. Date of publication 7 April 2023; date of current version 6 June 2023. This work was supported in part by the Natural Science Foundation of Jiangsu Province under Grant BK20221487 and in part by the National Natural Science Foundation of China under Grant 62172228. (*Corresponding author: Jianfeng Lu.*)

The authors are with the School of Computer Science and Engineering, Nanjing University of Science and Technology, Nanjing 210094, China (e-mail: xzhengwang@njust.edu.cn; taozhou.dreams@ gmail.com; saqibmamoon@njust.edu.cn; am775901039@gmail.com; lujf@njust.edu.cn).

Digital Object Identifier 10.1109/JBHI.2023.3265364

A brain network can be viewed as a collection of nodes and edges, where each node represents a brain region defined by the physiological template, and each edge represents the relationship between two brain regions. Recent years have witnessed an endless stream of research on functional brain network modeling. In general, the relationships represented by edges in these modeling methods can be divided into two categories: functional connectivity (FC) and effective connectivity (EC) [6]. FC reflects a statistical dependence between functional magnetic resonance imaging (fMRI) signals from distinct brain regions, which assess the direct or indirect interactions between brain regions. Unlike FC, EC is defined as the causal influence one brain region exerts over another, the purpose of which is to determine the direction of information flow. Due to the lack of directional information in FC, brain network models based on FC may yield suboptimal results in identifying abnormal patterns caused by brain diseases. For example, Li et al. [7] constructed a novel effective connectivity network for the diagnosis of mild cognitive impairment (MCI), and the experimental results show that the EC-based method achieves a significant improvement over the FC-based method. Chen et al. [8] designed a message-passing algorithm to estimate the direction of information flow between brain regions, and the result similarly demonstrated that the ECbased method significantly outperforms FC-based methods in detecting disease-related neuroimaging biomarkers. Overall, the results suggest that EC can provide more effective information for discriminating brain disorders than traditional FC.

Recently, various computational methods have been proposed to estimate the direction of information flow between brain regions, which model the problem from different perspectives [9], [10], [11]. However, these methods have their own limitations and they cannot accurately infer the direction of information flow between brain regions in some cases [12]. For example, Granger causality (GC), one of the most popular methods for inferring causality between brain regions, assumes that the upstream signal will be repeated by the downstream signal with a certain temporal-lag. However, in practice, this assumption is often invalid in fMRI due to the presence of noise during data acquisition [9]. Bayesian network (BN), a widely used algorithm for constructing brain effective connectivity network (ECN), restricts the candidate graph structure to be a directed acyclic graph (DAG). However, accumulating studies show that

2168-2194 © 2023 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See https://www.ieee.org/publications/rights/index.html for more information.



Fig. 1. Illustration of the proposed framework for brain disorder identification. The framework mainly consists of three steps, including data preprocessing, structure learning, and classification. The data preprocessing step is used to extract the signal values of each brain region from fMRI images. The structure learning step is adopted to estimate causal effects and temporal-lag values between brain regions. In order to better model the interaction between brain regions, we introduce three mechanisms to guide the modeling of brain networks. The classification step is utilized to obtain the final decision result.

functional interactions in the brain are not acyclic due to reciprocal polysynaptic connections [13], [14]. Linear non-Gaussian acyclic model (LiNGAM), a data-driven approach to infer the direction of information flow between brain regions from fMRI data, is based on independent component analysis (ICA) to search for solutions. However, ICA requires acquiring a large number of time series data, and the performance of LiNGAM is often not ideal when the fMRI data sample is small [11]. Based on the above analysis, it can be seen that the previous modeling methods may only be suitable for some specific situations, while several strong assumptions cannot be guaranteed to hold.

In addition to the shortcomings mentioned above, many existing effective connectivity estimation methods ignore a common problem: the temporal-lag value of the information flow between different brain regions should be different. For instance, GC method considers the influence of temporal-lag when inferring the causality between brain regions, but it sets the temporal-lag value to a fixed value, equivalent to the time of a complete brain sampling [15]. BN is based on the Bayesian scoring metric to search for the potential optimal graph structure, and the time delay of information transmission is not considered in the search process [16]. LiNGAM assumes that the time series signal of each brain region is a linear combination of all other brain regions, and there is no time lag between them [11]. Obviously, none of these assumptions correspond to the actual situation. For example, numerous studies have found highly reproducible temporal-lag patterns in the blood oxygen level-dependent (BOLD) signal of healthy subjects [17], [18].

Raatikainen et al. have proved that there are large variations in temporal-lag of information propagation within the brain [19].

To overcome the shortcomings of traditional ECN estimation methods, we propose a novel deep learning model, named effective temporal-lag neural network (ETLN), to simultaneously infer the direction (effective connectivity network, ECN) and the temporal-lag value (temporal-lag connectivity network, TCN) of information flow between brain regions. Different from existing methods, our method does not impose unrealistic constraints on the underlying graph structure. Our framework can consider the influence of temporal-lag while inferring the effective connectivity structure, while several methods often ignore this point. Fig. 1 illustrates a schematic diagram of the proposed framework. Specifically, we first extract the blood oxygen signal of each brain region from fMRI. Then, data extracted from fMRI are fed into the deep learning model ETLN to estimate the causality and temporal-lag value between brain regions. Finally, a classifier is trained based on the constructed brain network to realize the recognition of abnormal patterns of brain activity. More importantly, we introduce three new mechanisms to obtain more precise estimates of causal and temporal-lag effects between brain regions. For the relationship between ECN and TCN, we introduce the local consistency mechanism to ensure the correspondence between the two candidate graphs. The core idea of the local consistency mechanism is that if there is no causality between the two brain regions in ECN, their corresponding temporal-lag value in TCN should be 0. In order to determine whether there is a causal relationship between the two brain regions in the candidate ECN, we introduce the adaptive mechanism. Furthermore, we have an additional constraint on the data distribution of candidate TCN, that is, the longer the transmission distance between two brain regions, the larger the temporal-lag value of the information flow between them.

The main contributions of this paper can be summarized as follows. First, we propose a new strategy to simultaneously estimate the direction and temporal-lag values of information flow between brain regions via a deep learning pipeline. Second, our brain network modeling method can characterize the nonlinear interaction between brain regions, rather than the traditional linear interaction. Third, we introduce three new mechanisms into the proposed ETLN to guide the construction of the brain network model, which allows for a more accurate assessment of causal and temporal-lag effects between brain regions. The proposed method is verified on the public database (Alzheimer's Disease Neuroimaging Initiative, ADNI) and achieves promising performance compared with other popular benchmark approaches.

