

Ordinal Pattern: A New Descriptor for Brain Connectivity Networks

Daoqiang Zhang, Jiashuang Huang, Biao Jie, Junqiang Du, Liyang Tu, and Mingxia Liu¹

Abstract—Brain connectivity networks based on magnetic resonance imaging (MRI) or functional MRI (fMRI) data provide a straightforward way to quantify the structural or functional systems of the brain. Currently, there are several network descriptors developed for representing and analyzing brain connectivity networks. However, most of them are designed for unweighted networks, regardless of the valuable weight information of edges, or do not take advantage of the ordinal relationship of weighted edges (even though they are designed for weighted networks). In this paper, we propose a new network descriptor (*i.e.*, *ordinal pattern* that contains a sequence of weighted edges) for brain connectivity network analysis. Compared with previous network properties, the proposed ordinal patterns cannot only take advantage of the weight information of edges but also explicitly model the ordinal relationship of weighted edges in brain connectivity networks. We further develop an ordinal pattern-based learning framework for brain disease diagnosis using resting-state fMRI data. Specifically, we first construct a set of brain functional connectivity networks, where each network is corresponding to a particular subject. We then develop an algorithm to identify ordinal patterns that frequently appear in brain connectivity networks of patients and normal controls. We further perform discriminative ordinal pattern selection and extract feature representations for subjects based on the selected ordinal patterns, followed by a learning model for automated brain disease diagnosis. Experimental results on both Alzheimer's Disease Neuroimaging Initia-

tive and attention deficit hyperactivity disorder-200 data sets demonstrate that our method outperforms the several state-of-the-art approaches in the tasks of disease classification and clinical score regression.

Index Terms—Connectivity network, network descriptor, brain disease diagnosis, classification, regression.

I. INTRODUCTION

OUR brain is a complex network, containing a large number of structurally/functionally interconnected regions. Through a continuous integration of information across different regions of the brain, the structural/functional communication between brain regions plays a key role in the complicated cognitive process. Thus, it is essential to examine the structural/functional connectivity in the brain, which may provide interesting insights into the core organization of the brain [1], [2]. In the context of neuroimaging, magnetic resonance imaging (MRI) and functional magnetic resonance imaging (fMRI) provide efficient and non-invasive solutions to map structural/functional connectivity patterns of the brain [3]–[9], respectively. Based on fMRI data, recent studies [10]–[14] have shown that brain functional connectivity networks are effective in helping understand the pathology of brain diseases, *e.g.*, Alzheimer's disease (AD), mild cognitive impairment (MCI), and attention deficit hyperactivity disorder (ADHD).

In a brain structural/functional connectivity network, we generally treat each anatomical brain region as a particular node and regard the structural/functional connections between brain regions as edges. To facilitate the automated diagnosis for brain diseases, one usually first constructs a brain structural/functional connectivity network for each subject based on MRI/fMRI and then extract network properties (or descriptors) as feature representation to represent this subject for subsequent learning models. Although many properties have been proposed to describe brain structural/functional connectivity networks [10], [11], [13], most of them are designed for unweighted networks (ignoring edge weights), or do not take advantage of the ordinal relationship of weighted edges (even though they are designed for weighted networks).

Recently, diffusion model of communication is reported to be more appropriate for the analysis of brain networks [15]–[17]. Specifically, a diffusion process means that the flow of information in a network is dispersive by propagating simultaneously along multiple “fronts,” rather than pointing to a single destination. With diffusion,

Manuscript received December 6, 2017; revised January 18, 2018 and January 22, 2018; accepted January 23, 2018. Date of publication January 25, 2018; date of current version June 30, 2018. This work was supported in part by the National Natural Science Foundation of China under Grant 61703301, Grant 61573023, Grant 61422204, and Grant 61473149, in part by the University Science and Technology Project of Shandong Province under Grant J17KA086, and in part by the NUA Fundamental Research Funds under Grant NE2013105. (Corresponding author: Mingxia Liu; Daoqiang Zhang.)

D. Zhang, J. Huang, J. Du, and L. Tu are with the Department of Computer Science and Engineering, Nanjing University of Aeronautics and Astronautics, Nanjing 210016, China (e-mail: dqzhang@nuaa.edu.cn).

B. Jie is with the Department of Computer Science and Engineering, Nanjing University of Aeronautics and Astronautics, Nanjing 210016, China, and also with the Department of Computer Science and Technology, Anhui Normal University, Wuhu 241003, China.

M. Liu is with the Department of Computer Science and Engineering, Nanjing University of Aeronautics and Astronautics, Nanjing 210016, China, and also with the Department of Information Science and Technology, Taishan University, Tai'an 271000, China (e-mail: mingxialiu@nuaa.edu.cn).

Supplementary materials are available in the supplementary files/multimedia tab.

Color versions of one or more of the figures in this paper are available online at <http://ieeexplore.ieee.org>.

Digital Object Identifier 10.1109/TMI.2018.2798500

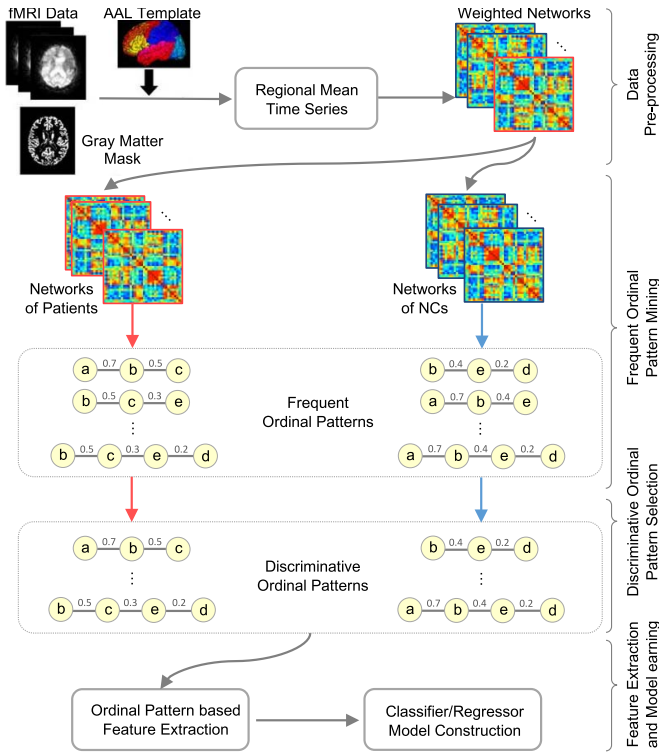


Fig. 1. Illustration of our proposed ordinal pattern based learning framework for brain disease diagnosis, including 1) imaging data pre-processing, 2) frequent ordinal pattern mining, 3) selection of discriminative ordinal patterns, and 4) feature extraction and model construction.

information sent from a single node can reach any of a number of destinations via any of multiple paths that are possibly biased by edge weights [15]. That is, the weighted edges in these paths may have the ordinal relationship in the diffusion model of communication for brain network analysis. Hence, descriptors that consider both the weight information and the ordinal relationship among weighted edges can represent the network more precisely, and thus may bring better learning performance for brain connectivity network based methods.

To this end, we propose a new network descriptor (called *ordinal pattern* [18]) for representing brain connectivity networks, and also apply it to computer-aided brain disease diagnosis. In an ordinal pattern, we construct a sequence of weighted edges for a brain connectivity network, to make use of both edge weights and the ordinal relationship among weighted edges. Compared with conventional network properties, our proposed ordinal patterns are directly designed on weighted networks, which can naturally take advantage of the weights of edges and the ordinal relationship among weighted edges in brain connectivity networks. Furthermore, we develop an ordinal pattern based brain disease diagnosis framework using resting-state fMRI (rs-fMRI) data, with a schematic diagram given in Fig. 1. Specifically, we first pre-process rs-fMR images and construct a set of brain connectivity networks, where each network is corresponding to a particular subject. We then propose to identify the frequent ordinal patterns from the network sets of patients and normal controls (NCs) separately. We further select the most discriminative

ordinal patterns and extract features for brain connectivity networks (*w.r.t.* subjects). Finally, we construct a classification/regression model using our ordinal pattern based representations. The preliminary work of this method was reported on MICCAI [18]. In this journal paper, we have offered new contributions in the following aspects: 1) investigating the learned ordinal pattern based feature representations, 2) describing the ordinal pattern mining algorithm and releasing the code, 3) applying the proposed method to regression tasks for clinical scores, and 4) studying the influence of three parameters on the performance of our method.

The major contributions of this paper are three-fold. *First*, a new network descriptor (ordinal pattern) is developed to represent brain connectivity networks, as well as a mining algorithm to identify the most discriminative ordinal patterns in a data-driven manner. To the best of our knowledge, the proposed ordinal pattern is among the first attempts to utilize the ordinal relationship among weighted edges in brain connectivity networks. *Second*, we develop an ordinal pattern based brain diagnosis framework based on rs-fMRI data, where the ordinal pattern based features are extracted from the brain connectivity networks for representing subjects. *Third*, we evaluate the proposed method on two datasets, with results demonstrating the effectiveness of our proposed method.

The rest of the paper is organized as follows. In Section II, we first briefly review relevant studies. We then introduce materials and describe the proposed method in Section III. In Section IV, we present the experimental settings and results, and investigate the influence of parameters. In Section V, we further study the frequently identified ordinal patterns and the impact of the ordinal relationship, and present the limitations of our method as well as possible research directions. We conclude this paper in Section VI.

II. RELATED WORK

Thanks to the development of neuroimaging technologies, recent studies have shown that brain structural/functional connectivity networks constructed on structural/functional MRI have great promise in revealing the pathology of brain diseases [10]–[14], [19], [20]. In a typical computer-aided brain disease diagnosis system, we usually first build a particular brain connectivity network for each subject using structural/functional MRI data, and then extract measures from the brain connectivity network for the subsequent classification/regression task. Previous studies have developed various network properties for representing brain structural/functional connectivity networks, such as node strength [10], clustering coefficients [11], and subnetworks [13]. For instance, Wee *et al.* [11] extracted clustering coefficients from brain connectivity networks to build classifiers for AD diagnosis. Durante *et al.* [21] exploited the network information through low-rank factorizations to identify patients characterized by the AD disease. Fei *et al.* [13] mined discriminative subnetworks from brain connectivity networks for MCI conversion prediction. Solmaz *et al.* [22] applied the bag-of-words based approach to represent each subject as a histogram for ADHD classification.

