2024 IEEE International Conference on Bioinformatics and Biomedicine (BIBM) | 979-8-3503-8622-6/24/$31.00 ©2024 IEEE | DOI: 10.1109/BIBM62325.2024.10821849

Integrating Low-order and High-order Functional Connectivity for Meta-stable State Transition based Brain Disorder Identification

Shengbing Pei[1], Yan Wang[1], Fan He[1], Zhao Lv[1], Chao Zhang[1,*], Jihong Guan[2,*]
[1] School of Computer Science and Technology, Anhui University, Anhui, China
[2] Department of Computer Science and Technology, Tongji University, Shanghai, China
shengbingpei@ahu.edu.cn, 14042@ahu.edu.cn, jhguan@tongji.edu.cn

***Abstract*—Dynamic functional connectivity network (FCN) can effectively mine meta-stable state transition within the period of data acquisition time, which is related to neurological diseases. However, conventional FCN directly describes the correlation between two brain regions in a meta-stable state, it is low-order FCN. In fact, the connection between two brain regions within the meta-stable state also has a changing pattern, which can reveal the functional consistency between two connections over time, we denote the changing pattern between two regions as high-order FCN. Here, we propose an end-to-end method that integrates low-order and high-order dynamic FCNs for better brain disorder identification. First, a sliding window operation is adopted to capture meta-stable states. Then, a matrix variate normal distribution based approach is employed to construct the low-order and high-order FCNs for each meta-stable state. Finally, a two-stage Transformer is designed to extract meta-stable state transition feature for classification. Experimental results on the ADHD-200 and ABIDE datasets indicate that: 1) our proposed method integrate multi-level connectivity information of dynamic brain functions, thereby effectively improving the identification of brain disorders; 2) the proposed two-stage Transformer is more effective in feature extraction than the well-known CNN-LSTM architecture; 3) high-order FCN help locate biomarkers that low-order FCN cannot be determine, including brain regions as well as functional connections between brain regions, which contribute significantly to the diagnosis of brain disorders.**

***Index Terms*—Functional magnetic resonance imaging, state transition, functional connectivity network, two-stage Transformer, brain disorder identification**

I. Introduction

Neurological diseases, such as autism spectrum disorder (ASD) [1], attention deficit hyperactivity disorder (ADHD) [2] and Alzheimer's disease, have become a significant threat to life, and early diagnosis is urgently required. Neuroimaging has developed into an important tool to aid in diagnosis, among which functional magnetic resonance imaging (fMRI) is one of the most widely used techniques due to its non-invasiveness and balanced spatiotemporal resolution [3].

To better apply the whole brain information, functional connectivity network (FCN) [4] is constructed based on fMRI data. Specifically, the brain is divided into multiple regions based on a brain function template, e.g., Anatomical Automatic Labeling (AAL) template, each brain region is called a region of interest (ROI) and its signal is obtained by averaging the voxel-wise time series within it. Then, a FCN can be constructed through calculating the functional connectivity between ROIs. FCN-based identification methods focus on mining the static information within the FCN. For instance, Wang et al. [5] proposed a distribution-guided network thresholding method for FCN analysis, which can determine the connectivity between two ROIs based on an adaptive threshold.

The FCN-based method assumes that functional connectivity keeps stationary throughout the entire fMRI scanning period. However, existing studies show that the functional connectivity pattern is also dynamically changing [6], that is, there is a meta-stable state transition within the period of acquisition time, which is related to neurological diseases. To explore the dynamic property of functional connectivity pattern, dynamic FCN (D-FCN) is studied, that is, firstly, divide the entire acquired data into multiple consecutive time windows, then construct FCNs in each window, and finally study the dynamic changes of FCNs. For instance, Xing et al. [7] employed the sliding window operation to construct dynamic graphs, and proposed a spectral graph convolution-based long short term memory network to identify mild cognitive impairment.

Moreover, the functional connectivity also has multi-level information, i.e, there exist low-order and high-order FCNs.

The low-order FCN based methods focus on the interaction between pairs of ROIs, which can record the complex network characteristics of brain function. Zhang et al. [8] construct a local-to-global graph neural network (GNN) to analyze and classify brain diseases through functional connectivity between ROIs, the local ROI-GNN learns embeddings in localized regions of the brain and identifies biomarkers, the global Subject-GNN learns the relationship between subjects with the embedded and non-imaging information generated by the local ROI-GNN.

