Constructing Multi-view High-order Functional Connectivity Networks for Diagnosis of Autism Spectrum Disorder

Feng Zhao, Xiangfei Zhang, Kim-Han Thung, Ning Mao, Seong-Whan Lee, Dinggang Shen, *Fellow, IEEE*

***Abstract*— Brain functional connectivity network (FCN) based on resting-state functional magnetic resonance imaging (rs-fMRI) has been widely used to identify neuropsychiatric disorders such as autism spectrum disorder (ASD). Most existing FCN-based methods only estimate the correlation between brain regions of interest (ROIs), without exploring more informative higher-level interactions among multiple ROIs which could be beneficial to disease diagnosis. To fully explore the discriminative information provided by different brain networks, a cluster-based multi-view high-order FCN (Ho-FCN) framework is proposed in this paper. Specifically, we first group the functional connectivity (FC) time series into different clusters and compute the multi-order central moment series for the FC time series in each cluster. Then we utilize the correlation of central moment series between different clusters to reveal the high-order FC relationships among multiple ROIs. In addition, to address the phase mismatch issue in conventional FCNs, we also adopt the central moments of the correlation time series as the temporal-invariance features to capture the dynamic characteristics of low-order dynamic FCN (Lo-D-FCN). Experimental results on the ABIDE dataset validate that: 1) the proposed multi-view Ho-FCNs is able to explore rich discriminative information for ASD diagnosis; 2) the phase mismatch issue can be well circumvented by using central moments; and 3) the combination of different types of FCNs can significantly improve the diagnostic accuracy of ASD (86.2%).**

***Index Terms*— Autism spectrum disorder, high-order functional connectivity, resting-state fMRI, central moment**

I. Introduction

Autism Spectrum Disorder (ASD) is a childhood nervous system developmental disorder which is characterized by clinical manifestations including stereotypic behavior, narrow interests, and communication disorders [1]. A recent report from the Centers for Disease Control and Prevention reveals that 1 out of 54 American children aged 8 years suffers from ASD with varying degrees, which causes substantial socio-economic burden to the society [2], [3]. Thus, early identification and treatment of ASD is of great clinical value [4]–[6].

Functional connectivity network (FCN) [7]–[13] calculated from resting-state functional magnetic resonance imaging (rs-fMRI) data has been widely used for computer-aided ASD diagnosis due to its effectiveness in finding the brain functional tissues and biomarkers of neuropsychiatric diseases. Generally, an FCN can be regarded as a graph, where a vertex corresponds to a brain region of interest (ROI), while an edge connecting two vertices represents the functional connectivity (e.g., correlation) between the two corresponding ROIs. Different types of FCNs have been proposed to explore functional connectivity (FC) among ROIs [14]–[19]. Conventional FCN (C-FCN) utilizes the whole time series of rs-fMRI data to measure FC between ROIs [7], [11], [19] based on the assumption that FC is stationary during the entire rs-fMRI scan. However, brain connectivity patterns are always changing, and increasing evidences have shown that the dynamic variations of FC are associated with the neurological diseases [14]–[16].

In order to explore the brain dynamic FC information, many dynamic FCN (D-FCN) models [18]–[22] have been proposed, among which the sliding-window-based D-FCN models are the most popular. In these models, rs-fMRI time series are first partitioned into multiple overlapping time segments by using a fixed length sliding window. Then a subnetwork is constructed for each time segment by calculating the short-term correlation between the time series segment of each pair of ROIs. The construction of the subnetwork is similar to that of C-FCN, where each vertex of the subnetwork corresponds to a specific ROI, while the edge connecting two vertices measures the short-term FC between the corresponding two ROIs. Thus, D-FCN is a sequential collection of subnetworks which contain dynamic variations of FC throughout the scan time and provide rich discriminative information for disease diagnosis. Despite the effectiveness in capturing dynamic FC information, D-FCN still suffers from the following limitations.

***First*, D-FCN cannot reveal complex and high-order FC information across multiple ROIs.** Similar to C-FCN, D-FCN only captures the dynamic variations of low-order (or first order) correlation between time series of each pair of ROIs [19]. For notation uniformity, we refer this D-FCN as low-order D-FCN (Lo-D-FCN) throughout this paper. Contrarily, FCN which can reflect more complex FC patterns among multiple ROIs is referred to as high-order FC network (Ho-FCN) in this paper.

In order to explore high-order FC patterns among brain regions, many Ho-FCNs [18]–[20] have been proposed. For example, in our previous works [23], [24], we proposed multilevel Ho-FCNs based on the "correlation of correlations" strategy. Specifically, we first computed the Pearson's correlation coefficients (PCCs) between the raw rs-fMRI time series of each two ROIs for obtaining the first-level FCN. Then, the second-, third-, and other higher-level FCNs were subsequently derived by computing the correlation between each pair of respective FC profiles from the lower-level FCN. Furthermore, some scholars use hypergraph theory to construct Ho-FCNs [25]–[28]. For instance, Jie *et al*. constructed connectivity hyper-networks from the rs-fMRI time series by using sparse representation to estimate the functional connection among multiple ROIs [25]; Guo *et al*. proposed two improved hyper-networks: the elastic net and the group lasso [26]; Wang *et al*. proposed a framework based on hyper-network to use structural voxel information and network connectivity information as intermediate traits that bridge genetic risk factors and

F. Zhao was supported in part by the National Natural Science Foundation of China (61773244, 61976125, 61873177, 61976124, 82001775), Yantai Key Research and Development Program of China (2017ZH065, 2019XDHZ081), Shandong Provincial Key Research and Development Program of China (2019GGX101069) and Doctoral Scientific Research Foundation of Shandong Technology and Business University (BS202016) *(Corresponding author: Dinggang Shen)*.

F. Zhao and X. Zhang are with the School of Computer Science and Technology, Shandong Technology and Business University, Yantai, China.

K. Thung is with the Department of Radiology and BRIC, University of North Carolina at Chapel Hill, Chapel Hill, NC, USA.

N. Mao is with the Department of Radiology, Yantai Yuhuangding Hospital, Yantai, China.

S. Lee is also with the Department of Brain and Cognitive Engineering, Korea University, Seoul, South Korea. He is also with the Department of Artificial Intelligence, Korea University, Seoul, south Korea.

D. Shen is with the School of Biomedical Engineering, ShanghaiTech University, Shanghai, China. He is also with the Shanghai United Imaging Intelligence Co., Ltd., Shanghai, China and with the Department of Artificial Intelligence, Korea University, Seoul, South Korea (e-mail: dinggang.shen@gmail.com).

disease status [27]; Xiao *et al*. generated hyperedges from fMRI time series based on sparse representation, and then used hypergraph learning to adaptively learn hyperedge weights. Finally, it defined a hypergraph similarity matrix to represent the FCN [28]. Although these methods can provide high-order FC information across multiple ROIs, the dynamic or temporal information of FC is often ignored.

To address the limitations mentioned above, we propose a multi-view Ho-FCNs framework, which considers both temporal and high-order properties of FC to fully explore the fundamental properties of brain network and the underpinnings of disordered brain connectivity changes. Specifically, we first employ sliding window approach to construct Lo-D-FCN and regard it as a set of FC time series, each of which is associated with a specific pair ROIs and characterizes the variation of their FC across the whole scanning time (see Fig. 4). Then, the FC time series are grouped into several clusters and the d-th order central moment series (e.g., mean as d = 1, variance as d = 2, the third-order central moment as d = 3, and so on) within the cluster is computed. Subsequently, the correlation between the d-th order central moment series are calculated to represent the high-order FC between any two clusters.

The two key components of our proposed method are (1) *the clustering strategy*, which significantly reduces the dimension of Ho-FCN. If we directly using the FC series of Lo-D-FCN to calculate high-order FC as in [18]–[20], we will end up with an undesirably large-scale Ho-FCN. In comparison, by grouping the FC series within Lo-D-FCN into several clusters, and computing the correlation of central moment series between clusters as high-order FC, we obtain a relatively small size Ho-FCN that takes into account the dynamic property of FC. We will discuss this in details in Section III-C; (2) *the multiple order central moments*, which could extract different statistical information to accurately measure the similarity between the FC time series from two clusters. For example, Fig. 1 shows three different groups of sequences. The sequences in Fig. 1(a) and Fig. 1(b) have the same mean (red dotted line) and variance (blue dotted line), while in Fig. 1(a) and Fig. 1(c) have the same mean but different variances. It can be seen that the sequences in Fig. 1(a) are similar to the sequences in Fig. 1(b), but different to the sequences in Fig. 1(c). Thus, when more orders of the central moment series between two groups are similar, the more likely the sequences in the two groups are similar. Based on the above observation, we use multiple order central moments to construct multi-view Ho-FCNs to improve the discriminative power of our proposed model. In summary, the proposed multi-view Ho-FCN has the following characteristics:

(1) Our Ho-FCN is derived from the correlation of central moment series between different clusters of FC time series. This is different from our previous works [23], [24], which derived Ho-FCN by computing correlations of FC time series. Furthermore, by clustering similar FC time series together, and aggregating the FC time series within a cluster to a central moment series, we effectively reduce the dimensionality of the Ho-FCN.

(2) Our multi-view Ho-FCN is derived from multiple order central moment series. We adopt different central moments to capture more subtle disruptions of brain FC induced by neuropsychiatric disorders. For example, the first-order, second-order, and third-order central moment series respectively reflect the mean, variance, and skewness of all the FC time series within a cluster. By computing the correlation of central moment series of different orders between different clusters of FC time series, we obtain different views of high-order FC interactions. We hypothesize that integrating multiple Ho-FCNs can comprehensively capture more informative high-order FC patterns among brain regions, which could be useful in discovering high-order FC biomarkers for disease diagnosis.

(3) Our high-order FC contains interactions among more than four ROIs. The central moment method is used to transform the FC time series in each cluster into a specific order central moment series, which makes a central moment series possibly integrate the low-order FC of multiple ROIs. As shown in Fig. 2, assume there are two clusters. We calculate the central moment series of each cluster to capture the statistical information of all ROI pairs in this cluster. Thus, the correlation between two central moment series, as the high-order FC between two clusters, can better exploit complex interaction relationships among brain regions. In other words, a vertex of Ho-FCN represents all the FC time series from one cluster involving more than four ROIs, and thus its edges can characterize the complex FC relationship across multiple ROIs.

It should be noted that the proposed Ho-FCN is significantly different from the method proposed in [20]. In [20], the authors just use the correlation between the mean correlation time series of different clusters, while our method can reflect high-level interaction among ROIs from multiple perspectives which can better explore high-order connectional biomarkers that influenced by ASD.

***Second*, Lo-D-FCN is sensitive to the chronological order of its subnetworks, which makes it difficult to make consistent and meaningful comparisons among different subjects.** Specifically, as the mental activity is unconstrained in resting state, the computed FC subnetworks have no temporal correspondence across different subjects, and across different scan sessions of the same subject. Moreover, switching positions of any two FC subnetworks will significantly alter the Lo-D-FCN. Due to these issues, Lo-D-FCN has limited use in comparative studies.

We address the cross-subject phase mismatch issue mentioned above by extracting central-moment features [19] from the Lo-D-FCN. This is because central-moment features are temporal-invariant and not sensitive to the chronological order of the time series. Specifically, we extract the central-moment features (e.g., mean, variance) for each FC time series in Lo-D-FCN for subsequent classification model training.

Finally, we integrate multiple FCNs, including C-FCN, Lo-D-FCN, and Multi-view Ho-FCNs, to yield the final classification results. We hypothesize that the brain network in individuals with ASD may be altered by miswiring due to abnormal development. Such miswiring may affect not only static and dynamic FCs, but also low- and high-order FCs. Therefore, it is helpful to extract and integrate features from a variety of networks to improve the accuracy of ASD diagnosis.

