Randomizing Human Brain Function Representation for Brain Disease Diagnosis

Mengjun Liu, Huifeng Zhang, Mianxin Liu, Dongdong Chen, Zixu Zhuang, Xin Wang, Lichi Zhang, Daihui Peng, and Qian Wang

***Abstract*— Resting-state fMRI (rs-fMRI) is an effective tool for quantifying functional connectivity (FC), which plays a crucial role in exploring various brain diseases. Due to the high dimensionality of fMRI data, FC is typically computed based on the region of interest (ROI), whose parcellation relies on a pre-defined atlas. However, utilizing the brain atlas poses several challenges including (1) subjective selection bias in choosing from various brain atlases, (2) parcellation of each subject's brain with the same atlas yet disregarding individual specificity; (3) lack of interaction between brain region parcellation and downstream ROI-based FC analysis. To address these limitations, we propose a novel randomizing strategy for generating brain function representation to facilitate neural disease diagnosis. Specifically, we randomly sample brain patches, thus avoiding ROI parcellations of the brain atlas. Then, we introduce a new brain function representation framework for the sampled patches. Each patch has its function description by referring to anchor patches, as well as the position description. Furthermore, we design an adaptive-selection-assisted Transformer network to optimize and integrate the function representations of all sampled patches within each brain for neural disease diagnosis. To validate our framework, we conduct extensive evaluations on three datasets, and the experimental results establish the effectiveness and generality of our proposed method, offering a promising avenue for advancing neural disease diagnosis beyond the confines of traditional atlas-based methods. Our code is available at https://github.com/mjliu2020/RandomFR.**

***Index Terms*— Randomizing, function representation, adaptive selection module, Transformer, brain disease diagnosis.**

I. Introduction

BRAIN disorders, such as autism spectrum disease (ASD), depressive disorders, and mild cognitive impairment (MCI), are increasingly prevalent [1]. The impact of ASD, depressive disorders, and MCI extends beyond the affected individuals, imposing significant burdens on families and society at large [2], [3]. While our understanding of the physiological and pathological mechanisms underlying these brain disorders remains incomplete, researches have associated brain diseases with alterations in brain functions [4]. Consequently, it is desired to establish a framework to represent brain functions for the diagnosis of brain disorders.

Resting-state functional magnetic resonance imaging (rs-fMRI) has emerged as a great tool for investigating brain functions, owing to its noninvasive nature and high spatial resolution [5]. Specifically, rs-fMRI enables the detection of blood-oxygen-level-dependent (BOLD) signals in the resting state, which serve as proxies for neural and metabolic activity [6]. The statistical dependency among BOLD signals across distinct brain regions serves as a quantitative measure of functional connectivity (FC), thereby providing a means of assessing and understanding brain functions [7]. Moreover, it has been extensively demonstrated that alterations in FC are closely associated with a myriad of brain disorders [8].

Currently, it is common to perform region-of-interest (ROI) based FC analysis upon fMRI [9], [10]. In brief, one can utilize an atlas and parcellate a subject brain into multiple ROIs, i.e., by registering the atlas with the subject. Subsequently, the FC network per subject is modeled by a graph, where each node indicates a certain ROI and the edge corresponds to FC. By examining the statistical characteristics of the networks, subtle distinctions among populations can be discerned [11]. Furthermore, the integration of sophisticated deep-learning techniques offers the chance to achieve subject-level classification for disease diagnosis [12].

The aforementioned paradigm heavily relies on the utilization of the atlas. It is expected that high-quality ROI parcellation is available when constructing the atlas. In addition, precise registration of the atlas with each subject (and thus parcellation) should be guaranteed. However, it gives rise to three non-trivial issues that warrant careful consideration.

1) The selection of an appropriate atlas for ROI parcellation presents a notable challenge [13], [14]. Extensive efforts have been dedicated to the construction of high-quality brain atlases with a particular emphasis on ROI parcellation, including Automated Anatomical Labeling (AAL) [15], Schaefer [16], Harvard Oxford (HO) [17], etc. However, the utilization of different atlases can introduce variations in the number, size, and spatial

Mengjun Liu, Dongdong Chen, Zixu Zhuang, Xin Wang, and Lichi Zhang are with the School of Biomedical Engineering, Shanghai Jiao Tong University, Shanghai 200030, China. (e-mail: mjliu2020@sjtu.edu.cn).