II. RELATED WORK

A. Brain Network Estimation Methods

Recently, many studies have found that neurological disorders are associated with abnormal functional integration between some brain regions [3], [4]. Over the past few years, several brain network modeling algorithms have been developed for the diagnosis of neurological diseases [7], [20], [21], [22]. These typical brain network construction methods can be divided into two categories from the meaning of functional connection, including 1) correlation-based methods, and 2) causality-based methods.

In the domain of brain network analysis, Pearson correlation is the most popular measure for constructing brain networks, which is widely used to define the strength of functional connectivity, i.e., the temporal correlation between signals from different brain regions. While many meaningful discoveries have been obtained based on correlation brain networks, it is undeniable that the correlation brain network model has its inherent limitations. First, correlation-based brain networks are often too dense to clearly reveal which functional connections are most relevant to brain disorder [23]. Second, compared with EC, FC lacks the direction of information flow. It may yield suboptimal results for brain disorder identification if ignoring the direction of information flow [24].

Besides, methods to construct effective connectivity networks can be divided into two categories: multivariate methods with constraints and multivariate methods without constraints. Bayesian network (BN) is a typical constraint-type method, which has an underlying assumption that the latent graph structure must be a directed acyclic graph [13]. However, recent research shows that functional connections between brain regions are not acyclic due to reciprocal polysynaptic connections [14]. It seems that those methods without constraints are superior to those constraint-type methods. But, those methods without constraints almost ignore the temporal-lag value of information flow [17], [18]. More details about temporal-lag will be given in the next subsection.

TABLE I								
Demographic	INFORMATION	OF	THE	Used	DATASET	ī		

DEMOGRAPHIC IN	FORMATION	OF THE USE	D DATASET.
Group	NC	eMCI	LMCI
Male/Female	22/29	21/37	24/16
Age(mean±STD)	$75.2 {\pm} 6.9$	72.2 ± 6.9	72.4 ± 8.0

B. Temporal-Lag Value Estimation Methods

In the past few decades, there have been few studies concerning the estimation of information propagation lag between brain regions. Granger causality (GC) is a popular effective connectivity estimation method to measure the causal effect of one brain region on another brain region [9]. It is based on the idea that the cause of an event cannot come after its consequence, which is one of the causal discovery methods that takes into account the influence of temporal-lag. Unfortunately, the temporal-lag value of GC is set to a fixed value when searching the causal relationship between brain regions [15]. However, many scholars have found that there is lag variability among different brain regions, that is, the temporal-lag values are not consistent between different brain regions, and this phenomenon has been demonstrated by many studies [25], [26]. In addition to GC, there is another class of studies that define the temporal-lag value across brain regions as a cross-correlation function between two BOLD signals [17], [27]. This type of approach also has its own drawbacks. The first controversial point is that the definition of temporal-lag may be too simple and does not correspond to the real situation. The second point is that this method separates the calculation of temporal-lag value from the identification of causality. The temporal-lag should be zero when there is no causality between two brain regions, which is not considered in the definition of such methods.

III. DATA ACQUISITION AND PROCESSING

In this study, we used a total of 149 subjects' resting state fMRI data from the publicly available Alzheimer's Disease Neuroimaging Intiative (ADNI) database,¹ which includes normal controls (NC), early MCI (eMCI), and late MCI (LMCI). Notably, some subjects in the ADNI dataset were recruited at regular intervals, which lead to the existence of different subjects belonging to the same subject. As discussed [28], such case can bring the data leakage issue. Because the samples of the same subject collected at different times could be divided into the training set and test set, respectively, which brings an unfair evaluation. Therefore, in order to make the subsequent evaluation results more convincing, we deduplicate the dataset to ensure that there are no two samples from the same subject. The demographic information can be listed in Table I.

In this dataset, each subject signed the written informed consent form after a full written and verbal explanation. This research was approved by the Research Ethics Board of ADNI.² All subjects were scanned with the same protocol using 3.0 T Philips Achieva scanners. The scanning parameters are as follows: repetition time (TR) = 3000 ms, echo time (TE) = 30 ms,

¹[Online]. Available: http://adni.loni.ucla.edu

²[Online]. Available: http://adni.loni.usc.edu/study-design/ongoing-investi gations/

flip angle = 80° , imaging matrix = 64×64 , slices = 48 and slice thickness = 3.3 mm. For fMRI data, we apply the standard procedures as follows. First, the first 5 volumes of each subject were discarded before preprocessing to avoid noise signals, and then the remaining 135 volumes were reserved for the subsequent analysis. All the functional images were registered to the first image and transformed into the Montreal Neurological Institute (MNI) space with a resample voxel size of $3 \times 3 \times 3$ mm³. Subsequently, Conn Toolbox 20b,³ a Statistical Parametric Mapping (SPM12) based preprocessing pipeline, was used to perform outlier detection, direction segmentation and normalization, linear detrending, and functional smoothing with a Gaussian kernel of 8 mm full width half maximum (FWHM), etc. Finally, the time series of each brain region is extracted from the preprocessed images based on the AAL atlas.

IV. METHODOLOGY

Previous fMRI studies for inferring causality in brain regions suffer from two shortcomings. First, existing methods often impose certain restrictive assumptions on candidate graphs when inferring the causal relationship between brain regions, and the plausibility of these methods has been questioned [13]. Second, most methods ignore the influence of temporal-lag, or simply set the temporal-lag value of all brain regions to a fixed value [15]. To overcome the two issues, we propose a novel GAN-based neural network, named effective temporal-lag network (ETLN), to simultaneously learn causality and temporal-lag values between brain regions. We provide the details of the proposed ETLN below.

