Diagnosis of Mild Cognitive Impairment With Ordinal Pattern Kernel

Kai Ma, Shuo Huang, and Daoqiang Zhang, *Senior Member, IEEE*

***Abstract*— Mild cognitive impairment (MCI) belongs to the prodromal stage of Alzheimer's disease (AD). Accurate diagnosis of MCI is very important for possibly deferring AD progression. Graph kernels, which measure the similarity between paired brain connectivity networks, have been widely used to diagnose brain diseases (e.g., MCI) and yielded promising classification performance. However, most of the existing graph kernels are based on unweighted graphs, and neglect the valuable weighted information of the edges in brain connectivity networks where edge weights convey the strengths of fiber connection or temporal correlation between paired brain regions. Accordingly, in this paper, we propose a new graph kernel called ordinal pattern kernel for measuring brain connectivity network similarity and apply it to brain disease classification tasks. Different from the existing graph kernels which measure the topological similarity of the unweighted graphs, our proposed ordinal pattern kernel can not only calculate the similarity of paired brain connectivity networks, but also capture the ordinal pattern relationship of edge weights in brain connectivity networks. To appraise the effectiveness of our proposed method, we perform extensive experiments in functional magnetic resonance imaging data of brain disease from Alzheimer's Disease Neuroimaging Initiative database. The experimental results show that our proposed ordinal pattern kernel outperforms the state-of-the-art graph kernels in the classification tasks of MCI.**

***Index Terms*— Graph kernel, ordinal pattern, brain network, classification, mild cognitive impairment.**

I. Introduction

ALZHEIMER'S disease (AD) is a neurodegenerative brain disease and brings substantial irreversible neuron damage to human brain. AD usually develops very slowly at the beginning and becomes worse as time goes on, which eventually leads to death. The clinical manifestations of AD include memory loss, cognitive dysfunction, and behavioral disorders. At the prodromal stage of AD, which is called mild cognitive impairment (MCI), patients usually have less neuron damage compared to AD. MCI patients whose clinical condition is between normal aging and AD still have some brain cognitive functions. Nevertheless, they have high probability of progressing to AD. In every year, there are about 10%∼15% MCI patients who will convert to AD, while there are about 1%∼2% normal controls (NC) who will progress to AD [1]. The related studies showed that disease-modifying therapies for MCI patients could delay the onset of clinical dementia expression to a certain extent and help patients preserve some brain cognitive functions [2]. Hence, it is very important for possibly deferring AD progression to accurately diagnose MCI.

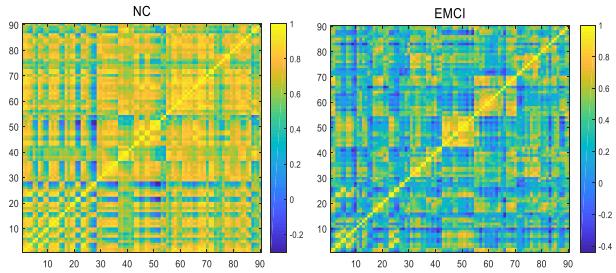

Fig. 1. Connectivity matrices of NC and early mild cognitive impairment patients. Compared with NC, the connectivities among brain regions in EMCI are changed. EMCI: early mild cognitive impairment. NC: normal controls.

A great deal of evidence from neuroscience studies indicate that brain cognitive function is closely related to neural activities between pairwise brain regions [3]. With the emergence of new medical imaging techniques, some non-invasive tools involving functional magnetic resonance imaging (fMRI) [4], electroencephalogram (EEG) [5] and functional near-infrared spectroscopy (fNIRS) [6], [7] can effectively capture the interaction patterns among brain neural activities. These interaction patterns in the brain can be described as a brain connectivity network with graph theory, which helps us better investigate the neuropathological mechanism of brain disease by analyzing the connectivities in brain functional network [8]. For example, the research based on brain functional connectivity network constructed from fMRI indicated that AD patients were significantly different from normal aging controls in the brain functional connections related to the hippocampus [9].

Manuscript received September 27, 2021; revised January 28, 2022 and March 7, 2022; accepted March 25, 2022. Date of publication April 11, 2022; date of current version April 22, 2022. This work was supported in part by the National Natural Science Foundation of China under Grant 62136004, Grant 61732006, and Grant 61876082 and in part by the National Key Research and Development Program of China under Grant 2018YFC2001600 and Grant 2018YFC2001602. *(Corresponding author: Daoqiang Zhang.)*

The authors are with the MIIT Key Laboratory of Pattern Analysis and Machine Intelligence, College of Computer Science and Technology, Nanjing University of Aeronautics and Astronautics, Nanjing, Jiangsu 210016, China (e-mail: dqzhang@nuaa.edu.cn).

This article has supplementary downloadable material available at https://doi.org/10.1109/TNSRE.2022.3166560, provided by the authors.
Digital Object Identifier 10.1109/TNSRE.2022.3166560

Brain connectivity network abstractly characterizes the structural or functional interaction of human brain, where brain regions correspond to nodes and functional or anatomical associations between nodes are considered as edges [10]. Brain connectivity network is widely applied to the researches of brain disease classification, including AD [11], attention deficit hyperactivity disorder [12], schizophrenia [13], and major depressive disorder [14]. In these studies, various network descriptors, e.g., degree, clustering coefficient [15], [16], are first extracted from connectivity networks as feature vectors. Then, these feature vectors are applied to machine learning classifiers to classify brain diseases. However, the brain connectivity network is a spatial topological structure. Representing the brain connectivity network as feature vectors will ignore the topological structure information of the brain connectivity network. Meantime, it is a challenge for the methods based on feature vector [17] to measure the topological similarity between a pair of networks (or graphs). To process this problem, in the last decade, many methods are proposed to measure the similarity among networks.

In various approaches, kernel method, especially graph kernel, offers a powerful framework for measuring the similarity of pairwise networks. Graph kernel is a kind of kernel function defined on graphs (i.e., networks), which has been widely applied to brain connectivity network analysis. However, most of the existing graph kernels are based on unweighted graph with edge presenting or not, and thus ignore the valuable weighted information of edges in brain connectivity network where edge weights convey the strengths of fiber connection or temporal correlation between paired brain regions. After transforming brain images into brain connectivity networks, each node in the network corresponds to a specific brain region. Brain regions are connected by different weighted edges (i.e., edges with weighted values). According to the different weighted values in network edges, there are specific ordinal pattern relationships among these edge weights. Previous graph kernels rarely take into account the weighted information of edge and the ordinal pattern relationship of edge weights, due to the inconvenience of measuring brain connectivity network similarity. A large amount of researches have indicated that the neurodegenerative brain diseases (e.g., MCI and AD) are associated with the connectivities among the specific brain regions [18]. The neurodegenerative disorders change the connectivities between the pairwise brain regions in brain networks of patients, which result in the different connectivity matrices compared to NC. We plot the connectivity matrices of NC and early MCI patients, as shown in Fig. 1. From Fig. 1, we can find that NC are different from early MCI patients in the connectivity matrices of brain networks. Meantime, the changed connectivities between brain regions also affect the ordinal pattern relationships of edge weights which result in the changed ordinal patterns. We plot the ordinal pattern whose start node is left amygdala (AMYG.L) in NC and early MCI patients, as shown in Fig. 2. Fig. 2 shows that the ordinal pattern from AMYG.L in NC is different from that in early MCI patients. Therefore, edge weight information and ordinal pattern relationships are very significant for brain connectivity network analysis.

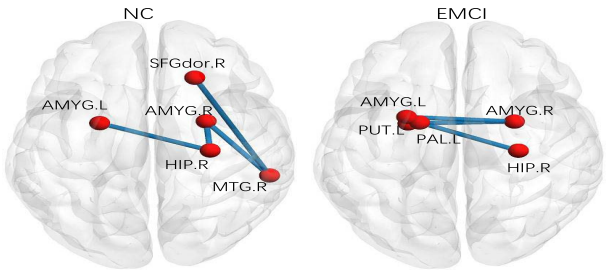

Fig. 2. Ordinal patterns of NC and early mild cognitive impairment patients. The ordinal pattern in NC is: *AMYG.L* → *HIP.R* → *AMYG.R* → *MTG.R* → *SFGdor.R*. The ordinal pattern in EMCI is: *AMYG.L* → *AMYG.R* → *PUT.L* → *PAL.L* → *HIP.R*. Compared to NC, the ordinal pattern from AMYG.L in EMCI is changed. EMCI: early mild cognitive impairment. NC: normal controls. The full names of brain regions are shown in Table IV.

In the previous work [19], ordinal pattern as a new descriptor for brain connectivity network was proposed, which could take advantage of the weight information of edges. In this work, the frequent ordinal patterns in brain connectivity networks were firstly identified with the frequent ordinal pattern mining algorithm. Then, the discriminative ordinal pattern selection was performed and feature representations were extracted based on the selected ordinal patterns. At last, these feature representations were applied to support vector machine (SVM) for automated brain disease diagnosis. Adopting the ordinal pattern method in [19] for classification task needs to perform frequent ordinal pattern mining and discriminative ordinal pattern selection for each brain region. Obviously, this method will cost much time. Meantime, this method regards the ordinal patterns as feature vectors and ignores the topological structure information of ordinal pattern in brain connectivity network, as shown in Fig. 2.