II. Materials and Data Preprocessing

In this study, we conducted experiments on the Autism Brain Imaging Data Exchange (ABIDE) database [29] that consists of 17 international imaging sites. In order to prevent the effects of the heterogeneity of multi-site data on the results during the comparison, we chose 45 ASD patients (36 males and 9 females) and 47 NC subjects (36 males and 11 females) aged between 7 and 15 years old, scanned at New York University (NYU) Langone Medical Center, that is, the dataset with the largest sample size. These subjects were sociodemographic-matched, where there are no significant differences ($p > 0.05$) in gender, age, and FIQ between the ASD and NC groups. The demographic information is summarized in Table I. In this work, only rs-fMRI data were utilized for diagnostic study. More details on the data collection, exclusion criteria, and scan parameters can be obtained from the ABIDE website (http://fcon_1000.projects.nitrc.org/indi/abide/abide_i.html).

All subjects underwent a 6-minute scan using a 3T Siemens Allegra scanner at NYU Langone Medical Center. During rs-fMRI scans, most subjects were asked to relax with their eyes open and gaze at a white fixed cross in the middle of a black background projected onto

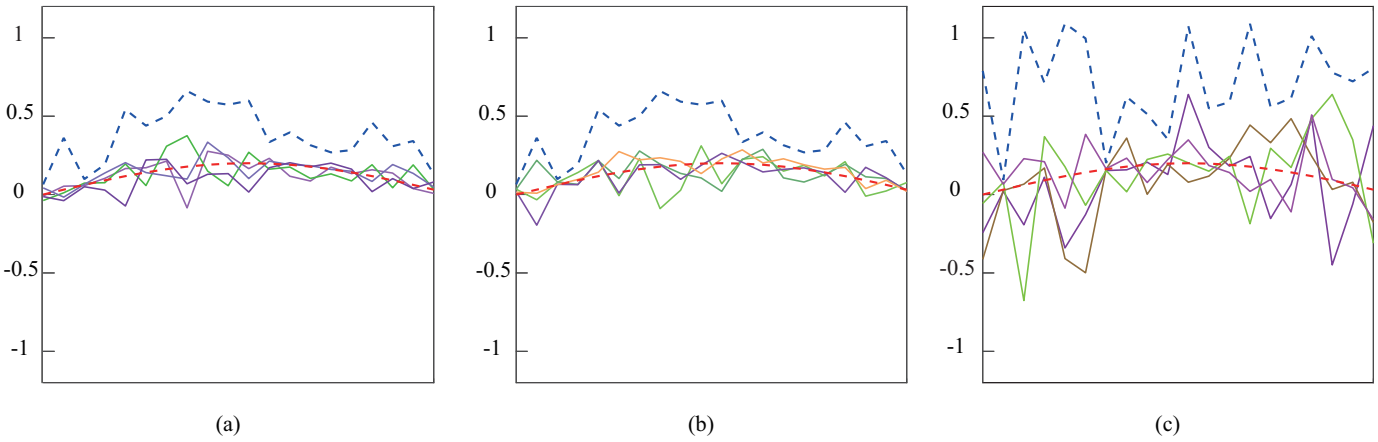


Fig. 1. Example for central moment features of sequences. The solid lines represent the sequences, and the dotted lines represent the central moment series (red dotted lines denote mean, blue dotted lines denote variance), where two groups of different sequences with same mean and variance are shown in (a) and (b); a group of sequences shown in (c) has the same mean but different variance compared with that in (a).

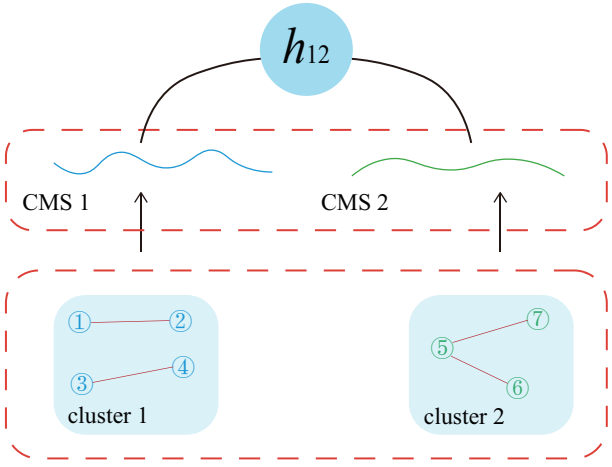


Fig. 2. Schematic diagram of an example of our proposed high-order FC. Cluster 1 contains FC time series for ROI pairs that involve ROI 1 to 4, while cluster 2 contains FC time series for ROI pairs that involve ROI 5 to 7. CMS 1 and CMS 2 denote the respective central moment series for cluster 1 and cluster 2. The high-order FC is then denoted as $\boldsymbol{h_{12}}$, i.e., the correlation between CMS 1 and CMS 2.

a screen. During rs-fMRI scans, eye movement was monitored by an eye tracker. When acquiring images, the following parameters were used: TR/TE = 2000/15ms, flip angle = 90°, 33 slices per volume, 180 volumes per scan, voxel thickness of 4.0 mm. The mean frame-wise displacement was computed to describe head motion for each individual. Individuals with mean frame-wise displacement larger than 1 mm were excluded from analysis to reduce the negative effect of head motion [30], [31].

Statistical Parametric Mapping (SPM8) software was adopted to preprocess the acquired rs-fMRI data. The first 10 rs-fMRI images of each subject were discarded for magnetization equilibrium. Then, the remaining images were spatially normalized into the Montreal Neurological Institute (MNI) template space with resolution of $3 \times 3 \times 3$ mm^3. Other corrections were further conducted, including the regression of nuisance signals (ventricle, white matter, global signals, and head motion with Friston 24-parameter model), signal detrending, and band-pass filtering (0.01–0.08Hz) [32]–[36]. Each brain image was then parcellated into 116 regions according to the Automated Anatomical Labeling (AAL) atlas [37]. Finally, the average of rs-fMRI time series within each ROI was calculated, which was treated as the data matrix $\boldsymbol{X} \in \mathbb{R}^{170\times116}$ for subsequent processing, where 170 denotes the total number of temporal image volumes and 116 denotes the total number of all brain ROIs.

TABLE I
DEMOGRAPHIC INFORMATION OF THE SUBJECTS

	ASD	NC	*p*-values
Gender (M/F)	36/9	36/11	0.2135[a]
Age (mean±SD)	11.1±2.3	11.0±2.3	0.773[b]
FIQ (mean±SD)	106.8±17.4	113.3±4.1	0.0510[b]
ADI-R (mean±SD)	32.2±4.3[c]	–	–
ADOS (mean±SD)	13.7±5.0	–	–

ASD, Autism Spectrum Disorders; NC, normal control; M, male; F, female; FIQ, Full Intelligence Quotient; ADI-R, Autism Diagnostic Interview-Revised; ADO, Autism Diagnostic Observation Schedule.
[a]The *p*-value was obtained by χ^2-test; [b]The p-value was obtained by two-sample two-tailed *t*-test; [c]Two patients do not have the ADI-R score.

III. METHOD

Our proposed framework consists of four steps, including (1) constructing various FCNs, i.e., C-FCN, Lo-D-FCN and multi-view Ho-FCNs; (2) extracting and selecting discriminative features from each constructed FCN; (3) constructing a support vector machine (SVM) model for each type of FCN based on the selected features; (4) fusing the output of different SVM classifiers to predict the target class label (ASD or NC) for each subject.

In this paper, we use lower case letter (e.g., x) to denote scalar, lower case bold letter (e.g., $\boldsymbol{x}$) to denote vector, and upper case bold letter (e.g., $\boldsymbol{G}$) to denote matrix or FCN. All FCNs are stored in matrices, where each column denotes a vertex of the corresponding FCN, and each matrix element denotes the edge between two vertices associated with the two columns of the matrix.

A. C-FCN Construction

Let $\boldsymbol{x}_i^l = (x_{i1}^l, x_{i2}^l, \cdots, x_{iM}^l)^T$ denote the rs-fMRI time series associated with the i-th ROI of the l-th subject, where M is the total number of image volumes during the whole scanning period.

In C-FCN, each vertex corresponds to an ROI, and an edge represents the FC between two ROIs in terms of Pearson's correlation coefficient (PCC). Specifically, the PCC between $\boldsymbol{x}_i^l$ and $\boldsymbol{x}_j^l$ is given as follows:

$$r_{ij}^l = corr(\boldsymbol{x}_i^l, \boldsymbol{x}_j^l) \quad (1)$$

Then, a C-FCN for the l-th subject can be estimated as $\boldsymbol{G}_C^l = \left(\left\{\boldsymbol{x}_i^l\right\}, \left\{r_{ij}^l\right\}\right)$, where $\left\{\boldsymbol{x}_i^l\right\}$ denotes the set of vertices, and $\left\{r_{ij}^l\right\}$ denotes the set of weights (i.e., FC) connecting each pair of vertices. An overall illustration of the C-FCN construction is shown in Fig. 3.

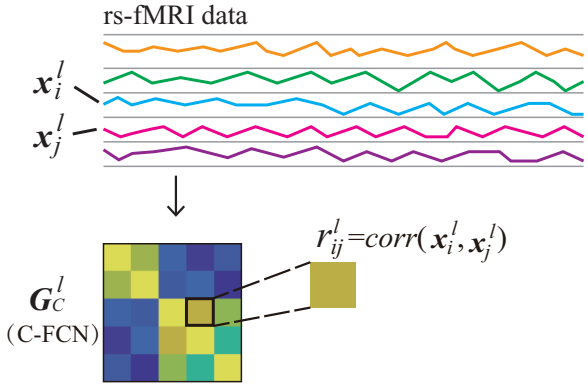


Fig. 3. The illustration of the C-FCN construction.

B. Lo-D-FCN Construction

As shown in Fig. 4, to construct Lo-D-FCN, the sliding window strategy is employed to partition the entire rs-fMRI time series into $K = \lfloor (M-W)/s \rfloor + 1$ overlapped segments, where W is the length of sliding window and s is the step size of sliding window. Let $\boldsymbol{x}_i^l(k)$ and $\boldsymbol{x}_j^l(k)$ denote the sub-series of the i-th and j-th ROIs in the k-th window, respectively. Then similar to Eq. (1), the FC $\rho_{ij}^l(k)$ between the two ROIs in the k-th window is calculated as follows:

$$\rho_{ij}^l(k) = corr(\boldsymbol{x}_i^l(k), \boldsymbol{x}_j^l(k)) \tag{2}$$

where a larger value of $\left|\rho_{ij}^l(k)\right|$ indicates a stronger correlation between the i-th and j-th ROIs in the k-th window. The k-th subnetwork of Lo-D-FCN is then estimated as $\boldsymbol{G}_{Lo}^l(k) = \left(\left\{\boldsymbol{x}_i^l(k)\right\}, \left\{\rho_{ij}^l(k)\right\}\right)$, where $\left\{\boldsymbol{x}_i^l(k)\right\}$ denotes the set of vertices and $\left\{\rho_{ij}^l(k)\right\}$ denotes the set of edge weights connecting each pair of vertices. After this procedure, K subnetworks have been constructed for each subject. Then, those K subnetworks are arranged in chronological order to form a Lo-D-FCN, i.e., $\boldsymbol{G}_{Lo}^l = \left[\boldsymbol{G}_{Lo}^l(1), \boldsymbol{G}_{Lo}^l(2), \ldots, \boldsymbol{G}_{Lo}^l(K)\right]$, as shown in Fig. 4.