However, existing properties of brain connectivity networks commonly ignore the valuable weight information of edges. Actually, in the construction process of brain connectivity networks, one usually assigns a weight to each edge to quantitatively measure the connectivity strength between a pair of nodes (*i.e.*, a pair of brain regions). For simplicity, lots of studies often first adopt a simple thresholding strategy to transform the original weighted brain connectivity networks into unweighted ones and then extract properties from such unweighted networks as feature representations for subjects. Recently, Jie *et al.* [12] proposed a multi-thresholding strategy for extracting properties from brain connectivity networks using fMRI data. Specifically, they first thresholded each functional connectivity network using different thresholds, and then extracted clustering coefficients of each region-of-interest (ROI) with the remaining ROIs from each thresholded connectivity network as features for feature selection and MCI conversion prediction. Although this method can achieve better performance than single-thresholding based approaches, it still cannot take advantage of the weight information of edges.

Recently, several descriptors based on weighted networks have been proposed for brain network analysis. For instance, Zhou *et al.* [23] used connectivity strength between ROIs to analyze divergent network connectivity changes in behavioral variant frontotemporal dementia and AD. Abraham *et al.* [24] and Casanova *et al.* [25] used the correlation feature between signals of ROIs for the prediction of autism spectrum disorder and the classification of sex, respectively. Challis *et al.* [26] calculated the covariance between signals of ROIs as features of subjects for AD/MCI classification. Wang *et al.* [27] extracted regional homogeneity (*i.e.*, Kendall's coefficient concordance) from rs-fMRI signals for ADHD classification. Dey *et al.* [28] computed a local descriptor comprising of a set of properties of the nodes for ADHD detection. Jie *et al.* [29] integrated local connectivity properties (*i.e.*, local weighted clustering coefficients) and global topological properties for MCI classification. Chung *et al.* [30] defined a heat-kernel-based feature representation to capture salient network properties that can be used to discriminate between different network topologies, and applied it to the analysis of preterm babies.

Although the network descriptors mentioned above are designed on the weight connectivity networks, most of them ignore the ordinal relationship of weighted edges in connectivity networks. Recent studies have shown that many brain diseases (*e.g.*, AD, MCI, and ADHD) are not only related to single brain regions but also corresponding to local modular structures [31]. For instance, many pieces of evidence [32]–[34] have declared that the normal local modular structures of brain connectivity networks are disrupted in AD and MCI. In the research of ADHD, altered functional networks were demonstrated in the brains of ADHD patients when compared with the NCs, where increased local efficiencies combined with a decreasing tendency in global efficiencies are found in ADHD [35], [36]. Unfortunately, it is hard to precisely describe such disorders in local structures caused by brain diseases using existing properties that ignores the ordinal relationship of edges in brain connectivity networks. To this end, we propose *ordinal patterns* to describe brain

TABLE I
DEMOGRAPHIC AND CLINICAL INFORMATION OF SUBJECTS FROM 2 DATASETS. WE REPORT CORRESPONDING VALUES AS MEAN \pm STANDARD DEVIATION; M/F: MALE/FEMALE; MMSE: MINI-MENTAL STATE EXAMINATION. ALL THE RECRUITED SUBJECTS WERE DIAGNOSED BY EXPERT CONSENSUS PANELS

Datasets	Class	Subject #	Age	Gender (M/F)	MMSE
ADNI	AD	34	72.5 \pm 7.1	16/18	21.04 \pm 3.55
	MCI	99	71.1 \pm 7.4	47/52	27.24 \pm 2.12
	NC	50	75.0 \pm 6.9	21/29	28.60 \pm 2.72
ADHD-200	ADHD	118	11.2 \pm 2.7	25/93	-
	NC	98	12.2 \pm 2.1	51/47	-

connectivity networks, which can model the ordinal relationship of weighted edges. We further develop an ordinal pattern based brain disease diagnosis framework, by identifying the most discriminative ordinal patterns in a data-driven manner.

III. MATERIALS AND METHOD

In this section, we first introduce the materials and the image pre-processing approach used in this study. We then introduce the definitions of ordinal pattern and frequent ordinal pattern, and the proposed ordinal pattern based learning framework for computer-aided brain disease diagnosis.

A. Materials and Image Pre-Processing

1) Data Acquisition: Two datasets are used in this study. The first dataset contains all subjects with rs-fMRI data from the baseline Alzheimer's Disease Neuroimaging Initiative (ADNI) database [37], including 34 AD patients, 99 MCI patients, and 50 NCs. Data acquisition for this dataset was performed as follows: 34 axial slices, matrix size 64 \times 64, 4 mm slice thickness, 0 mm spacing, 256 mm FOV, TR = 2 s, TE = 32 ms, flip angle = 77. The mini-mental state examination (MMSE) scores for subjects in ADNI are available. Another dataset is ADHD-200 with the Athena¹ pre-processed rs-fMRI data acquired from the New York University Child Study Center (NYU), containing 118 ADHD patients and 98 NCs. The acquisition of rs-fMRI data in ADHD-200 was performed as follows: 47 axial slices, matrix size 49 \times 58, 4 mm slice thickness, 0 mm spacing, 240 mm FOV, TR = 2 s, TE = 15 ms, flip angle = 90, voxel size 3 \times 3 \times 4 mm. The demographic and clinical information of subjects is listed in Table I.

2) Image Pre-Processing: Imaging pre-processing is performed using Statistical Parametric Mapping (SPM) software package.² Specifically, we first removed the first ten acquired fMR images of each subject in ADNI, and the first four fMR images of each subject in ADHD-200. Then, we performed slice timing correction for the remaining images, followed by re-aligning the remaining images to the first volume for head motion correction. Here, we only used the blood-oxygen-level-dependent (BOLD) signals extracted from gray matter (GM) tissue to construct functional connectivity networks, since the regions of ventricles and white matter (WM) contain a relatively high proportion of noise caused by the cardiac

¹<http://www.nitrc.org/plugins/mwiki/index.php/neurobureau:AthenaPipeline>

²<http://www.fil.ion.ucl.ac.uk/spm/>

and respiratory cycles [38]. Accordingly, we first segmented the T1-weighted MR image of each subject into GM, WM, and cerebrospinalfluid (CSF). Then, we used the GM tissue of each subject to mask their corresponding fMRI images to eliminate the possible effect of WM and CSF in the fMRI time series.

Then, we co-registered the first scan of remaining images to the T1-weighted MR image of the same subject, and the estimated transformation was also applied to the remaining fMRI scans of the same subject. We then partitioned the brain space of fMRI scans into 90 regions of interesting (ROIs) by warping the Automated Anatomical Labeling (AAL) [39] template to the subject space by using a deformable registration method [40]. For each subject, the mean time series of each ROI was computed by averaging the GM-masked fMRI time series over all voxels in the particular ROI. For each subject in the ADNI dataset, the GM-masked mean time series of each region is band-pass filtered within the frequency interval of $[0.025 \text{ Hz}, 0.100 \text{ Hz}]$, and the spatial smoothing was performed using a 6 mm full width at half maximum (FWHM) Gaussian filter. For each subject in the ADHD-200 dataset, the range of frequencies for band-pass filtering is $[0.009 \text{ Hz}, 0.080 \text{ Hz}]$, and the kernel for spatial smoothing is a 6 mm full width at half maximum (FWHM) Gaussian filter.

3) Network Construction: For each subject, the GM-masked fMRI time series of all voxels within each ROI were averaged to be the mean time series of the particular ROI. Denote t as the length of time series, and $\mathbf{X} = [x_1, x_2, \dots, x_q, \dots, x_Q] \in \mathbf{R}^{Q \times t}$ as a training subject with a total of Q ROIs, where $x_q \in \mathbf{R}^t$ is the regional mean time series of the q -th ROI. With each ROI as a node, we then computed the Pearson correlation coefficients between the i -th and j -th ROIs as the connectivity weight for the nodes i and j ($i, j = 1, \dots, Q$). Thus, we can construct a fully-connected weighted network for each subject. Given multiple subjects, we finally obtained a set of brain networks.

B. Ordinal Pattern

Definition 1 (Weighted Network): For a weighted network $\mathcal{G} = \{\mathcal{V}, \mathcal{E}, \mathbf{w}\}$, we denote \mathcal{V} as a set of nodes, \mathcal{E} as a set of edges, and \mathbf{w} as a weight vector for edges. Let $w(e_i)$ represent the i -th element of \mathbf{w} , which is the weight of the edge e_i .

Definition 2 (Ordinal Pattern): Given a weighted network $\mathcal{G} = \{\mathcal{V}, \mathcal{E}, \mathbf{w}\}$, if $w(e_i) > w(e_j)$ for all $0 < i < j \leq M$, $op = \{e_1, e_2, \dots, e_M\} \subseteq \mathcal{E}$ is called an *ordinal pattern*. Here, M is the number of edges in the ordinal pattern op .