To address these problems, we develop a new graph kernel called ordinal pattern kernel for measuring the similarities between brain connectivity networks. In this work, according to the definition of ordinal pattern, we firstly introduce our proposed general ordinal pattern (GOP) kernel and provide the theoretical foundations and proofs for it. Then, we find that computing the general ordinal pattern kernel is non-deterministic polynomial hard (NP-hard), which is computationally intractable. To address this problem, we further propose a depth-first-based ordinal pattern (DOP) kernel, which measures the similarity of paired brain connectivity networks by matching the depth-first-based ordinal pattern of each node. Different from the existing graph kernels, the proposed ordinal pattern kernel can not only calculate the similarity of brain connectivity networks via comparing node ordinal patterns, but also capture the ordinal pattern relationship of edge weights. We evaluate our ordinal pattern kernel in the network data of brain diseases from Alzheimer's Disease Neuroimaging Initiative (ADNI)[1] database.

There are three main contributions in this paper: (1) The proposed ordinal pattern kernel can make full use of the weighted information of edge in brain network and outperforms the existing state-of-the-art graph kernels in the

[1] http://adni.loni.usc.edu/

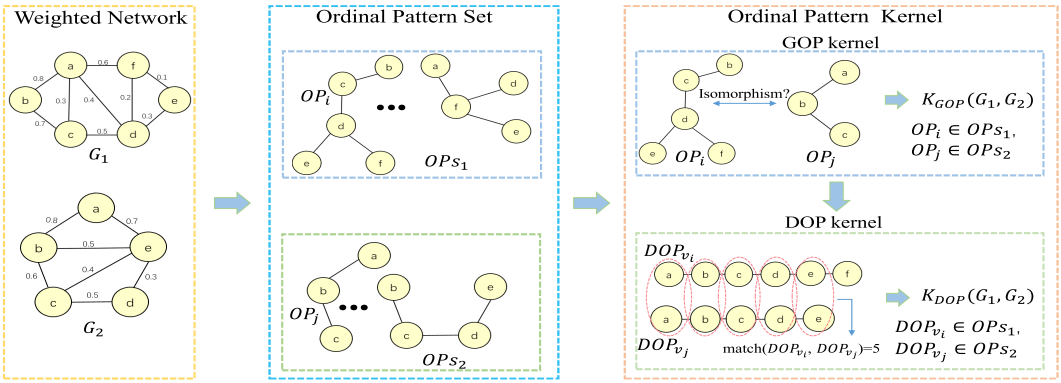

Fig. 3. Illustration of the ordinal patterns and ordinal pattern kernel defined on weighted networks. The left are weighted network G_1 and G_2. The middle are ordinal pattern sets from network G_1 and G_2. The right are ordinal pattern kernel defined on weighted networks. OP_i and OP_j are two ordinal patterns from G_1 and G_2. OPs_1 and OPs_2 are two ordinal pattern sets decomposed from weighted network G_1 and G_2. DOP_{v_i} and DOP_{v_j} are two depth-first-based ordinal patterns with starting node v_i and v_j which are respectively from G_1 and G_2. GOP kernel is general ordinal pattern kernel. DOP kernel is depth-first-based ordinal pattern kernel.

TABLE I
CHARACTERISTIC OF THE SUBJECT (MMSE±STANDARD DEVIATION). MMSE: MINI-MENTAL STATE EXAMINATION. EMCI: EARLY MILD COGNITIVE IMPAIRMENT. LMCI: LATE MILD COGNITIVE IMPAIRMENT. NC: NORMAL CONTROLS

Group	NC	EMCI	LMCI
Male/Female	21/29	26/30	20/23
Age(Mean±SD)	75.0±6.9	72.8±6.4	73.1±6.2
MMSE(Mean±SD)	28.9±1.6	28.1±2.3	27.7±1.5

classification accuracy; (2) The proposed ordinal pattern kernel has strong robustness. When brain networks have missing data (i.e., network edges have null values), our method can still acquire the best classification accuracy. (3) The proposed ordinal pattern kernel can be used to investigate the discriminative ordinal patterns of patients with brain disease, compared with NC.

The rest of the paper is organized as follows. In Section II, we introduce materials and describe the proposed method. In Section III, we present the experimental settings and results. In Section IV, we present the discussion for the experimental results. At last, we summarize this paper in Section V.

II. MATERIALS AND METHOD

A. Subjects

The network data of brain diseases used in the experiments are constructed from the resting state fMRI data which are from ADNI. There are 56 early mild cognitive impairment (EMCI) patients, 43 late mild cognitive impairment (LMCI) patients, and 50 NC in our research. All subjects were scanned on 3.0 Tesla Philips scanners. The image resolution in X and Y dimensions is from 2.29 mm to 3.31 mm. The slice thickness is 3.31 mm. TE is 30 ms and TR is from 2.2 s to 3.1 s. We report the subjects' demographic and clinical information in Table I.

B. Image Processing and Brain Network Construction

All the acquired fMRI datasets were preprocessed with data processing assistant for resting-state fMRI (rs-fMRI) [20], which is an integrated toolbox consisting of statistical parametric mapping[2] and rs-fMRI analysis toolkit.[3] The DICOM data of all subjects were transformed into the file format of neuroimaging informatics technology initiative. The full name of "DICOM" is digital imaging and communications in medicine. In order to enable all subjects adapt to the machine noise, the first ten time points of all data were removed for signal equilibrium. Slice-timing correction was executed by utilizing the middle slice as a reference frame. Head motion correction was performed for correcting the brain of each subject in the same position of every image. Functional brain images of all subjects were spatially mapped into a normalized Montreal Neurological Institute space by utilizing an echo-planar imaging template. The fMRI data of all subjects were smoothed via utilizing a Gaussian kernel which has a full-width at half-maximum of 6 mm. The frequency ranges of band pass filtering and global drift removal were selected from 0.01Hz to 0.1Hz.

We utilize Automated Anatomical Labeling (AAL) atlas [21] to divide brain structures of all subjects into 90 brain regions. In our research, we regard each brain region as a region-of-interest (ROI). Within the specific ROI, we average the signals of blood oxygen level-dependent based on all voxels, and then calculate mean rs-fMRI time series for each ROI. In order to acquire the functional connectivity between a pair of ROIs, we compute the linear correlation (i.e., Pearson correlation) of mean time series between paired ROIs. Accordingly, we obtain a 90 × 90 brain functional network of each subject. In this network, each node corresponds to one ROI and the edge weight corresponds to functional connection strength.

C. Ordinal Pattern and Isomorphism

In the previous work [19], ordinal pattern was regarded as a new descriptor for brain connectivity networks, which provided ordinal edge sequences for each node. This work regarded the ordinal patterns as feature vectors and ignored the topological structure information of ordinal pattern in brain

[2]http://www.fil.ion.ucl.ac.uk/spm
[3]http://www.restfmri.net

connectivity network. To address this problem, we extend ordinal pattern into graph structure and redefine it with graph theory. In this section, we mainly introduce the concepts of ordinal pattern in graph structure and ordinal pattern isomorphism.

1) Ordinal Pattern: A weighted network or graph G consists of a set of nodes V, edges E and weight vectors W, $G = (V, E, W)$. W is the weight vector for those edges with the i-th element $W(e_i)$ representing the connection strength of the edge e_i, $e_i \in E$. The ordinal pattern (OP) defined in graph G is a set including ordinal nodes and ordinal edges $OP = (V_{op}, E_{op})$. E_{op} is an ordinal edge set, $E_{op} = \{e_1, e_2, \cdots, e_i, e_j, \cdots, e_M\} \subseteq E$, all $0 < i < j \leq M$, $W(e_i) > W(e_j)$, e_i and e_j are called ordinal edges. V_{op} is a vertex set where vertexes are connected by ordinal edges included in E_{op}. The illustration of ordinal patterns can be seen in Fig. 3. OP_i and OP_j are two ordinal patterns.