In $\boldsymbol{G}_{Lo}^l$, $\rho_{ij}^l(k)(1 \leqslant k \leqslant K)$ reflects the FC between the i-th and j-th ROIs during a relatively short time period, thus the FC time series $\boldsymbol{\rho}_{ij}^l = \left[\rho_{ij}^l(1), \rho_{ij}^l(2), \ldots, \rho_{ij}^l(K)\right]$ can reflect the temporal changes of correlations (i.e., FCs) between the i-th and j-th ROIs during the whole scan period (a brief illustration is shown in Fig. 4).

C. Clustering for constructing Ho-FCNs

Recall that in low-order FCN (i.e., C-FCN or Lo-D-FCN), the vertex corresponds to ROI while the edge represents the low-order FC between a pair of ROIs. To capture more complex and higher-level interactions among multiple ROIs, we can adopt "correlation of correlations" strategy to compute PCC between the FC time series of two pairs of vertices as high-order FC, which is denoted as:

$$h_{ij,pq}^l = corr(\boldsymbol{\rho}_{ij}^l, \boldsymbol{\rho}_{pq}^l) \tag{3}$$

where $\boldsymbol{\rho}_{ij}^l$ is the correlation time series between the i-th and j-th ROIs, and $\boldsymbol{\rho}_{pq}^l$ represents the correlation time series between the p-th and q-th ROIs. Thus, the high-order FC, denoted as $h_{ij,pq}^l$, can represent the

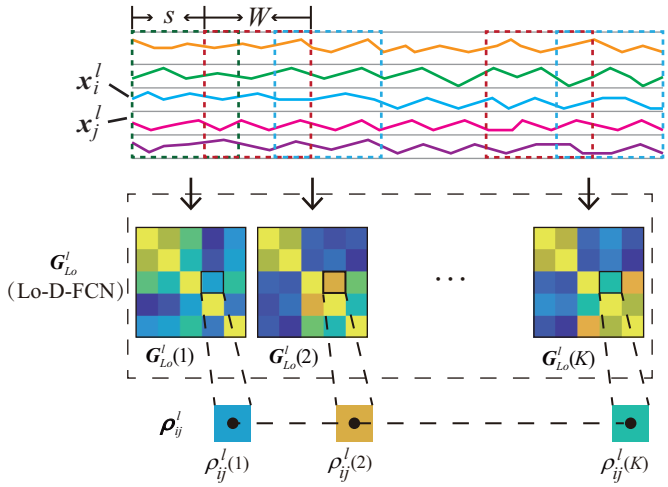


Fig. 4. The illustration of the Lo-D-FCN construction.

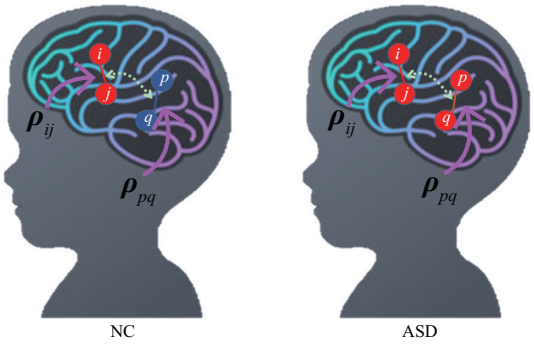


Fig. 5. A simple example for illustrating the interaction between brain ROI pairs, where $\boldsymbol{\rho}_{ij}$ denotes the correlation time series between i-th ROI and j-th ROI, and $\boldsymbol{\rho}_{pq}$ for p-th ROI and q-th ROI.

interaction of four ROIs for the l-th subject, i.e., how the correlation of ij-th ROI pairs interacts with the correlation of pq-th ROI pairs $(1 \leqslant i, j, p, q \leqslant R)$. Physiologically, the interaction differences between different ROI pairs (as revealed by high-order FC) among different subjects could be used to identify Normal Control (NC) and ASD. Fig. 5 shows a simple example of high-order FC differences between two subjects of different disease cohorts (i.e., NC and ASD). We use four brain ROIs (i, j, p, q) in this example, where the red and blue colors respectively indicate a strong and weak correlation between two ROIs. As shown in Fig. 5, the correlation (i.e., high-order FC) between $\boldsymbol{\rho}_{ij}$ and $\boldsymbol{\rho}_{pq}$ could be different for subjects with different disease cohorts. This correlation difference could give some discriminative information which is beneficial for ASD classification.

However, the Ho-FCN constructed using the method above has serious curse of dimensionality issue. For instance, given U ROIs, the scale of Lo-D-FCN is $U \times U$, while the corresponding scale of this Ho-FCN is $(U \times U)^2$. In our case, $U = 116$, thus the scale of this Ho-FCN will be 116^4, containing at least thousands of vertices and millions of edges, which will lead to two terrible drawbacks. First, the computational complexity is very high in the stage of network construction and subsequent feature extraction. Second, the generalization performance of this large-scale network learning system will be negatively affected.

To tackle this problem, we propose to first perform a clustering algorithm to divide the FC time series $\left\{\boldsymbol{\rho}_{ij}^l\right\}_{1 \leqslant i,j \leqslant U}$ into several clusters, and then calculate the correlation between any two clusters as high-order FC. Specifically, to ensure the consistency of clustering

results across different subjects, we first stack temporal FC matrices of all subjects together (as shown Fig. 6) so that the FC time series for the same pair of ROIs for all the subjects are concatenated into a long vector. For example, the FC time series between the i-th and j-th ROIs of all subjects, i.e., $\left\{\boldsymbol{\rho}_{ij}^{l}\right\}_{1\leqslant l\leqslant L}$, are concatenated together into a long vector as $\boldsymbol{\rho}_{ij}=\left[\boldsymbol{\rho}_{ij}^{1},\boldsymbol{\rho}_{ij}^{2},\ldots,\boldsymbol{\rho}_{ij}^{L}\right]$, where L is the total number of all subjects. After that, we use Ward's linkage clustering [38], a widely used hierarchical clustering algorithm, to group all the concatenated FC time series of any pair of ROIs, i.e., $\left\{\boldsymbol{\rho}_{ij}\right\}_{1\leqslant i,j\leqslant U}$, into different clusters $\{\Phi_n\}_{1\leqslant n\leqslant N}$, where Φ_n consists of the index pair (i, j) if $\boldsymbol{\rho}_{ij}$ is included in the n-th cluster, and N is the total number of clusters. Then, multi-order central moment series of Φ_n are calculated to represent different statistical views of each cluster. The first-order central moment, i.e., mean, of the k-th correlations in the n-th cluster for the l-th subject is calculated as follows:

$$\overline{\rho}_{n}^{l}\left(k\right)=\frac{\sum_{(i,j)\in\Phi_n}\rho_{ij}^{l}\left(k\right)}{|\Phi_n|}\quad{}_{1\leqslant k\leqslant K} \tag{4}$$

where $|\Phi_n|$ denotes the number of FC time series in cluster Φ_n. Then, the dynamic changes of mean FC time series $\left\{\overline{\rho}_{n}^{l}(k)\right\}_{1\leqslant k\leqslant K}$ for cluster Φ_n is given as $\overline{\boldsymbol{\rho}}_{n}^{l}=\left[\overline{\rho}_{n}^{l}(1),\overline{\rho}_{n}^{l}(2),\cdots,\overline{\rho}_{n}^{l}(K)\right]$. Furthermore, the d-th order ($d>1$) central moment of the k-th correlations in the n-th cluster for the l-th subject is calculated as:

$$m_{n}^{l}\left(d\right)\left(k\right)=\sqrt[d]{\frac{\sum_{(i,j)\in\Phi_n}\left(\rho_{ij}^{l}(k)-\overline{\rho}_{n}^{l}(k)\right)^{d}}{|\Phi_n|}}\quad{}_{2\leqslant d\leqslant D,1\leqslant k\leqslant K} \tag{5}$$

where D denotes the highest order number in our study. Then, the d-th order ($d>1$) central moment series of the n-th cluster for the l-th subject is given as $\boldsymbol{m}_{n}^{l}(d)=\left[m_{n}^{l}(d)(1),m_{n}^{l}(d)(2),\ldots,m_{n}^{l}(d)(K)\right]$. It represents the d-th order central moment characteristic over time for all the FC time series in the n-th cluster. The construction of $\boldsymbol{m}_{n}^{l}(d)$ is illustrated in Fig. 7. We describe our Ho-FCN construction in the following subsection.

D. Multi-view Ho-FCNs Construction

We then compute correlation of central moment series between two clusters as high-order FC. Specifically, the PCC value for the d-th order central moment series between the u-th and v-th clusters of the l-th subject is calculated as follows:

$$h_{u,v}^{l}\left(d\right)=corr\left(\boldsymbol{m}_{u}^{l}\left(d\right),\boldsymbol{m}_{v}^{l}\left(d\right)\right)_{2\leqslant d\leqslant D} \tag{6}$$

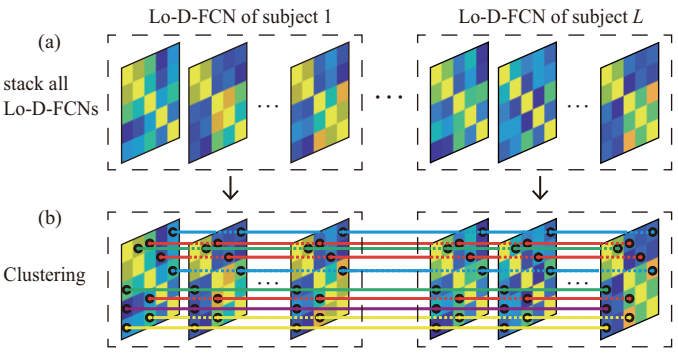


Fig. 6. The illustration of clustering. (a) stack Lo-D-FCNs of all subject together; (b) the result of clustering, lines with the same color in the graph indicate that they are in the same cluster.

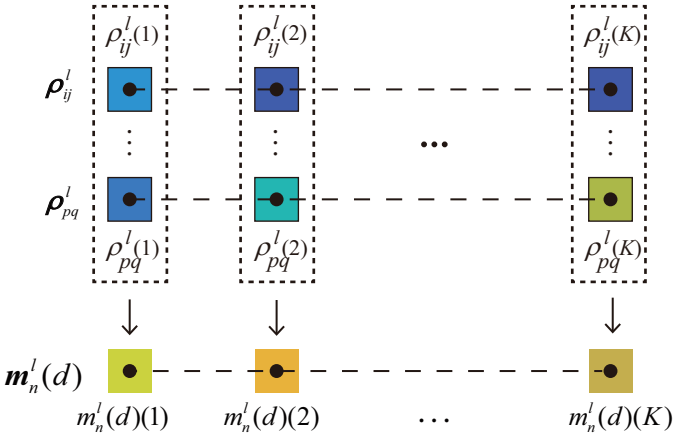


Fig. 7. Schematic diagram about central moment series ($\boldsymbol{m}_{n}^{l}(d)$) of n-th cluster, where $\left\{\boldsymbol{\rho}_{ij}^{l},\ldots,\boldsymbol{\rho}_{pq}^{l}\right\}$ belongs to n-th cluster, i.e., $\boldsymbol{\Phi_n}=\left\{\boldsymbol{\rho}_{ij}^{l},\ldots,\boldsymbol{\rho}_{pq}^{l}\right\}$.

where $h_{u,v}^{l}(d)$ is the corresponding high-order FC that captures complex and high-level interactions among multiple ROIs included in the u-th and v-th clusters. By grouping high-order FCs computed from every pair of clusters, we obtain a Ho-FCN $\boldsymbol{G}_{Ho}^{l}(d)=\left(\left\{\boldsymbol{m}_{n}^{l}(d)\right\},\left\{h_{u,v}^{l}(d)\right\}\right)$, where $\boldsymbol{m}_{n}^{l}(d)$ represents the n-th vertex and $h_{u,v}^{l}(d)$ represents the weight of the edge linking vertices $\boldsymbol{m}_{u}^{l}(d)$ and $\boldsymbol{m}_{v}^{l}(d)$. Note that, by using this "correlation of cluster moments" strategy, we effectively reduce the dimensionality of Ho-FCN from U^4 (i.e., for Ho-FCN constructed using "correlation of FC time series" strategy) to N^2, where N is the number of clusters, thus averting the curse of dimensionality issue of previous Ho-FCN construction.