From this definition, we can see that an ordinal pattern is constructed by a sequence of weighted edges, where these edges are ordered according to their corresponding weights. In Fig. 2, we show an illustration of the proposed ordinal pattern of a weighted network (containing five nodes and seven weighted edges). We can construct ordinal patterns that contain two edges, e.g., $op^1 = \{e_{a-b}, e_{b-c}\}$ and $op^3 = \{e_{d-c}, e_{c-e}\}$. We can see that the ordinal pattern op^1 denotes the ordinal relationship $w(e_{a-b}) > w(e_{b-c})$ between a pair of edges (i.e., e_{a-b} and e_{b-c}). Then, we can construct ordinal patterns that contain three edges, e.g., $op^4 = \{e_{a-b}, e_{b-c}, e_{c-e}\}$ and $op^5 = \{e_{b-c}, e_{c-d}, e_{c-e}\}$. That is, our proposed ordinal

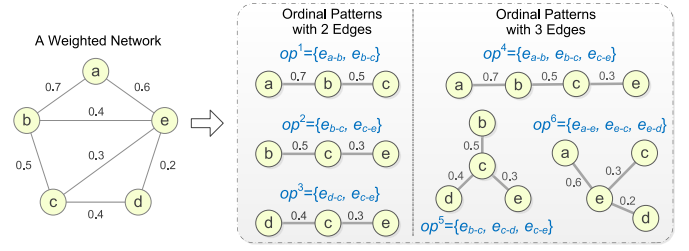


Fig. 2. Illustration of the proposed ordinal patterns defined on a weighted network corresponding to a specific subject. Here, we show two types of ordinal patterns that contain two and three edges, respectively.

pattern is a specific combination of the ordinal relationship among weighted edges. Given an ordinal pattern $op^i = \{e_{a-b}, e_{b-c}, \dots, e_{m-n}\}$, any valid ordinal pattern $op^j = \{e_{a-b}, e_{b-c}, \dots, e_{m-n}, e_{n-o}\}$ is called a *child* of op^i , and op^i is called the *parent* of op^j . Apparently, to construct a valid ordinal pattern, e_{n-o} must be an edge which weight is smaller than that of the edge e_{m-n} . On the other hand, from Fig. 2, we can observe that the proposed ordinal pattern would follow a tree structure, where these patterns are defined on undirected graphs.

Compared with previous properties of brain connectivity networks, the proposed ordinal patterns have at least two advantages. *First*, the ordinal pattern is defined on a weighted network directly, which considers both weights of edges and the ordinal relationship among weighted edges. In contrast, most of the conventional properties of brain connectivity networks do not utilize such ordinal relationship among edges. *Second*, as a particular subnetwork, an ordinal pattern can model the ordinal relationship of weighted edges in a brain network, which can preserve the local topology of the network.

C. Frequent Ordinal Pattern

We further propose to identify the ordinal patterns that frequently appear in brain connectivity networks, which is called *frequent ordinal pattern* in this study. Denote a set of weighted networks as $\mathcal{D} = \{\mathcal{G}_1, \mathcal{G}_2, \dots, \mathcal{G}_n, \dots, \mathcal{G}_N\}$, where N is the number of weighted networks in \mathcal{D} . For clarity, we first introduce several definitions.

Definition 3 (Frequency Ratio): Given an ordinal pattern $op \in \mathcal{D}$, its frequency ratio $FR(op|\mathcal{D})$ is defined as

$$FR(op|\mathcal{D}) = \frac{1}{N} \sum_{n=1}^N a_n \quad (1)$$

where $a_n = 1$ if op is an ordinal pattern of \mathcal{G}_n ($\mathcal{G}_n \in \mathcal{D}$), and $a_n = 0$ otherwise. Here, N is the number of networks in \mathcal{D} .

Definition 4 (Frequent Ordinal Pattern): Give a set of weighted networks $\mathcal{D} = \{\mathcal{G}_1, \mathcal{G}_2, \dots, \mathcal{G}_n, \dots, \mathcal{G}_N\}$ and a pre-defined threshold value θ , if $FR(op|\mathcal{D}) > \theta$, the ordinal pattern op is a *frequent ordinal pattern* of \mathcal{D} .

From this definition, we can see that frequent ordinal patterns are ordinal patterns that frequently appear in a set of weighted networks. For a set of brain networks, a frequent ordinal pattern may reflect the functional information that is common to all subjects. It is easy to show that the frequent ordinal patterns have several appealing properties.

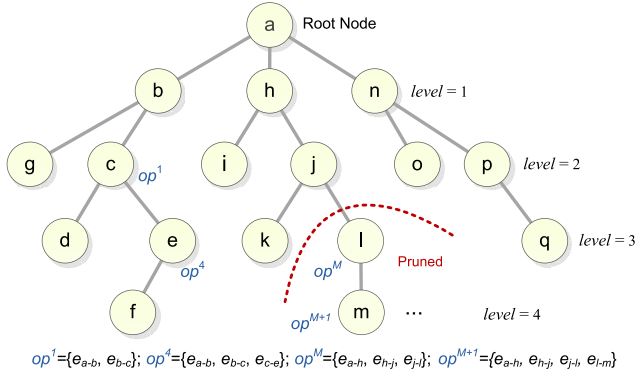


Fig. 3. Illustration of a deep first search (DFS) tree for mining frequent ordinal patterns. Here, a child of the ordinal pattern $op^1 = \{e_{a-b}, e_{b-c}\}$ is $op^4 = \{e_{a-b}, e_{b-c}, e_{c-e}\}$, and the parent of op^{M+1} is op^M .

Property 1: For an ordinal pattern, its frequency ratio is no larger than the frequency ratio of its parent.

Property 2: If an ordinal pattern is not a frequent ordinal pattern, its children and descendants are not frequent ordinal patterns, either; if an ordinal pattern is a frequent ordinal pattern, its parents must be frequent ordinal patterns.

As illustrated in Fig. 3, $op^4 = \{e_{a-b}, e_{b-c}, e_{c-e}\}$ is a child of $op^1 = \{e_{a-b}, e_{b-c}\}$, and op^1 is the parent of op^4 . Based on Eq. 1, we can obtain $FR(op^1) > FR(op^4)$. Besides, if op^M is not a frequent ordinal pattern, its child op^{M+1} is not a frequent ordinal pattern, either. This is an appealing property, which helps us significantly reduce the searching space in the following algorithm for mining frequent ordinal patterns.

D. Frequent Ordinal Pattern Mining

Motivated by the idea of frequent subgraph mining approach [13], [41], we propose a frequent ordinal pattern mining algorithm³ based on a deep first search (DFS) tree [42]. Specifically, as illustrated in Fig. 3, we first randomly choose an edge (e.g., e_{a-b}) that connects the first node (e.g., a that is corresponding to the first brain region in the template space). If the frequency ratio of e_{a-b} is larger than a threshold θ , we treat the node a as the root node. Then, the path from the root node to the current node forms a specific ordinal pattern, e.g., $op^1 = \{e_{a-b}, e_{b-c}\}$. Given a set of connectivity networks, we can record the number of occurrences of this ordinal pattern, as well as compute its frequency ratio. According to **Property 1**, if an ordinal pattern (e.g., op^1) is a frequent ordinal pattern (i.e., its frequency ratio is larger than θ), we can further search its children (e.g., op^4). In contrast, if an ordinal pattern (e.g., op^M) is not a frequent ordinal pattern, we can directly prune its children and descendants (e.g., op^{M+1}). The max number of edges in a frequent ordinal pattern is limited by the depth (denoted as *level*) of a DFS tree. For instance, if *level* = 3, there will be at most 3 edges in an identified frequent ordinal pattern. That is, a larger *level* will generate more frequent ordinal patterns and more running time of the proposed algorithm. The pseudo codes of our proposed algorithm for frequent ordinal pattern mining are given in

Algorithms 1-2, while the computational complexity analysis is shown in the *Supplementary Materials*.

Algorithm 1: Pseudo Code of Frequent Ordinal Pattern Mining Algorithm

Input: A weighted network set \mathcal{D} ; The max depth of the DFS tree *level*; The threshold θ for frequency ratio in Eq. (1)

Output: The frequent ordinal pattern set OP

- 1 Compute the frequency ratios of all edges in \mathcal{E} appeared in \mathcal{D} ;
- 2 Remove infrequent edges in \mathcal{E} according to **Definition 4**;
- 3 **foreach** $e \in \mathcal{E}$ **do**
- 4 Initialize *op* with e ;
- 5 $OPM(\mathcal{D}, level, \theta, op)$;
- 6 **endfor**

Algorithm 2: $OPM(\mathcal{D}, level, \theta, op)$

Input: $\mathcal{D}, level, \theta, op$

Output: OP

- 1 Construct \mathcal{E}_c by finding edges that are connected to edges in op ;
- 2 **foreach** $e \in \mathcal{E}_c$ **do**
- 3 **if** $e \notin op$ **then**
- 4 $op_{new} = \{op, e\}$;
- 5 Compute the frequency ratio of op_{new} as $FR(op_{new})$ according to Eq. (1);
- 6 **if** $FR(op_{new}) > \theta$ **then**
- 7 Add op_{new} into OP ;
- 8 **if** $|op_{new}| < level$ **then**
- 9 $OPM(\mathcal{D}, level, \theta, op_{new})$;
- 10 **endif**
- 11 **endif**
- 12 **endif**
- 13 **endfor**

We now analyze the off-line computational complexity of the proposed algorithm (Algorithms 1-2) for mining the frequent ordinal patterns. In Algorithms 1-2, two key factors determine the computational complexity, including the number of edges and the max depth (i.e., *level*) of the DFS tree. Given a fully-connected weighted network containing n_e edges and a fixed *level*, the computational complexity of mining the frequent ordinal patterns is $\mathcal{O}(\frac{n_e!}{(n_e-level)!})$. Given N training subjects, the overall computational complexity of Algorithms 1 and 2 is $\mathcal{O}(N \times \frac{n_e!}{(n_e-level)!})$. It is worth noting that, according to **Property 1**, many edges will not be searched in the proposed algorithm, and hence the real computational cost of our method is less than $\mathcal{O}(N \times \frac{n_e!}{(n_e-level)!})$. Using a computer with the processor of Intel(R) Core(TM)2 i3-3220HQ 3.30GHz, the proposed method requires approximately 10 minutes to mine frequent ordinal patterns from a training set containing 152 networks (with each network

³<http://ibrain.nuaa.edu.cn>

corresponding to a specific subject), and the implementation of the proposed algorithm is based on Matlab. In particular, since the training process is performed *off-line*, the proposed method would be scalable to large databases in the application/testing stage.