Huifeng Zhang and Daihui Peng are with Shanghai Mental Health Center, Shanghai Jiao Tong University School of Medicine, Shanghai 200030, China. (e-mail: zhanghuifengsh@126.com).

Mianxin Liu is with Shanghai Artificial Intelligence Laboratory, Shanghai 200232, China. (e-mail: 16483073@life.hkbu.edu.hk).

Qian Wang is with the School of Biomedical Engineering, ShanghaiTech University, Shanghai 201210, China, and with Shanghai Clinical Research and Trial Center, Shanghai 201210, China. (e-mail: wangqian2@shanghaitech.edu.cn).

Mengjun Liu and Huifeng Zhang contribute equally to this manuscript. Daihui Peng and Qian Wang are correspondence authors.

locations of brain regions, consequently resulting in considerable inconsistencies in computational experiments [18].

2) The application of a common atlas for parcellating diverse brains inadvertently neglects individual specificity [19]. Empirically, different subjects may exhibit distinct functional parcellations, resulting in variations in the boundaries of brain regions [20]. The distinctive functional architecture of each individual's brain assumes particular importance, especially in personalized medicine [21].
3) The atlas construction and FC-based classification are two separate steps, which inherently restrict diagnostic performance due to the absence of effective communication between them. Typically, a pre-defined atlas is an image of either a single healthy subject or an average, lacking in incorporating population-specific variations observed in diagnosing diseased brains. Moreover, the absence of a unified optimization objective in training for both the parcellation of brain regions and the ROI-based FC analysis further impedes disease diagnosis.

Numerous studies have recognized and endeavored to address the aforementioned limitations. For instance, multi-atlas approaches have emerged to amalgamate the image features derived from multiple atlases, thus enhancing brain disease diagnosis [12]. Moreover, the proposition of individual atlases has been put forward to facilitate personalized brain disease diagnosis [21], [22]. Additionally, investigations into data-driven brain atlases have been conducted to advance brain disease diagnosis [23]. Despite these endeavors have boosted the diagnostic performance by partially mitigating the limitations associated with atlas-based FC, the challenges at hand have not yet been entirely resolved.

Hence, in this paper, we introduce an innovative strategy to represent all spatial locations in an fMRI and then sample the FC network, to generate brain function representation for neural disease diagnosis. Our proposed approach opens up promising directions as it significantly diverges from traditional atlas-based methods.

1) We propose to sample the brain FC network. That is, we randomly sample a set of patches from each brain. Each patch is signified by not only the BOLD signals inside the patch of rs-fMRI, but also the center position of the patch. By adjusting the patch size and number, we can acquire a set of patches that covers the entire brain and delivers a comprehensive representation of the brain. As the patches are randomly sampled, the above strategy avoids using the ROI parcellation of a brain atlas, and also alleviates concerns about registering atlas with the subject precisely.
2) We construct the representation of these randomly sampled patches through joint *function description* and *position description*. Specifically, we designate a set of anchor patches as reference to construct the function description. Furthermore, the position coordinate of each sampled patch is important and can be properly transformed into the position description. By combining the two descriptions, we design a holistic representation for each sampled patch.
3) We design a Transformer network to integrate the randomly sampled patches for the diagnosis of brain diseases. We further introduce an adaptive selection strategy to identify the patches exhibiting robust discriminative capabilities. The Transformer network can thus effectively model the sampled patches into a global representation of the brain, while reducing the computational cost of attention in the Transformer network and enabling interpretability exploration.

We conduct systematic experiments on three datasets, and demonstrate superior performance over competing methods for the diagnosis of various brain diseases. The rest of this paper is organized as follows. In Section II, we provide a concise overview of the related work. In Section III, we elaborate on the datasets and methods in detail. In Section IV, we present the experimental setup and results. In Section V, we analyze the key modules and parameters of our method. Finally, we conclude this paper in Section VI.

II. Related Work

Constructing precise ROI-based brain function representation necessitates accurate brain parcellation. In this section, we briefly review two primary categories of brain parcellation methods employed in brain disease diagnosis: atlas-based methods and data-driven methods.