2) Isomorphism: A graph or network can be decomposed into multiple ordinal patterns. Here, we call the set consisting of all ordinal patterns ordinal pattern set (OPs). OPs_1 and OPs_2 are two ordinal pattern sets of the graph G_1 and G_2, $OP_1 = (V_{op_1}, E_{op_1})$ and $OP_2 = (V_{op_2}, E_{op_2})$ are two ordinal patterns, $OP_1 \in OPs_1$, $OP_2 \in OPs_2$. An ordinal pattern isomorphism between two ordinal patterns OP_1 and OP_2 is a bijective mapping $\varphi : V_{op_1} \to V_{op_2}$, i.e., $\forall v_{op_1}, v'_{op_1} \in V_{op_1} : (v_{op_1}, v'_{op_1}) \in E_{op_1} \Leftrightarrow \varphi(v_{op_1}), \varphi(v'_{op_1}) \in V_{op_2}$, $(\varphi(v_{op_1}), \varphi(v'_{op_1})) \in E_{op_2}$. OP_1 and OP_2 are isomorphic, written as $OP_1 \cong OP_2$. $V'_1 \subseteq V_{op_1}$ and $V'_2 \subseteq V_{op_2}$ are subsets of ordinal pattern vertices. An ordinal pattern isomorphism τ of $OP_1[V'_1]$ and $OP_2[V'_2]$ is called sub-ordinal pattern isomorphism ($SOPI$) of OP_1 and OP_2.

D. GOP Kernel

We suppose that a graph G has N ordinal patterns, the ordinal pattern set of graph G is $OPs = \{OP_1, \cdots, OP_i, \cdots, OP_N\}$, OP_i is the *i-th* ordinal pattern, $1 \leq i \leq N$. Two graphs G_1 and G_2 have their own ordinal pattern set which is OPs_1 and OPs_2, respectively. φ_{ij} is the isomorphism mapping from ordinal pattern OP_i to ordinal pattern OP_j, $OP_i \in OPs_1$, $OP_j \in OPs_2$. Let $\Psi(OP_i, OP_j)$ refer to the set which includes all sub-ordinal pattern isomorphisms of OP_i and OP_j. Υ: $\Psi(OP_i, OP_j) \to \mathbb{R}^+$ is a weight function calculating the node number in sub-ordinal patterns, when sub-ordinal patterns are isomorphic. The sub-ordinal pattern isomorphism kernel is defined as:

$$K_{SOPI}(OP_i, OP_j) = \sum_{\tau \in \Psi(OP_i, OP_j)} \Upsilon(\tau) \tag{1}$$

where K_{SOPI} is a kernel function, and measure the similarity between ordinal pattern OP_i and OP_j by using isomorphic sub-ordinal patterns in OP_i and OP_j. If these sub-ordinal patterns are isomorphic, Υ return the node number of sub-ordinal patterns, otherwise, Υ return zero.

Theorem 1: Sub-ordinal pattern isomorphism kernel K_{SOPI} in Eq.(1) is a positive semidefinite (p.s.d) kernel.

Proof: The sub-ordinal pattern can be regarded as a kind of the sub-structure of the graph. The kernel measuring the similarity between graph sub-structures is *p.s.d* [22]. Hence, the K_{SOPI} is *p.s.d*.

The sub-ordinal pattern isomorphism kernel measures the similarity between two ordinal patterns by counting the number of sub-ordinal pattern isomorphisms. Then, we get the GOP kernel between two graphs G_1 and G_2, defined as:

$$K_{GOP}(G_1, G_2) = \sum_{OP_i \in OPs_1} \sum_{OP_j \in OPs_2} \textit{iso-count}(OP_i, OP_j) \tag{2}$$

where OPs_1 and OPs_2 are the ordinal pattern sets of G_1 and G_2, *iso-count*$(\cdot)$ is a function calculating isomorphism:

$$\textit{iso-count}(OP_i, OP_j) = \begin{cases} \lambda_o(OP_i), & if\ OP_i,\ OP_j\ isomorphic \\ K_{SOPI}(OP_i, OP_j), & otherwise. \end{cases} \tag{3}$$

where λ_o is a weight function: $OPs \to \mathbb{R}^+$, which counts the node number when two ordinal patterns are isomorphic.

Theorem 2: The general ordinal pattern kernel K_{GOP} in Eq.(2) is p.s.d.

Proof: The *iso-count*(OP_i, OP_j) is a *p.s.d* function, K_{GOP} is the summation of *iso-count*(OP_i, OP_j), and hence K_{GOP} is *p.s.d*.

Although we have known how to calculate general ordinal pattern kernel, there still exists one problem that the kernel based on ordinal pattern isomorphism is NP-hard.

Theorem 3: Computing the kernel K_{GOP} is NP-hard.

Proof: K_{GOP} is a kernel based on ordinal pattern isomorphism which is a kind of subgraph isomorphism belonging to NP-hard problem [22]. Hence, K_{GOP} is NP-hard.

E. DOP Kernel

There are two problems in the above computation of GOP kernel. One problem is that computing the kernel K_{GOP} is NP-hard, which is computationally intractable. The other problem is that an ordinal pattern (e.g., OP_1) may be the sub-ordinal pattern of another ordinal pattern (e.g., OP_2) in one graph. For example, in Fig. 3, given two ordinal patterns OP_1 and OP_2, OP_1 and OP_2 are from weighted network G_1. $OP_1 = (V_{op_1}, E_{op_1})$, $V_{op_1} = \{\text{a,b,c}\}$, $E_{op_1} = \{e_{ab}, e_{bc}\}$, $OP_2 = (V_{op_2}, E_{op_2})$, $V_{op_2} = \{\text{a,b,c,d}\}$, $E_{op_2} = \{e_{ab}, e_{bc}, e_{cd}\}$, OP_1 is the sub-structure of OP_2. This problem brings redundant calculations for general ordinal pattern kernel. In order to address these two problems, we further propose a modified ordinal pattern kernel based on depth-first search called depth-first-based ordinal pattern kernel. We adopt the depth-first search algorithm to seek the deepest ordinal pattern for each node in the graph and then design the relevant ordinal pattern kernel on them. The ordinal pattern constructed by depth-first search is called depth-first-based ordinal pattern. The DOP is a linear structure, and hence the isomorphism problem in DOP can be regarded as a matching problem.

1) DOP: Here, we introduce the establishment process of depth-first-based ordinal pattern in detail. A weighted network or graph G consists of a set of nodes V, edges E and weight vectors W, $G = (V, E, W)$. $\forall u \in V$, the neighborhood vertex

Algorithm 1 Depth-First-Based Ordinal Pattern of Node v_0: $DOP(W, v_0)$

Input: Weight matrix W of graph G, start node v_0
Output: The deepest ordinal pattern of node v_0: $DOP(v_0)$

$visitnode[1] = v_0$ % Mark the node v_0 as the visited node
$nextnode = V(max\{\mathbb{W}(v_0)\})$ % V() is node function, return the node 'nextnode'
$visitnode[2] = nextnode$
$v = v_0$; $i = 2$
while !isempty ($\delta(nextnode)$) **do**
  j=1
  $W_{des} = descend(\mathbb{W}(nextnode))$ % Sort the set $\mathbb{W}(nextnode)$ in descending order
  Len-W_{des}=Length (W_{des}) % Calculate the number of edges in W_{des}
  while $j \leq$Len-W_{des} **do**
    if $W_{des}[j] < W(nextnode, v)$ **then**
      $v = nextnode$
      $nextnode = V(W_{des}[j])$
      $i = i + 1$
      $visitnode[i] = nextnode$
      break
    else
      j=j+1
return $visitnode_{v_0}$ % Depth-first-based ordinal pattern sequence of node v_0

set of a vertex u: $\delta(u) = \{v : (u, v) \in E, v \in V\}$, the edge weight set between vertex u and its neighborhood vertices: $\mathbb{W}(u) = \{W(u, v) : v \in \delta(u), (u, v) \in E\}$. In graph G, we choose a node as the start node v_0 and use the depth-first search algorithm to seek the depth-first-based ordinal pattern of node v_0, as shown in Algorithm 1.

2) DOP Kernel: We use Algorithm 1 to calculate the DOP for each node in the graphs G_1 and G_2. Subsequently, we will utilize these ordinal patterns to construct the depth-first-based ordinal pattern kernel K_{DOP} between the graphs G_1 and G_2. In our research, G_1 and G_2 represent the brain networks of two subjects. The depth-first-based ordinal pattern kernel is defined as follows:

$$K_{DOP}(G_1, G_2) = \sum_{v_i \in G_1} \sum_{v_j \in G_2} match(DOP(v_i), DOP(v_j)) \tag{4}$$

where $DOP(v_i)$ and $DOP(v_j)$ are the depth-first-based ordinal patterns of node v_i and v_j, and

$$match\left(DOP\left(v_i\right), DOP\left(v_j\right)\right) = \sum_{\substack{p \subseteq DOP(v_j) \\ q \subseteq DOP(v_i)}} \lambda |\{p : p \cong q\}| \tag{5}$$

where λ is a sequence $\lambda_1, \lambda_2, \ldots, \lambda_N$ of weights ($\lambda_i \in \mathbb{R}; \lambda_i > 0 \; for \; all \; i \in N$). Because, $DOP(v_i)$ and $DOP(v_j)$ are linear, $match\left(DOP\left(v_i\right), DOP\left(v_j\right)\right)$ can be calculated by matching the node numbers between $DOP\left(v_i\right)$ and $DOP\left(v_j\right)$, as shown in Fig. 3. The computational process of DOP kernel is shown in Algorithm 2.