E. Ordinal Pattern Based Learning Framework

1) *Network Construction and Frequent Ordinal Pattern Mining*: Using the above-mentioned pre-processing method for rs-fMRI data, we can construct one brain functional connectivity network for each subject, where each node denotes a particular ROI, and the weight of an edge is the Pearson correlation coefficient between a pair of ROIs. Given all training subjects, we can obtain two sets of brain networks. The first set contains all networks from patients, and the second one includes all networks from NCs. We then construct ordinal patterns on both sets of brain networks of patients and NCs separately. We further mine the frequently ordinal patterns from both sets using Algorithms 1-2.

2) *Discriminative Ordinal Pattern Selection*: Since not all identified frequent ordinal patterns are discriminative, we further perform ordinal pattern selection to preserve the most informative frequent ordinal patterns and discard those with less discriminative power. Currently, there exist various measures to evaluate the discriminative power of a feature, such as *t*-test [43] and ratio score [44]. In this study, we utilize the ratio score (RS) [44] to measure the discriminative capability of all identified frequent ordinal patterns.

Definition 5 (Ratio Score): Denote the network set of patients as \mathcal{D}^+ (containing N^+ networks), and that of NCs as \mathcal{D}^- (containing N^- networks). Given a frequent ordinal pattern op^i from \mathcal{D}^+ , its ratio score (RS) is defined as

$$RS(op^i|\mathcal{D}) = \log \frac{\sum_{n=1}^{N^+} \alpha_n}{\sum_{m=1}^{N^-} \beta_m + \epsilon} \times \frac{N^-}{N^+} \quad (2)$$

where $\alpha_n = 1$ if op^i is an ordinal pattern of $\mathcal{G}_n \in \mathcal{D}^+$, and $\alpha_n = 0$ otherwise; $\beta_m = 1$ if op^i is an ordinal pattern of $\mathcal{G}_m \in \mathcal{D}^-$, and $\beta_m = 0$ otherwise. Here, ϵ is a small value to prevent the denominator to be 0. Similarly, for the frequent ordinal pattern op^j from the network set of NCs (*i.e.*, \mathcal{D}^-), its ratio score is defined as

$$RS(op^j|\mathcal{D}) = \log \frac{\sum_{m=1}^{N^-} \beta_m}{\sum_{n=1}^{N^+} \alpha_n + \epsilon} \times \frac{N^+}{N^-} \quad (3)$$

where $\beta_m = 1$ if op^j is an ordinal pattern of $\mathcal{G}_m \in \mathcal{D}^-$, and $\beta_m = 0$ otherwise; $\alpha_n = 1$ if op^j is an ordinal pattern of $\mathcal{G}_n \in \mathcal{D}^+$, and $\alpha_n = 0$ otherwise. From Eqs. 2 and 3, we can see that a large RS value indicates that the frequent ordinal pattern has a strong discriminative ability, and vice versa. With ratio scores, we can rank ordinal patterns in descending order and select the top ones as the discriminative ordinal patterns.

3) *Feature Extraction and Model Construction*: Based on ratio scores, we first select a total of k discriminative ordinal patterns from training data, where half of them are from the brain network set of patients and the other half are from that of NCs. Then, the discriminative ordinal patterns are combined to a feature matrix for representing the original rs-fMRI data of subjects. Given N brain networks (with each one corresponding to a particular subject) and k ordinal patterns, the feature matrix is denoted as $\mathbf{F} \in \mathbb{R}^{N \times k}$, with the element $F_{i,j}$ denoting the j -th feature of the i -th subject. If the connectivity network of the i -th subject owns the j -th discriminative ordinal pattern, $F_{i,j} = 1$; and $F_{i,j} = 0$, otherwise. Based on the feature matrix \mathbf{F} of training subjects, we further construct a linear support vector machine (SVM) classifier and a linear support vector regressor (SVR) to perform tasks of classification and regression, respectively.

Given a new testing subject with rs-fMRI data, we first pre-process the imaging data, construct a weighted functional connectivity network, and build ordinal patterns on this network. We then extract the feature representation $z \in \mathbb{R}^k$ for this subject based on the k discriminative ordinal patterns (identified from training subjects). Finally, we feed the feature vector z into the trained classifier and regressor for class label prediction and clinical score estimation, respectively.

IV. EXPERIMENTS

In this section, we first present experimental settings, results of brain disease diagnosis and clinical score regression. We then analyze the influence of parameters and compare our method with several state-of-the-art approaches in brain disease diagnosis using fMRI data.

A. Experimental Settings

In this work, we perform two groups of experiments, including 1) disease classification, and 2) clinical score regression. In the first group of experiments, with a linear SVM and a 10-fold cross-validation strategy, three classification tasks are performed, including 1) AD vs. NC classification, 2) MCI vs. NC classification, and 3) ADHD vs. NC classification. Note that only training subjects are used for identifying discriminative ordinal patterns. We evaluate the classification performance by accuracy (ACC), sensitivity (SEN), specificity (SPE), and the area under the receiver operating characteristic (ROC) curve (AUC). In the second group of experiments, we adopt a linear SVR as the regressor to estimate MMSE scores for subjects in the ADNI dataset. We evaluate the performance of regression by both correlation coefficient (Cor) and root mean square error (RMSE) between estimated clinical scores and actual clinical scores.

We first compare our method with a baseline method (denoted as **Baseline**) to investigate the influence of ordinal relationship among weighted edges in functional connectivity networks. Specifically, in the Baseline method, we simply vectorize the weights of edges in the original weighted network as a feature vector for each subject, without considering the ordinal relationship among different edges. In contrast, we explicitly utilize such ordinal relationship among weighted

TABLE II

CLASSIFICATION RESULTS ACHIEVED BY SIX DIFFERENT METHODS IN THREE CLASSIFICATION TASKS. ACC: ACCURACY; SEN: SENSITIVITY; SPE: SPECIFICITY; AUC: AREA UNDER THE ROC CURVE

Method	AD vs. NC				MCI vs. NC				ADHD vs. NC			
	ACC (%)	SEN (%)	SPE (%)	AUC (%)	ACC (%)	SEN (%)	SPE (%)	AUC (%)	ACC (%)	SEN (%)	SPE (%)	AUC (%)
Baseline	61.25	46.18	72.04	59.86	63.83	74.45	44.02	65.64	57.55	61.78	52.49	52.36
CC	72.62	73.53	67.94	70.94	71.14	72.73	68.00	68.69	71.29	72.03	70.41	70.51
CCMT	80.95	82.35	80.00	76.35	74.50	75.76	72.00	74.79	74.53	75.43	73.47	77.64
DS	76.19	76.47	76.00	75.59	77.18	78.79	74.00	74.89	81.01	81.36	80.61	80.82
DSMT	85.71	85.29	86.00	87.59	79.19	80.81	76.00	76.99	83.79	84.74	82.65	84.63
Proposed	94.05	96.77	92.45	96.35*	88.59	87.27	92.31	84.57*	87.50	88.89	85.85	87.37*

edges in the proposed method. We further compare our method with two network properties that are widely used in brain network analysis, including cluster coefficients [11] and discriminative subnetworks [13]. Since both properties require a thresholding process, we adopt a single thresholding strategy and a multi-thresholding [12], [31] strategy to transform the original weighted networks to unweighted ones, respectively. That is, besides Baseline, there are four competing network properties, including 1) clustering coefficients with the single thresholding strategy (CC), 2) clustering coefficient using the multi-thresholding strategy (CCMT), 3) discriminative subnetworks with the single thresholding strategy (DS), and 4) discriminative subnetworks using the multi-thresholding strategy (DSMT).

A linear SVM/SVR is used as the classifier/regressor in the proposed method and those five competing methods, with a default parameter (*i.e.*, $C = 1$). We adopt the default parameters given by the authors for four competing methods (*i.e.*, CC [11], CCMT [11], [12], DS [13], and DSMT [12], [13]). For our ordinal pattern based approach, the parameter θ in defining frequent ordinal patterns (see Definition 4) is empirically set as 0.7, while the parameter ϵ in Eqs. (2) and (3) is empirically set as 0.1. With an inner cross-validation strategy, the parameter *level* in our frequent ordinal pattern mining algorithm is chosen from the range [1, 5] with the step size of 1, and the number of discriminative ordinal patterns (*i.e.*, k) is chosen from [10, 120] with the step size of 10. We further investigate the influence of three important parameters (*i.e.*, *level*, k , and θ in Definition 4) on the performance of our method in Section IV-D.

B. Results of Brain Disease Classification

1) Classification Results: We conduct three classification tasks in this group of experiments, including (1) AD vs. NC classification, (2) MCI vs. NC classification, and (3) ADHD vs. NC classification. The experimental results are reported in Table II. To test whether the results of our method and those of each competing method are statistically different, we perform the paired Delong's test [45] (with the significance level at 0.05) on the AUC values achieved by our method and each competing method. The AUC values in Table II further marked by * indicate that our method achieves significantly different results compared with the competing methods.

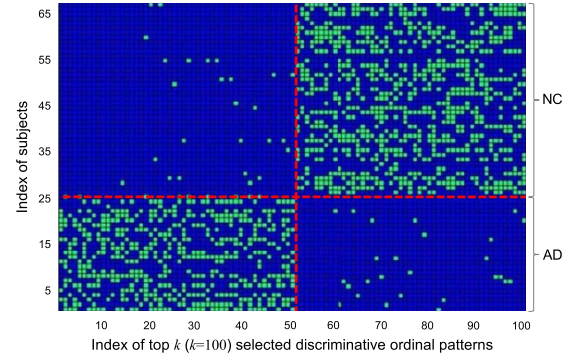


Fig. 4. Illustration of the proposed ordinal pattern based feature representations for AD and NC subjects in AD vs. NC classification. Each row denotes a subject, and each column is corresponding to a discriminative ordinal pattern. Thus, the proposed ordinal pattern based representation for each subject is a k -dimensional vector. Here, green and blue indicate 1 and 0, respectively.