A. Atlas-based Methods

The construction of FC based on atlas is the predominant strategy in the field. Atlas-based methods for brain disease analysis can be categorized into three main types: utilizing a single atlas, combining multiple atlases, and constructing individual-specific atlas.

1) Utilizing a Single Atlas: This form of atlas-based FC analysis is reliant on a single atlas. For instance, Li *et al.* [9] proposed an interpretable brain graph neural network for fMRI analysis, utilizing the Desikan-Killiany atlas [17] to parcellate the brain for the diagnosis of ASD. Qin *et al.* [24] employed a graph convolutional network to characterize individuals with major depressive disorder across multiple imaging sites, which parceled the entire brain into 160 ROIs according to Dosenbach's atlas [25]. Huang *et al.* [10] proposed an attention-diffusion-bilinear neural network for brain network analysis, which adopted AAL atlas to parcellate the brain into 90 ROIs. Although different brain atlases can be employed for diagnosing brain diseases, there is no standardized strategy for selecting brain atlas and different brain atlases will lead to inconsistent diagnosis.

2) Combining Multiple Atlases: This form of atlas-based FC analysis is predicated on multiple atlases. For instance, Yao *et al.* [12] proposed a mutual multi-scale triplet graph convolutional network for the classification of brain disorders. This approach employed four atlases to construct multi-scale brain networks, namely AAL atlas, Craddock atlas [26], Brainnetome atlas [27], and Bootstrap Analysis of Stable Clusters atlas [28]. Cui *et al.* [29] combined two atlas-based

FCs, namely personalized functional connectivity (PFC) atlas and AAL atlas, to achieve superior performance compared to single-atlas methods (only AAL or PFC atlas). Although the fusion of multiple atlases has demonstrated improvements in the diagnosis, challenges remain regarding atlas selection and determining the optimal number of brain atlases.

3) Constructing Individual-Specific Atlas: This form of atlas-based FC analysis relies on individual-specific atlas. Zhang *et al.* [22] developed an individualized atlas by implementing a three-step transformation-based method on the standard AAL atlas for improving MCI diagnosis. Cui *et al.* [29] proposed personalized FC based on individual-level brain atlas for MCI identification. Wang *et al.* [30] utilized simple linear iterative clustering to generate individualized brain atlases, which could be used for personalized medicine. Although the construction of individual-specific atlas facilitates personalized brain disease diagnosis, these methods necessitate an initial atlas or involve several intricate iterations to derive the individual-specific atlas.

B. Data-Driven Methods

In addition to the pre-defined brain atlases, several data-driven methods for brain parcellation have been proposed to aid in the diagnosis of brain diseases. For instance, Liu *et al.* [31] proposed a landmark-based deep feature representation approach for brain disease diagnosis. They identified landmarks that had statistically significant differences between AD and NC subjects. Subsequently, a pre-trained landmark detection model was employed to automatically detect landmarks. Lian *et al.* [32] devised a fully convolutional network backbone to automatically localize discriminative brain regions for dementia diagnosis. Abraham *et al.* [23] developed data-driven atlases using four strategies for ASD diagnosis, including K-Means clustering, Ward's clustering, independent component analysis (ICA), and multi-subject dictionary learning (MSDL).

These aforementioned methods have demonstrated that a pre-defined brain atlas is not essential for the diagnosis of brain diseases. However, most of these studies primarily focus on structural images, with limited investigations into rs-fMRI. Additionally, the data-driven brain atlas generation and subsequent ROI-based analysis are often treated as two separate steps, indicating the potential for improvement by exploring joint optimization between these two stages.

III. Materials and Methods

The architecture of the proposed method is depicted in Fig. 1. We first randomly sample brain patches from fMRI, avoiding the use of ROI parcellation of brain atlas. Then, we design a representation framework for the sampled patches, which contains the functional and structural information of the brain regions comprehensively. Subsequently, we introduce an adaptive-selection-assisted Transformer network to optimize and integrate the aforementioned patch representation for brain disease diagnosis. Although the idea of splitting an original image into patches and attaching positional embeddings to them is similar with Vision Transformer (ViT) [33], we design new strategies in patch extraction, patch representation and patch integration to realize brain disease diagnosis without relying on brain atlas.