A. Overview of ETLN

The detailed structure of ETLN is shown in Fig. 2, which embeds the causal relationship and temporal-lag values into a generator network as parameters to be optimized. To clearly describe the proposed model, we first give notations as follows. Let $X \in \mathbb{R}^{v \times t}$ denote the obtained time series data from preprocessed fMRI images, where v and t represent the number of brain regions and the length of the time series, respectively. $\boldsymbol{D} \in \mathbb{R}^{v imes (t-1)}$ is the first-order differences of the time series data X. Notably, X and D are jointly used as the input of the generator network for training, and the dimensions of both are $v \times (t-1)$. The reason why the dimension of X is not $v \times t$ will be given in Section IV-B. Based on the two inputs, the generator network will generate fake time-series data \hat{X} that can match the real time-series data X as closely as possible. The discriminator network takes real fMRI time series data X and fake time series data \hat{X} as inputs and tries to find a mapping that can distinguish them. The generator network is composed of v causal structure inference modules, where each module is designed to search for the direct causes of the corresponding brain region and the temporal-lag values from these causes to the brain region. In the causal structure inference module, we embed two gates, namely the causal gate $C_{:,i} \in \mathbb{R}^{v \times 1}$ and lag gate $L_{:,i} \in \mathbb{R}^{v \times 1}$, to preserve the causal effects and temporal-lag values from other brain regions to the *i*-th brain region. After the model is well-trained, the causal relationship and temporal-lag values between v brain regions can be obtained from the parameters of the generator network. As shown in Fig. 2, by splicing the causal parameters and lag parameters of all brain regions, i.e., the causal and lag gates of each brain region, two matrices with dimensions equal to $v \times v$ can be obtained, which is the solution we have been looking for, namely the effective connectivity network $C \in \mathbb{R}^{v \times v}$ and the temporal-lag connectivity network $L \in \mathbb{R}^{v \times v}$.

B. Formulation of Structure Inference

The current mainstream methods are based on the idea of multivariate regression to estimate the causal relationship between variables [29], [30]. The main advantage of these methods is that they are data-driven and do not take any restrictive assumptions about the structure of the causal graph. Such methods can retrieve excitatory, inhibitory, and bidirectional connections between variables. The core idea behind them is that each univariate component $X_{i,:}$ is a mixture of the remaining components $X_{i,:}$, $j \neq i$, and their general form can be defined by

$$\underset{\boldsymbol{C}_{:,i}}{\arg\min} \parallel \boldsymbol{X}_{i,:} - \boldsymbol{C}_{:,i}^T \boldsymbol{X} \parallel_2^2, \tag{1}$$

where $X \in \mathbb{R}^{v \times t}$ is the time series data extracted from fMRI, and $X_{i,:}$ denotes the *i*-th row of X. $C_{:,i} = [C_{1,i}, \ldots, C_{v,i}]$ is the regression coefficients, which denotes the causal effect of all other brain region on the *i*-th brain region. The superscript^T denotes the transpose operation. Besides, $C_{i,i}$ is set 0 to avoid self-loops.

However, such methods ignore the effects of the temporal-lag. The transmission of information between two brain regions takes a certain amount of time, even if the spatial distance between the two brain regions is very close. Therefore, considering the influence of the temporal-lag, (1) can be reformulated as follows:

$$\underset{\boldsymbol{C}_{:,i}, \boldsymbol{L}_{:,i}}{\arg\min} \| \boldsymbol{X}_{i,:} - f_i(\boldsymbol{C}_{:,i}, \boldsymbol{X}, \boldsymbol{L}_{:,i}, \boldsymbol{D}) \|_2^2,$$
(2)

where the output of $f_i(\cdot)$ is the predicted signal value of *i*-th brain region. The specific definition form of $f_i(\cdot)$ will be given below. X is used to model the causal relationship between brain regions, and matrix D is the first-order difference time series of X, which is used to model the temporal-lag relationship between brain regions. The reason why D can model the temporal-lag relationship between brain regions will be discussed later. Similar to the above-mentioned methods, the core idea behind our approach is that the signal value of each brain region can be jointly generated by the causal and temporal-lag effects of all other brain regions on that brain region. Likewise, $L_{i,i}$ is set to 0 to avoid self loops.

A key question is how to simultaneously model causal and temporal-lag relationships between brain regions, namely the specific form of f_i in (2). For clarity, we show the schematic diagram in Fig. 3. Suppose $X_{i,:}$ and $X_{j,:}$ represent the time series data of two brain regions, and there is a causal link from $X_{j,:}$ to $X_{i,:}$, which can be denoted as $X_{j,:} \rightarrow X_{i,:}$. A schematic diagram of the information transmission between two brain

³[Online]. Available: https://web.conn-toolbox.org/



Detailed structure of ETLN

Fig. 2. The detailed structure of ETLN. The time series data X and its first-order difference D of each subject are fed into the generator network to estimate causal effects and temporal-lag values between v brain regions. Based on the two inputs, the generator network will generate fake time-series data \hat{X} . The objective of the discriminator network is to distinguish whether the input samples are from real data X or fake data \hat{X} . It is worth noting that there are differences in splicing causal and lag gates across brain regions. According to (4), it can be found that there is a transpose relationship between C and L. Therefore, we splice the causal gate in a row-first manner, and the lag gate in a column-first manner.

regions is shown on the right side of Fig. 3. However, these methods often ignore the influence of temporal-lag, and the causal effect between two brain regions is defined as $X_{j,p} \rightarrow X_{i,p}$. $X_{i,p}$ represents the signal value of the *i*-th brain region at the *p*-th time point. This type of methods assumes that the cause and effect can be produced at the same moment, i.e., there is no time delay in the transmission of information between two brain regions, also known as an instantaneous effect.

Several studies have demonstrated that there is a temporal-lag in the transmission of information across brain regions [19], [31], and the range of temporal-lag is shorter than the measurement rate of fMRI [32]. Therefore, the true causal effect between two brain regions should be defined as $X_{j,p} \rightarrow X_{i,(p+L_{j,i})}$, the most crucial point is that the moment when the cause occurs should be before the moment when the effect occurs. $L_{j,i}$ denotes the temporal-lag value from the *j*-th brain region to the *i*-th brain region. In addition to instantaneous and true causal effects, we argue that there is also a temporal-lag effect. The temporal-lag effect should be from the instantaneous effect to the true causal effect along the time dimension, as shown by the red directed line in Fig. 3. According to the law of vector addition, we can deduce that the true causal effect should be equal to the sum of



Fig. 3. Schematic diagram for temporal-lag relationship description.