From Table II, we can observe that the proposed method outperforms the Baseline method regarding four evaluation measures (*i.e.*, ACC, SEN, SPE, and AUC) in three classification tasks. These results suggest that the ordinal relationship among weighted edges can be used to improve the performance of brain disease diagnosis in fMRI-based studies. Besides, Table II shows the proposed method consistently outperforms four popular network properties (*i.e.*, CC, CCMT, DS, and DSMT) in three classification tasks. For instance, our method yields an ACC of 94.05% in the task of AD vs. NC classification, which is better than the second best result (*i.e.*, 85.71%) yielded by DSMT. Also, the AUC values obtained by our method are better than that achieved by four competing methods in three tasks. On the other hand, the performance difference in terms of AUC between our method and each competing method is significant via the paired Delong's test. This demonstrates that the ordinal patterns are useful in identifying AD/MCI/ADHD patients from the population of normal controls, compared with the conventional network descriptors.

2) Ordinal Pattern Based Representations: We also illustrate the proposed ordinal pattern based feature representations for the original rs-fMR images of subjects. In Fig. 4, we show the feature matrix of AD and NC subjects in training set in a fold, where each row denotes a subject and each column is corresponding to a discriminative ordinal pattern. There are a total of $k = 100$ selected frequent ordinal patterns,

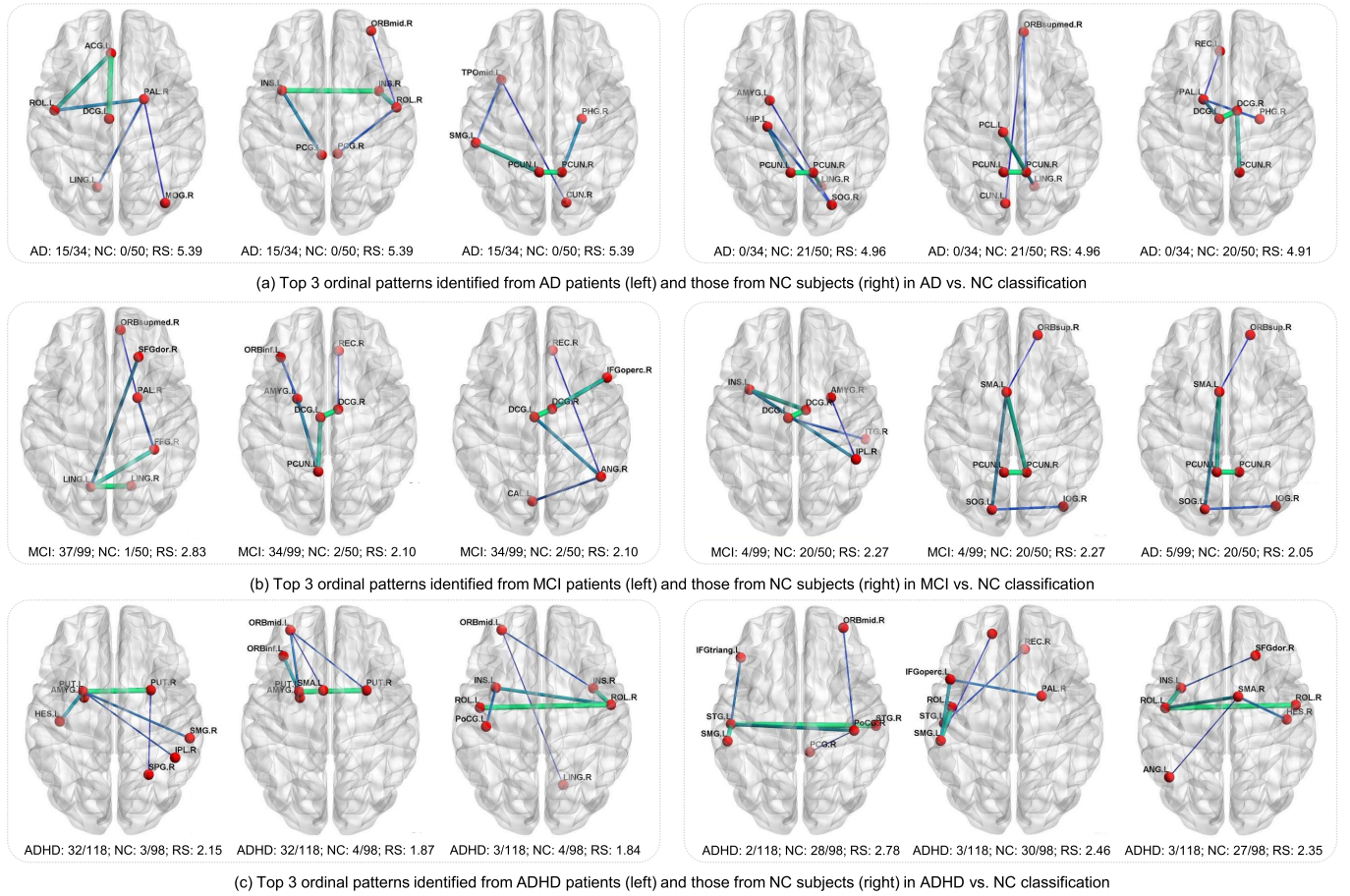


Fig. 5. Illustration of the most discriminative ordinal patterns discovered by our proposed method in the classification tasks of (a) AD vs. NC, (b) MCI vs. NC, and (c) ADHD vs. NC. For instance, the discriminative ordinal pattern in the left top corner can be recorded as $op = \{e_{DCG.L-ACG.L}, e_{ACG.L-ROL.L}, e_{ROL.L-PAL.R}, e_{PAL.R-LING.L}, e_{PAL.R-MOG.R}\}$. The bottom of each discriminative ordinal pattern lists the number of occurrences of a particular ordinal pattern in both network sets of patients and NCs, as well as its ratio score (RS).

including 50 patterns mined from AD patients (*i.e.*, the first 50 columns in Fig. 4) and 50 patterns identified from NC subjects (*i.e.*, the last 50 columns in Fig. 4). Thus, the proposed ordinal pattern based representation for each subject is a k -dimensional feature vector.

From Fig. 4, we can observe that there are many 1-values in both the bottom left and the top right corners, while the remaining parts are mostly filled with 0-values. It suggests that the ordinal patterns identified from AD patients are different from those mined from NCs, and thus such ordinal pattern based features could be discriminative for representing both AD patients and NCs. Also, due to the inter-subject variability in rs-fMRI data, the proposed representation for two subjects within the same group (*i.e.*, AD or NC) may have different values at the same ordinal pattern.

We further plot the top 3 discriminative ordinal patterns identified in three classification tasks in Fig. 5. As shown in the top left of Fig. 5 (a), the most discriminative ordinal pattern for AD patients in AD vs. NC classification is $op = \{e_{DCG.L-ACG.L}, e_{ACG.L-ROL.L}, e_{ROL.L-PAL.R}, e_{PAL.R-LING.L}, e_{PAL.R-MOG.R}\}$. This suggests that the functional communication of a series of ROIs, such as middle cingulate gyrus left (DCG.L), anterior cingulate

gyrus left (ACG.L), precuneus left (PCG.L), and precuneus right (PCG.R), are discriminative in the task of AD vs. NC classification. Also, these results imply that our ordinal patterns do reflect the local structures of brain connectivity networks. Besides, we also study the stability of ordinal patterns identified by our method across ten folds in three classification tasks, with results reported in Fig. S2 of the *Supplementary Materials*.

C. Results of Clinical Score Regression

We then perform the regression task for MMSE scores of subjects in the ADNI dataset. Specifically, six types of network properties are used in this group of experiments, including Baseline, CC [11], CCMT [11], [12], DS [13], DSMT [12], [13], and our proposed ordinal patterns. Based on a particular network descriptor, we learn a linear SVR on training subjects and estimate MMSE scores for testing subjects, with results shown in Table III. Besides, in Fig. 6, we visually show the scatter plots of estimated MMSE scores vs. actual MMSE scores achieved by different methods for subjects in ADNI.

As can be seen from Table III, in the task of MMSE score regression, our method achieves higher Cor and lower RMSE values, compared with Baseline. This further implies that the

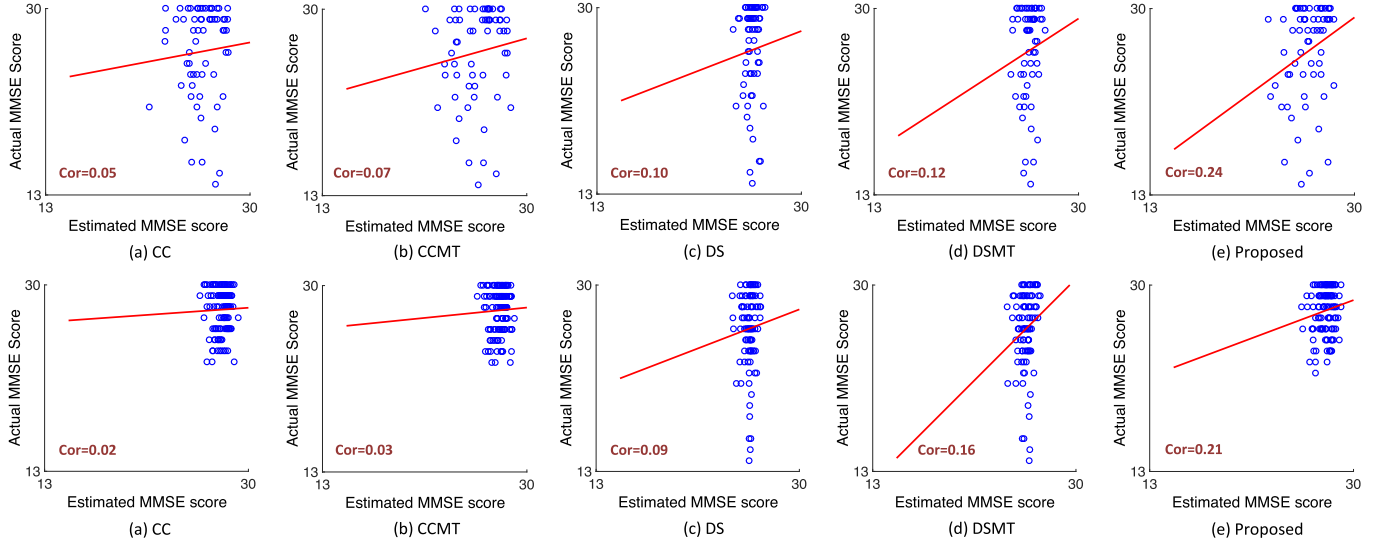


Fig. 6. Scatter plots of estimated MMSE scores vs. actual MMSE scores achieved by our method and 4 competing methods for (top) AD and NC subjects and (bottom) MCI and NC subjects in ADNI using linear SVR as the regressor. Cor: Correlation coefficient.