The general high-order FCN based methods focus on the coherence of cooperation patterns between brain region and other brain regions, which can record the high-order connection properties of brain function. Xiao et al. [9] reveal high-order relationships among multiple ROIs by constructing a hypergraph in which each edge can connect any number of ROIs, specifically, they first generate hyperedges with sparse representation, then adaptively obtain hyperedge weights with

* Corresponding authors

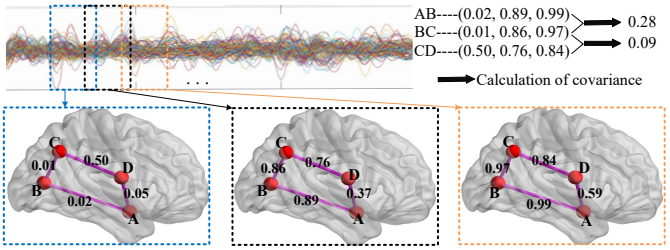

Fig. 1. Illustration of high-order functional connectivity.

hypergraph learning, and finally define a hypergraph similarity matrix to represent the FCN that reveals the high-order interactions among brain regions.

However, the general high-order FCNs focus on collaboration in space, ignoring collaboration patterns in the time dimension. As shown in Fig. 1, within a collection period, the connection values of ROI A and ROI B experience changes of 0.02, 0.89, and 0.99, the connection values of ROI B and ROI C experience changes of 0.01, 0.86 and 0.97, the connection values of ROI C and ROI D experience changes of 0.50, 0.76 and 0.84, this means the connection has a changing pattern, which can reveal the functional consistency between two connections over time, and in contrast, the connections of AB and BC have higher consistency (covariance is 0.28). Here, we introduce a metric between a pair of ROIs (temporal high-order functional connectivity), the product of the metrics of two pairs of connections can reflect their coherence in the time dimension. By our understanding, the spatial high-order information of FCN is included in the low-order FCN and can be extracted by effective feature extraction method. But the temporal high-order information of FCN needs to be established during sample construction, which is an additional information besides low-order FCN, we denote it as high-order FCN in this work.

Here, to better utilize multi-level information related to dynamic functional connectivity (meta-stable state transition), we propose a unified identification model that integrates the dynamic properties of low-order and high-order functional connectivity in the brain. The main contributions include:

- An end-to-end identification model based on fMRI data is proposed, which utilizes multi-level D-FCNs to effectively promote the recognition of brain disorders.
- A two-stage Transformer architecture is designed to extract features of high-order and low-order D-FCNs for fusion analysis.
- Key brain regions and functional connectivities are identified based on high-order D-FCN, which are not realized by the conventional low-order FCN analysis.

II. Methods

Fig. 2 illustrates the framework of our method, which includes the sliding window operation, construction of low-order and high-order FCNs, a two-stage Transformer, and classification with two fusion strategies.

A. Construction of Meta-stable States

The meta-stable state transition pattern based on D-FCN is related to brain disorder. To capture the meta-stable state within the acquired fMRI data, sliding window is a commonly used processing approach. As is shown in Fig. 2 (A), by setting fixed window size l and step size s, the time series data X of a subject is divided into T subsequences (W_1, W_2, $\cdots$, W_T) of the same length. The design of the sliding window is particularly important for the extraction of meta-stable states, according to [6], the combination of $l = 50$ and $s = 40$ can effectively achieves the acquisition of state transition pattern. Here, we follow this setting.

B. Low-order and High-order FCNs within A Meta-stable State

Each time window W_t is used to capture meta-stable state, here we mine its low-order and high-order FCNs. Inspired by [10], the low-order and high-order FCN can be obtained based on matrix variate normal distribution, as compared to directly calculating Pearson correlation, this starts from the distribution of data, has better anti-interference ability, and the mathematical model ensures the interpretability of high-order relationship derivation.

Specifically, applying sliding window operation to W_t again, window size is 10 and step size is 2, q microsequences of (W_{t1}, W_{t2}, $\cdots$, W_{tq}) are generated. For the o-th microsequence W_{to}, the functional connectivity ρ_{ij}^o of the i-th ROI and j-th ROI is calculated by

$$\rho_{ij}^o = corr\left(x_i^o, x_j^o\right) \tag{1}$$

where x_i^o and x_j^o are the time series of the i-th ROI and j-th ROI, respectively. $corr(\cdot,\cdot)$ is a function that calculates the correlation of two signals, typically using Pearson correlation coefficient. As a result, q FCNs are obtained for W_t.