the instantaneous effect and the temporal-lag effect. Based on this, we can derive the following equation:

$$\boldsymbol{C}_{j,i} \cdot \boldsymbol{X}_{j,p} = \boldsymbol{X}_{i,p} + \boldsymbol{L}_{j,i} \cdot \left(\boldsymbol{X}_{i,p+1} - \boldsymbol{X}_{i,p} \right), \qquad (3)$$

where the term $C_{j,i} \cdot X_{j,p}$ means the true causal effect. $X_{i,p}$ corresponds to the instantaneous effect, and the term $L_{j,i} \cdot (X_{i,p+1} - X_{i,p})$ represents the temporal-lag effect. Generalizing (3) to a multivariate case, we can obtain

$$\hat{\mathbf{X}} = \mathbf{C}^T \times \mathbf{X} - \mathbf{L} \times \mathbf{D},$$
$$\mathbf{D} = \frac{100}{TR} \cdot (\mathbf{X}_{:,2:t} - \mathbf{X}_{:,1:t-1}),$$
ith $\mathbf{C}_{i,i} = 0, \mathbf{L}_{i,i} = 0, i = [1, \dots, v],$ (4)

where X represents the predicted time series data, and D is the first-order difference of matrix X. $X_{:,2:t}$ denotes all samples of matrix X from column 2 to column t. The value ranges of causal gate C and lag gate L are set in the range of [-1, 1] and [0, 1]1], respectively. It is worth noting that the dimension of matrix **D** is $v \times (t-1)$. Since the causal structure inference module performs pairing training based on the samples of X and D, thus the dimension of the two needs to be consistent. In this case, the last column of the matrix X is discarded, so the dimension of X in Fig. 2 is $v \times (t-1)$. Besides, the numerical range of **D** is limited by multiplying by the constant $\frac{100}{TR}$. The reasons for this can be attributed to the following points. First, many studies based on other imaging techniques have reported that the signal transmission delay across brain regions is in the range of 0-100ms [12], [25], [26], which is much smaller than the measurement time interval of fMRI used in this paper. Second, this design is more conducive to the training of ETLN. Since the causal structure inference module is a two-branch network, if the lag gate is represented by the real numerical range of 0-100, it will lead to the numerical imbalance of the two branches, making the model difficult to train. After setting a $\frac{100}{TB}$ bound for **D**, the lag gate can be set in the range of 0-1, which helps to solve the imbalance problem between the two branches.

C. Constraint Mechanisms

W

In addition to giving a formal definition, we introduce three mechanisms: local consistency mechanism, adaptive mechanism, and spatial constraint mechanism, to better guide the modeling of brain networks. The details of the three mechanisms will be given below. It is worth noting that there is local consistency in the causal and temporal-lag relationships between brain regions. Specifically, for those brain regions with a causal link, the temporal-lag value between them is unclear. But for those brain regions without the causal link, the temporal-lag value between them should be 0. As no information is transferred between two brain regions, which indicates that no time delay between them. Therefore, a local consistency loss is introduced to ensure that the temporal-lag value of those brain regions without a causal link is 0, and its calculation can be defined as follows:

index =
$$\Psi(\boldsymbol{C}, \mu)$$
,
 $\mathcal{L}_{local} = CE\left(\vec{\boldsymbol{L}}_{index,index}, \vec{\boldsymbol{Z}}\right)$, (5)

where μ is a threshold for determining whether there is a causal link between brain regions. The absolute value of the weights in *C* less than μ is considered as no causal link between brain regions. $\Psi(\cdot)$ returns a set that includes the indices of those brain regions without causal links. $L_{index,index}$ is the induced subgraph of matrix *L*. *Z* is a zero matrix with the same dimension as $L_{index,index}$. The symbol \neg means to convert a matrix to a vector, and *CE* refers to the cross-entropy loss.

For the threshold μ in (5), it is usually defined in two ways. First, it can be set to a fixed value, and the other is to introduce a dynamic mechanism. Second, the value of μ is not fixed, and its value varies under different situations. Due to differences among individuals, we utilize the second way to set different values of μ . To achieve this, we introduce an adaptive mechanism to assess whether there is a causal link between brain regions, which is defined by

$$\mu = mean(\boldsymbol{C}),\tag{6}$$

where $mean(\cdot)$ is a function that returns the mean of all elements in matrix C. During the optimization process of the network ETLN, the weight of the parameter C is not fixed. Therefore, the value of μ is also constantly changing within the process of network optimization.

Existing studies have shown that the information transmission delay between the ipsilateral hemispheres is shorter, and the information transmission across hemispheres often requires more time [25], [33]. In general, the delay of information transmission may have a certain correlation with the transmission distance. Thus, we introduce the spatial distance prior information of the brain to constrain the data distribution of matrix L. Then, the first point is how to define the spatial distance prior information of the brain. For the convenience of illustration, we give a schematic diagram as shown in Fig. 4 to explain how to define the prior information. In Fig. 4, each node represents a brain region. Near the node, we give the abbreviation of the brain region and the corresponding center coordinates, respectively. Suppose there is a causal link from ANG.L to SFGmed.L. Since both brain regions are located in the left hemisphere, the transmission distance between them is defined as the distance between the center coordinates of the two brain region, i.e., $len_{ANG,L \rightarrow SFGmed,L} =$ $\sqrt{(-44 - (-5))^2 + (-61 - 49)^2 + (36 - 31)^2} = 116.82.$ Suppose there is a causal link from ANG.L to ANG.R. Since



Fig. 4. Schematic diagram for brain spatial distance.

the two brain regions are located in different hemispheres, the definition of the transmission distance between them is different from the previous one. Studies have shown that the interaction of the left and right hemispheres of the brain relies on the corpus callosum [34], which is located roughly in the center of the brain. In this study, the coordinate (0, 0, 0) is used to represent the center coordinate of the corpus callosum. $len_{ANG.L \to ANG.R} = \sqrt{(-44)^2 + (-61)^2 + 36^2} +$ Thus, $\sqrt{46^2 + (-60)^2 + 39^2} = 168.45$. Based on the above calculation method, the pairwise spatial transmission distances between all brain regions of the AAL atlas can be obtained, and the prior information is stored in the matrix **P**. The center coordinates of each brain region in the AAL atlas can be obtained from [35]. When the required prior information P is obtained, the spatial constraint loss is defined as follows:

$$\mathcal{L}_{spatial} = CE\left(\vec{L}_{\overline{index},\overline{index}}, \vec{P}_{\overline{index},\overline{index}}\right), \qquad (7)$$

where *index* is the complement of the set *index* in (5). The aim of the above formula is to constrain the distribution of temporallag values for those brain regions with causality.

D. Model Training and Implementation

The overall objective function of ETLN can be derived by combing the local consistency loss \mathcal{L}_{local} and the spatial constraint loss $\mathcal{L}_{spatial}$, as follows:

$$G_{loss} = MSE\left(\hat{\boldsymbol{X}}, \boldsymbol{X}\right) + \mathcal{L}_{local} + \mathcal{L}_{spatial},$$
$$D_{loss} = KL(\boldsymbol{X} \mid\mid \hat{\boldsymbol{X}}), \tag{8}$$

where G_{loss} is the generator loss, and $MSE(\cdot)$ is the mean square loss. D_{loss} is the discriminator loss, which minimizes the data distribution difference between the fake data \hat{X} and the real data X by Kullback-Leibler divergence.