Furthermore, since central moments of different orders can reflect different distinct statistical properties of each cluster, we propose to use multiple central moments to extract multi-view high-order FCs and construct multi-view Ho-FCNs. For instance, for $d=1$, $\boldsymbol{m}_{n}^{l}(1)=\overline{\boldsymbol{\rho}}_{n}^{l}$ is the mean FC time series for the n-th cluster, and the corresponding high-order FC between the u-th and v-th clusters of the l-th subject is calculated as:

$$h_{u,v}^{l}\left(1\right)=corr\left(\overline{\boldsymbol{\rho}}_{u}^{l},\overline{\boldsymbol{\rho}}_{v}^{l}\right) \tag{7}$$

and the corresponding Ho-FCN is $\boldsymbol{G}_{Ho}^{l}\left(1\right)=\left(\left\{\overline{\rho}_{n}^{l}\right\},\left\{h_{u,v}^{l}\left(1\right)\right\}\right)$.

E. Feature Extraction of Functional Connectivity Network

After constructing the low-order and high-order FCNs, it is essential to extract informative features from these networks to train a classifier. For C-FCN and Ho-FCNs, we use the FC matrices $\boldsymbol{G}_{C}^{l}$ and $\boldsymbol{G}_{Ho}^{l}(d)$ as the raw features. However, we cannot directly use the FC matrices $\boldsymbol{G}_{Lo}^{l}$ from Lo-D-FCN for classification, as the dynamic features are not exactly corresponding for different subjects due to possible phase mismatch. For example, the time series in the k-th window for two different subjects could be out of phase (i.e., not corresponding), due to the unconstrained mental activities when acquiring the rs-fMRI signals. As the phases of the constructed subnetworks in Lo-D-FCN are unknown, it is hard to ensure temporal correspondence for the subnetworks between different subjects. To avoid this phase mismatch issue, we propose to extract central-moment features from Lo-D-FCN, considering their temporal-invariant attribute.

Similar with Eq. (5), the e-th order central moment $m_{ij}(e)$ of the FC time series $\boldsymbol{\rho}_{ij}^{l}$ of the l-th subject is given as follows:

$$m_{ij}^{l}(e)=\sqrt[e]{\frac{\sum_{k=1}^{K}\left(\rho_{ij}^{l}(k)-\overline{\boldsymbol{\rho}}_{ij}^{l}\right)^{e}}{K}}\ (e=2,\cdots,E) \tag{8}$$

where E is the largest moment order and $\overline{\boldsymbol{\rho}}_{ij}^{l}$ denotes the average of all elements in $\boldsymbol{\rho}_{ij}^{l}$, which is also equivalent to $m_{ij}^{l}(1)$, i.e., e = 1, as given below:

$$m_{ij}^{l}(1)\ =\ \frac{\sum_{k=1}^{K}\boldsymbol{\rho}_{ij}^{l}(k)}{K} \tag{9}$$

Comparing Eq. (8)-(9) with Eq. (4)-(5), we can see that the central moments described here and in Section III.C are computed across different dimension of data. Here, the central moments are computed across temporal dimension for each FC time series, thus summarizing each FC time series $\left(\in U^{K}\right)$ with temporal-invariant moment values $\left(\in U^{K}\right)$. In contrast, the central moments in Section III.C are computed across FC time series within a cluster, thus summarizing each cluster of FC time series $\left(\in U^{|\phi_n|\times K}\right)$ with a small set of moment series $\left(\in U^{D\times K}\right)$. Finally, by computing the e-th order central moments for all the FC time series in Lo-D-FCN (represented by $\boldsymbol{G}_{Lo}^{l}$, we obtain the e-th order central moment matrix $\boldsymbol{M}_{Lo}^{l}(e)(e=1,2,\ldots,E)$ as follows:

$$\boldsymbol{M}_{Lo}^{l}(e)=\left\{m_{ij}(e)\right\}_{1\leqslant i,j\leqslant U} \tag{10}$$

where $\boldsymbol{M}_{Lo}^{l}(e)$ is a $U\times U$ symmetric matrix, which will be subsequently used to train a classifier, as described in the following subsection.

F. Feature selection, Classification and Fusing

Three types of FCNs, i.e., C-FCN, Lo-D-FCN, and multi-view Ho-FCNs, are constructed for each subject in this study. To avoid redundancy, we only use lower off-diagonal triangular part of these FCNs matrices as features, and vectorize them into a feature vector. For convenience, we denote the feature vectors for the l-th subject as $\boldsymbol{y}_{C}^{l}$ for C-FCN, $\boldsymbol{y}_{Lo}^{l}(e)$ for Lo-D-FCN, and $\boldsymbol{y}_{Ho}^{l}(d)$ for Ho-FCNs, where e $(1\leqslant e\leqslant E)$ denotes the order number of the central moments used to extract features from Lo-D-FCN, while d $(1\leqslant d\leqslant D)$ denotes the order number of the central moments used to construct the multi-view Ho-FCNs.

We then employ a two-stage feature selection strategy to select discriminative features from the feature vector of each FCN. Firstly, the t-test is performed on the training dataset to select a set of preliminary features that significantly discriminate from $\boldsymbol{y}_{C}^{l}$, $\boldsymbol{y}_{Lo}^{l}(e)$ and $\boldsymbol{y}_{Ho}^{l}(d)$. These selected features, which are denoted as $\hat{\boldsymbol{y}}_{C}^{l}$, $\hat{\boldsymbol{y}}_{Lo}^{l}(e)$ and $\hat{\boldsymbol{y}}_{Ho}^{l}(d)$, respectively, are highly correlated to class labels. Then, we adopt a l_1-norm regularized least squares regression, i.e., LASSO [39], to remove redundant features and further select a smaller set of crucial features related to ASD. Specifically, let I^{l} denote the class label of the l-th subject (i.e., $I^{l}=+1$ for ASD and $I^{l}=-1$ for NC), $\boldsymbol{y}^{l}$ denotes the feature vector selected by t-test, and $\boldsymbol{\omega}$ denotes a weight vector, then the LASSO feature selection model can be formalized as follows:

$$\min_{\boldsymbol{\omega}}\frac{1}{2}\sum_{l=1}^{L}\left\|I^{l}-<\boldsymbol{y}^{l},\boldsymbol{\omega}>\right\|_{2}^{2}+\lambda\|\boldsymbol{\omega}\|_{1} \tag{11}$$

where $<\cdot,\cdot>$ denotes the inner product operator, and λ is a regularization term. A larger value of $\boldsymbol{\lambda}$ can make the solution $\boldsymbol{\omega}$ sparser. By setting a proper value for λ, we can achieve sparse feature selection, where features corresponding with the non-zero elements of $\boldsymbol{\omega}$ are retained. For simplicity, let $\hat{\hat{\boldsymbol{y}}}_{C}^{l}$, $\hat{\hat{\boldsymbol{y}}}_{Lo}^{l}(e)$ and $\hat{\hat{\boldsymbol{y}}}_{Ho}^{l}(d)$ represent the final feature sets selected from the feature vectorsy $\hat{\boldsymbol{y}}_{C}^{l}$, $\hat{\boldsymbol{y}}_{Lo}^{l}(e)$ and $\hat{\boldsymbol{y}}_{Ho}^{l}(d)$, respectively.

After the feature selection, we train a SVM [40] for disease identification. Since we have designed three types of networks, one SVM classifier is trained for each type of features ($\hat{\hat{\boldsymbol{y}}}_{C}^{l}$, $\hat{\hat{\boldsymbol{y}}}_{Lo}^{l}(e)$ and $\hat{\hat{\boldsymbol{y}}}_{Ho}^{l}(d)$). Finally, the outputs of the SVM models are fused together to give the final classification result. Specifically, each SVM will output an associated decision score, indicating the probability of that subject belonging to a class. Then, the decision scores from all SVM models are fused linearly (by a weighting parameter $\alpha\in[0.1,0.2,\ldots,0.9]$ tuned for each SVM, when fusing F classifiers, we set F weights, and $\alpha_1+\alpha_2+\cdots+\alpha_F=1$) to output the final predicted label for the target subject.

IV. Experimeetal Analysis

We adopted the nested six-fold cross-validation (CV) strategy to evaluate the effectiveness of our proposed method. It consists of two loops of six-fold CV, i.e., an outer loop of six-fold CV to verify the model performance, and an inner loop of six-fold CV to determine the model hyperparameters for each fold of the outer loop. Specifically, in the outer loop six-fold CV, all data were first randomly divided into six subsets of roughly equal size, where one subset was used as the testing set, while the other five subsets were used as the training set. In the inner loop, six-fold CV was performed on the training set in the outer loop to tune the model hyperparameters (i.e., t, λ and C). The hyperparameters obtained from the inner loop was then applied to the training set of the outer loop to train our model, which was subsequently used to evaluate our model in the testing set. This process was repeated six times where each of the six subsets was used exactly one time as testing set as in six-fold CV. We also repeated this nested six-fold CV for ten times and obtained the average performance as our final result. Therefore, the dataset used to test performance never engages in the parameters selection thus avoiding overfitting. To avoid selection bias, the validation process was conducted 10 times, with random partitions each time. The average classification performance was reported as the final evaluation results. Four widely used metrics were adopted for performance evaluation, including classification accuracy (ACC), sensitivity or true positive rate (TPR), specificity or true negative rate (TNR), and F1 score. In our experiments, ASD and NC were considered as positive and negative classes, respectively.

A. Influence of Parameters on FCNs

As described in the Method section, the sliding window strategy is employed to construct Lo-D-FCN and Ho-FCNs. Thus, the free parameters such as sliding window length (W), step size (s) and the order of central moments (i.e., e for Lo-D-FCN, and d for Ho-FCN) may affect the identification accuracy. We optimize each network by adjusting the values of parameters W, s, e and d. The range of these parameters are $W\in[30,40,\ldots,80]$, $s\in[2,4,\ldots,12]$, $e\in[1,2,\ldots,7]$, and $d\in[1,2,\ldots,7]$. In addition, the performance of our approach also depends on some hyperparameters, such as t in t-test, λ in LASSO and C in SVM model. In our experiment, we set their ranges to $t\in[0.01,0.02,\ldots,0.05]$, $\lambda\in[0.1,0.2,\ldots,0.9]$, and $C\in[2^{-4},2^{-3},\ldots,2^{4}]$. Moreover, for Ho-FCN, the value of clustering number N also affects the diagnosis accuracy of Ho-FCN. Thus, we also change the value of N in the range of $N\in[200,300,\ldots,500]$ to investigate its effect on the classification performance.