Assume that the FCN M^t of microsequence in W_t follows a normal distribution, that is

$$M^{to} = \left(\rho_{ij}^{to}\right)_{1\le i,j\le n} \tag{2}$$

$$M^t \sim \mathcal{MN}_{n\times n}\left(E^t, U^t, V^t\right) \tag{3}$$

where M^{to} is the FCN corresponding to the o-th microsequence in the t-th window, $\mathcal{MN}(\cdot)$ is matrix normal distribution, n is the number of ROIs, $E^t \in \mathbb{R}^{n\times n}$ is the expectation of M^t, $U^t \in \mathbb{R}^{n\times n}$ and $V^t \in \mathbb{R}^{n\times n}$ are the corresponding covariance matrices. In the context of matrix normal distribution, the covariance matrix serves as a crucial component, encapsulating the linear relationships among different elements of the random matrix, particularly their correlations. In this study, we employ the covariance matrix to construct a high-order FCN that effectively represents the intricate, global interactions within the network, thereby capturing complex high-order functional relationships.

Since the q FCNs are samples of the distribution M^t, maximum likelihood estimation (MLE) can be adopted to calculate E^t, U^t, V^t. Furthermore, since matrix M^t representing FCN is symmetric about the diagonal, we assume that M^t has the

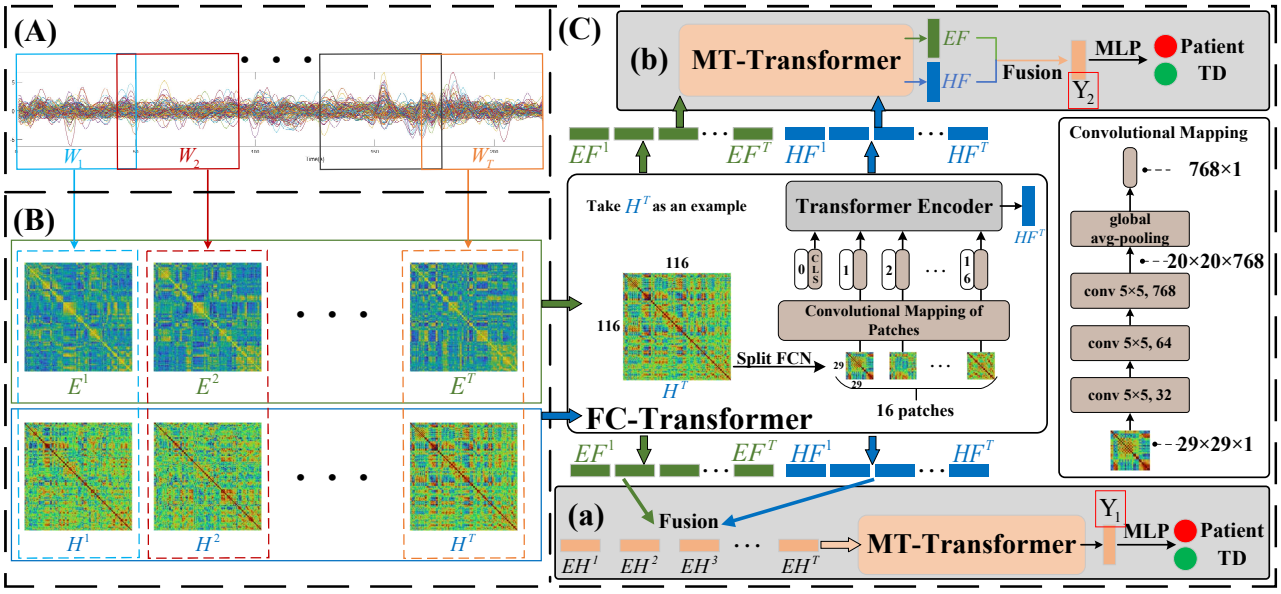

Fig. 2. Framework of the proposed model, which contains three modules. (A) Sliding window is to study the meta-stable state transition. (B) Construction of low-order and high-order FCNs is to mine the multi-level representations of meta-stable state. (C) Two-stage Transformer and classification are to extract the multi-level dynamic information for brain disorder identification, where (a) and (b) are two kinds of fusion strategies.

form of Kronecker product decomposition. Then the column covariance matrix is equal to the row covariance matrix, i.e. $U^t = V^t$, we can further deduce:

$$M^t \sim \mathcal{MN}_{n\times n}\left(E^t, H^t, H^t\right) \tag{4}$$

where positive semi-definite matrix $H^t = U^t = V^t \in \mathbb{R}^{n\times n}$ is the covariance matrix. With the assumption, the number of parameters needed to be estimated reduces from $n^2 \times \left(n^2-1\right)/2$ to $n \times (n-1)/2$. This is also why matrix normal distribution and multivariate normal distribution can be equivalent, but we choose to construct a complex matrix normal distribution.