The detailed structure of the ETLN network is as follows: each causal structure inference consists of a 1-hidden layer with 200 Tanh units. The discriminator network consists of a 2-hidden layer with 200 hidden LeakyReLU units. The causal and lag gates are initialized to 1, except the self-loop term is set to 0. A standard Adam optimizer with a learning rate of 0.001 is used to optimize the model. The epochs of both generator and discriminator are set to 1000. Our model is implemented in PyTorch using an NVIDIA GeForce 1080Ti GPU with 11 GB memory. In order to obtain more accurate modeling results, the

modeling is repeated 32 times for each subject, and the average value is taken as the final result.

V. EXPERIMENTS

In this section, Section V-A first briefly describes the details of the 7 comparison methods. Then, the specific details of the classification step are provided in Section V-B. Finally, Sections V-C and V-D present the experimental results of comparative methods and ablation study, respectively.

A. Comparison Methods

To validate the effectiveness of the proposed method, we compare it with many popular methods. According to whether the brain network contains directional information, these methods can be divided into two types: correlation-based methods (i.e., Pearson correlation-based method (PC) [4], sparse representation method (SR) [5], low-rank representation method (LR) [21] and sparse low-rank representation method (SLR) [21]) and causality-based methods (i.e., Granger causality based method (GC) [9], linear non-Gaussian acyclic model (LiNGAM) [11] and transfer entropy (TE) [36]). In fact, causality-based methods should also include Bayesian network (BN) [13]. However, BN is not suitable for modeling large-scale (more than 20 nodes) data. The optimization time of BN grows exponentially as the number of nodes increases. Since BN is too time-consuming, we did not add it for comparison. We now briefly summarize these competing methods as follows.

- 1) PC method [4]. In the PC method, the edge weight is defined as the Pearson correlation coefficient between distinct brain regions.
- 2) SR method [5]. This method adds an L_1 regularization constraint to the graph to exclude confounding effects between brain regions.
- 3) LR method [21]. This model introduces modular prior knowledge in brain network construction via a low-rank constraint.
- 4) SLR method [21]. This method utilizes both sparse and low-rank constraints to the brain network weight matrix.
- 5) GC method [9]. This method determines whether there is a causal link from the *i*-th brain region to the *j*-th brain region based on the Granger causality test.
- 6) LiNGAM method [13]. The core idea of this method is that the blood oxygen signal data of each brain region is a linear combination of the signals of all other brain regions with no time delay. The regression coefficients are regarded as the edge weights of the brain network.
- 7) TE method [36]. This method estimates the direction of information flow between brain regions based on the concept of Shannon entropy.

B. Experimental Settings

For a fair comparison, the same feature selection strategy and classifier are used to test the performance of each brain network model. Specifically, we take the weight of network edges as features, and employ the recursive feature elimination (RFE) [37]

	Mathada	NC vs. eMCI			NC vs. LMCI				
	Wiethous	ACC	SEN	SPE	F1	ACC	SEN	SPE	F1
Correlation based methods	PC	68.81	66.67	70.69	66.67	70.33	76.47	62.50	74.29
	SR	67.89	62.75	72.41	64.65	68.13	72.55	62.50	71.84
	LR	69.72	66.67	72.41	67.33	74.73	80.39	67.50	78.10
	SLR	59.63	52.94	65.52	55.10	64.84	78.43	47.50	71.43
Causality based methods	GC	62.39	50.98	72.41	55.91	67.03	76.47	55.00	72.22
	LiNGAM	54.13	47.06	60.34	48.98	58.24	68.63	45.00	64.81
	TE	56.88	56.86	56.90	55.24	64.84	76.47	50.00	70.91
	$ETLN^E$	74.31	70.59	77.59	72.00	78.02	82.35	72.50	80.77
	\mathbf{ETLN}^T	71.56	64.71	77.59	68.04	72.53	84.31	57.50	77.48
	Methode	eMCI vs. LMCI			NC vs. eMCI vs. LMCI				
	Methods	ACC	SEN	SPE	F1	ACC	SEN	SPE	F1
	PC	65.31	79.31	45.00	73.02	70.92	56.15	78.09	55.96
Correlation based methods	SR	64.29	75.86	47.50	71.54	67.34	51.06	75.28	50.85
	LR	67.35	75.86	55.00	73.33	69.13	51.42	76.05	50.95
	SLR	62.24	77.59	40.00	70.87	60.63	38.27	69.09	37.06
Causality based methods	GC	61.22	75.86	40.00	69.84	63.76	44.26	71.88	44.53
	LiNGAM	58.16	75.86	32.50	68.22	59.28	37.78	68.99	37.81
	TE	51.02	63.79	32.50	60.66	59.28	38.35	69.13	38.39
	$ETLN^E$	73.47	82.76	60.00	78.69	74.05	59.74	79.91	60.26
	$\mathbf{E}\mathbf{T}\mathbf{L}\mathbf{N}^{T}$	69.39	84.48	47.50	76.56	70.92	57.13	78.35	56.47

TABLE II PERFORMANCE COMPARISON OF DIFFERENT METHODS

strategy to perform feature selection. Finally, a support vector machine (SVM) [38] is utilized for classification. The number of selected features by RFE is set to 1000. The parameters of SVM are set as follows: the number of iterations is 1000, the kernel is 'linear,' and the regularization parameter is set to 1.0.

In this study, four classification tasks (i.e., NC vs. eMCI, NC vs. LMCI, eMCI vs. LMCI and NC vs. eMCI vs. LMCI) are conducted to evaluate the performance of the proposed method. Prediction accuracy (ACC), sensitivity (SEN), specificity (SPE), and F1 score are used as evaluation metrics. It is worth noting that our method can obtain two kinds of brain network models, including the effective connectivity network (ECN) and the temporal-lag connectivity network (TCN), while other methods can only obtain one kind of brain network model. For a fair comparison, instead of fusing the features of both, we feed ECN and TCN into the classifier separately to evaluate their classification performance. Besides, we conduct a standard 10-fold cross-validation strategy to evaluate the classification performance of all brain network models.