Algorithm 2 $K_{DOP}(G_1, G_2)$

Input: Brain network G_1, G_2
Output: The depth-first-based ordinal pattern kernel K_{DOP} between G_1 and G_2

$G_1NodeNum = size(G_1, 1)$ % The node number of graph G_1
$G_2NodeNum = size(G_2, 1)$ % The node number of graph G_2
$K_{DOP} = 0$
for i=1 to $G_1NodeNum$ **do**
  for j=1 to $G_2NodeNum$ **do**
    $DOP(v_i) \leftarrow DOP(G_1, v_i)$
    $DOP(v_j) \leftarrow DOP(G_2, v_j)$
    $K_{match} = match(DOP(v_i), DOP(v_j))$ % K_{match} is a function measuring the similarity between $DOP(v_i)$ and $DOP(v_j)$. $match()$ function is defined in Eq.(5).
    $K_{DOP} = K_{DOP} + K_{match}$ % K_{DOP} is a function measuring the similarity between G_1 and G_2, which is defined in Eq.(4).
return K_{DOP}

Theorem 4: Depth-first-based ordinal pattern kernel K_{DOP} in Eq.(4) is p.s.d.

Proof: According to the definitions of the sub-ordinal pattern isomorphism kernel and the literature [23], we know that $match\left(DOP\left(v_i\right), DOP\left(v_j\right)\right)$ is a *p.s.d* kernel function measuring the similarity between $DOP(v_i)$ and $DOP(v_j)$. This kernel function is calculated by matching the node numbers between $DOP(v_i)$ and $DOP(v_j)$. Hence, K_{DOP} is *p.s.d.*

Computation Complexity: We suppose that each brain network has N nodes, then the complexity for calculating depth-first-based ordinal pattern kernel between a pair of brain network is $\mathcal{O}(N^2)$.

F. Ordinal Pattern Kernel Based Learning

We use the image processing method mentioned in materials and method to process the resting state fMRI data of all subjects and construct brain functional connectivity networks for all subjects. In brain functional connectivity network, each node represents a brain region, and the weighted edge calculated by the Pearson correlation coefficient denotes the functional connection between paired brain regions. Given the brain connectivity networks of all subjects, we compute our proposed ordinal pattern kernel on them with Algorithm 2 and apply the kernel matrix to SVM classifier for brain disease classification. Suppose that we have N subjects, then the size of the kernel matrix is $N \times N$.

III. Results

A. Methods for Comparison

We compare the proposed ordinal pattern kernel with the state-of-the-art graph kernels involving Weisfeiler-Lehman

TABLE II
CLASSIFICATION RESULTS OF ALL METHODS. DOP IS THE PROPOSED METHOD. ACC: CLASSIFICATION ACCURACY, AUC: AREA UNDER RECEIVER OPERATING CHARACTERISTIC CURVE. EMCI: EARLY MILD COGNITIVE IMPAIRMENT. LMCI: LATE MILD COGNITIVE IMPAIRMENT. NC: NORMAL CONTROLS

Method	EMCI vs. NC		EMCI vs. LMCI		LMCI vs. NC	
	ACC(%)	AUC	ACC(%)	AUC	ACC(%)	AUC
SP	61.38	0.602	63.57	0.618	67.72	0.652
WL-ST	69.41	0.653	71.37	0.683	73.43	0.714
WL-SP	63.18	0.607	64.15	0.623	69.64	0.683
RW	66.87	0.641	70.18	0.691	71.94	0.702
PM	70.24	0.701	72.34	0.707	73.42	0.717
WWL	71.13	0.695	73.19	0.712	79.85	0.722
GH	65.56	0.649	66.23	0.654	69.74	0.651
Tree++	61.29	0.602	57.64	0.551	66.72	0.643
SKL	71.36	0.706	74.81	0.726	82.56	0.732
DOP	**75.47**	**0.737**	**77.42**	**0.743**	**89.25**	**0.858**

subtree kernel (WL-ST) [24], shortest path kernel (SP) [25], Weisfeiler-Lehman shortest path kernel (WL-SP) [24], random walk kernel (RW) [26], pyramid match kernel (PM) [27], Wasserstein Weisfeiler-Lehman graph kernel (WWL) [28], GraphHopper kernel (GH) [29], truncated tree based graph kernels (Tree++) [30], and sub-network kernels (SKL) [31]. WL-ST and WL-SP are two parts of graph Weisfeiler-Lehman framework which are proposed by Shervashidze *et al.* [24]. WWL is proposed by Togninalli et al [28]. They utilize the Wasserstein distance to measure the similarity of pairwise graphs based on Weisfeiler-Lehman. PM measures the graph similarity on hierarchical structures which look like pyramid. SP, GH and Tree++ are based on path kernel. SP computes the matching shortest path for measuring the similarity of pairwise graphs. GH counts sub-path similarities for pairwise graphs. Tree++ uses super path containing truncated breadth-first search trees rooted at the vertices in a path to measure the graph similarity. RW counts the number of matching random walks for measuring the similarity between two graphs. SKL measures the similarity between paired networks by calculating the similarity of sub-networks from nodes.

B. Experimental Setting

We perform three kinds of binary classification tasks in the experiments involving 1) EMCI vs. NC classification, 2) LMCI vs. NC classification, and 3) EMCI vs. LMCI classification. The experiments are performed in weighted functional networks of all subjects. SVM as the final classifier is exploited to conduct the classification experiment, and is based on a SVM tool called LIBSVM library [32]. The tradeoff parameter C in the SVM is selected from $\{10^{-3}, 10^{-2}, \ldots, 10^{3}\}$. We utilize the default parameters offered by authors in the competing methods. We implement leave-one-out cross-validation for all the classifications. We evaluate the performances of all methods with classification accuracy (ACC) and area under receiver operating characteristic curve (AUC).

We use kernel two-sample test (KTST) [33], [34] to analyze the discriminative ordinal pattern, the significant level is 0.05. The alternative hypothesis is that the ordinal patterns of NC have differences with those of patients (e.g., EMCI and LMCI). In KTST, the input information is weighted functional network. We use BrainNet Viewer toolbox (version: 1.61) [35] to plot the brain connectivity figures (e.g., Fig.2 and Fig.5) and brain region figure (e.g., Fig.6). We use MATLAB (version: 2018a) to plot the deep ordinal pattern sequences in Table III.

C. Classification Performance

We compare our ordinal pattern kernel with the state-of-the-art graph kernels in three classification tasks, including (1) EMCI vs. NC classification, (2) EMCI vs. LMCI classification, and (3) LMCI vs. NC classification. The classification results of all methods are summarized in Table II. From Table II, we can find that our proposed graph kernel outperforms the control methods in three classification tasks of brain diseases. Specifically, our proposed method obtain the accuracies of 75.47%, 77.42%, and 89.25% for EMCI vs. NC, EMCI vs. LMCI, and LMCI vs. NC classification, respectively, while the best accuracies of the control methods are 71.36%, 74.81%, and 82.56%, respectively. In addition, the AUC of our proposed method is respectively 0.737, 0.743, and 0.858 in three classification tasks of brain diseases, which demonstrates the excellent diagnostic power, meanwhile the best AUC from the competing methods is respectively 0.706, 0.726, and 0.732. These classification results demonstrates that our proposed ordinal pattern kernel is good at distinguishing EMCI and LMCI from NC, and distinguishing EMCI from LMCI, compared to the state-of-the-art kernels.

D. Robustness Analysis

Information missing [36], [37], which often exists in the research field of brain fMRI, will bring challenges for brain network classification. Hence, the robust methods are very important for the classification task based on brain connectivity network. In order to verify the robustness of our proposed ordinal pattern kernel in brain connectivity network with missing information, we randomly throw out the specific percentage of connectivity edges (5%, 15%, 25%, 35%, 45%, respectively) in the brain network of each subject, i.e., the specific percentage of network edge values are set as null values. Then, we measure the similarities among these specific brain networks with our ordinal pattern kernel. We perform the classification experiments on these network data with SVM classifier. The classification results are shown in Fig. 4.