TABLE I
The Demographic Statistics of Datasets. M: Male, F: Female

Datasets	Subgroup	Number	Gender (M/F)	Age (years)
ABIDE	ASD	73	64/9	14.9±7.1
	NC	98	72/26	15.7±6.2
ADNI	MCI	60	31/29	71.8±7.3
	NC	88	33/55	73.0±6.6
Depression Subtype	Melancholic	46	6/40	25.8±5.6
	Atypical	42	8/34	23.5±4.3

A. Tasks and Data Preparation

We evaluate our method using three distinct tasks pertaining to different brain image cohorts, including the Autism Brain Imaging Data Exchange I (ABIDE) dataset, the Alzheimer's Disease Neuroimaging Initiative (ADNI) dataset, and the private depression subtype dataset. The demographic statistics of these datasets are provided in Table I.

For the ASD diagnosis task, we utilize the publicly available ABIDE dataset. Specifically, we focus on the largest site (NYU) in ABIDE to avoid multi-site confounding. We exclude the subjects diagnosed with Asperger or PDD-NOS [34]. Consequently, our analysis involves rs-fMRI data from a total of 171 subjects, including 73 ASDs and 98 NCs. We select the data preprocessed through Data Processing Assistant for Resting-State fMRI (DPARSF) toolbox [35]. The first 4 volumes are dropped for signal equilibrium, followed by slice timing correction, motion correction, covariates removal, temporal filtering, MNI space normalization, and spatial smoothing.

To diagnose MCI from normal, we employ the publicly available ADNI dataset. We specifically choose data from the six sites with the largest cohort sizes. We select a total of 333 rs-fMRI data from 148 distinct subjects, encompassing 60 MCIs and 88 NCs. It is worth noting that the ADNI dataset consists of multi-site and multi-scan data. To make full use of the data, we treat multiple scans as data augmentation, as the training-test data split is at the subject level to prevent data leakage. The rs-fMRI data follows a standardized preprocessing pipeline using Analysis of Functional NeuroImages (AFNI) toolbox [36], which discards the first 10 volumes for signal equilibrium, followed by slice timing correction, motion correction, covariates removal, temporal filtering, MNI space normalization, and spatial smoothing.

For the depression subtype diagnosis, we utilize our private dataset comprising 88 rs-fMRI data, consisting of 46 melancholic and 42 atypical patients. The rs-fMRI data is collected at Shanghai Mental Health Center affiliated to Shanghai Jiao Tong University School of Medicine, where the Ethics Review Board has reviewed and approved this research. Written informed consent is obtained from all participants before their enrollment in the study. The scanning parameters for the rs-fMRI data are as follows: TR = 2000ms, TE = 30ms, field of view (FOV) = 220mm×220mm, slice thickness = 3.0mm, and voxel size = 3.0mm×3.0mm×3.0mm. The rs-fMRI data is preprocessed using the DPARSF toolbox. The first 10 volumes

This article has been accepted for publication in IEEE Transactions on Medical Imaging. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TMI.2024.3368064