C. Classification Results

The comparison results of all methods are summarized in Table II, and the best scores are highlighted in bold. From Table II, it can be seen that our method achieves the best performance on all three tasks, which demonstrates the effectiveness of our method. Furthermore, we can find that SLR has the worst performance among the four correlation-based methods, while the classification performance of the other three methods is roughly the same. The reasons for this can be attributed to the following two points. First, SLR introduces two types of constraints at the same time, which makes the constructed brain network too sparse. As a result, it is difficult to extract effective features for the classifier training, so SLR shows the worst classification performance. Secondly, sparsity or low-rank constraints may not be very effective in eliminating the effects of confounding factors. Compared with PC, although SR and LR remove some pseudo-functional connections, the number of effective features retained by the two methods may be roughly the same as the number of effective features extracted from the brain networks constructed by the PC method. Therefore, their classification performance in these three tasks is roughly the same.

However, the classification results of GC, LiNGAM, and TE are not satisfactory in comparison to correlation-based methods. The reasons for the poor performance of the three methods are different. The reasons for the failure of GC mainly lie in two points: on the one hand, the brain network it constructs is a binary graph, which can obtain fewer effective features than the weight graph; on the other hand, it is because the temporal-lag values between all brain regions are set to a fixed value, which is obviously contrary to the actual situation. Similarly, the reasons for the failure of LiNGAM can also be attributed to two points: one is the addition of acyclic constraint to the brain network, resulting in the constructed brain network being too sparse and not having enough features to train the classifier; the other is that it can only model the linear relationship between brain regions, and cannot fully explore the deeper nonlinear relationship. The reason for the failure of TE is similar to that of GC. The range of temporal-lag values for TE is at least one full TR, which is clearly inconsistent with previous findings [19], [32].

D. Ablation Study

The main contribution of this paper is to design a dual-branch model, which can effectively estimate causal and temporal-lag effects between brain regions. Besides, in order to obtain more accurate estimation results, we also introduce three constraint rules to guide the modeling of the brain network. To verify whether each of these innovative components contributes to the excellent performance, we design four degraded networks in the ablation study, including 1) we remove the temporal-lag



Fig. 5. Recognition performance for ablation studies.

inference branch, denoted "ETLN_LagGate," 2) we remove the local consistency loss from the loss function of ETLN, denoted "ETLN_local," 3) we replace the adaptive mechanism with a fixed threshold of 0.5, denoted "ETLN_adaptive," and 4) we remove the spatial constraint loss from the loss function of ETLN, denoted "ETLN_spatial".

Fig. 5 show the experimental results of the ablation study. It can be seen that ETLN_LagGate obtains the worst performance among all variant methods due to ignoring the effect of temporal-lag. This result suggests that considering the effect of temporal-lag helps to construct better brain network models. Comparing the classification results of the other three variants with ETLN, it can be found that the variant that removes the spatial constraint mechanism has the most performance degradation, followed by the local consistency constraint mechanism, and the adaptive mechanism has the least impact on the result. This result shows that the three constraint mechanisms can be helpful to improve the classification performance, but the spatial constraint mechanism is more helpful, followed by the consistency constraint mechanism, and the adaptive mechanism has the least impact by the consistency constraint mechanism is more helpful, followed by the consistency constraint mechanism, and the adaptive mechanism has the least impact by the consistency constraint mechanism.

As we all known, in order to achieve whole-brain coverage, complete brain data (i.e., volume) is obtained by sequentially and repeatedly acquiring images of individual slices. This will result in a time offset between slices within the 3D volume. The purpose of slice time correction is to restore time-shifted slices in the same volume to the same moment. Without slice time correction, this would result in a time offset between slices within the same volume. However, the time interval of corresponding slices between adjacent volumes is fixed (equal to TR), which seems to have little effect on our research objective: estimating the average temporal-lag effect between volumes. In order to further study the impact of this preprocessing step on the results, we have performed a new set of comparative experiments. Specifically, we adopt two preprocessing pipelines for the obtained fMRI data, the only difference between them is whether to perform slice time correction. We denote the experiment without slice time correction as "ETLN ST".

Fig. 6 shows the classification results of the comparative experiment on four tasks. From the results, it can be seen that



Fig. 6. Classification results of the comparative experiment under four tasks.

the classification accuracy of the corresponding brain network models in both sets of experiments is relatively close to each task, and there are no significant differences. This indicates that the slice time correction operation has little effect on the classification results. For this result, we think there may be two reasons. On the one hand, as described in the previous paragraph, the slice time correction operation does not conflict with the objective of this research, and should have little effect on the results in theory. On the other hand, ETLN, the brain network modeling algorithm proposed in this paper, is a deep learning method with powerful fitting ability. Therefore, fMRI data can either perform or not perform the slice time correction operation when the ETLN is executed.

As many comparison methods adopted a standard preprocessing pipeline (including slice time correction) [4], [21], [36], we also apply the pipeline to our method for a fair comparison. Finally, the results reported in the rest of the paper are obtained based on the standard preprocessing pipeline.

VI. DISCUSSION

A. Visualization of the Optimization Process

In order to verify whether the GAN model ETLN constructed in this paper has an overfitting problem, we show the visualization results of the brain network with epochs in Fig. 7. The first column is the initialization of ECN and TCN, both of which are a matrix of size 116×116 . They both set the diagonal elements to 0 and the other elements to 1. The last column, namely epoch = 1000, is the finally obtained brain causal and temporal-lag networks. It can be seen from Fig. 7 that when the epoch reaches 600, the GAN model ETLN has converged, and there is almost no change in the follow-up.

Comparing ECN and TCN, it can be found that the two maintain good local consistency under different epochs, and there are some differences at the same time. This may be attributed to the local consistency loss introduced in the loss function. Besides, it can be seen from Fig. 7 that in both ECN and TCN, the value of a small area is significantly greater than that of the adjacent areas. This may because these brain regions are the hubs of the brain, and they undertake more tasks of information



Fig. 7. Visualization results of the average causal and temporal-lag brain networks of all normal subjects with different epochs.



Fig. 8. Visualization results of the most discriminating causal and lag patterns among the three tasks.

transmission and forwarding. This not only leads to a high causal effect across these brain regions, but also significantly increases the time required to transmit information among them.