From Fig. 4, we can find that our ordinal pattern kernel can acquire robust and excellent classification accuracies in these brain network data with missing information. For example, in the classification task of missing data (i.e., missing rate $=$ $\{5\%, 15\%, 25\%, 35\%, 45\%\}$) between EMCI and NC, the classification accuracy obtained from our ordinal pattern kernel is 75.38%, 75.06%, 74.63%, 73.82%, 73.23%, respectively, and the AUC is 0.741, 0.73, 0.721, 0.71, 0.709, respectively, which are better than the second best results (i.e., 70.89%, 70.62%, 70.05%, 68.66%, 66.37% in ACC, 0.702, 0.694, 0.671, 0.664, 0.646 in AUC). In the classification task of missing data between EMCI and LMCI, the classification accuracy obtained from our ordinal pattern kernel is 76.68%, 76.01%, 75.88%, 75.16%, 74.64%, respectively, and the AUC

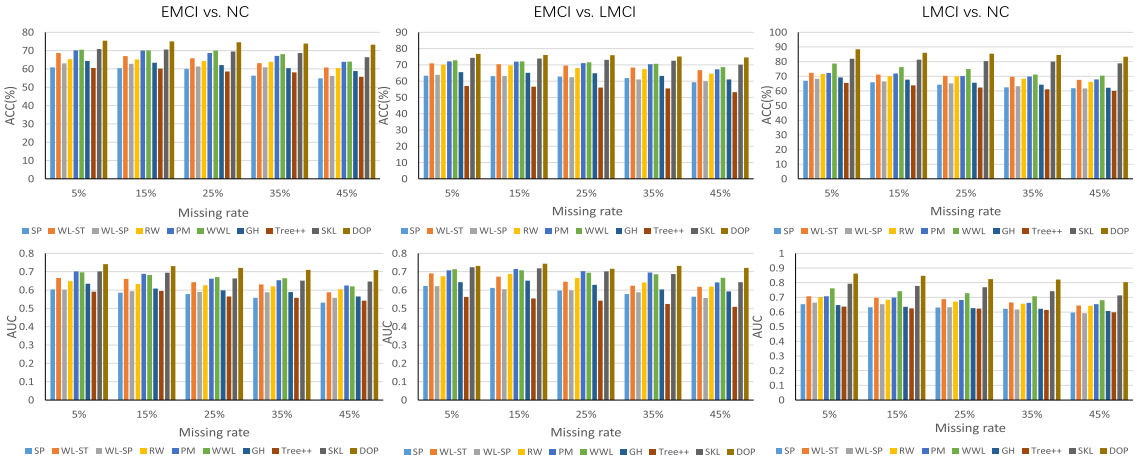

Fig. 4. Classification results in brain connectivity networks with different missing rates (*i.e.*, *missing rate* $= \{5\%, 15\%, 25\%, 35\%, 45\%\}$). ACC: classification accuracy. AUC: area under receiver operating characteristic curve. EMCI: early mild cognitive impairment. LMCI: late mild cognitive impairment. NC: normal controls.

is 0.732, 0.743, 0.716, 0.732, 0.721, respectively, which are better than the second best results (i.e., 74.25%, 73.89%, 73.12%, 72.49%, 70.18% in ACC, 0.724, 0.718, 0.703, 0.695, 0.667 in AUC). In the classification task of missing data between LMCI and NC, the classification accuracy obtained from our ordinal pattern kernel is 88.36%, 86.04%, 85.41%, 84.58%, 83.38%, respectively, and the AUC is 0.862, 0.847, 0.825, 0.821, 0.804, respectively, which are better than the second best results (i.e., 82.01%, 81.38%, 80.27%, 80.06%, 78.89% in ACC, 0.793, 0.778, 0.769, 0.742, 0.713 in AUC). Fig. 4 shows that the classification accuracies and AUCs obtained by our proposed method are better than the competing graph kernels in the brain connectivity network with missing data. These classification results verify the robustness of our proposed ordinal pattern kernel, and indicate that our proposed method is good at classifying EMCI and LMCI from NC, and classifying EMCI from LMCI in brain connectivity networks with missing data, compared to the state-of-the-art kernels..

E. Discriminative Ordinal Patterns

In this section, we further investigate the discriminative substructure [39] (i.e., DOP) in brain networks of patients and NC with our proposed ordinal pattern kernel. The matched function $match(DOP(v_i), DOP(v_j))$ defined in Eq. (5) measures the similarity between two DOPs, which is also a ordinal pattern kernel called matched kernel. We apply the matched kernel to KTST [33], [34], which is a two-sample test based on kernel function determining whether two samples are drawn from different distributions, and use this two-sample test method to identify some discriminative DOPs.

We calculate the DOP for each brain region in brain connectivity network of each subject with Algorithm 1. Then, we calculate the matched kernels for these DOPs between paired brain networks and apply these kernels to KTST to determine whether EMCI or LMCI patients are significantly different from NC in these DOPs. If DOPs in EMCI or LMCI patients are significantly different (i.e., p value is smaller than 0.05, when the significant level is 0.05) from those in NC, these DOPs are regarded as the discriminative DOPs.

We rank these discriminative DOPs according to their p values and then select the top 4 DOPs with the smallest p

TABLE III
DISCRIMINATIVE ORDINAL PATTERNS. L IS LEFT. R IS RIGHT. THE NODES LISTED IN TABLE ARE THE ABBREVIATIONS OF BRAIN REGIONS. THE FULL NAME OF BRAIN REGION CAN REFER TO THE ILLUSTRATION OF AAL ATLAS [38]. EMCI: EARLY MILD COGNITIVE IMPAIRMENT. LMCI: LATE MILD COGNITIVE IMPAIRMENT. NC: NORMAL CONTROLS. THE FULL NAMES OF BRAIN REGIONS ARE SHOWN IN TABLE IV

	Start node is left hippocampus (HIP.L)
Subject	Deep ordinal pattern sequences
NC	HIP.L→ HIP.R→ THA.R→ CAU.R→ INS.R→HES.R
EMCI	HIP.L→ HIP.R→ PHG.R→TPOsup.R→ TPOmid.R→TPOmid.L
LMCI	HIP.L→ HIP.R→ PHG.R→FFG.R→ SMA.L→STG.L
	Start node is right hippocampus (HIP.R)
Subject	Deep ordinal pattern sequences
NC	HIP.R→ HIP.L→ THA.L→ DCG.R→ THA.R→DCG.L
EMCI	HIP.R→ HIP.L→ PHG.L→TPOsup.L→ STG.L→PoCG.L
LMCI	HIP.R→ HIP.L→ PHG.L→TPOsup.L→ TPOmid.L→TPOmid.R
	Start node is left posterior cingulate gyrus (PCG.L)
Subject	Deep ordinal pattern sequences
NC	PCG.L→PCG.R→ PCUN.R→PoCG.L→ STG.L→THA.L
EMCI	PCG.L→PCG.R→PCUN.R→CAL.R→ STG.L→FFG.L
LMCI	PCG.L→PCG.R→ PCUN.R→THA.L→ STG.L→PoCG.R
	Start node is right posterior cingulate gyrus (PCG.R)
Subject	Deep ordinal pattern sequences
NC	PCG.R→PCG.L→ PCUN.L→ACG.R→ PUT.R→CAU.L
EMCI	PCG.R→PCG.L→PCUN.L→PreCG.R→ PCUN.R→THA.R
LMCI	PCG.R→PCG.L→ ACG.R→PUT.R→ ACG.L→STG.R

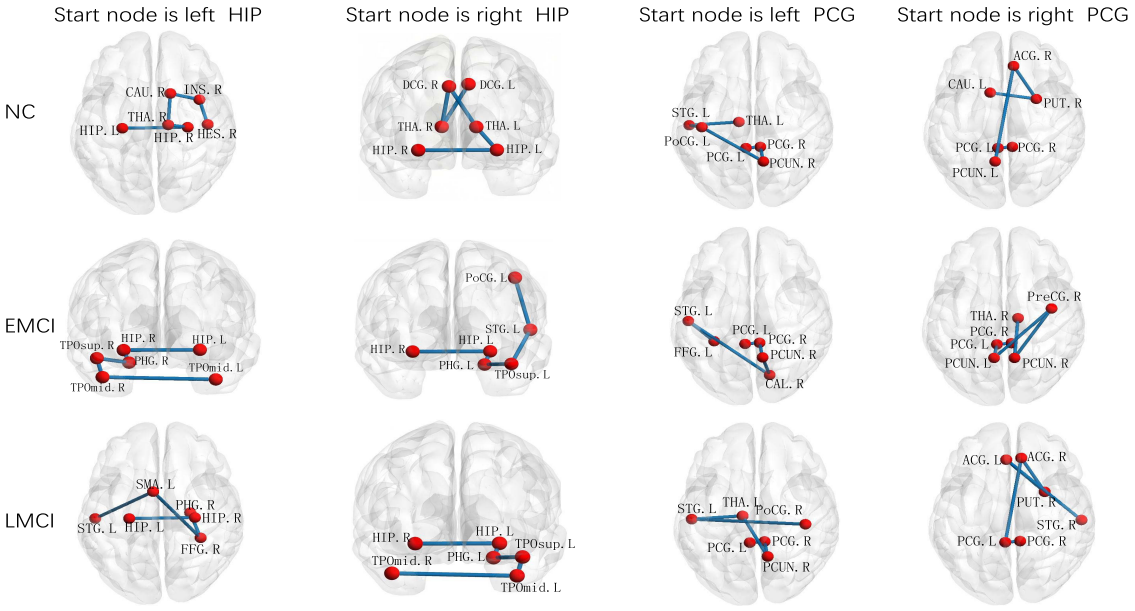

Fig. 5. Discriminative ordinal patterns. In order to visually show the ordinal pattern paths, we plot the brain regions and edges in different view directions. EMCI: early mild cognitive impairment. LMCI: late mild cognitive impairment. NC: normal controls. The significant level is 0.05. The full names of brain regions are shown in Table IV.

value. In order to expediently show the DOPs, we plot the DOPs of length 6 which have 6 brain regions in average brain connectivity network of EMCI, LMCI, and NC in Fig. 5 and describe their paths in Table III. The start node of DOP shown in Fig. 5 and Table III is left hippocampus (HIP), right HIP, left posterior cingulate gyrus (PCG) and right PCG, respectively.