The probability density function of matrix M^t following normal distribution is as follows:

$$f(M^t) = (2\pi)^{-n^2/2}|H^t|^{-n}\cdot \exp\left(-\frac{1}{2}\mathrm{tr}\left[(H^t)^{-1}\left(M^t-E^t\right)^\top (H^t)^{-1}\left(M^t-E^t\right)\right]\right) \tag{5}$$

where $tr(\cdot)$ is the trace operator. E^t (containing the interaction information between pairs of ROIs) is the low-order FCN and H^t (containing the interaction information between different functional connectivity patterns) is the high-order FCN. By utilizing $(M^{t1}, M^{t2}, \cdots, M^{tq})$ and employing MLE, E^t and H^t can be iteratively calculated as follows:

$$E^t := \frac{1}{q}\sum_{i=1}^{q}(M^{ti}) \tag{6}$$

$$H^t := \frac{1}{nq}\sum_{i=1}^{q}(M^{ti}-E^t)(H^t)^{-1}(M^{ti}-E^t)^T \tag{7}$$

Furthermore, by constructing the low-order and high-order FCNs for all the time windows $(W_1, W_2, \cdots, W_T)$ of a subject, we obtain $(E^1, E^2, \cdots, E^T)$ and $(H^1, H^2, \cdots, H^T)$.

C. Two-stage Transformer and Fusion Strategy

To extract mode information of meta-stable state transition within $(E^1, E^2, \cdots, E^T)$ and $(H^1, H^2, \cdots, H^T)$, we design a two-stage Transformer architecture, including a FC-Transformer and a MT-Transformer. Meanwhile, we design two fusion strategies to integrate low-order and high-order D-FCNs for classification.

FC-Transformer is to extract multi-level functional connectivity features of meta-stable states. It is able to capture long-range dependencies and global information in multi-level FCNs. Specifically, both E^i and H^i are split into 16 patches with size of 29×29. Each patch undergoes sequential convolution by three layers of convolution kernels, all sized 5×5, with the number of kernels being 32, 64, and 768 respectively. After convolutional processing, each patch is encoded into a 3D tensor of size $20 \times 20 \times 768$, which is then reduced to a 768-dimensional representation via global average pooling. FC-Transformer encoder consists of 8 blocks, each including multi-head attention, a multi-layer perceptron, layer normalization, and residual connections. The class token serves as the final output of the FC-Transformer, and denoted as EF^i (for E^i) and HF^i (for H^i). This encodes each low-order and high-order FCN into a 768-dimensional functional connectivity feature.

MT-Transformer extracts state transition patterns from meta-stable state representations. And the operations of MT-Transformer are the same as those of FC-Transformer. The self-attention mechanism can capture contextual information and identifies specific relationships among meta-stable states.

Fig. 2 (a) shows a fusion strategy, which we denote as Strategy 1, where high-order and low-order FCN features are fused within a meta-stable state before being input to

TABLE I
DIVISION OF PKU, NYU AND NI IN ADHD-200 DATASET.

	Training set		Testing set	
	TD	ADHD	TD	ADHD
PKU	116	78	24	27
NYU	99	123	12	29
NI	23	25	14	11

TABLE II
SUBJECT DEMOGRAPHICS OF 14 SITES IN ABIDE DATASET.

Site	ASD	TD	Site	ASD	TD
KKI	20	28	SDSU	10	19
Leuven	29	34	Stanford	19	20
MaxMun	24	28	Trinity	22	24
NYU	75	100	UCLA	54	44
OHSU	12	14	UM	66	74
OLIN	19	15	USM	46	25
PITT	20	19	Yale	28	28

MT-Transformer. Firstly, the low-order functional connectivity features (EF^1, EF^2, $\cdots$, EF^T) and the high-order functional connectivity features (HF^1, HF^2, $\cdots$, HF^T) are integrated to obtain the fused functional connectivity features (EH^1, EH^2, $\cdots$, EH^T), EH^i is calculated as follows:

$$EH^i = Fusion(EF^i, HF^i) \tag{8}$$

where $Fusion(\cdot,\cdot)$ represents the averaging operation. Then, MT-Transformer is used to extract temporal feature from (EH^1, EH^2, $\cdots$, EH^T) to obtain multi-level meta-stable state transition information Y_1. Finally, Y_1 undergoes a 128-32-2 multi-layer perceptron for brain disorder classification.