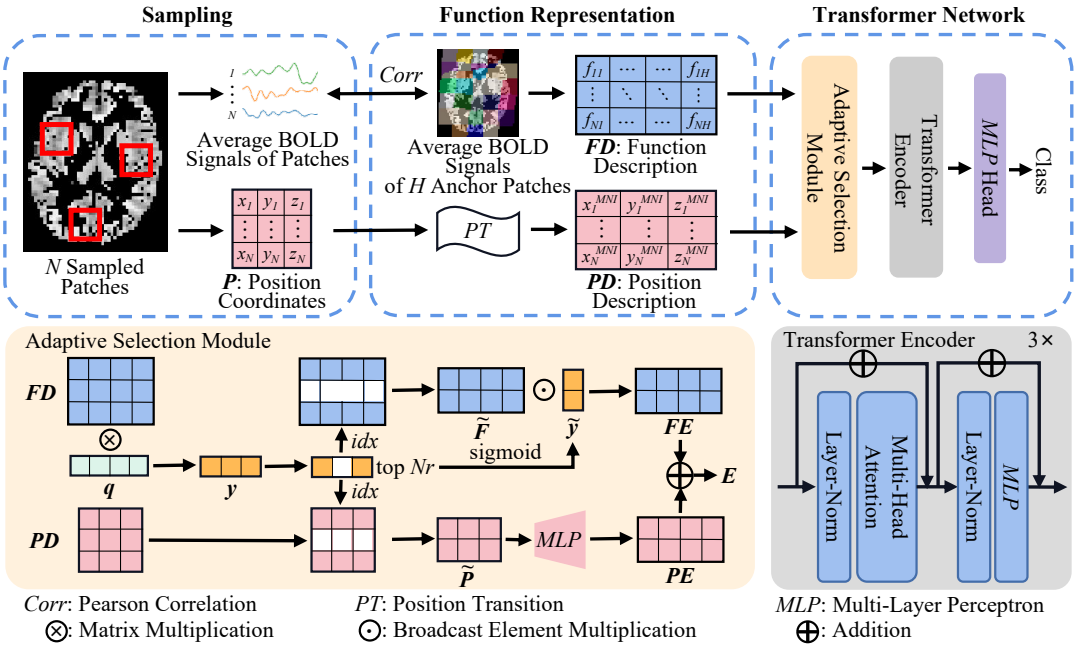


Fig. 1. The architecture of the proposed randomizing brain function representation method for brain disease diagnosis. In the **Sampling** module, we first randomly sample position coordinates in gray matter and obtain their corresponding *N* sampled patches, which are signified by average BOLD signals and position coordinates. Then, we design a **Function Representation** framework for the sampled patches consisting of function descriptions $\boldsymbol{FD}$ and position descriptions $\boldsymbol{PD}$. Finally, we propose an **Adaptive-Selection-Assisted Transformer Network** to integrate the function representations of sampled patches for disease diagnosis.

are discarded, followed by slice timing correction, motion correction, covariates removal, and temporal filtering.

B. Patch Sampling via Randomizing

In our study, we adopt a randomizing strategy by randomly sampling patches. Specifically, we sample voxel locations within the gray matter of each subject. The gray matter mask is obtained by segmenting the structural image aligned with the functional image. Subsequently, we extract the corresponding gray matter voxels in the patch surrounding each sampled location.

Depending on respective dataset configurations, we develop two specific ways to sample patches.

1) Sampling Patches in Individual Space: Given a subject fMRI in our private dataset, it is straightforward to employ randomizing and randomly sample position coordinates (x, y, z) from its own image space. And then we extract 3D patches with a size of $S \times S \times S$ surrounding these sampled position coordinates. Notably, our method only utilizes gray matter within the patches, as gray matter has been demonstrated to be closely related to brain function. The position coordinates $p_i = (x_i, y_i, z_i)$ $(i = 1, 2, ..., N)$ are chosen to represent their respective locations of sampled patches in the brain. N denotes the number of patches sampled in each subject, which is sufficiently large to ensure comprehensive coverage of the entire gray matter.

2) Sampling Patches in Standard Space: Many publicly available datasets (such as ABIDE and ADNI used in our study) comprise preprocessed brain images that have already been registered to the standard MNI space. We directly sample the position coordinates $p_i^{MNI} = (x_i^{MNI}, y_i^{MNI}, z_i^{MNI})$ and obtain the corresponding patches in the MNI space.

For both cases above, we can sample a set of patches to represent each subject brain. As depicted in the Sampling module in Fig. 1, the sampled patches are then signified by average BOLD signals of rs-fMRI together with position coordinates. It is important to note that the randomly sampled patches differ across subjects since the sampling process is independent for each subject. Consequently, the random sampling strategy retains individual specificity.

C. Function Representation

We present a novel brain representation, which combines function descriptions and position descriptions of the sampled patches, for the characterization of brain function. This framework is proposed for the first time and aims to enhance our understanding of brain function in the context of random sampling.