TABLE II
THE BEST PERFORMANCES OF LO-D-FCN USING DIFFERENT ORDER OF CENTRAL MOMENTS

Feature type	W	s	ACC (%)	TPR (%)	TNR (%)	F1 (%)
$\boldsymbol{G}_{Lo}(1)$	40	8	73.9±0.24	71.4±0.23	74.6±0.01	71.7±0.07
$\boldsymbol{G}_{Lo}(2)$	60	10	78.9±0.15	80.8±0.26	77.7±0.11	78.1±0.51
$\boldsymbol{G}_{Lo}(3)$	40	10	65.5±0.17	63.0±0.19	67.9±0.13	62.3±0.23
$\boldsymbol{G}_{Lo}(4)$	60	2	79.3±0.47	76.9±0.32	79.2±0.41	77.4±0.14
$\boldsymbol{G}_{Lo}(5)$	40	12	67.4±0.20	65.5±0.36	69.0±0.23	65.0±0.16
$\boldsymbol{G}_{Lo}(6)$	70	10	76.7±0.45	71.8±0.41	80.1±0.16	73.7±0.12
$\boldsymbol{G}_{Lo}(7)$	40	12	72.6±0.52	69.2±0.35	71.7±0.18	69.1±0.44

TABLE III
THE BEST PERFORMANCES OF HO-FCNS USING DIFFERENT ORDER OF CENTRAL MOMENTS

Feature type	W	s	ACC (%)	TPR (%)	TNR (%)	F1 (%)
$\boldsymbol{G}_{Ho}(1)$	40	10	76.1±0.02	75.4±0.07	74.9±0.06	74.4±0.04
$\boldsymbol{G}_{Ho}(2)$	30	12	71.3±0.035	68.1±0.05	74.7±0.05	69.4±0.03
$\boldsymbol{G}_{Ho}(3)$	30	2	74.9±0.01	66.1±0.05	79.2±0.03	69.4±0.03
$\boldsymbol{G}_{Ho}(4)$	30	6	69.5±0.02	68.6±0.05	70.3±0.04	68.3±0.03
$\boldsymbol{G}_{Ho}(5)$	30	2	72.9±0.02	70.5±0.03	75.1±0.03	71.3±0.02
$\boldsymbol{G}_{Ho}(6)$	80	2	68.5±0.02	68.4±0.04	68.7±0.06	67.6±0.10
$\boldsymbol{G}_{Ho}(7)$	30	2	75.7±0.05	70.7±0.06	76.6±0.05	71.9±0.04

TABLE IV
THE BEST PERFORMANCES OF C-FCN

Feature type	ACC (%)	TPR (%)	TNR (%)	F1 (%)
$\boldsymbol{G}_C$	74.2±0.24	71.4±0.05	76.9±0.08	72.4±0.12

Detailed results of average ACC using features from Lo-D-FCN and Ho-FCN with different values of W, s and d are reported in the supplementary materials (Fig. S1 and Fig. S2). Fig. S1 shows the average ACC using features from Lo-D-FCN with different values of W, s and e. It can be seen that the maximum value of ACC is 79.3% when W = 60, s = 2, e = 4, and the minimum value of ACC is 54.0% when W = 30, s = 10, e = 3. Furthermore, Fig. S2 shows the average ACC value for each Ho-FCN with different values of W, s, d and N. In Fig. S2, we can see that the maximum value of ACC is 76.1% when W = 40, s = 10, d = 1, N = 300, and the minimum value of ACC is 45.7% when W = 70, s = 10, d = 4, N = 400. In subsequent experiments, we set the cluster number of Ho-FCN to 300, that is, N = 300.

Table II shows the best performances of Lo-D-FCN in terms of different feature types, where $\boldsymbol{G}_{Lo}(e)$ denotes the features extracted with the e-th order central moments (e = 1, 2, …, 7), and Table III shows the best performances of Ho-FCNs built with central moments (N = 300), where $\boldsymbol{G}_{Ho}(d)$ denotes the Ho-FCN constructed with the d-th order central moments (d = 1, 2, …, 7).

Based on the results, we can derive the following empirical findings: (1) the classification preformation is rather sensitive to the free parameters, so it is necessary to adjust them to obtain the best performance; (2) Ho-FCNs constructed with different order of central moments have different performance, indicating that each view of Ho-FCNs provides different and useful information for ASD identification.

In addition, the performance of C-FCN is not affected by the free parameters. C-FCN emerges the strength and directionality (i.e., positive or negative) of FC from a static view.

The positive correlation reflects the synchronous activities between brain regions, while the negative correlation represents a competitive relationship between brain regions. Both positive and negative correlations reflect genuine physiological processes [41]–[43].

C-FCN, as a traditional functional connectivity network, provides a static view for the interaction relationship between different brain regions, and is usually used as a baseline for comparison with other networks. In this study, we list its performance in Table IV.

B. The relationship between clustering and high-order FC

In order to analyze the relationship between clustering and high-order FC, an example is given in Supplementary Materials Fig. S3. Fig. S3(a) and Fig. S3(d) show the FC time series from one NC subject, while Fig. S3(b) and Fig. S3(e) show the FC time series from one ASD subject. In addition, the FC time series in Fig. S3(a) and Fig. S3(b) belong to the same cluster, while the FC time series in Fig. S3(d) and Fig. S3(e) belong to another cluster. Fig. S3(c) and Fig. S3(f) respectively show the connections of ROI pairs contained in these two clusters. From Fig. S3, we can see that the FC time series within the same cluster have similar trends, and the brain regions contained in each cluster have similar functions, for example $CAL^{41,42}$, $MOG^{49,50}$, $CUN^{43,44}$, $IOG^{51,52}$, etc., (the digits represent the indexes of ROIs, see Supplementary Materials: Table I) are related to visual [44]–[46]; on the other hand, for the same subject, whether it is NC or ASD, the trends of FC time series from different clusters show significant differences.

We also show the correlation of central moment series of the two clusters (i.e., high-order FC) in Fig. 8. In Fig. 8, the blue and red bars denote the correlation between the central moment series of clusters from NC (Fig. S3(a) and Fig. S3(d)) and ASD (Fig. S3(b) and Fig. S3(e)) subjects, respectively. As shown in Fig. 8, (1) there are significant differences in these high-order FCs between the NC and ASD subjects, especially for the correlation of seventh-order central moment sequences of the two clusters, showing a great difference in positive and negative directions. This indicates that the Ho-FCNs constructed with different order of central moment series may contain different discriminative information which can be used for ASD diagnosis. (2) For the same subject, the correlations of central moment series from the two clusters are greatly differing for different orders, indicating that the information of high-order FC revealed by each central moment is different.

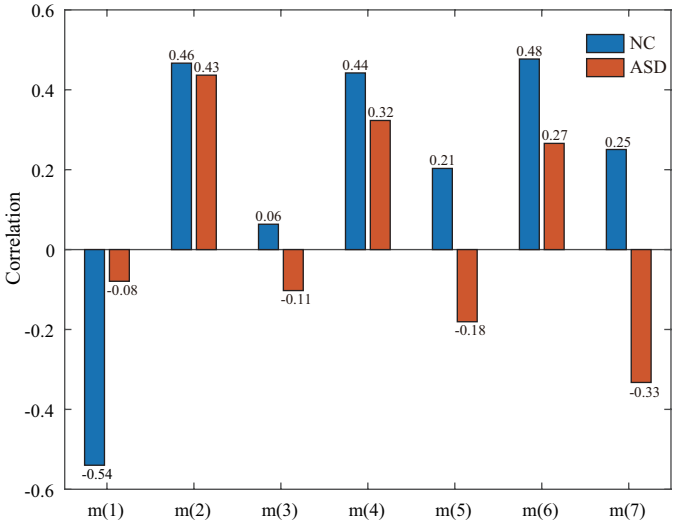


Fig. 8. The high-order correlation of NC is compared with that of ASD, where $m(i)$ denotes the correlation calculated using the FC time series of i-order central moment, i = 1, 2,…, 7. The NC and ASD subjects are chosen randomly in the dataset.

C. Most Discriminative Features for ASD Diagnosis

Since we use two-stage feature selection strategy, the number of features selected depends on two hyperparameter variables, i.e., t in

the t-test, and λ in LASSO. We performed ten repetitions of six-fold CV to investigate how these two hyperparameter values affect the number of selected features. The average number of features selected in ten repetitions of six-fold CV using different combinations of hyperparameter values are shown in Supplementary Materials Fig. S4 and Fig. S5. From these two figures, we can see that the number of selected features decrease with the increase of t value and the decrease of λ value. In our experiment, on average, when the values of λ and t equal 0.2 and 0.01, respectively, the classification accuracy reaches the highest. At this time, the number of selected features fed to SVM is between 55 and 100 since in the procedure of CV the training dataset in each fold is different.

To find the most discriminative features for ASD diagnosis, we computed the frequency of features that were selected in the CV and 10 features with the highest selection frequency were selected. In this section, we analyze the most discriminative features extracted from Lo-D-FCN and Ho-FCNs, whose ACC are greater than C-FCN. In Lo-D-FCN, the features are the moment statistics of the low-order dynamic FC, so one feature represents the connection of an ROI pair, while in Ho-FCN, the features are extracted from the Ho-FCN matrix, so one feature represents the connection of a cluster pair.

The connectograms of the 10 most frequently selected connections from C-FCN and Lo-D-FCN are shown in Fig. 9 and Supplementary Materials Fig. S6, respectively. Each connection represents the correlation between two ROIs, and the thickness of the connection indicates the selected frequency, i.e., a thicker line indicates a higher frequency.

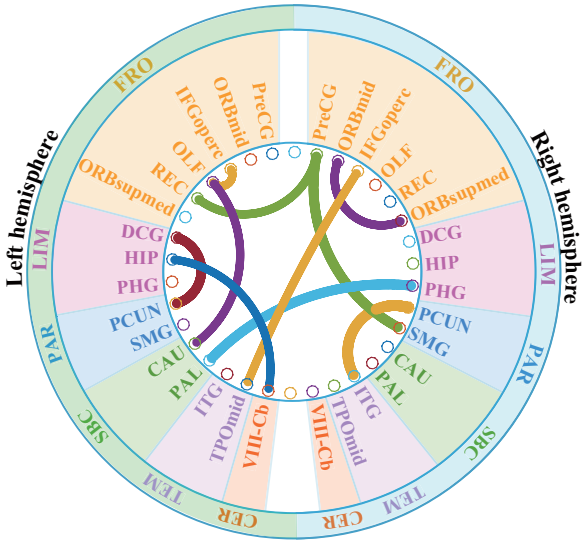


Fig. 9. Connectogram of the 10 most frequently selected connections in C-FCN.

Through comparison of Fig. 9 and Fig. S6, it can be seen that all their connection lines are different. This implies that features selected from C-FCN and Lo-D-FCN are significantly different, which indicates that C-FCN and Lo-D-FCN can provide complementary information for diagnosis. The reason is that C-FCN generally captures the static connectional feature since its FC is measured using the whole scanning time in rs-fMRI series for each pair of ROIs, while Lo-D-FCN reveals the dynamic connectional relationship for ROI pairs as each of its FCs is measured using a short segment in rs-fMRI series.

In order to find the relationship between the features selected from Lo-D-FCN and Ho-FCNs, we show whether the most discriminative ROI pairs selected from the Lo-D-FCN are included in the most discriminative cluster pairs selected from the Ho-FCN in Table V. It can be seen that the ROI pairs selected by Lo-D-FCN seldomly appear in the clusters selected by Ho-FCNs. This can be explained by the fact that the information captured by these two types of networks is significantly different, i.e., Lo-D-FCN only captures simple low-order FC between two ROIs, while Ho-FCNs capture high-order FC among multiple ROI pairs. This indicates that the information extracted from Lo-D-FCN and Ho-FCN is complementary and can be used together for more accurate ASD diagnosis.