Fig. 2 (b) shows another fusion strategy, which we denote as Strategy 2, where high-order and low-order D-FCN features are fused after passing through MT-Transformer. Firstly, MT-Transformer is adopted to extract the low-order state transition feature EF from low-order functional connectivity features ($EF^1, EF^2, \cdots, EF^T$), and high-order state transition feature HF from high-order functional connectivity features (HF^1, HF^2, $\cdots$, HF^T), respectively. Then, the mode features of meta-stable state transition within EF and HF are linearly fused to obtain Y_2, which is calculated as follows:

$$Y_2 = Fusion(EF, HF) \tag{9}$$

Finally Y_2 undergoes a 128-32-2 multi-layer perceptron for brain disorder classification.

III. DATASETS

We validate our proposed method on two brain disease public datasets, namely the ADHD-200 dataset (http://fcon_1000.projects.nitrc.org/indi/adhd200/) and the ABIDE dataset (http://fcon_1000.projects.nitrc.org/indi/abide). They are both multi-site datasets. The ADHD dataset is released in the 2011 ADHD-200 Global Competition, the dataset division by the organization is shown in Table I. TD denotes typically developing control. The ABIDE dataset has aggregated data collected from laboratories around the world. The subject demographic of multiple sites is shown in Table II.

TABLE III
THE ACCURACY (%) COMPARISON OF EXISTING METHODS AT PKU, NYU AND NI SITES ON ADHD-200 DATASET. "STRATEGY 1" AND "STRATEGY 2" REPRESENT THE PROPOSED METHODS USING TWO KINDS OF FUSION STRATEGIES, RESPECTIVELY.

Site	STAAE [11]	DeepFMRI [12]	HLGSNet [13]	Riaz [14]	EM-MI [15]	**Strategy 1**	**Strategy 2**
PKU	65.20	62.70	-	64.70	70.60	74.51	**76.47**
NYU	59.50	73.10	68.29	60.90	63.40	**75.61**	**75.61**
NI	61.00	67.90	68.00	44.00	-	**72.00**	**72.00**

TABLE IV
THE ACCURACY (%) COMPARISON OF EXISTING METHODS ON ABIDE DATASET. "STRATEGY 1" AND "STRATEGY 2" REPRESENT THE PROPOSED METHODS USING TWO KINDS OF FUSION STRATEGIES, RESPECTIVELY.

Site	Zhang [16]	SAENet [17]	Diagnet [18]	Heinsfeld [19]	**Strategy 1**	**Strategy 2**
KKI	61.24	72.60	69.50	58.20	**72.67**	71.43
Leuven	67.13	64.60	61.30	51.80	72.95	**73.62**
MaxMun	40.44	47.50	48.60	54.30	69.46	**70.12**
NYU	71.45	72.00	68.00	64.50	**74.29**	73.64
OHSU	74.06	72.00	**82.00**	74.00	79.99	78.91
OLIN	72.85	66.60	65.10	44.00	76.19	**76.86**
PITT	63.36	73.10	67.80	59.80	**75.15**	74.08
SDSU	63.42	64.20	63.00	63.60	**74.64**	**74.64**
Stanford	65.92	53.20	64.20	48.50	66.79	**68.44**
Trinity	55.46	57.50	54.10	61.00	73.33	**75.12**
UCLA	74.84	68.30	73.20	57.70	**76.42**	75.96
UM	66.60	67.80	63.80	57.60	**77.86**	74.76
USM	72.00	70.00	68.20	62.00	77.52	**78.16**
Yale	68.14	66.00	63.60	56.10	73.03	**74.43**

IV. RESULTS

A. Performance on ADHD-200 and ABIDE Datasets

Table III and Table IV show the accuracy comparison with existing methods on ADHD-200 dataset and ABIDE dataset, respectively. The "Strategy 1" and "Strategy 2" are our proposed methods, which represent the two fusion strategies in Section II-C.

Compare "Strategy 1" and "Strategy 2" with existing methods, respectively. As can be seen, for the ADHD-200 dataset, both "Strategy 1" and "Strategy 2" show good improvement; for ABIDE dataset, except for OHSU site, both "Strategy 1" and "Strategy 2" show improvement. And on OHSU site, "Strategy 1" is 2.01% less than the highest value, "Strategy 2" is 2.21% less than the highest value, by our understanding, this may due to the size of dataset is relatively small, which makes the performance more susceptible.

Overall, the results indicate that the proposed methods of integrating multi-level dynamic functional connectivity can help extract more discriminative features, and therefore improve the recognition ability of brain disorders.