1) Function Description: FC serves as a common indicator of brain function [37], so we calculate FCs relating to the randomly sampled patches to initialize their function descriptions. To ensure comparability for the FCs of the sampled patches, we introduce anchor patches as their common reference. That is, we utilize equidistant grids to partition the entire gray matter to generate anchor masks as shown in Fig. 2. The anchors can thus be generated straightforwardly and consistently for all subjects. Note that, similar to the sampled patch, each anchor patch only considers gray matter voxels inside.

Specifically, we compute the average BOLD signals of gray matter within each randomly sampled patch and each anchor patch, resulting in the respective mean time series.

Then, we employ Pearson correlation (denoted as $Corr$ in Fig. 1) to calculate the FC matrix $\boldsymbol{FD} = \{f_{ij}\}$ ($\boldsymbol{FD} \in \mathbb{R}^{N\times H},\ i = 1, 2, ..., N,\ j = 1, 2, ..., H$), where f_{ij} represents the correlation between the randomly sampled patch v_i and the anchor patch a_j. The parameter N represents the number of patches sampled in each subject, and H denotes the number of anchor patches. As depicted in the Function Representation module in Fig. 1, each row of the FC matrix represents the initial function description of the corresponding randomly sampled patch.

2) Position Description: In addition to the function description, we find it important to have an additional position description for each patch. As our patches are randomly sampled across different subjects, we can ensure the intra- and inter-subject comparability of the patches with the help of position descriptions.

Specifically, we obtain position descriptions of the sampled patches via the standard MNI space. For the position coordinates $\boldsymbol{P} = \{p_i\}$ ($\boldsymbol{P} \in \mathbb{R}^{N\times 3},\ i = 1, 2, ..., N$) of the sampled patches in the private dataset, we leverage the T1 image as an intermediary to register the MNI template with the functional image. This registration process enables the calculation of the standard space coordinates for each voxel in the individual space of the functional image. Consequently, we can obtain the position description of the sampled patch:

$$\boldsymbol{PD} = norm(\boldsymbol{P^{MNI}}) = norm(\phi^{-1}(\boldsymbol{P})), \tag{1}$$

where ϕ^{-1} is the inverse deformation derived from the aforementioned registration. For the position coordinates $\boldsymbol{P^{MNI}} = \{p_i^{MNI}\}$ ($\boldsymbol{P^{MNI}} \in \mathbb{R}^{N\times 3},\ i = 1, 2, ..., N$) of the sampled patches in the publicly available dataset, they are already in the standard MNI space, eliminating the need for additional registration procedures:

$$\boldsymbol{PD} = norm(\boldsymbol{P^{MNI}}). \tag{2}$$

Here $norm$ is employed for normalizing coordinates to the range [0, 1].

It is worth noting that the above process is very different from registration-based ROI parcellation. The position description lets each patch have a global awareness of its position in the brain. And it establishes the basis to mutually compare individual patches across different subjects. On the contrary, if using an atlas to parcellate brain ROIs, one may have to complete image registration precisely, especially at the boundaries of the ROIs.

D. Adaptive-Selection-Assisted Transformer Network

We design a Transformer network to optimize and integrate the function and position descriptions of all sampled patches. It is powered by the adaptive selection strategy (c.f. Adaptive Selection Module in Fig. 1), which aims to identify the most discriminative regions for diagnosis [38].

1) Adaptive Selection Module: The adaptive selection module serves a dual purpose. It optimizes the computational efficiency of the Transformer by reducing input dimensions. And it enables the interpretability of brain disease diagnosis. Fig. 2 provides examples of the initially sampled patches and the adaptively selected patches within a subject. Furthermore, we can conduct statistical analysis on these adaptively selected patches at the group level to identify specific brain regions that are primarily associated with the brain disease.

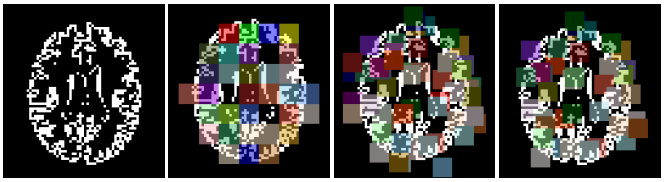


Fig. 2. The gray matter mask, anchor patches, and examples of initially sampled patches together with adaptively selected patches in a subject on ABIDE dataset. The sampled patches may overlap due to random sampling. Only the axial plane of the 3D patches is depicted.