B. Most Discriminative Patterns

Fig. 8 displays the top 20 most discriminative causal and temporal-lag patterns across the three tasks, respectively. There are four points to note about the meaning of the circos graph. First, the color of each arc in the circos graph is randomly assigned. Second, the wider the width of each arc, the more important the corresponding connection. Then, the motion direction of the ball in each arc represents the causal relationship between two brain regions (from cause to effect). Finally, the value corresponding to each connection represents the average causal effect or temporal-lag value for the two groups of participants. For example, for the circos graph in the upper left corner, the value corresponding to the link SFGdor.R \rightarrow Vermis10 is 0.296 / 0.249. This means that the average causal effect of the connection in NC group is 0.296, and the average causal effect in eMCI group is 0.249. For the circos graph in the lower right corner, the value corresponding to the link Vermis3 \rightarrow Vermis45

Comparing the most discriminative causal patterns identified by the three tasks, we can draw some interesting observations. First, many brain regions were jointly recognized as potential biomarkers for dementia identification by the three tasks, including the Vermis10, the Vermis12, the left superior parietal gyrus (SPG.L), the right caudate nucleus (CAU.R) and the amygdala (AMYG). This suggests that these brain regions may be potential biomarkers for dementia recognition. Second, all three tasks identified Vermis10 as the brain region containing the most discriminative information. The reason for this may be that this brain region is relatively close to the corpus callosum, and much communication information between brain regions needs to be further transmitted through this brain region. Third, it can be found that the classifier tends to give higher discriminative weights to those functional connections that are spatially distant, especially the connections between the brain and cerebellar regions. We think this phenomenon is reasonable. The transmission of information flow between distant brain regions may need to span multiple brain regions, and abnormal alteration in any node in the transmission path may lead to changes in the strength of functional connectivity.

Similarly, comparing the most discriminative lag patterns identified by the three tasks, we can also find some interesting phenomena. Likewise, many brain regions were also jointly identified as biomarkers by the three tasks, including the Vermis10, the Vermis9, the Vermis3 and the left inferior temporal gyrus (ITG.L). But unlike the causal patterns, the most discriminative brain region of the three lag patterns was not consistent. It can also be observed that the classifier prefers to assign higher discriminative weights to those functional connections that are farther away.

Comparing the causal pattern and the lag pattern of the same task, many functional connections are jointly identified as discriminative features. It is clear that these functional connections may play a pivotal role in the determination of classification results. Specifically, for the task NC vs. eMCI, the path from SFGdor.R to Vermis10 and the path from SMG.L to Vermis10 are jointly identified as abnormal connections. For the task NC vs. LMCI, the common abnormal connections include the path from CAU.L to CRBL9.R, the path from CAU.L to CRBL9.L and the path from HIP.R to Vermis6. The reason for this may be due to the introduction of the local consistency mechanism as well as the spatial constraint mechanism. For those brain regions without causality, the local consistency mechanism constrains the temporal-lag values between them to zero. For those brain regions with causality, the spatial constraint mechanism constrains the data distribution of temporal-lag values between them. They both guarantee the local similarity between causal and lag patterns.

In addition, it can be observed that all six circos graphs contain some connections in which the causal effects seem to be strengthened or the temporal-lag values are shortened. This result seems to be contrary to the clinical observations, but we believe that this situation is reasonable. For example, suppose there are two paths from brain region A to brain region B, one with a shorter total distance and one with a longer total distance. Previously, much information was transmitted over the shorter path. If the shorter path is abnormal, much information will be transmitted over the longer path. This may lead to the enhancement of the causal effect between the related brain regions of the longer path. As the interactions between brain regions belonging to the longer path become active, this may lead to shorter temporal-lag values between them. Nevertheless, our results are consistent with clinical experience in terms of the overall change trend. For example, compared with the NC group, the causal effect of 10 connections in the eMCI group was weakened, and the temporal-lag value of 10 connections was prolonged. The LMCI group was more severe, the causal effect of 14 connections was weakened, and the temporal-lag value of 16 connections was prolonged. This result is consistent with clinical experience, that is, the more severe the disease, the more severe the cognitive decline.

VII. CONCLUSION

In this paper, we propose a novel framework for inferring the causal effects and the temporal-lag values between brain regions. The first point is the design of the network structure, which embeds the target of the solution into the network model as the parameters to be learned. The second point is the introduction of three mechanisms to guide the modeling of brain networks. The whole network is trained in an end-to-end manner and achieved excellent performance on the ADNI dataset. The proposed method not only improves the classification performance, but also provides a new solution for the inference of causal and temporal-lag relationships among large-scale nodes.

REFERENCES

- J. L. Ji, M. Spronk, K. Kulkarni, G. Repovš, A. Anticevic, and M. W. Cole, "Mapping the human brain's cortical-subcortical functional network organization," *Neuroimage*, vol. 185, pp. 35–57, 2019.
- [2] J. M. Shine and R. A. Poldrack, "Principles of dynamic network reconfiguration across diverse brain states," *Neuroimage*, vol. 180, pp. 396–405, 2018.
- [3] J. Huang, L. Zhou, L. Wang, and D. Zhang, "Attention-diffusion-bilinear neural network for brain network analysis," *IEEE Trans. Med. Imag.*, vol. 39, no. 7, pp. 2541–2552, Jul. 2020.
- [4] F. Zhao, H. Zhang, I. Rekik, Z. An, and D. Shen, "Diagnosis of autism spectrum disorders using multi-level high-order functional networks derived from resting-state functional MRI," *Front. Hum. Neurosci.*, vol. 12, 2018, Art. no. 184.
- [5] J. Peng, P. Wang, N. Zhou, and J. Zhu, "Partial correlation estimation by joint sparse regression models," *J. Amer. Stat. Assoc.*, vol. 104, no. 486, pp. 735–746, 2009.
- [6] K. J. Friston, "Functional and effective connectivity in neuroimaging: A synthesis," *Hum. Brain Mapping*, vol. 2, no. 1/2, pp. 56–78, 1994.
- [7] Y. Li et al., "Novel effective connectivity network inference for MCI identification," in *Proc. Int. Workshop Mach. Learn. Med. Imag.*, 2017, pp. 316–324.
- [8] N. Chen, M. Guo, Y. Li, X. Hu, Z. Yao, and B. Hu, "Estimation of discriminative multimodal brain network connectivity using messagepassing-based nonlinear network fusion," *IEEE/ACM Trans. Comput. Biol. Bioinf.*, early access, Dec. 23, 2021, doi: 10.1109/TCBB.2021.3137498.
- [9] S. L. Bressler and A. K. Seth, "Wiener–Granger causality: A well established methodology," *Neuroimage*, vol. 58, no. 2, pp. 323–329, 2011.
- [10] B. J. Frey and N. Jojic, "A comparison of algorithms for inference and learning in probabilistic graphical models," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 27, no. 9, pp. 1392–1416, Sep. 2005.