From Fig. 5 and Table III, we can find that the DOPs in brain disease patients are different from those in NC. The first two brain regions in DOPs are same among EMCI, LMCI, and NC. In DOP whose start node is right HIP, the first four brain regions in EMCI are same to those in LMCI, seeing the second and third figures in the second column in Fig. 5 and Table III. From EMCI to LMCI, the DOP of right HIP changes slowly compared to the DOP of left HIP. In the DOP of PCG including left PCG and right PCG, we can find that the first three brain regions (i.e., PCG.L, PCG.R, PCUN.R in the DOPs of left PCG and PCG.R, PCG.L, PCUN.L in the DOP of right PCG) are same in NC and EMCI, seeing the third and fourth column in Fig. 5 and Table III. In the DOP of left PCG, the first three brain regions are same among EMCI, LMCI, and NC, as shown in the third column in Fig. 5 and Table III. These results indicate that our proposed method can be used to identify the discriminative ordinal pattern structures which are challenges for the existing methods in brain network analysis.

F. Important Brain Regions

When investigating the discriminative ordinal patterns, we find that some brain regions frequently appear in many depth-first-based ordinal patterns. For example, when the start node is left hippocampus, right hippocampus occurs in the depth-first-based ordinal patterns of NC, EMCI, and LMCI,

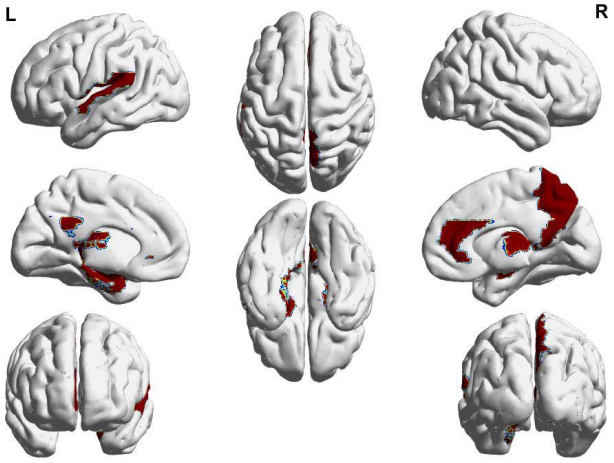

Fig. 6. The important ROIs (i.e., brain regions) selected by the proposed ordinal pattern kernel. The color represents the frequency of occurrence of the important brain regions (i.e., 10 ROIs) in the depth-first-based ordinal patterns of all brain regions.

as shown in Table III. This result indicates that the right hippocampus may play an important role in the depth-first based ordinal pattern of the left hippocampus. Here, we rank these ROIs (i.e., brain regions) according to their frequency of occurrence in the depth-first-based ordinal patterns of all brain regions and then select the top 10 ROIs using the highest frequency. The calculation steps include: (1) Find out all discriminative ordinal patterns; (2) Calculate the occurrence frequency of each ROI in the discriminative ordinal patterns; (3) Rank all brain regions according to their frequency of occurrence; (4) Select the top 10 ROIs.

The selected ROIs include left thalamus(THA.L), right thalamus (THA.R), right precuneus (PCUN.R), left

TABLE IV
FULL NAME AND ABBREVIATION OF BRAIN REGIONS

Index	ROI	Full name
1	AMYG.L	Left amygdala
2	AMYG.R	Right amygdala
3	HIP.L	Left hippocampus
4	HIP.R	Right hippocampus
5	MTG.R	Right middle temporal gyrus
6	SFGdor.R	Right superior frontal gyrus (dorsolateral)
7	PUT.L	Left lenticular nucleus (putamen)
8	PAL.L	Left lenticular nucleus (pallidum)
9	THA.L	Left thalamus
10	THA.R	Right thalamus
11	CAU.L	Left caudate nucleus
12	CAU.R	Right caudate nucleus
13	INS.R	Right insula
14	HES.R	Right heschl gyrus
15	PHG.R	Right parahippocampal gyrus
16	TPOsup.L	Left temporal pole (superior temporal gyrus)
17	TPOsup.R	Right temporal pole (superior temporal gyrus)
18	TPOmid.L	Left temporal pole (middle temporal gyrus)
19	TPOmid.R	Right temporal pole (middle temporal gyrus)
20	FFG.L	Left fusiform gyrus
21	FFG.R	Right fusiform gyrus
22	SMA.L	Left supplementary motor area
23	STG.L	Left superior temporal gyrus
24	STG.R	Right superior temporal gyrus
25	DCG.L	Left median cingulate and paracingulate gyri
26	DCG.R	Right median cingulate and paracingulate gyri
27	PHG.L	Left parahippocampal gyrus
28	PoCG.L	Left postcentral gyrus
29	PoCG.R	Right postcentral gyrus
30	PCG.L	Left posterior cingulate gyrus
31	PCG.R	Right posterior cingulate gyrus
32	PCUN.L	Left precuneus
33	PCUN.R	Right precuneus
34	ACG.L	Left anterior cingulate and paracingulate gyri
35	ACG.R	Right anterior cingulate and paracingulate gyri
36	PreCG.R	Right precental gyrus
37	CAL.R	Right calcarine fissure and surrounding cortex
38	PUT.R	Right Lenticular nucleus (putamen)

hippocampus (HIP.L), right fusiform gyrus (FFG.R), left parahippocampal gyrus (PHG.L), left superior temporal gyrus (STG.L), left posterior cingulate gyrus (PCG.L), right anterior cingulate-paracingulate gyri (ACG.R), and left amygdala (AMYG.L), as shown in Fig. 6. These ROIs are consistent with those revealed in the previous researches related to MCI.

IV. DISCUSSION

A. Significance of Results

Brain network has been widely applied to brain disease analysis. In brain network analysis, how to measure network similarity is a challenging task. In various methods, graph kernels provide an effective framework for addressing this problem. However, most of the existing graph kernels are based on unweighted graph and neglect the valuable weighted information of the edges in brain connectivity network. In this paper, we have built a graph kernel based on the weighted information of network and applied it to the classification tasks of brain diseases. The experimental results indicate that, compared with the competing methods, our proposed graph kernel can significantly improve the classification performances in MCI classification tasks.

In addition, some important ROIs frequently appear in ordinal patterns, when we investigate the discriminative ordinal patterns. These ROIs include thalamus, precuneus, hippocampus, fusiform gyrus, parahippocampal gyrus, superior temporal gyrus, posterior cingulate gyrus, anterior cingulate-paracingulate gyri, and amygdala, which have been reported in the previous studies related to MCI and AD. For example, in the left hippocampus and left thalamus, NC have significant differences in shape compared with AD. NC are significantly different from MCI in the left and right thalamus in shape [40]. According to the evidence of cortical beta-amyloid, precuneus can be regarded as a witness for excessive beta-amyloid gathering in AD and MCI pathology [41].

The related researches indicate that the connectivity strengths (i.e., edge weights) between identified brain regions in the MCI group are significantly different from those in the NC group. For example, Yao *et al.* [42] have found that AD and MCI subjects had an alteration in the functional connectivity of the amygdala compared with NC subjects. Bokde *et al.* [43] found that functional connectivities of the fusiform gyrus were different from those of NC.

These changes in the connectivity strengths suggest that some weighted information of edge in brain connectivity networks may be disrupted by brain disease. These disrupted weighted information may lead to the changed ordinal patterns between brain regions. Our experimental results show that the proposed ordinal pattern kernel can effectively classify EMCI and LMCI from NC, and can also classify EMCI from LMCI, by making full use of the weighted information of edge and the ordinal pattern relationships of edge weights in brain connectivity networks, which also provide empirical evidence for disorder network connectivity patterns in EMCI and LMCI.

B. Influence of Different Thresholds

The brain functional connectivity networks used in our experiment are constructed from fMRI data by calculating Pearson correlations between mean time series of pairs of ROIs. In the experiment, we directly apply the dense functional connectivity networks to the proposed ordinal pattern kernel and do not threshold them. The dense brain connectivity networks usually involve some unnecessary information. Hence, we need to threshold the connectivity networks. Different thresholds will form different network structures. The related research demonstrated that the networks with average connection density interval of [25%,75%] can obtain higher classification performance [44]. The average connectivity density of each threshold in the set $T = \{0.3, 0.35, 0.4, 0.45, 0.5\}$ is in the interval [30%,70%] [31]. Hence, we choose five different threshold values (i.e., $T = \{0.3, 0.35, 0.4, 0.45, 0.5\}$) to construct brain connectivity networks.