The adaptive selection module involves a series of operations. As illustrated in Fig. 1, we first derive the importance $\boldsymbol{y}$ of each sampled patch computed as:

$$\boldsymbol{y} = \boldsymbol{FDq}/\|\boldsymbol{q}\|, \tag{3}$$

where $\boldsymbol{q}$ is a learnable projection vector, $\boldsymbol{FD}$ represents the function description. Then, a $rank(\cdot)$ function ranks and returns the indices idx of the largest Nr values in $\boldsymbol{y}$:

$$idx = rank(\boldsymbol{y}, Nr), \tag{4}$$

where r represents the adaptive selection ratio (set to a predefined percentage of 80% following Section V-D), and N represents the number of patches sampled in each subject. Next, the selected patches $\widetilde{\boldsymbol{F}} \in \mathbb{R}^{C\times H}$ and their corresponding position descriptions $\widetilde{\boldsymbol{P}} \in \mathbb{R}^{C\times 3}$ are retained based on the indices idx:

$$\begin{aligned} \widetilde{\boldsymbol{F}} &= \boldsymbol{FD}(idx, :), \\ \widetilde{\boldsymbol{P}} &= \boldsymbol{PD}(idx, :). \end{aligned} \tag{5}$$

Subsequently, the function embedding $\boldsymbol{FE}$ and the position embedding $\boldsymbol{PE}$ can be obtained:

$$\begin{aligned} \widetilde{\boldsymbol{y}} &= sigmoid(\boldsymbol{y}(idx)), \\ \boldsymbol{FE} &= \widetilde{\boldsymbol{F}} \odot (\widetilde{\boldsymbol{y}}\boldsymbol{1_H^T}), \\ \boldsymbol{PE} &= MLP(\widetilde{\boldsymbol{P}}), \end{aligned} \tag{6}$$

where the symbol $\odot$ denotes the Hadamard product, and $\boldsymbol{1}_H$ represents a vector of size H with all components equal to 1. The penultimate term in Eq. (6) makes the projection vector $\boldsymbol{q}$ in Eq. (3) trainable through back-propagation. The MLP stands for a projection layer to transform dimensions and enhance model representation. Finally, the function embedding $\boldsymbol{FE}$ and the corresponding position embedding $\boldsymbol{PE}$ of the selected patches are fused via addition to obtain the representation $\boldsymbol{E}$ of patches:

$$\boldsymbol{E} = \boldsymbol{FE} + \boldsymbol{PE}, \tag{7}$$

where $\boldsymbol{E}$ is input to Transformer encoder.

2) Implementation of Transformer Network: To integrate the function and position embeddings of the sampled patches in each brain, we employ Transformer to obtain a global representation of brain function. The Transformer encoders comprise multi-head self-attention (MSA) and multi-layer perceptron (MLP) layers. Additionally, before and after them there are layer-norm (LN) and residual connection respectively. The Transformer encoders receive fused function representations $\boldsymbol{E}$ of patches as input, encompassing both the functional and structural information of the brain.

The self-attention (SA) mechanism is employed to learn the relationships among features from the sampled patches:

$$SA = softmax(\frac{\boldsymbol{EW_Q}(\boldsymbol{EW_K})^{\boldsymbol{T}}}{\sqrt{D}})\boldsymbol{EW_V}, \tag{8}$$

where $\boldsymbol{W_Q} \in \mathbb{R}^{H\times D}$, $\boldsymbol{W_K} \in \mathbb{R}^{H\times D}$, $\boldsymbol{W_V} \in \mathbb{R}^{H\times M}$ are learnable linear transformation matrices. M represents the dimension of the output feature in a single attention head. The numbers of D and M are set to 32 following [39]. Furthermore, MSA can be obtained by averaging or concatenating the outputs of multiple SAs. In our approach, the model concatenates two-head SAs as:

$$MSA = concat(SA_1, SA_2)\boldsymbol{W_{MSA}}. \tag{9}$$

$\boldsymbol{W_{MSA}} \in \mathbb{R}^{2M\times O}$ is a learnable linear transformation matrix. Our model stacks three Transformer encoders to derive the global function representation, which is fed into an MLP Head for disease diagnosis.