- [11] S. Shimizu, P. O. Hoyer, A. Hyvärinen, A. Kerminen, and M. Jordan, "A linear non-Gaussian acyclic model for causal discovery," *J. Mach. Learn. Res.*, vol. 7, no. 10, pp. 2003–2030, 2006.
- [12] S. M. Smith et al., "Network modelling methods for fMRI," *Neuroimage*, vol. 54, no. 2, pp. 875–891, 2011.
- [13] K. J. Friston, "Functional and effective connectivity: A review," Brain Connect, vol. 1, no. 1, pp. 13–36, 2011.
- [14] N.T. Markov et al., "A weighted and directed interareal connectivity matrix for macaque cerebral cortex," *Cereb. Cortex*, vol. 24, no. 1, pp. 17–36, 2014.
- [15] A. K. Seth, A. B. Barrett, and L. Barnett, "Granger causality analysis in neuroscience and neuroimaging," *J. Neurosci*, vol. 35, no. 8, pp. 3293–3297, 2015.
- [16] J. Ramsey, M. Glymour, R. Sanchez-Romero, and C. Glymour, "A million variables and more: The fast greedy equivalence search algorithm for learning high-dimensional graphical causal models, with an application to functional magnetic resonance images," *Int. J. Data Sci. Anal.*, vol. 3, no. 2, pp. 121–129, 2017.
- [17] A. Mitra, A. Z. Snyder, C. D. Hacker, and M. E. Raichle, "Lag structure in resting-state fMRI," *J. Neurophysiol.*, vol. 111, no. 11, pp. 2374–2391, 2014.
- [18] S. K. Bandt et al., "Connectivity strength, time lag structure and the epilepsy network in resting-state fMRI," *Neuroimage Clin.*, vol. 24, 2019, Art. no. 102035.
- [19] V. Raatikainen et al., "Dynamic lag analysis reveals atypical brain information flow in autism spectrum disorder," *Autism Res.*, vol. 13, no. 2, pp. 244–258, 2020.
- [20] Y. Zhang, H. Zhang, X. Chen, M. Liu, X. Zhu, and D. Shen, "Inter-subject similarity guided brain network modeling for MCI diagnosis," in *Proc. Mach. Learn. Med. Imag. Workshop*, 2017, pp. 168–175.
- [21] L. Qiao, H. Zhang, M. Kim, S. Teng, L. Zhang, and D. Shen, "Estimating functional brain networks by incorporating a modularity prior," *Neuroimage*, vol. 141, pp. 399–407, 2016.
- [22] X. Wang, R. Wang, F. Li, Q. Lin, X. Zhao, and Z. Hu, "Large-scale Granger causal brain network based on resting-state fMRI data," *Neuroscience*, vol. 425, pp. 169–180, 2020.
- [23] C.-Y. Wee, P.-T. Yap, D. Zhang, L. Wang, and D. Shen, "Group-constrained sparse fMRI connectivity modeling for mild cognitive impairment identification," *Brain Struct. Function*, vol. 219, no. 2, pp. 641–656, 2014.
- [24] N. Z. Bielczyk, S. Uithol, T. V. Mourik, P. Anderson, J. C. Glennon, and J. K. Buitelaar, "Disentangling causal webs in the brain using functional magnetic resonance imaging: A review of current approaches," *Netw. Neurosci.*, vol. 3, no. 2, pp. 237–273, 2019.

- [25] J. L. Ringo, R. W. Doty, S. Demeter, and P. Y. Simard, "Time is of the essence: A conjecture that hemispheric specialization arises from interhemispheric conduction delay," *Cereb. Cortex*, vol. 4, no. 4, pp. 331–343, 1994.
- [26] F. D. Pasquale et al., "Temporal dynamics of spontaneous MEG activity in brain networks," *Proc. Nat. Acad. Sci.*, vol. 107, no. 13, pp. 6040–6045, 2010.
- [27] A. Mitra, A. Z. Snyder, J. N. Constantino, and M. E. Raichle, "The lag structure of intrinsic activity is focally altered in high functioning adults with autism," *Cereb. Cortex*, vol. 27, no. 2, 2015, Art. no. bhv294.
- [28] J. Wen et al., "Convolutional neural networks for classification of Alzheimer's disease: Overview and reproducible evaluation," *Med. Image Anal.*, vol. 63, 2020, Art. no. 101694.
- [29] A. Carballedo et al., "Functional connectivity of emotional processing in depression," J. Affect. Disord., vol. 134, no. 1/3, pp. 272–279, 2011.
- [30] D. Kalainathan, O. Goudet, I. Guyon, D. Lopez-Paz, and M. Sebag, "Structural agnostic modeling: Adversarial learning of causal graphs," J. Mach. Learn. Res., vol. 23, no. 219, pp. 1–62, 2022.
- [31] V. Raatikainen et al., "Combined spatiotemporal ICA (stICA) for continuous and dynamic lag structure analysis of MREG data," *Neuroimage*, vol. 148, pp. 352–363, 2017.
- [32] Z. Rajna, J. Kananen, A. Keskinarkaus, T. Seppänen, and V. Kiviniemi, "Detection of short-term activity avalanches in human brain default mode network with ultrafast MR encephalography," *Front. Hum. Neurosci.*, vol. 9, 2015, Art. no. 448.
- [33] R. J. Meszlényi, P. Hermann, K. Buza, V. Gál, and Z. Vidnyánszky, "Resting state fMRI functional connectivity analysis using dynamic time warping," *Front. Neurosci.*, vol. 11, 2017, Art. no. 75.
- [34] L. J. van der Knaap and I. J. van der Ham, "How does the corpus callosum mediate interhemispheric transfer? a review," *Behav. Brain Res.*, vol. 223, no. 1, pp. 211–221, 2011.
- [35] B. Yu et al., "Aberrant whole-brain functional connectivity and intelligence structure in children with primary nocturnal enuresis," *PLoS One*, vol. 8, no. 1, 2013, Art. no. e51924.
- [36] F. Parente and A. Colosimo, "Modelling a multiplex brain network by local transfer entropy," *Sci. Rep.*, vol. 11, no. 1, pp. 1–23, 2021.
- [37] I. Guyon, J. Weston, S. Barnhill, and V. Vapnik, "Gene selection for cancer classification using support vector machines," *Mach. Learn.*, vol. 46, no. 1, pp. 389–422, 2002.
- [38] C. Cortes and V. Vapnik, "Support-vector networks," *Mach. Learn.*, vol. 20, no. 3, pp. 273–297, 1995.