Furthermore, we compare "Strategy 1" and "Strategy 2" to explore the impact of the fusion strategy. For the ADHD-200 dataset, the identification accuracy of "Strategy 2" is greater than or equal to "Strategy 1". For ABIDE dataset, "Strategy 1" and "Strategy 2" have different performances on different sites. The sites where "Strategy 1" has higher accuracy than "Strategy 2" are KKI, NYU, OHSU, PITT, UCLA, and UM. While the sites where "Strategy 2" has higher accuracy than

TABLE V
THE ACCURACY (%) COMPARISON OF ABLATION EXPERIMENTS AT PKU, NYU AND NI SITES ON ADHD-200 DATASET.

Site	Low-order	High-order	**Strategy 2**
PKU	70.59	72.55	**76.47**
NYU	70.73	73.17	**75.61**
NI	68.00	68.00	**72.00**

TABLE VI
THE ACCURACY (%) COMPARISON OF ABLATION EXPERIMENTS ON ABIDE DATASET.

Site	Low-order	High-order	**Strategy 2**
KKI	66.45	68.67	**71.43**
Leuven	69.74	71.28	**73.62**
MaxMun	67.09	67.64	**70.12**
NYU	69.14	72.57	**73.64**
OHSU	72.67	77.33	**78.91**
OLIN	70.47	73.81	**76.86**
PITT	71.67	**76.97**	74.08
SDSU	63.93	72.14	**74.64**
Stanford	61.79	64.27	**68.44**
Trinity	69.33	71.33	**75.12**
UCLA	67.32	69.32	**75.96**
UM	70.72	71.43	**74.76**
USM	73.14	76.09	**78.16**
Yale	67.73	71.21	**74.43**

"Strategy 1" are Leuven, MaxMun, OLIN, Stanford, Trinity, USM and Yale. In conclusion, "Strategy 2" has the highest value on more sites than "Strategy 1", which means that it is helpful to observe the dynamic changes of low-order functional connectivity and high-order functional connectivity separately within the framework of meta-stable state transition. Here we select "Strategy 2" as our best model and perform the following analysis.

B. Effect of Fusing Low-order and High-order D-FCNs

Here we conduct ablation experiments on "Strategy 2" model to study effect of fusing low-order and high-order D-FCNs. Tables V and VI show the accuracy comparison of ablation experiments on ADHD-200 dataset and ABIDE dataset, respectively. "Low-order" and "High-order" are the proposed method that only use low-order D-FCNs and high-order D-FCNs, respectively.

Comparing "Strategy 2" with "Low-order" and "High-order", we can see that except for PITT site on ABIDE dataset, "Strategy 2" outperforms "Low-order" and "High-order" on both ADHD-200 and ABIDE datasets. This indicates that multi-level D-FCNs contain complementary brain functional features, and fusion helps improve the classification ability of features.

Meanwhile, comparing "Low-order" and "High-order" in Tables V and VI, we notice that "High-order" outperforms "Low-order" on both datasets, which validate the recognition ability of high-order D-FCN. In fact, high-order D-FCN describes the coordinated changes of functional connectivity patterns, which can more reliably simulate brain functional patterns, thus enhancing disease recognition capabilities.

TABLE VII
THE ACCURACY (%) COMPARISON OF TWO-STAGE TRANSFORMER AND CNN-LSTM ON ADHD-200 DATASET.

Operate	Site		
	PKU	NYU	NI
CNN-LSTM	72.55	70.73	68.00
Two-stage Transformer ("Strategy 2")	76.47	75.61	72.00

C. Two-stage Transformer vs CNN-LSTM

The structure of convolutional neural network (CNN) and long short term mermory (LSTM) network is commonly used to study the dynamic of brain function. For examples, Mao et al. [20] utilize 3D CNN to extract features from individual fMRI frames, and subsequently input the extracted features sequentially into LSTM for feature fusion.

Here, we compare the proposed two-stage Transformer ("Strategy 2") with CNN-LSTM structure to investigate its feature extraction capability. In the CNN-LSTM structure, CNN is employed to replace FC-Transformer and LSTM is adopted to replace MT-Transformer. The CNN architecture consists of two convolutional layers, the first is with 16 kernels of size 116×1 and the second is with 32 kernels of size 1×116. These kernels are utilized to extract both local ROI features and global brain functional connectivity features for each FCN. The LSTM is employed to extract meta-stable state transition features. The comparative experiments are conducted on the ADHD-200 dataset, the results are presented in Table VII. As can be seen, the two-stage Transformer significantly outperforms the CNN-LSTM structure across all the three sites. This is due to the Transformer's ability to model global relationships in FCNs and capture complex dependencies. By fully utilizing global context information, the Transformer effectively captures multi-level functional connectivity features and meta-stable state transition features for identifying brain disorders.