In addition, we also calculate the contribution of the 10 most discriminative features selected from Lo-D-FCN. The contribution is defined as the ratio of occurrence for ROI pairs in 10 six-fold CVs. For example, if a feature is selected 41 times, then its contribution is $41/60 = 0.68$, the related detailed information is shown in Table VI. It can be seen that the contribution of the same ROI pair to the classification performance is different in Lo-D-FCN with different orders. This suggests that they may contain complementary information for classification, and integrating them may comprehensively capture subtle disruptions of brain functional organization induced by neuropsychiatric disorders.

Fig. S7 shows the connectogram of the 10 most frequently selected connections from the Ho-FCNs (N = 300) (see Fig. S7(a)–(c) in the supplementary materials). It can be seen that each line connects two different clusters, and each cluster contains multiple ROI pairs. The connection lines in Fig. S7(a)–(c) differ greatly, indicating that features selected from $\boldsymbol{G}_{Ho}(1)$, $\boldsymbol{G}_{Ho}(3)$ and $\boldsymbol{G}_{Ho}(7)$ are greatly different.

The main reason could be: (1) the presented multi-view Ho-FCNs are based on clustering which depends on the parameters of constructing Lo-D-FCN (i.e., W, s). Therefore, even if the number of clusters is the same, the different parameters will still lead to different higher-order network structure; (2) as we mentioned above, high-order FC calculated with different central moment series can reflect the FC patterns among multiple ROIs from different aspects. In this way, multi-view Ho-FCNs can provide complementary FC information.

In summary, different types of FCNs and different features of the same FCN can provide complementary information, which can be used jointly to significantly improve the classification performance.

TABLE V
THE CORRESPONDENCE OF ROI PAIRS SELECTED BY LO-D-FCN AND THE CLUSTERS SELECTED IN HO-FCN

ROI pair	Corresponding cluster		
	$\boldsymbol{G}_{Ho}(1)$	$\boldsymbol{G}_{Ho}(3)$	$\boldsymbol{G}_{Ho}(7)$
PreCG.L & VI-VER		•	
IFGoperc.R & TPOmid.R			
PCL.R & FFG.L			
REC.R & III-VER	•		
IOG.L & TPOsup.L			
FFG.L & IPL.R		•	
PCUN.R & IX-Cb			
MTG.R & VIII-VER			
ITG.L & VII-VER	•		
IX-VER & X-Cb.L	•	•	•
SMA.R & X-Cb.L			
SMA.L & VIIB-Cb.R			

the solid dot "•" means that the ROI pair (of the corresponding row) is contained in the cluster (of the corresponding column). For example, the solid dot in the first row means that the ROI pair PreCG.L & VI-VER which is selected from Lo-D-FCN is also included in the cluster selected from $\boldsymbol{G}_{Ho}(3)$ of Ho-FCN, but not in $\boldsymbol{G}_{Ho}(1)$ and $\boldsymbol{G}_{Ho}(7)$ of Ho-FCNs.

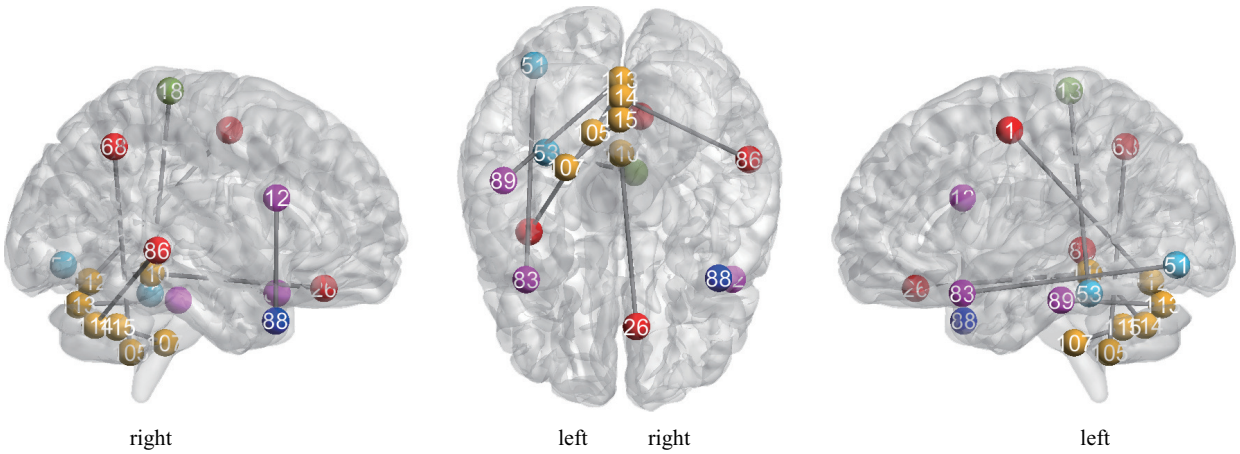


Fig. 10. Visualization of the common connections selected from $\boldsymbol{G_{Lo}(2)}$, $\boldsymbol{G_{Lo}(4)}$ and $\boldsymbol{G_{Lo}(6)}$, RIOs corresponding to the indexes are shown in the Supplementary Materials: Table I.

D. Analysis of related ROIs with ASD

All the ROIs are grouped into different resting-state networks represented by different colors according to previous studies [47]–[51], including default mode network (DMN), executive attention network (EAN), visual network (VN), sensorimotor network (SMN), subcortical regions (SR) and cerebellum (CER). The DMN is known to activate at rest and deactivate during task performance. Studies have shown that the ASD children are unable to modulate the deactivation of the DMN and suppress the mental activity during rest [52]–[55]. The CER involving in the fine motor function plays an important role in higher cognitive functions such as language [56]. Some recent studies have implicated cerebellar connectivity deficits in ASD patients [57]–[59].

All the ROIs associated with the connections in Fig. S6(d) are shown in Fig. 10 which are colored according to the resting state network that they belong to. From Fig. 10 and Fig. S6(d), we have the following findings: (1) the selected connections and brain regions are distributed across both hemispheres and different lobes, indicating the distributed pattern of functional abnormalities over the whole brains of ASD patients; (2) the selected brain regions were distributed in six resting state networks, and one third of them belong to CER, which is highly associated with cognitive functions. These findings are consistent with previous reports [58], [59].

To reveal the relationship between the ROIs selected by Ho-FCNs and ASD, we also visualized the ROI pairs connected by the connections with the highest selected frequency (see Fig. S7(d)–(g) in the supplementary material). From these figures, we can draw the following conclusions: (1) the distribution of ROIs in the same cluster are highly symmetric. This is probably because the activation state of ROIs in the left hemisphere is usually similar to that in the right hemisphere during resting state [60]; (2) generally, the selected ROIs are highly related to ASD. For instance, most of the ROIs contained in cluster pair (23, 64) (in Fig. S7(d) and Fig. S7(f)) belong to DMN, and many ROIs in the 102th cluster (Fig. S7(c)) belong to CER. Obviously, both DMN and CER are highly related to ASD; (3) these ROIs in the same cluster usually belong to the same resting state network, indicating that they have the similar functions. This also implies that the clustering strategy we adopted can maintain the purity of each cluster, thus making the high-order correlation reliable.

TABLE VI
THE DISCRIMINATIVE FEATURES OBTAINED FROM LO-D-FCN AND THEIR CONTRIBUTIONS

ROI pair	Contribution		
	$\boldsymbol{G}_{Lo}(2)$	$\boldsymbol{G}_{Lo}(4)$	$\boldsymbol{G}_{Lo}(6)$
PreCG.L & VI-VER	0.68	0.93	0.83
IFGoperc.R & TPOmid.R	0.73	0.68	0.71
FFG.L & PCL.R	0.98	0.95	0.98
REC.R & III-VER	0.90	0.96	0.85
IOG.L & TPOsup.L	0.53	0.85	0.88
IX-Cb.L & PCUN.R	0.91	0.90	0.85
MTG.R & VIII-VER	0.75	0.75	1.00
ITG.L & VII-VER	0.45	0.48	0.98
X-Cb.L & IX-VER	0.71	0.95	0.70
FFG.L & IPL.R	0.45	-	-
X-Cb.L & SMA.R	-	0.43	-
SMA.L & VIIB-Cb.R	-	-	0.85

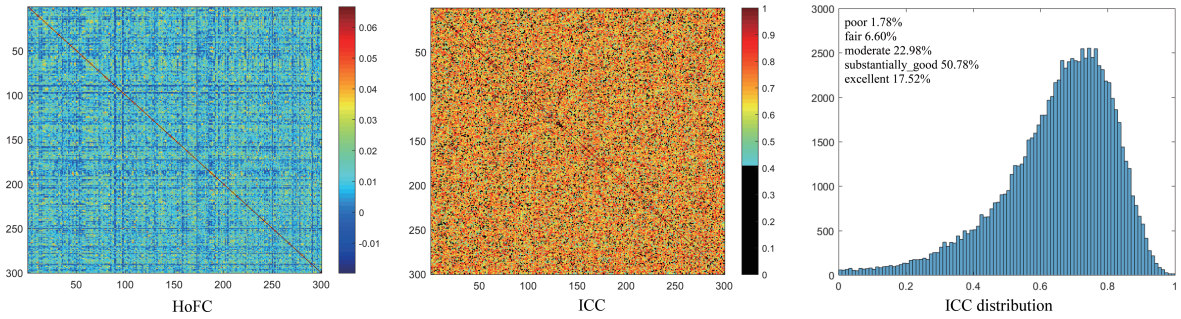
E. Fusion Results of FCNs

Firstly, we train three SVMs for each type of our FC networks, i.e., C-FCN, Lo-D-FCN, and Ho-FCNs, respectively. The parameters make the highest ACC were selected for each SVM model using only the training data. In our experiment, the selected parameters are $W = 60$, $s = 2$, $e = 4$ for Lo-D-FCN, $W = 40$, $s = 10$, $d = 1$, $N = 300$ for Ho-FCN. Then, we ensemble the classification scores from these SVMs to yield the final classification result. Note that only the training data are adopted to avoid biased results. Secondly, based the above analysis (i.e., *Most Discriminative Features for ASD Diagnosis*), the central moment features with different orders in the same type of FCN, including both Lo-D-FCN and Ho-FCNs, can provide complementary information. Therefore, we also integrate all the orders of the same type of FCN to further improve the classification performance.