To verify the robustness of the proposed graph kernel on different network structures from the different thresholds, we examine the classification performance of the proposed ordinal pattern kernel in these network structures. We calculate the ordinal pattern kernels and the competing graph kernels on the brain connectivity networks built with different threshold values and use SVM technique for classification. The classification results are shown in Fig. 7. From Fig. 7, we can find that our proposed ordinal pattern kernel outperforms

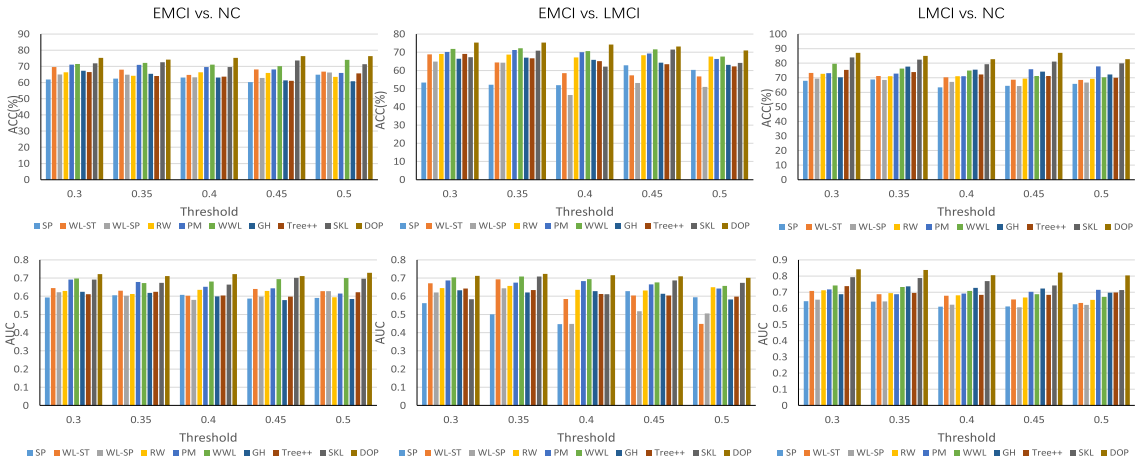

Fig. 7. Classification results in brain connectivity networks with different thresholds (*i.e.*, $T = \{0.3, 0.35, 0.4, 0.45, 0.5\}$). ACC: classification accuracy. AUC: area under receiver operating characteristic curve. EMCI: early mild cognitive impairment. LMCI: late mild cognitive impairment. NC: normal controls.

the state-of-the-art kernels in brain connectivity networks of different thresholds. For instance, our method yields accuracies of 75.27%, 74.19%, 75.27%, 76.34%, 76.34% in T = {0.3, 0.35, 0.4, 0.45, 0.5} in EMCI vs. NC classification, which are better than the second best results (i.e., 71.89%, 72.62%, 71.05%, 73.66%, 74.04%). We also perform the supplementary classification experiments in the brain networks with higher threshold (i.e., T = {0.55, 0.6, 0.65, 0.7, 0.75}), the classification results are shown in Fig. S1 in the Supplementary Materials. From Fig. S1, we can find that the proposed method still outperforms the competing methods.

C. Comparison With Other Methods

In this section, we discuss the superiority of the proposed method when compared to the other methods including convolutional neural network (CNN), random forest, and ordinal pattern descriptor. In the studies related to CNN, Yang *et al.* [45], [46] diagnosed MCI by combining fNIRS with CNN. They applied oxygenated hemoglobin change maps, temporal feature maps, and image biomarkers to CNN and trained the CNN model. Khatun *et al.* [47] used a single-channel EEG-based approach to detect MCI. They extracted 590 features from the event-related potential of the collected EEG signals and ranked these features by using random forest. Then, the top 25 features were applied to classification models. In MCI diagnosis tasks, Yang *et al.* [45], [46] and Khatun *et al.* [47] used novel non-invasive neuroimaging techniques and a neural network model to investigate MCI, and obtained higher classification results. These studies offered new approaches for MCI diagnosis. However, these methods ignored the spatial structure information of brain regions. Different from these methods, our proposed ordinal pattern kernel is built on the brain network and can make full use of the structure information of ordinal patterns between brain regions.

In the previous work [19], ordinal pattern as a new descriptor for brain connectivity network was proposed. Then, frequency ratio was used to calculate the frequent ordinal patterns, which frequently appeared in a set of brain connectivity networks. On the frequent ordinal patterns, discriminative ordinal pattern selection and feature selection were performed. At last, the discriminate ordinal pattern based learning framework was constructed for brain disease classification. In this framework, the frequent ordinal pattern mining and discriminative ordinal pattern selection are two key steps for discriminative ordinal patterns. According to [19], in given M subjects, the computational complexity of frequent ordinal pattern mining is $\mathcal{O}(M \times \frac{n_e!}{(n_e - level)!})$, n_e is the edges in brain connectivity and level is the max depth of depth-first search tree. In discriminative ordinal pattern selection, ratio score is calculated for each frequent ordinal pattern and these ordinal patterns are ranked in descending order. Here, we hypothesize that there are L frequent ordinal patterns. Hence, the computational complexity of the discriminate ordinal pattern based learning framework is $\mathcal{O}(L \times M \times \frac{n_e!}{(n_e - level)!})$. In [19] and our work, the used connectivity network is nearly full-connected. Hence, $n_e \approx \frac{N(N-1}{)}2$, N is the node number in the brain network. The computational complexity of the method in [19] is $\mathcal{O}(L \times M \times \frac{(N^2)!}{(N^2 - level)!})$. The computational complexity of our method is $\mathcal{O}(N^2)$, $\mathcal{O}(N^2) < \mathcal{O}(L \times M \times \frac{(N^2)!}{(N^2 - level)!})$, indicating that our proposed ordinal pattern kernel is better than the method in [19] in computational complexity. Meantime, in three tasks of EMCI vs. NC classification, EMCI vs. LMCI classification, and LCMI vs. NC classification, the classification accuracy obtained by the method in [19] is respectively 69.86%, 72.34%, and 81.35%, which are not better than our proposed ordinal pattern kernel.

V. Conclusion

Measuring the network similarity is a challenging task in brain network analysis. In this paper, we propose an ordinal pattern kernel to measure the similarity between paired brain networks. Different from the existing graph kernels built on unweighted graph, our proposed graph kernel can measure network similarity by making full use of the weighted information of edge and can also capture ordinal pattern relationship of edge weights in brain connectivity network. We have further constructed a learning framework based on the propossed ordinal pattern kernel for MCI classification using resting state

fMRI data, with the classification results demonstrating the effectiveness of the proposed kernel.

REFERENCES

[1] R. C. Petersen *et al.*, "Current concepts in mild cognitive impairment," *Arch. Neurol.*, vol. 58, no. 12, pp. 1985–1992, 2001.

[2] U. Gangishetti *et al.*, "Validating non-amyloid, non-tau CSF biomarkers for Alzheimer's disease in the pre-symptomatic, MCI, and dementia stages: A multi-center study," *Alzheimers Dementia*, vol. 13, no. 7, pp. P1513–P1514, 2017.

[3] O. Sporns, "Contributions and challenges for network models in cognitive neuroscience," *Nature Neurosci.*, vol. 17, no. 5, pp. 652–660, 2014.

[4] N. K. Logothetis, J. Pauls, M. Augath, T. Trinath, and A. Oeltermann, "Neurophysiological investigation of the basis of the fMRI signal," *Nature*, vol. 412, no. 6843, pp. 150–157, 2001.

[5] P. Amodio *et al.*, "Prevalence and prognostic value of quantified electroencephalogram (EEG) alterations in cirrhotic patients," *J. Hepatol.*, vol. 35, no. 1, pp. 37–45, Jul. 2001.

[6] K.-S. Hong and M. A. Yaqub, "Application of functional near-infrared spectroscopy in the healthcare industry: A review," *J. Innov. Opt. Health Sci.*, vol. 12, no. 6, 2019, Art. no. 1930012.

[7] N. Naseer and K.-S. Hong, "fNIRS-based brain-computer interfaces: A review," *Frontiers Hum. Neurosci.*, vol. 9, pp. 1–15, Jan. 2015.

[8] Z. Xia, T. Zhou, S. Mamoon, and J. Lu, "Recognition of dementia biomarkers with deep finer-DBN," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 29, pp. 1926–1935, 2021.

[9] Z. Wang, M. Zhang, Y. Han, H. Song, R. Guo, and K. Li, "Differentially disrupted functional connectivity of the subregions of the amygdala in Alzheimer's disease," *J. X-Ray Sci. Technol.*, vol. 24, no. 2, pp. 329–342, Mar. 2016.

[10] H. Wang *et al.*, "The reorganization of resting-state brain networks associated with motor imagery training in chronic stroke patients," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 27, no. 10, pp. 2237–2245, Oct. 2019.