D. Effect of High-order FCN in Locating Biomarkers

Here, we delve into the crucial ROIs that significantly contribute to the diagnosis of brain disorders, specifically analyzing both low-order and high-order FCNs. For all subjects, including patients and healthy controls, we commence by computing the local clustering coefficient for each node within an FCN, effectively transforming the FCN into a 116-dimensional vector. Subsequently, we construct an observation matrix, $X_{nT\times116}$, along with its corresponding label vector, $Y_{nT\times1}$, where n represents the total number of subjects and T denotes the count of time windows. Utilizing these matrices, we establish an L1-norm regularized least squares regression model. The regression coefficients, encapsulated in $S_{116\times1}$, offer insights into the significance of each ROI in the identification process.

We take the PKU site of the ADHD-200 dataset as an example. Fig. 3 (A) shows the key ROIs identified separately by low-order and high-order FCNs. Notably, the low-order FCNs locate 9 ROIs, while the high-order FCNs reveal 16

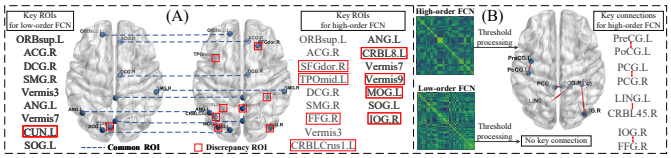

Fig. 3. Comparision of low-order FCN and high-order FCN. (A) Key brain regions in low-order FCN and high-order FCN. (B) Additional key connections in high-order FCN.

ROIs. Remarkably, 8 ROIs were common to both approaches, highlighting their shared significance. However, specific ROIs such as SFGdor.R, MOG.L, IOG.R, FFG.R, TPOmid.L, CRBLCrus1.L, CRBL8.L, and Vermis9 emerged exclusively in the high-order FCNs, underscoring the supplementary information captured by our approach. This observation not only underscores the complementarity of information between the two FCN types but also demonstrates the added value of high-order FCNs in extracting additional diagnostic insights.

Meanwhile, the mean low-order FCN and mean high-order FCN for all subjects are separately calculated, which is shown in Fig. 3 (B). The values in the FCN reflect the strength of connections. As can be seen, the FCN matrices of mean low-order FCN and mean high-order FCN show different pattern, indicating the different perspectives on high-order and low-order connections. Moreover, we set a threshold of 0.8 to capture the key connections with absolute values ≥ 0.8, after processing, the low-order FCN has 0 key connection, while the high-order FCN retains 4 connections, which shows some consistency with the key regions in Fig. 3 (A). This also demonstrates that the high-order FCN reveals additional information as compared to the low-order FCN.

V. Conclusion

In this work, we mine the meta-stable state transition feature of the brain from the perspective of functional connectivity to assist in disease diagnosis, and propose an end-to-end identification model that integrates low-order and high-order D-FCN features. The experimental results demonstrate that learning meta-stable state transition representation with the two-stage Transformer based on integration of multi-level D-FCN is effective in promoting disease identification.

VI. Acknowledgment

This work was supported in part by the Natural Science Research Project of Anhui Educational Committee under Grant 2023AH050081; in part by the Collaborative Innovation Project of Key Laboratory of Philosophy and Social Sciences in Anhui Province under Grant HZ2302; in part by the Scientific Research Project of Colleges and Universities in Anhui Province under Grant 2024AH040115.

References

[1] W. Yin, L. Li, and F.-X. Wu, "A graph attention neural network for diagnosing ASD with fMRI data," in *2021 IEEE International Conference on Bioinformatics and Biomedicine (BIBM)*. IEEE, 2021, pp. 1131–1136.

[2] Y. Li, Z. Lian, M. Li, Z. Liu, L. Xiao, and Z. Wei, "ELM-based classification of ADHD patients using a novel local feature extraction method," in *2016 IEEE International Conference on Bioinformatics and Biomedicine (BIBM)*. IEEE, 2016, pp. 489–492.

[3] Q. Mai, U. Nakarmi, and M. Huang, "BrainVGAE: end-to-end graph neural networks for noisy fMRI dataset," in *2022 IEEE International Conference on Bioinformatics and Biomedicine (BIBM)*. IEEE, 2022, pp. 3852–3855.

[4] W. Wang, S. Zhang, Z. Wang, X. Luo, P. Luan, A. Hramov, J. Kurths, C. He, and J. Li, "Diagnosis of early mild cognitive impairment based on associated high-order functional connection network generated by multi-modal MRI," *IEEE Transactions on Cognitive and Developmental Systems*, 2023.