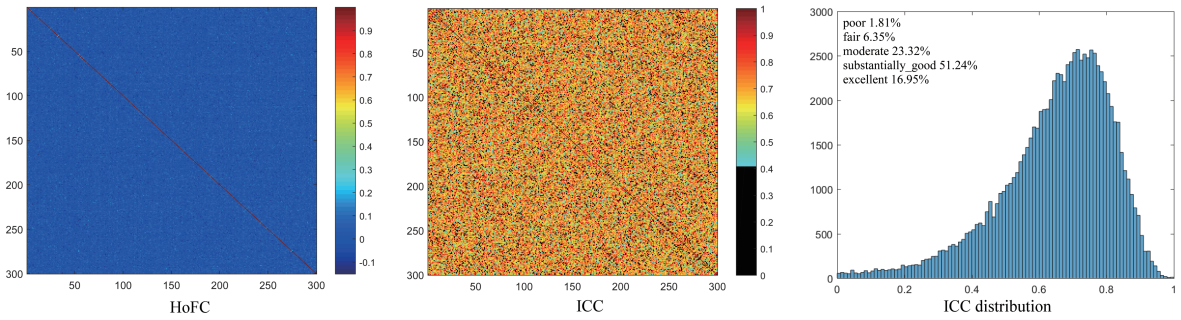
The average classification performances for each compared feature type are shown in Table VII, where $\boldsymbol{G}_C$ denotes the features obtained from C-FCN, $\boldsymbol{G}_C + \boldsymbol{G}_{Lo}$ denotes the features obtained from C-FCN

TABLE VII
ASD CLASSIFICATION USING DIFFERENT COMBINATIONS OF FEATURE TYPES

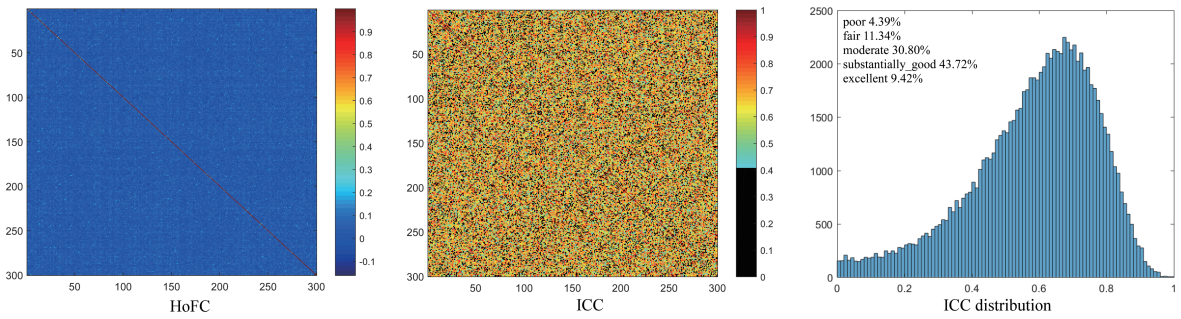
Feature type	ACC (%)	TPR (%)	TNR (%)	F1 (%)
$\boldsymbol{G}_C$+$\boldsymbol{G}_{Lo}(1)$	80.2	80.4	80.0	79.8
$\boldsymbol{G}_C$+$\boldsymbol{G}_{Ho}(1)$	76.8	77.3	76.3	76.5
$\boldsymbol{G}_{Lo}(4)$+$\boldsymbol{G}_{Ho}(1)$	81.3	81.3	81.2	80.9
$\boldsymbol{G}_C$+$\boldsymbol{G}_{Lo}(4)$+$\boldsymbol{G}_{Ho}(1)$	82.6	83.3	81.9	82.4
$\boldsymbol{G}_{Lo}(1\text{-}7)$	80.0	75.5	84.2	78.6
$\boldsymbol{G}_{Ho}(1\text{-}7)$	84.5	80.8	80.1	83.6
$\boldsymbol{G}_{Lo}(1\text{-}7)$+$\boldsymbol{G}_{Ho}(1\text{-}7)$	85.0	86.0	84.0	84.8
$\boldsymbol{G}_{ALL}$	**86.2**	**86.4**	**85.9**	**85.9**



(a) Ho-FCN with 1-*st* order central moment



(b) Ho-FCN with 3-*rd* order central moment



(c) Ho-FCN with 7-*th* order central moment

Fig. 11. ICC value for each connection of Ho-FCN with 1-*st*, 3-*rd*, and 7-*th* order central moment (window length = 60, step size = 4). From left to right are the ICC matrix without thresholding, ICC matrix shows the connectivities with thresholded by ICC $> \mathbf{0.4}$, and the ICC distribution for all connections. In the right of each row, the percentage of the connection with different reliability are also shown.

and Lo-D-FCN, the $\boldsymbol{G}_{Lo}(1\text{-}7)$ denotes fusing all type features for Lo-D-FCN, the meaning of $\boldsymbol{G}_{Ho}(1\text{-}7)$ is similar to $\boldsymbol{G}_{Lo}(1\text{-}7)$; $\boldsymbol{G}_{Lo}(1\text{-}7)$ + $\boldsymbol{G}_{Ho}(1\text{-}7)$ denotes fusing all type features for Lo-D-FCN and Ho-FCNs; $\boldsymbol{G}_{ALL}$ denotes fusing all type features for C-FCN, Lo-D-FCN and Ho-FCNs.

According to Table VII and the previous analysis, we can draw the following conclusions: (1) for Lo-D-FCN, compared with the performance of single type feature classification, the ACC of $\boldsymbol{G}_{Lo}(1\text{-}7)$ is only a little higher than that of $\boldsymbol{G}_{Lo}(4)$, which indicates that the fusion of different features of Lo-D-FCN cannot significantly improve the classification performance. This is probably because different features of Lo-D-FCN contain more redundant and less complementary information for ASD classification; (2) for multi-view Ho-FCNs, the ACC of $\boldsymbol{G}_{Ho}(1\text{-}7)$ is at least 8.4% higher than that of single Ho-FCN. The may imply that the classification information among Ho-FCNs is highly complementary; (3) the ensembled classifier consistently outperforms that of single feature type, which supports the assumption that integrating multiple order connectional features can boost classification accuracy; (4) $\boldsymbol{G}_{ALL}$ achieves the best classification performance, indicating that features derived from different FCNs can provide complementary information for ASD diagnosis, and the fusion of this information can further improve the classification performance.

F. The reliability analysis of Ho-FCN

In order to analyze the reliability of our proposed Ho-FCN, we used dedicated rs-fMRI data as part of the Consortium for Reliability and Reproducibility (CoRR) [61] for test-retest measurement. The dataset is publicly available (http://dx.doi.org/10.15387/fcp_indi.corr.hnu1). The original dataset contains 30 healthy adults (aged 20–30 years old, 15 females) with 10 repeated rs-fMRI scans (sessions) within 1 month. For more detailed data information including the scanner equipment, resolution, etc., please refer to the data release and CoRR websites. Following the previous test-retest reliability works [62], [63], we used two sessions for reliability assessment. The dataset was preprocessed by using DPARSF v5.2 [64] based on previous studies [65]. In order to measure the test-retest reliability of our proposed Ho-FCN connections, we adopt a common index, which is termed

as ICC [66]. ICC divides the total sum of variance across subjects and repeated rs-fMRI scans into two parts: between-subject (σ_b^2) and within-subject (or inter-session variance, σ_w^2) sum of variance. The definition of ICC is given as:

$$ICC = \frac{\sigma_w^2}{\sigma_b^2 + \sigma_w^2} \tag{12}$$

In practice, the formulation to estimate ICC based on samples is given by:

$$ICC = \frac{MS_b - MS_w}{MS_b + (k-1) \times MS_w} \tag{13}$$

where MS_b denotes the mean square of between-subject sum of variance, MS_w denotes the mean square of within-subject sum of variance, and k denotes the number of repeated scans (here k = 2). Following previous works [67], [68], the reliability can be divided into five classes according to the calculated ICC: poor ($0 <$ ICC < 0.2), fair ($0.2 <$ ICC < 0.4), moderate ($0.4 <$ ICC < 0.6), substantially good ($0.6 <$ ICC < 0.8), and excellent ($0.8 <$ ICC). In the experiments, for Ho-FCN, different parameters including window length, step size, the number of clusters, will affect the constructed Ho-FCN. To simplify the analysis, we fix the number of clusters to 300 as reported in the paper and generate multiple networks of our proposed networks $G_{Ho}(1)$, $G_{Ho}(3)$ and $G_{Ho}(7)$ for different combinations of window length and step size. Specifically, we change window length and step size in [50, 60, 70] and [2, 4, 6], respectively. The experimental result of $G_{Ho}(1)$, $G_{Ho}(3)$ and $G_{Ho}(7)$ when window length equals 60 and step size equals 4 is shown in Fig. 11(a), (b), and (c), respectively. The other results with different window length and step size are shown in the supplementary materials (Fig. S8–S11) of this paper.

From the results, we can see that the proposed Ho-FCNs are reliable metrics. For example, for Ho-FCN with 1-*st* order central moment $G_{Ho}(1)$, the amount of connections with ICC > 0.4 is 91.28% of all possible connections. In fact, a dedicated study of reliability carried out by [69] have confirmed the reliability of some Ho-FCN, including aHo-FCN (associated Ho-FCN), tHo-FCN (topographical profile similarity-based Ho-FCN), and dHo-FCN (dynamics-based Ho-FCN). Therefore, our results are consistent with this previous works and confirm the reliability of Ho-FCN again.

V. Conclusion

In this article, we propose a clustering-based multi-view high-order FCNs framework for ASD diagnosis, which is motivated by the observation that multi-view high-order FCNs can provide complementary identification information from different aspects. The experimental results provide the following findings: (1) The proposed Ho-FCNs is helpful for mining the relevant information for ASD diagnosis; (2) different types of FCNs can provide complementary information at different levels, and fusion of these information in a proper way can improve the classification accuracy of ASD; (3) the central-moment method can solve the phase mismatch problem in the low-order dynamic FC networks, allowing the dynamic features to be compared among different subjects.

One limitation of this work is that the Ho-FCN cannot accurately locate the brain regions that have significant impacts on ASD. A feasible solution to handle this problem is to construct a high-order network using the most discriminative clusters, and then select the brain regions highly related to ASD. In the future work, we will tackle this issue to further improve the performance of the model.

References

[1] E. Christine *et al*., "Describing the brain in autism in five dimensions–magnetic resonance imaging-assisted diagnosis of autism spectrum disorder using a multiparameter classification approach," *Journal of Neuroscience*, vol. 30, no. 32, pp. 10612-10623, 2010.

[2] M. J. Maenner *et al*. "Prevalence of Autism Spectrum Disorder Among Children Aged 8 Years — Autism and Developmental Disabilities Monitoring Network, 11 Sites, United States, 2016". MMWR Surveill Summ, vol. 69, no. 4, pp. 1-12, 2020.

[3] T. A. Lavelle *et al*, "Economic burden of childhood autism spectrum disorders". Pediatrics, vol. 133, no. 3, pp. e520-529, 2014.

[4] J. J. Wolff *et al*., "Differences in white matter fiber tract development present from 6 to 24 months in infants with autism," *American Journal of Psychiatry*, vol. 169, no. 6, 2012.

[5] Y. Jin *et al*., "Identification of infants at high-risk for autism spectrum disorder using multiparameter multiscale white matter connectivity networks," *Human Brain Mapping*, vol. 36, no. 12, pp. 4880-4896, 2016.

[6] L. Zwaigenbaum *et al*., "Early Identification of Autism Spectrum Disorder: Recommendations for Practice and Research," *Pediatrics*, vol. 136, pp. 41-49, 2015.

[7] J. Zhang *et al*., "Functional Brain Network Classification With Compact Representation of SICE Matrices," *IEEE Transactions on Biomedical Engineering*, vol. 62, no. 6, pp. 1623-1634, 2015.

[8] B. Jie *et al*., "Brain connectivity hyper-network for MCI classification," *Springer International Publishing*, vol. 17, no. Pt 2, pp. 724-732, 2014.

[9] C. Y. Wee *et al*., "Temporally Dynamic Resting-State Functional Connectivity Networks for Early MCI Identification," *International Workshop on Machine Learning in Medical Imaging*, 2013.

[10] B. Cai *et al*., "Capturing Dynamic Connectivity From Resting State fMRI Using Time-Varying Graphical Lasso," *IEEE Transactions on Biomedical Engineering*, vol. 66, no. 7, pp. 1852-1862, 2019.

[11] L. Xiao *et al*., "Alternating Diffusion Map Based Fusion of Multimodal Brain Connectivity Networks for IQ Prediction," *IEEE Transactions on Biomedical Engineering*, vol. 66, no. 8, pp. 2140-2151, 2019.

[12] P. Angkoon *et al*., "Resting-State fMRI Functional Connectivity: Big Data Preprocessing Pipelines and Topological Data Analysis," *IEEE TRANSACTIONS ON BIG DATA*, vol. 3, no. 4, pp. 415-428, 2017.