TABLE III

REGRESSION RESULTS OF MMSE SCORES ACHIEVED BY DIFFERENT METHODS ON THE ADNI DATASET. COR: CORRELATION COEFFICIENT; RMSE: ROOT MEAN SQUARE ERROR

Method	AD vs. NC		MCI vs. NC	
	Cor	RMSE	Cor	RMSE
Baseline	0.03	5.67	0.01	4.52
CC	0.05	4.42	0.02	3.49
CCMT	0.07	4.30	0.03	3.43
DS	0.10	4.33	0.09	2.59
DSMT	0.12	4.25	0.16	2.28
Proposed	0.24	4.07	0.21	1.97

modeling the ordinal relationship among weighted edges helps boost the learning performance of brain disease diagnosis based on fMRI data. Also, from Table III and Fig. 6, one can observe that our method consistently performs better than CC, CCMT, DS, and DSMT, in terms of Cor and RMSE. For instance, the Cor value of our method for AD and NC subjects is 0.24, which is much higher than the second best one (*i.e.*, 0.12 achieved by DSMT). These results demonstrate that the proposed ordinal pattern provides an effective solution for representing brain connectivity networks.

D. Influence of Parameters

Now we investigate the influence of two parameters (*i.e.*, *level* and *k*) on the performances of our method in three classification tasks. Specifically, we vary the number of *level* in the set $\{1, 2, \dots, 5\}$. Also, we vary the value of *k* in the set $\{10, 22, \dots, 118\}$ for both AD vs. NC and MCI vs. NC classification, and in the set $\{10, 30, \dots, 190\}$ for ADHD vs. NC classification, respectively. The AUC values achieved by our method using different parameters are shown in Fig. 8, while the ACC values are reported in . S1 of the Supplementary Materials.

It can be observed from Fig. 8 that the results achieved by the proposed method are stable when the number of selected

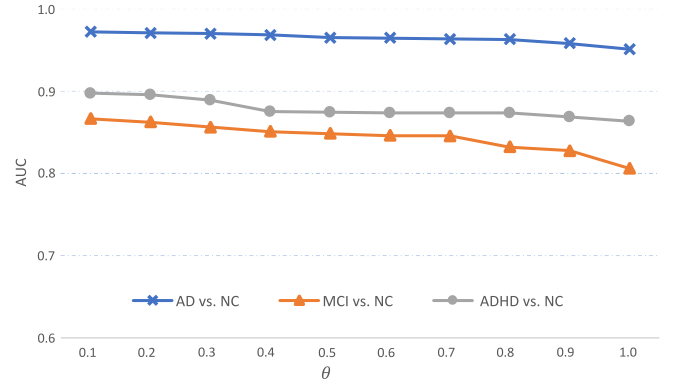


Fig. 7. Influence of the threshold θ in Definition 4 on AUC values achieved by the proposed method in three classification tasks.

ordinal patterns is larger than 40 (*i.e.*, $k > 40$) in three tasks. Also, our method produces the overall good performance when *level* = 4 in both tasks of AD vs. NC and MCI vs. NC classification, and when *level* = 5 in the task of ADHD vs. NC classification. The possible reason could be that the ADHD patients would exhibit more significant functional connectivity compared with normal controls, and such abnormal increased functional connectivity may alternate the ordinal patterns of brain connectivity on more brain regions [46]. In such a case, our ordinal patterns would require a larger value of *level* to include more brain regions for identifying ADHD patients, compared with that used in AD and MCI diagnosis.

We further analyze the influence of the parameter θ in Definition 4. Specifically, we vary the value of θ in $\{0.1, 0.2, \dots, 1\}$ and record the AUC values achieved by our method in three classification tasks in Fig. 7. From Fig. 7, we can see that the performance of our method is relatively stable when $\theta \leq 0.8$ in three tasks, and the performance slightly slows down with the increase of θ when $\theta > 0.8$. Note that, with a small θ , the number of ordinal patterns that need to be searched is large, which brings a huge computational

TABLE IV

COMPARISON WITH STATE-OF-THE-ART METHODS IN THREE CLASSIFICATION TASKS USING RS-FMRI DATA. ROI-VC: ROI VARIANCECOVARIANCE; REHO: REGIONAL HOMOGENEITY; LDA: LINEAR DISCRIMINANT ANALYSIS; GP-LR: BAYESIAN GAUSSIAN PROCESS LOGISTIC REGRESSION; SVM: SUPPORT VECTOR MACHINE. NOTE THAT DIFFERENT METHODS USED DIFFERENT SUBJECT SUBSETS FROM THE ADNI AND ADHD-200 DATASETS

Classification Task	Method	Features	Classifier	Subjects	Accuracy (%)
AD vs. NC	Zhou <i>et al.</i> [23]	ROI-based difference	LDA	AD: 12; NC: 12	92.00
	Challis <i>et al.</i> [26]	ROI-VC	GP-LR	AD: 27; NC: 39	85.00
	Proposed	Ordinal pattern	SVM	AD: 34; NC: 50	94.05
MCI vs. NC	Challis <i>et al.</i> [26]	ROI-VC	GP-LR	MCI: 50; NC: 21	75.00
	Jie <i>et al.</i> [29]	Clustering coefficients	SVM	MCI: 12; NC: 21	91.90
	Proposed	Ordinal pattern	SVM	MCI: 99; NC: 50	88.59
ADHD vs. NC	Wang <i>et al.</i> [27]	ReHo	SVM	ADHD: 12; NC: 23	85.00
	Dey <i>et al.</i> [28]	Graph-based measure	SVM	ADHD: 307; NC: 307	73.50
	Proposed	Ordinal pattern	SVM	ADHD: 118; NC: 98	87.50

burden. In contrast, with a large θ , we can significantly reduce the searching space by cutting a lot of ordinal patterns to be mined. Hence, we empirically set $\theta = 0.7$ in the experiments.

E. Comparison With State-of-the-Art Approaches

We further compare the results achieved by our method with those of the state-of-the-art approaches using rs-fMRI data for brain disease classification. We list the details of these methods and their respective results in three classification tasks in Table IV. Note that results in Table IV were achieved by five competing methods, where each method utilized different subsets of the ADNI and ADHD-200 databases.

From Table IV, we can see that our method obtains good performance in most cases. Specifically, compared with the other methods, our method achieves much higher accuracies in both tasks of AD vs. NC classification and ADHD vs. NC classification, and obtains the comparable result in MCI vs. NC classification. It is worth noting that MCI has a variety of subcategories, such as MCI Converters (MCI-C) and MCI Non-converters (MCI-NC) [37]. These complex patient conditions have brought significant challenges for MCI classification, especially in the case of large datasets. Since we use a relatively larger dataset for MCI vs. NC classification than that used in previous studies [26], [29], our method yields reasonable results in MCI vs. NC classification.

V. DISCUSSION

In this section, we first show the most frequently identified brain regions achieved by the proposed method, and then present the limitations of our method as well as possible future research directions.

A. Frequently Identified Brain Regions

In a brain connectivity network, each node is corresponding to a specific ROI defined in the AAL template space. We now record the number of occurrences of each ROI among all identified discriminative ordinal patterns in three classification tasks (*i.e.*, AD vs. NC, MCI vs. NC, and ADHD vs. NC classification). We visually plot those top 10 frequently identified ROIs in three classification tasks in Fig. 9, with

TABLE V

THE MOST FREQUENT 10 ROIs APPEARING IN DISCRIMINATIVE ORDINAL PATTERNS OF SUBJECTS IN THREE CLASSIFICATION TASKS. THE ROIs ARE LISTED IN DESCENDING ORDER ACCORDING TO THEIR FREQUENCIES

AD vs. NC			
Index	ROI	Full Name	Related Studies
1	PCUN.L	Precuneus left	Alexander <i>et al.</i> [47]
2	PCUN.R	Precuneus right	Frisoni <i>et al.</i> [48]
3	DCG.L	Middle cingulate gyrus left	Supekar <i>et al.</i> [49]
4	DCG.R	Middle cingulate gyrus right	Supekar <i>et al.</i> [49]
5	IOG.R	Inferior occipital gyrus right	Supekar <i>et al.</i> [49]
6	LING.R	Lingual gyrus right	Supekar <i>et al.</i> [49]
7	IFGoperc.L	Inferior frontal gyrus (opercular) left	Supekar <i>et al.</i> [49]
8	ACG.L	Anterior cingulate gyrus left	Frisoni <i>et al.</i> [48]
9	AMYG.R	Amygdala right	Supekar <i>et al.</i> [49]
10	IOG.L	Inferior occipital gyrus left	Supekar <i>et al.</i> [49]
MCI vs. NC			
Index	ROI	Full Name	Related Studies
1	DCG.R	Middle cingulate gyrus right	Feng <i>et al.</i> [50]
2	PCUN.L	Precuneus left	Zhang <i>et al.</i> [51]
3	PCUN.R	Precuneus right	Zhang <i>et al.</i> [51]
4	DCG.L	Middle cingulate gyrus left	Feng <i>et al.</i> [50]
5	LING.L	Lingual gyrus left	Zhang <i>et al.</i> [51]
6	LING.R	Lingual gyrus right	Zhang <i>et al.</i> [51]
7	PAL.L	Pallidum left	Bai <i>et al.</i> [52]
8	IOG.L	Inferior occipital gyrus left	Zhang <i>et al.</i> [51]
9	IOG.R	Inferior occipital gyrus right	Liu <i>et al.</i> [53]
10	SMA.R	Supplementary motor area right	Zhang <i>et al.</i> [51]
ADHD vs. NC			
Index	ROI	Full Name	Related Studies
1	IFGoperc.R	Inferior frontal gyrus (opercular) right	Alexander <i>et al.</i> [47]
2	IFGtriang.R	Inferior frontal gyrus (triangular) right	Frisoni <i>et al.</i> [48]
3	ROL.L	Rolandic operculum left	Supekar <i>et al.</i> [49]
4	ROL.R	Rolandic operculum right	Supekar <i>et al.</i> [49]
5	SFGmed.L	Superior frontal gyrus (medial) left	Supekar <i>et al.</i> [49]
6	SFGmed.R	Superior frontal gyrus (medial) right	Supekar <i>et al.</i> [49]
7	ORBmed.L	Orbitofrontal cortex (medial) left	Supekar <i>et al.</i> [49]
8	INS.L	Insula left	Frisoni <i>et al.</i> [48]
9	INS.R	Insula right	Supekar <i>et al.</i> [49]
10	PUT.R	Putamen right	Supekar <i>et al.</i> [49]

their names listed in Table V. From Fig. 9 and Table V, we can observe that the most frequent ROIs identified by our method are consistent with those reported in previous studies [20], [47]–[51]. For instance, in AD vs. NC classification, the most frequently identified ROIs include precuneus left [47], precuneus right [48], middle cingulate gyrus left [49], middle cingulate gyrus right [49], and inferior occipital gyrus right [49].