[11] S.- Z. Si, X. Liu, J.-F. Wang, B. Wang, and H. Zhao, "Brain networks modeling for studying the mechanism underlying the development of Alzheimer's disease," *Neural Regener. Res.*, vol. 14, no. 10, pp. 1805–1813, 2019.

[12] A. H. Ghaderi *et al.*, "Functional brain connectivity differences between different ADHD presentations: Impaired functional segregation in ADHD-combined presentation but not in ADHD-inattentive presentation," *Basic Clin. Neurosci. J.*, vol. 8, no. 4, pp. 267–278, Jul. 2017.

[13] Q. Yu, J. Sui, K. A. Kiehl, G. Pearlson, and V. D. Calhoun, "State-related functional integration and functional segregation brain networks in schizophrenia," *Schizophrenia Res.*, vol. 150, nos. 2–3, pp. 450–458, Nov. 2013.

[14] C. Fee, M. Banasr, and E. Sibille, "Somatostatin-positive gamma-aminobutyric acid interneuron deficits in depression: Cortical microcircuit and therapeutic perspectives," *Biol. Psychiatry*, vol. 82, no. 8, pp. 549–559, Oct. 2017.

[15] M. Rubinov and O. Sporns, "Complex network measures of brain connectivity: Uses and interpretations," *NeuroImage*, vol. 52, no. 3, pp. 1059–1069, 2010.

[16] K. Ma, J. Yu, W. Shao, X. Xu, Z. Zhang, and D. Zhang, "Functional overlaps exist in neurological and psychiatric disorders: A proof from brain network analysis," *Neuroscience*, vol. 425, pp. 39–48, Jan. 2020.

[17] Y.-H. Liu, C.-C. Lin, W.-H. Lin, and F. Chang, "Accelerating feature-vector matching using multiple-tree and sub-vector methods," *Pattern Recognit.*, vol. 40, no. 9, pp. 2392–2399, Sep. 2007.

[18] M. Montembeault *et al.*, "Differential language network functional connectivity alterations in Alzheimer's disease and the semantic variant of primary progressive aphasia," *Cortex*, vol. 117, pp. 284–298, Aug. 2019.

[19] D. Zhang, J. Huang, B. Jie, J. Du, L. Tu, and M. Liu, "Ordinal pattern: A new descriptor for brain connectivity networks," *IEEE Trans. Med. Imag.*, vol. 37, no. 7, pp. 1711–1722, Jul. 2018.

[20] C. Yan and Y. Zang, "DPARSF: A MATLAB toolbox for 'pipeline' data analysis of resting-state fMRI," *Frontiers Syst. Neurosci.*, vol. 13, no. 4, p. 13, 2010.

[21] Z. Zhuo, X. Mo, X. Ma, Y. Han, and H. Li, "Identifying aMCI with functional connectivity network characteristics based on subtle AAL atlas," *Brain Res.*, vol. 1696, pp. 81–90, Oct. 2018.

[22] N. M. Kriege and P. Mutzel, "Subgraph matching kernels for attributed graphs," in *Proc. Int. Conf. Mach. Learn.*, 2012, pp. 291–298.

[23] T. Gartner, P. A. Flach, and S. Wrobel, "On graph kernels: Hardness results and efficient alternatives," in *Learning Theory and Kernel Machines*. Berlin, Germany: Springer, 2003, pp. 129–143.

[24] N. Shervashidze, P. Schweitzer, E. J. van Leeuwen, K. Mehlhorn, and K. M. Borgwardt, "Weisfeiler-Lehman graph kernels," *J. Mach. Learn. Res.*, vol. 12, no. 3, pp. 2539–2561, Sep. 2011.

[25] K. M. Borgwardt and H. Kriegel, "Shortest-path kernels on graphs," in *Proc. 5th IEEE Int. Conf. Data Mining (ICDM)*, Nov. 2005, pp. 74–81.

[26] S. V. N. Vishwanathan, N. N. Schraudolph, I. R. Kondor, and K. M. Borgwardt, "Graph kernels," *J. Mach. Learn. Res.*, vol. 11, no. 2, pp. 1201–1242, Mar. 2010.

[27] G. Nikolentzos, P. Meladianos, and M. Vazirgiannis, "Matching node embeddings for graph similarity," in *Proc. 31st AAAI Conf. Artif. Intell.*, 2017, pp. 2429–2435.

[28] M. Togninalli, E. Ghisu, F. Llinares-Lpez, B. Rieck, and K. Borgwardt, "Wasserstein Weisfeiler-Lehman graph kernels," in *Proc. Adv. Neural Inf. Process. Syst.*, 2019, pp. 6439–6449.

[29] A. Feragen, N. Kasenburg, J. Petersen, M. De Bruijne, and K. M. Borgwardt, "Scalable kernels for graphs with continuous attributes," in *Proc. Adv. Neural Inf. Process. Syst.*, 2013, pp. 216–224.

[30] W. Ye, Z. Wang, R. Redberg, and A. Singh, "Tree++: Truncated tree based graph kernels," *IEEE Trans. Knowl. Data Eng.*, vol. 33, no. 4, pp. 1778–1789, Apr. 2021.

[31] B. Jie, M. Liu, D. Zhang, and D. Shen, "Sub-network kernels for measuring similarity of brain connectivity networks in disease diagnosis," *IEEE Trans. Image Process.*, vol. 27, no. 5, pp. 2340–2353, May 2018.

[32] C.-C. Chang and C.-J. Lin, "LIBSVM: A library for support vector machines," *ACM Trans. Intell. Syst. Technol.*, vol. 2, no. 3, pp. 1–27, Apr. 2011.

[33] A. Gretton, K. M. Borgwardt, M. J. Rasch, B. Schölkopf, and A. Smola, "A kernel two-sample test," *J. Mach. Learn. Res.*, vol. 13, no. 1, pp. 723–773, Jan. 2012.

[34] K. Ma, W. Shao, Q. Zhu, and D. Zhang, "Kernel based statistic: Identifying topological differences in brain networks," *Intell. Med.*, vol. 2, no. 1, pp. 30–40, Feb. 2022. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S2667102621000164

[35] M. Xia, J. Wang, and Y. He, "BrainNet Viewer: A network visualization tool for human brain connectomics," *PLoS ONE*, vol. 8, no. 7, Jul. 2013, Art. no. e68910.

[36] H. M. Nguyen and G. H. Glover, "A modified generalized series approach: Application to sparsely sampled fMRI," *IEEE Trans. Biomed. Eng.*, vol. 60, no. 10, pp. 2867–2877, Oct. 2013.

[37] J. de Jong, S. O. Dumoulin, B. P. Klein, A. Keizer, and C. Dijkerman, "Hand position modulates visually-driven fMRI responses in premotor cortex," in *Proc. Soc. Neurosci.*, 2016, p. 1.

[38] R. T. Edmund, H. Chu-Chung, L. Ching-Po, F. Jiling, and M. Joliot, "Automated anatomical labelling atlas 3," *NeuroImage*, vol. 206, Feb. 2020, Art. no. 116189.

[39] 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.

[40] A. Bokde, D. Brennan, M. Hamill, J. Mcnulty, P. Mullins, and D. Farrell, "Shape differences of the hippocampus and thalamus in risk-stratified MCI groups," *Alzheimers Dementia*, vol. 10, no. 4, p. P901, 2014.

[41] G. Aghakhanyan *et al.*, "The precuneus—A witness for excessive Aβ gathering in Alzheimer's disease pathology," *Neurodegenerative Diseases*, vol. 18, nos. 5–6, pp. 302–309, 2018.

[42] H. Yao *et al.*, "Decreased functional connectivity of the amygdala in Alzheimer's disease revealed by resting-state fMRI," *Eur. J. Radiol.*, vol. 82, no. 9, pp. 1531–1538, Sep. 2013.

[43] A. L. W. Bokde *et al.*, "Functional connectivity of the fusiform gyrus during a face-matching task in subjects with mild cognitive impairment," *Brain*, vol. 129, no. 5, pp. 1113–1124, May 2006.

[44] M. Zanin *et al.*, "Optimizing functional network representation of multivariate time series," *Sci. Rep.*, vol. 2, no. 1, pp. 1–6, Dec. 2012.

[45] D. Yang, K.-S. Hong, S.-H. Yoo, and C.-S. Kim, "Evaluation of neural degeneration biomarkers in the prefrontal cortex for early identification of patients with mild cognitive impairment: An fNIRS study," *Frontiers Hum. Neurosci.*, vol. 13, no. 6, p. 317, Sep. 2019.

[46] D. Yang *et al.*, "Detection of mild cognitive impairment using convolutional neural network: Temporal-feature maps of functional near-infrared spectroscopy," *Frontiers Aging Neurosci.*, vol. 12, p. 141, May 2020.

[47] S. Khatun, B. I. Morshed, and G. M. Bidelman, "A single-channel EEG-based approach to detect mild cognitive impairment via speech-evoked brain responses," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 27, no. 5, pp. 1063–1070, May 2019.