[13] Q. Xie *et al*., "Constructing high-order functional connectivity network based on central moment features for diagnosis of autism spectrum disorder", *PeerJ*, vol. 9, pp. e11692, 2021.

[14] M. Dai *et al*., "Analyzing Dynamical Brain Functional Connectivity as Trajectories on Space of Covariance Matrices," *IEEE Transactions on Medical Imaging*, vol. 39, no. 3, pp. 611-620, 2020.

[15] X. Chen *et al*., "Ensemble Hierarchical High-Order Functional Connectivity Networks for MCI Classification," *Springer International Publishing AG*, pp. 18-25, 2016.

[16] Y. Zhang *et al*., "Hybrid High-order Functional Connectivity Networks Using Resting-state Functional MRI for Mild Cognitive Impairment Diagnosis," *Scientific Reports*, vol. 7, no. 1, pp. 6530-6544, 2017.

[17] C. Y. Wee *et al*., "Sparse temporally dynamic resting-state functional connectivity networks for early MCI identification," *Brain Imaging and Behavior*, vol. 10, no. 2, pp. 342-356, 2016.

[18] Z. Yueying *et al*., "Simultaneous Estimation of Low- and High-Order Functional Connectivity for Identifying Mild Cognitive Impairment," *Frontiers in Neuroinformatics*, vol. 12, 2018.

[19] F. Zhao *et al*., "Diagnosis of Autism Spectrum Disorder Using Central-Moment Features from Low- and High-Order Dynamic Resting-State Functional Connectivity Networks," *Frontiers in Neuroence*, vol. 14, 2020.

[20] X. Chen *et al*., "High-order resting-state functional connectivity network for MCI classification," *Human Brain Mapping*, vol. 37, no. 9, pp. 3282–3296, 2016.

[21] G. Hao *et al*., "Alzheimer Classification Using a Minimum Spanning Tree of High-Order Functional Network on fMRI Dataset," *Frontiers in neuroscience*, vol. 11, 2017.

[22] M. Dai *et al*., "Analyzing Dynamical Brain Functional Connectivity As Trajectories on Space of Covariance Matrices," *IEEE Transactions on Medical Imaging*, vol. 39, no. 3, pp. 611-620, 2019.

[23] F. Zhao *et al*., "Diagnosis of Autism Spectrum Disorders Using Multi-Level High-Order Functional Networks Derived from Resting-State Functional MRI," *Frontiers in Human Neuroence*, vol. 12, 2018.

[24] H. Zhang *et al*., "Topographical Information-Based High-Order Functional Connectivity and Its Application in Abnormality Detection for Mild Cognitive Impairment," *Journal of Alzheimer's Disease*, vol. 54, no.3, pp. 1095–1112, 2016.

[25] B. Jie *et al*., "Hyper-connectivity of functional networks for brain disease diagnosis", *Medical Image Analysis*, vol. 32, pp. 84–100, 2016.

[26] H. Guo *et al*., "Resting-State Brain Functional Hyper-Network Construction Based on Elastic Net and Group Lasso Methods", *Front. Neuroinform*, vol. 12, pp. 25, 2018.

[27] M. Wang *et al*., "Discovering network phenotype between genetic risk factors and disease status via diagnosis-aligned multi-modality regression method in Alzheimer's disease", *Bioinformatics*, vol. 35, no. 11, pp. 1948-1957, 2019.

[28] L. Xiao *et al*., "Multi-Hypergraph Learning-Based Brain Functional Connectivity Analysis in fMRI Data", *IEEE Transactions on Medical Imaging*, vol. 39, no. 5, pp. 1746-1758, 2020.

[29] A. Y. Di Martino *et al*., "The autism brain imaging data exchange: Towards a large-scale evaluation of the intrinsic brain architecture in autism," *Mol Psychiatry*, vol. 19, no. 6, pp. 659–667, 2014.

[30] H. Y. Lin *et al*., "Altered Resting-State Frontoparietal Control Network in Children with Attention-Deficit/Hyperactivity Disorder," *J Int Neuropsychol Soc*, vol. 21, no. 4, pp. 271-284, 2015.

[31] S. Ray *et al*., "Altered Functional Connectivity Strength in Abstinent Chronic Cocaine Smokers Compared to Healthy Controls," *Brain Connectivity*, vol. 5, no. 8, pp. 476–486, 2015.

[32] T. D. Satterthwaite *et al*., "An improved framework for confound regression and filtering for control of motion artifact in the preprocessing of resting-state functional connectivity data," *NeuroImage*, vol. 64, no. 1, pp. 240–256, 2013.

[33] C.-G. Yan *et al*., "A comprehensive assessment of regional variation in the impact of head micromovements on functional connectomics," *NeuroImage*, vol. 76. no. 1, pp. 183–201, 2013.

[34] D. Cordes *et al*., "Frequencies contributing to functional connectivity in the cerebral cortex in "resting-state" data," *American Journal of Neuroradiology*, vol. 22, no. 7, pp. 1326-1333, 2001.

[35] A. Sophie *et al*., "Fractal connectivity of long-memory networks," *Physical review E*, vol. 77, no. 3, pp. 036104, 2008.

[36] D. Tomasi and N. D. Volkow, "Functional connectivity density mapping," *Proceedings of the National Academy of Sciences*, vol. 107, no. 21, pp. 9885-9890, 2010.

[37] 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.

[38] H. J. Ward, "Hierarchical Grouping to Optimize an Objective Function," *Publications of the American Statal Association*, vol. 58, no. 301, pp. 236-244, 1963.

[39] R. Tibshirani, "Regression shrinkage and selection via the lasso: a retrospective," *Journal of the Royal Statal Society*, vol. 58, 1996.

[40] C. Cortes and V. Vapnik, "Support-vector networks," *Machine Learning*, vol. 20, no. 3, pp. 273-297, 1995.

[41] G. Goelman *et al*., "Rat12_bData from: Maximizing negative correlations in resting-state functional connectivity MRI by time-lag," *Plos One*, vol. 9, no. 11, p. e111554, 2014.

[42] H. Yu *et al*., "A deep residual computation model for heterogeneous data learning in smart Internet of Things", *Applied Soft Computing*, vol. 107. no. 5, pp. 107361, 2021.

[43] H. Yu *et al*., "Convolutional neural networks for medical image analysis: State-of-the-art, comparisons, improvement and perspectives", *Neurocomputing*, vol. 444, no. 9, pp. 92-110, 2021.

[44] E. Lauren *et al*., "The role of mirroring and mentalizing networks in mediating action intentions in autism," *Molecular autism*, vol. 5, no. 1, pp. 431-434, 2014.

[45] T. J. Perkins *et al*., "Increased premotor cortex activation in high functioning autism during action observation," *Journal of clinical neuroscience*, vol. 22, no. 4, pp. 664–669, 2015.

[46] S. Isabelle *et al*., "Autistic fluid intelligence: Increased reliance on visual functional connectivity with diminished modulation of coupling by task difficulty," *NeuroImage Clinical*, vol. 9, pp. 467–478, 2015.

[47] S. J. Broyd *et al*., "Default-mode brain dysfunction in mental disorders: a systematic review," *Neuroence & Biobehavioral Reviews*, vol. 33, no. 3, pp. 279-296, 2009.

[48] B. N. N. Dosenbach *et al*., "Prediction of Individual Brain Maturity Using fMRI," *Science*, vol. 329, no. 5997, pp. 1358-1361, 2010.

[49] A.-H. J. R, S. Jonathan, and S. R. Nathan, "The default network and self-generated thought: component processes, dynamic control, and clinical relevance," *Annals of the New York Academy of Sciences*, vol. 1316, no. 1, pp. 29–52, 2014.

[50] H.-I. Suk *et al*., "State-space model with deep learning for functional dynamics estimation in resting-state fMRI," *NeuroImage*, vol. 129, pp. 292-307, 2016.

[51] Z. Wang, *et al*., "Changes in the brain intrinsic organization in both on-task state and post-task resting state," *NeuroImage*, vol. 62, no. 1, pp. 394-407, 2012.

[52] D. P. Kennedy *et al*., "Failing to deactivate: Resting functional abnormalities in autism," *Proceedings of the National Academy of Sciences*, vol. 103, no. 21, pp. 8275-8280, 2006.

[53] S. D. Washington *et al*., "Dysmaturation of the default mode network in autism," *Human Brain Mapping*, vol. 35, no. 4, pp. 1284–1296, 2014.

[54] D. L. Murdaugh *et al*., "Differential Deactivation during Mentalizing and Classification of Autism Based on Default Mode Network Connectivity," *PLOS ONE*, vol. 7, no. 11, pp. e50064, 2012.

[55] M. Assaf *et al*., "Abnormal functional connectivity of default mode subnetworks in autism spectrum disorder patients," *Neuroimage*, vol. 53, no. 1, pp. 247-256, 2010.

[56] D. R. Hampson and G. J. Blatt, "Autism spectrum disorders and neuropathology of the cerebellum," *Frontiers in neuroscience*, vol. 9, 2015.

[57] Kajsa. M. Igelström, T. W. Webb, and M. S. A. Graziano, "Functional Connectivity Between the Temporoparietal Cortex and Cerebellum in Autism Spectrum Disorder," *Cerebral cortex*, vol. 27, no. 4, 2017.

[58] A. J. Khan *et al*., "Cerebro-cerebellar Resting-State Functional Connectivity in Children and Adolescents with Autism Spectrum Disorder," *Biological psychiatry*, vol. 78, no. 9, pp. 625-634, 2015.

[59] S. S.-H. Wang *et al*., "The Cerebellum, Sensitive Periods, and Autism," *Neuron*, vol. 83, no. 3, 2014.

[60] M. Aviv *et al*., "Cluster analysis of resting-state fMRI time series," *NeuroImage*, vol. 45, no. 4, pp. 1117-1125, 2009.

[61] X. N. Zuo *et al*., "An open science resource for establishing reliability and reproducibility in functional connectomics", *Scientific Data*, vol. 1, pp. 140049, 2014.

[62] H. Zhang *et al*., "Test-retest assessment of independent component analysisderived resting-state functional connectivity based on functional near-infrared spectroscopy", *Neuroimage*, vol. 55, no. 2, pp. 607–615, 2011.

[63] H. Zhang *et al*., "Is resting-state functional connectivity revealed by functional near-infrared spectroscopy test-retest reliable?", *Journal of Biomedical Optics*, vol. 16, no. 6, pp. 067008, 2011.

[64] C. G. Yan and Y. F. Zang, "DPARSF: a MATLAB toolbox for "Pipeline" data analysis of resting-state fMRI", *Front. Syst. Neurosci*, vol. 4, pp. 13, 2010.

[65] R. Yu *et al*., "Connectivity strength-weighted sparse group representation-based brain network construction for MCI classification", *Human brain mapping*, vol. 38, no. 5, pp. 2370-2383, 2017.

[66] P. E. Shrout and J. L. Fleiss, "Intraclass correlations: uses in assessing rater reliability", *Psychological bulletin*, vol. 86 no. 2, pp. 420-428, 1979.

[67] J. R. Landis and G. G. Koch, "The measurement of observer agreement for categorical data", *Biometrics*, vol. 33, no. 1, pp. 159–174, 1977.

[68] B. Chen *et al*., "Individual variability and test-retest reliability revealed by ten repeated resting-state brain scans over one month", *PLoS ONE*, vol. 10 no. 12, pp. e0144963.

[69] H. Zhang *et al*., "Test-Retest Reliability of "High-Order" Functional Connectivity in Young Healthy Adults", *Frontiers in Neuroscience*, vol. 11, pp. 439, 2017.