B. Limitations and Future Work

Although the proposed method achieves good results in both disease classification and clinical score regression, there are still several limitations to be considered in this study.

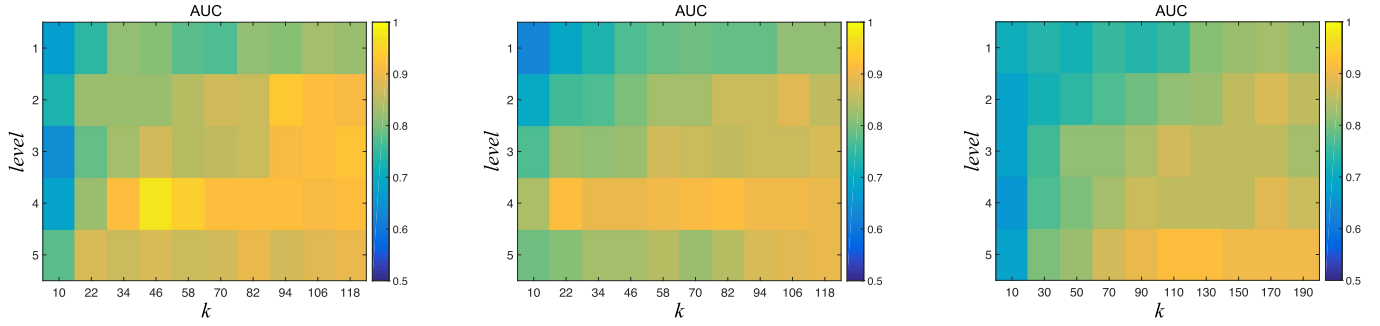


Fig. 8. AUC values achieved by the proposed method using different values of two parameters (*i.e.*, *level*, *k*) in the tasks of (left) AD vs. NC classification, (middle) MCI vs. NC classification, and (right) ADHD vs. NC classification.

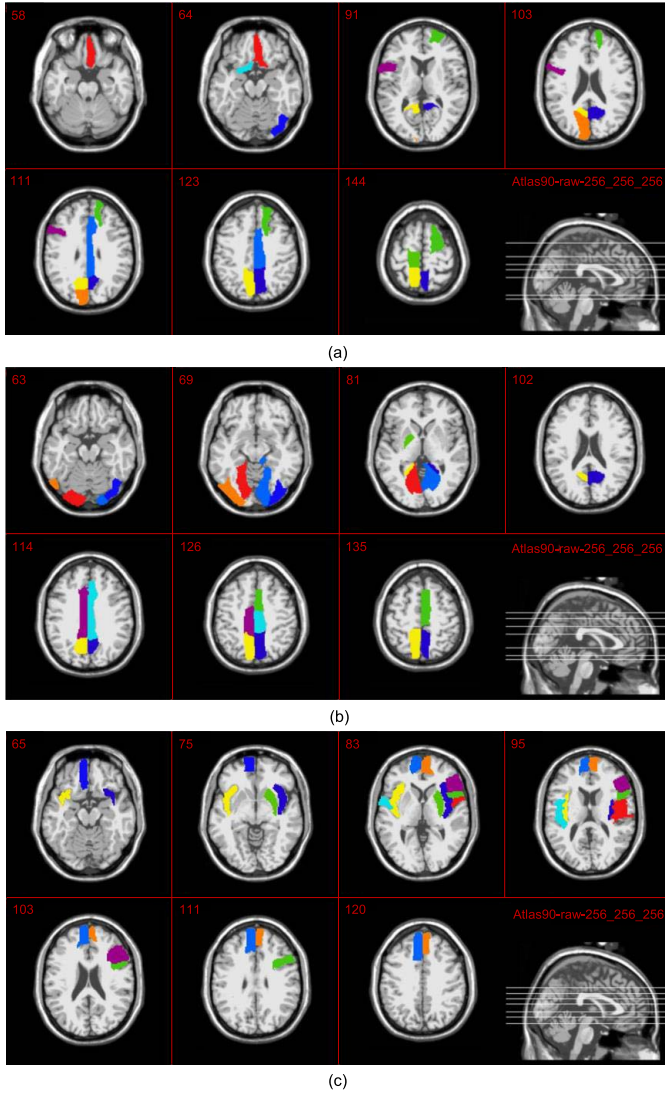


Fig. 9. Visual plots of the most frequent ROIs identified by the proposed method in three classification tasks of (a) AD vs. NC. (b) MCI vs. NC. (c) ADHD vs. NC.

First, in our ordinal pattern based binary feature matrix \mathbf{F} , we simply set its element $F_{i,j} = 1$ if the i -th brain connectivity network contains the j -th ordinal pattern; and $F_{i,j} = 0$, otherwise. Such binarization strategy may lead to the loss of the statistical properties of weights of edges in ordinal patterns. It is interesting to extract real-valued features based on the proposed discriminative ordinal patterns, which will be

one of our future work. Second, we only select the most discriminative ordinal patterns according to their discriminative powers (evaluated by ratio scores). Based on such selection criteria, the selected discriminative ordinal patterns could be redundant. Inspired by these studies [12], [54], we can apply feature selection strategies to refine the selected discriminative ordinal patterns, to further improve the performance of our method. Third, we only used ordinal patterns to represent brain networks constructed on rs-fMRI data. As a general network descriptor, the ordinal pattern can be applied to measure both functional and structural connectivity networks. Hence, it is straightforward to directly apply our method to problems having both MRI and fMRI data. Besides, we will consider the unbalanced data problem regarding the gender in ADHD-200 in the future.

VI. CONCLUSION

In this paper, we proposed a novel network descriptor (*i.e.*, *ordinal pattern*) for the analysis of brain connectivity network based on rs-fMRI data. The proposed ordinal patterns are directly defined on weighted networks, which can make use of the weights of edges and the ordinal relationship among weighted edges in brain connectivity networks. We further develop an ordinal pattern based learning framework for automated brain disease diagnosis. Experimental results on both ADNI and ADHD-200 databases demonstrate the effectiveness of our method.

REFERENCES

- [1] K. J. Friston, C. D. Frith, P. F. Liddle, and R. S. J. Frackowiak, "Functional connectivity: The principal-component analysis of large (PET) data sets," *J. Cerebral Blood Flow Metabolism*, vol. 13, no. 1, pp. 5–14, 1993.
- [2] A. M. Aertsen, G. L. Gerstein, M. K. Habib, and G. Palm, "Dynamics of neuronal firing correlation: Modulation of 'effective connectivity,'" *J. Neurophysiol.*, vol. 61, no. 5, pp. 900–917, 1989.
- [3] E. C. Robinson, A. Hammers, A. Ericsson, A. D. Edwards, and D. Rueckert, "Identifying population differences in whole-brain structural networks: A machine learning approach," *NeuroImage*, vol. 50, no. 3, pp. 910–919, 2010.
- [4] O. Sporns, "From simple graphs to the connectome: Networks in neuroimaging," *NeuroImage*, vol. 62, no. 2, pp. 881–886, 2012.
- [5] J. Zhang, M. Liu, L. An, Y. Gao, and D. Shen, "Alzheimer's disease diagnosis using landmark-based features from longitudinal structural MR images," *IEEE J. Biomed. Health Inform.*, vol. 21, no. 6, pp. 1607–1616, Nov. 2017.
- [6] M. E. Shenton, M. Kubicki, and N. Makris, "Understanding alterations in brain connectivity in attention-deficit/hyperactivity disorder using imaging connectomics," *Biol. Psychiatry*, vol. 76, no. 8, pp. 601–602, 2014.