[5] Z. Wang, B. Jie, C. Feng, T. Wang, W. Bian, X. Ding, W. Zhou, and M. Liu, "Distribution-Guided Network Thresholding for Functional Connectivity Analysis in fMRI-Based Brain Disorder Identification," *IEEE Journal of Biomedical and Health Informatics*, vol. 26, no. 4, pp. 1602–1613, 2021.

[6] S. Pei, F. He, S. Cao, W. Liang, J. Guan, and Z. Lv, "Learning Meta-Stable State Transition Representation of Brain Function for ADHD Identification," *IEEE Transactions on Instrumentation and Measurement*, vol. 72, pp. 1–13, 2023.

[7] X. Xing, Q. Li, H. Wei, M. Zhang, Y. Zhan, X. S. Zhou, Z. Xue, and F. Shi, "Dynamic spectral graph convolution networks with assistant task training for early MCI diagnosis," in *MICCAI*. Springer, 2019, pp. 639–646.

[8] H. Zhang, R. Song, L. Wang, L. Zhang, D. Wang, C. Wang, and W. Zhang, "Classification of brain disorders in rs-fMRI via local-to-global graph neural networks," *IEEE Transactions on Medical Imaging*, vol. 42, no. 2, pp. 444–455, 2022.

[9] L. Xiao, J. Wang, P. H. Kassani, Y. Zhang, Y. Bai, J. M. Stephen, T. W. Wilson, V. D. Calhoun, and Y.-P. Wang, "Multi-hypergraph learning-based brain functional connectivity analysis in fMRI data," *IEEE Transactions on Medical Imaging*, vol. 39, no. 5, pp. 1746–1758, 2019.

[10] Y. Zhou, L. Qiao, W. Li, L. Zhang, and D. Shen, "Simultaneous estimation of low-and high-order functional connectivity for identifying mild cognitive impairment," *Frontiers in Neuroinformatics*, vol. 12, p. 3, 2018.

[11] Q. Dong, N. Qiang, J. Lv, X. Li, T. Liu, and Q. Li, "Spatiotemporal Attention Autoencoder (STAAE) for ADHD Classification," in *MICCAI*. Springer, 2020, pp. 508–517.

[12] A. Riaz, M. Asad, S. M. R. Al Arif, E. Alonso, D. Dima, P. Corr, and G. Slabaugh, "Deep fMRI: an end-to-end deep network for classification of fMRI data," in *ISBI*. IEEE, 2018, pp. 1419–1422.

[13] R. R. Jha, A. Nigam, A. Bhavsar, G. Jaswal, S. K. Pathak, and R. Kumar, "HLGSNet: Hierarchical and Lightweight Graph Siamese Network with Triplet Loss for fMRI-based Classification of ADHD," in *IJCNN*. IEEE, 2020, pp. 1–7.

[14] A. Riaz, M. Asad, E. Alonso, and G. Slabaugh, "Fusion of fMRI and non-imaging data for ADHD classification," *Computerized Medical Imaging and Graphics*, vol. 65, pp. 115–128, 2018.

[15] C. Dou, S. Zhang, H. Wang, L. Sun, Y. Huang, and W. Yue, "ADHD fMRI short-time analysis method for edge computing based on multi-instance learning," *Journal of Systems Architecture*, vol. 111, p. 101834, 2020.

[16] J. Zhang, F. Feng, T. Han, X. Gong, and F. Duan, "Detection of autism spectrum disorder using fMRI functional connectivity with feature selection and deep learning," *Cognitive Computation*, pp. 1–12, 2022.

[17] F. Almuqhim and F. Saeed, "ASD-SAENet: a sparse autoencoder, and deep-neural network model for detecting autism spectrum disorder (ASD) using fMRI data," *Frontiers in Computational Neuroscience*, vol. 15, p. 654315, 2021.

[18] T. Eslami, V. Mirjalili, A. Fong, A. R. Laird, and F. Saeed, "ASD-DiagNet: a hybrid learning approach for detection of autism spectrum disorder using fMRI data," *Frontiers in Neuroinformatics*, vol. 13, p. 70, 2019.

[19] A. S. Heinsfeld, A. R. Franco, R. C. Craddock, A. Buchweitz, and F. Meneguzzi, "Identification of autism spectrum disorder using deep learning and the ABIDE dataset," *NeuroImage*, vol. 17, pp. 16–23, 2018.

[20] Z. Mao, Y. Su, G. Xu, X. Wang, Y. Huang, W. Yue, L. Sun, and N. Xiong, "Spatio-temporal deep learning method for ADHD fMRI classification," *Information Sciences*, vol. 499, pp. 1–11, 2019.