- [7] J. Mitra *et al.*, "Statistical machine learning to identify traumatic brain injury (TBI) from structural disconnections of white matter networks," *NeuroImage*, vol. 129, pp. 247–259, Apr. 2016.
- [8] M. Liu, J. Zhang, E. Adeli, and D. Shen, "Landmark-based deep multi-instance learning for brain disease diagnosis," *Med. Image Anal.*, vol. 43, pp. 157–168, Jan. 2018.
- [9] M. Liu, D. Zhang, and D. Shen, "Relationship induced multi-template learning for diagnosis of Alzheimer's disease and mild cognitive impairment," *IEEE Trans. Med. Imag.*, vol. 35, no. 6, pp. 1463–1474, Jun. 2016.
- [10] M. Rubinov and O. Sporns, "Complex network measures of brain connectivity: Uses and interpretations," *NeuroImage*, vol. 52, no. 3, pp. 1059–1069, 2010.
- [11] C.-Y. Wee *et al.*, "Enriched white matter connectivity networks for accurate identification of MCI patients," *NeuroImage*, vol. 54, no. 3, pp. 1812–1822, 2011.
- [12] B. Jie, D. Zhang, C.-Y. Wee, and D. Shen, "Topological graph kernel on multiple thresholded functional connectivity networks for mild cognitive impairment classification," *Hum. Brain Mapping*, vol. 35, no. 7, pp. 2876–2897, 2014.
- [13] F. Fei, B. Jie, and D. Zhang, "Frequent and discriminative subnetwork mining for mild cognitive impairment classification," *Brain Connectivity*, vol. 4, no. 5, pp. 347–360, 2014.
- [14] B. Jie, C.-Y. Wee, D. Shen, and D. Zhang, "Hyper-connectivity of functional networks for brain disease diagnosis," *Med. Image Anal.*, vol. 32, pp. 84–100, Aug. 2016.
- [15] A. Fornito, A. Zalesky, and E. Bullmore, *Fundamentals of Brain Network Analysis*. San Diego, CA, USA: Academic, 2016.
- [16] B. Mivsić *et al.*, "Cooperative and competitive spreading dynamics on the human connectome," *Neuron*, vol. 86, no. 6, pp. 1518–1529, 2015.
- [17] J. Goñi *et al.*, "Resting-brain functional connectivity predicted by analytic measures of network communication," *Proc. Nat. Acad. Sci. USA*, vol. 111, no. 2, pp. 833–838, 2014.
- [18] M. Liu, J. Du, B. Jie, and D. Zhang, "Ordinal patterns for connectivity networks in brain disease diagnosis," in *Proc. Int. Conf. Med. Image Comput. Comput.-Assisted Interv.*, 2016, pp. 1–9.
- [19] X. Peng, P. Lin, T. Zhang, and J. Wang, "Extreme learning machine-based classification of ADHD using brain structural MRI data," *PLoS ONE*, vol. 8, no. 11, p. e79476, 2013.
- [20] M. Liu, J. Zhang, P.-T. Yap, and D. Shen, "View-aligned hypergraph learning for Alzheimer's disease diagnosis with incomplete multi-modality data," *Med. Image Anal.*, vol. 36, pp. 123–134, Feb. 2017.
- [21] D. Durante, M. Daianu, N. Jahanshad, P. M. Thompson, and D. B. Dunson. (2015). "Unifying inference on brain network variations in neurological diseases: The Alzheimer's case." [Online]. Available: <https://arxiv.org/abs/1510.05391>
- [22] B. Solmaz, S. Dey, A. R. Rao, and M. Shah, "ADHD classification using bag of words approach on network features," *Proc. SPIE, Med. Imag., Image Process.*, vol. 8314, p. 83144T, Feb. 2012.
- [23] J. Zhou *et al.*, "Divergent network connectivity changes in behavioural variant frontotemporal dementia and Alzheimer's disease," *Brain*, vol. 133, no. 5, pp. 1352–1367, 2010.
- [24] A. Abraham *et al.*, "Deriving reproducible biomarkers from multi-site resting-state data: An autism-based example," *NeuroImage*, vol. 147, pp. 736–745, Feb. 2017.
- [25] R. Casanova, C. T. Whitlow, B. Wagner, M. A. Espeland, and J. A. Maldjian, "Combining graph and machine learning methods to analyze differences in functional connectivity across sex," *Open Neuroimag. J.*, vol. 6, pp. 1–9, Jan. 2012.
- [26] E. Challis, P. Hurley, L. Serra, M. Bozzali, S. Oliver, and M. Cercignani, "Gaussian process classification of Alzheimer's disease and mild cognitive impairment from resting-state fMRI," *NeuroImage*, vol. 112, pp. 232–243, May 2015.
- [27] X. Wang, Y. Jiao, T. Tang, H. Wang, and Z. Lu, "Altered regional homogeneity patterns in adults with attention-deficit hyperactivity disorder," *Eur. J. Radiol.*, vol. 82, no. 9, pp. 1552–1557, 2013.
- [28] S. Dey, A. R. Rao, and M. Shah, "Attributed graph distance measure for automatic detection of attention deficit hyperactive disordered subjects," *Front. Neural Circuits*, vol. 8, p. 64, Jun. 2014.
- [29] B. Jie, D. Zhang, W. Gao, Q. Wang, C.-Y. Wee, and D. Shen, "Integration of network topological and connectivity properties for neuroimaging classification," *IEEE Trans. Biomed. Eng.*, vol. 61, no. 2, pp. 576–589, Feb. 2014.
- [30] A. W. Chung *et al.*, "Characterising brain network topologies: A dynamic analysis approach using heat kernels," *NeuroImage*, vol. 141, pp. 490–501, Nov. 2016.
- [31] E. J. Sanz-Arigita *et al.*, "Loss of 'small-world' networks in Alzheimer's disease: Graph analysis of FMRI resting-state functional connectivity," *PLoS ONE*, vol. 5, no. 11, p. e13788, 2010.
- [32] W. De Haan, W. M. van der Flier, T. Koene, L. L. Smits, P. Scheltens, and C. J. Stam, "Disrupted modular brain dynamics reflect cognitive dysfunction in Alzheimer's disease," *NeuroImage*, vol. 59, no. 4, pp. 3085–3093, 2012.
- [33] G. Chen *et al.*, "Modular reorganization of brain resting state networks and its independent validation in Alzheimer's disease patients," *Front. Hum. Neurosci.*, vol. 7, p. 746, Aug. 2013.
- [34] M. R. Brier *et al.*, "Functional connectivity and graph theory in preclinical Alzheimer's disease," *Neurobiol. Aging*, vol. 35, no. 4, pp. 757–768, 2014.
- [35] K. Konrad and S. B. Eickhoff, "Is the ADHD brain wired differently? A review on structural and functional connectivity in attention deficit hyperactivity disorder," *Hum. Brain Mapping*, vol. 31, no. 6, pp. 904–916, 2010.
- [36] L. Wang *et al.*, "Altered small-world brain functional networks in children with attention-deficit/hyperactivity disorder," *Hum. Brain Mapping*, vol. 30, no. 2, pp. 638–649, 2009.
- [37] C. R. Jack *et al.*, "The Alzheimer's disease neuroimaging initiative (ADNI): MRI methods," *J. Magn. Reson. Imag.*, vol. 27, no. 4, pp. 685–691, 2008.
- [38] K. R. Van Dijk, T. Hedden, A. Venkataraman, K. C. Evans, S. W. Lazar, and R. L. Buckner, "Intrinsic functional connectivity as a tool for human connectomics: Theory, properties, and optimization," *J. Neurophysiol.*, vol. 103, no. 1, pp. 297–321, 2010.
- [39] N. Tzourio-Mazoyer *et al.*, "Automated anatomical labeling of activations in SPM using a macroscopic anatomical parcellation of the MNI MRI single-subject brain," *NeuroImage*, vol. 15, no. 1, pp. 273–289, 2002.
- [40] J.-P. Thirion, "Image matching as a diffusion process: An analogy with Maxwell's demons," *Med. Image Anal.*, vol. 2, no. 3, pp. 243–260, 1998.
- [41] X. Yan and J. Han, "gSpan: Graph-based substructure pattern mining," in *Proc. IEEE Int. Conf. Data Mining*, Dec. 2002, pp. 721–724.
- [42] R. Tarjan, "Depth-first search and linear graph algorithms," *SIAM J. Comput.*, vol. 1, no. 2, pp. 146–160, 1972.
- [43] T. Anderson, *Multivariate Statistical Analysis*. New York, NY, USA: Wiley, 1984.
- [44] X. Yan, H. Cheng, J. Han, and P. S. Yu, "Mining significant graph patterns by leap search," in *Proc. ACM SIGMOD Int. Conf. Manage. Data*, 2008, pp. 433–444.
- [45] E. R. DeLong, D. M. DeLong, and D. L. Clarke-Pearson, "Comparing the areas under two or more correlated receiver operating characteristic curves: A nonparametric approach," *Biometrics*, vol. 44, no. 3, pp. 837–845, 1988.
- [46] L. Tian *et al.*, "Enhanced resting-state brain activities in ADHD patients: A fMRI study," *Brain Develop.*, vol. 30, no. 5, pp. 342–348, 2008.
- [47] G. E. Alexander, K. Chen, P. Pietrini, S. I. Rapoport, and E. M. Reiman, "Longitudinal PET evaluation of cerebral metabolic decline in dementia: A potential outcome measure in Alzheimer's disease treatment studies," *Amer. J. Psychiatry*, vol. 159, no. 5, pp. 738–745, 2002.
- [48] G. B. Frisoni *et al.*, "Detection of grey matter loss in mild Alzheimer's disease with voxel based morphometry," *J. Neurol., Neurosurg. Psychiatry*, vol. 73, no. 6, pp. 657–664, 2002.
- [49] K. Supekar, V. Menon, D. Rubin, M. Musen, and M. D. Greicius, "Network analysis of intrinsic functional brain connectivity in Alzheimer's disease," *PLoS Comput. Biol.*, vol. 4, no. 6, p. e1000100, 2008.
- [50] Y. Feng *et al.*, "fMRI connectivity analysis of acupuncture effects on the whole brain network in mild cognitive impairment patients," *Magn. Reson. Imag.*, vol. 30, no. 5, pp. 672–682, 2012.
- [51] T. Zhang and C. Davatzikos, "Optimally-discriminative voxel-based morphometry significantly increases the ability to detect group differences in schizophrenia, mild cognitive impairment, and Alzheimer's disease," *NeuroImage*, vol. 79, pp. 94–110, Oct. 2013.
- [52] F. Bai *et al.*, "Abnormal whole-brain functional connection in amnesic mild cognitive impairment patients," *Behav. Brain Res.*, vol. 216, no. 2, pp. 666–672, 2011.
- [53] Z. Liu *et al.*, "Altered topological patterns of brain networks in mild cognitive impairment and Alzheimer's disease: A resting-state fMRI study," *Psychiatry Res., Neuroimag.*, vol. 202, no. 2, pp. 118–125, 2012.
- [54] Y. Zhu, J. X. Yu, H. Cheng, and L. Qin, "Graph classification: A diversified discriminative feature selection approach," in *Proc. 21st ACM Int. Conf. Inf. Knowl. Manage.*, 2012, pp. 205–214.