IV. Experiments and Results

A. Experimental Settings

The proposed model is implemented with PyTorch, trained on a single GPU (NVIDIA GeForce RTX 2080 SUPER with 8GB memory). We adopt a 5-fold cross-validation strategy to assess the model's performance. For classification tasks, we utilize the cross-entropy loss function. The Adam optimizer is employed during training. The model is trained for 100 epochs with a learning rate of 0.001. The classification performance is evaluated using four metrics, including accuracy (ACC), sensitivity (SEN), specificity (SPE), and area under the receiver operating characteristic curve (AUC).

B. Classification Results on Different Datasets

To assess the effectiveness of our proposed method, we compare our method with the following approaches in the diagnosis tasks of ASD, MCI, and depression subtypes. Support Vector Machine (SVM) is a traditional machine learning method widely employed for disease classification. BrainNetCNN [40] is a convolutional neural network specifically designed for brain network analysis. Graph Attention Network (GAT) [41] excels at leveraging topological information from brain functional networks to diagnose brain diseases. Graph Convolutional Network (GCN) [24] aims at identifying individuals with major depressive disorder. BrainGNN [9] is an interpretable brain graph neural network to enable ASD diagnosis. MVS-GCN [42] is a prior brain structure learning-guided multi-view graph convolution network for ASD diagnosis. KD-Transformer [39] is a diffusion kernel attention network based on Transformer to diagnose MCI.

1) Classification Results on ABIDE: Given the diagnostic performance on the ABIDE dataset, as presented in Table II, we can draw the following conclusions. Firstly, deep learning methods consistently outperform traditional approaches. SVM exhibits inferior diagnostic performance compared to all other methods. This discrepancy arises from the capability of deep learning models to extract discriminative information in an end-to-end fashion, thereby enhancing the diagnosis of brain diseases. Secondly, the models specifically designed for brain disease diagnosis demonstrate superior diagnostic performance. Notably, BrainGNN, MVS-GCN, and KD-Transformer achieve higher classification accuracy than general GAT. This observation highlights the importance of appropriately designing models to effectively capture the complexity of brain function for accurate diagnosis. Finally, our proposed method achieves the most favorable performance, e.g., with the highest ACC of 0.801 and the highest AUC of 0.783, These results substantiate the effectiveness of sampling fMRI patches for constructing brain function representation without relying on a brain atlas for ROI parcellation.

2) Classification Results on ADNI: Table II also presents the classification performance on the ADNI dataset. Similar conclusions can be drawn from the results. Notably, our method demonstrates superior performance compared to the other competing methods, e.g., achieving the highest ACC of 0.778 and the highest AUC of 0.753. Furthermore, it is worth noting that the ADNI dataset comprises data collected from multiple sites. The observed results further demonstrate the compatibility of our method with multi-site data, showcasing its improved performance in such scenarios.

3) Classification Results on Depression Subtype: We further verify our method on the challenging task of diagnosing depression subtypes. Subtype diagnosis of depression aims to distinguish more homogeneous groups within depression to facilitate optimized treatment, which is a more difficult classification task [43]. The classification results on the depression subtype dataset are summarized in Table II. Our method achieves the highest ACC of 0.762. This indicates the ability of our proposed method to effectively capture comprehensive brain information, thereby boosting the diagnosis of brain diseases.

Although our method does not achieve the highest value in sensitivity on the ABIDE and ADNI datasets, it achieves a better balance in sensitivity and specificity. It may be because that GAT and MVS-GCN pay too much attention to certain sub-populations, and dwarf the overall classification performance. In addition, our method obtains a suboptimal AUC value for depression subtype diagnosis. However, the margin is very close and our method tends to have better stability (i.e., smaller standard deviation). Overall, the above results prove the effectiveness and generalizability of our proposed method.

TABLE II
COMPARISON WITH DIFFERENT METHODS ON ABIDE, ADNI, AND DEPRESSION SUBTYPE DATASETS (MEAN±STD)

Metric	ABIDE							
	SVM	BrainNetCNN	GAT	GCN	BrainGNN	MVS-GCN	KD-Transformer	Ours
ACC	0.649±0.056	0.696±0.039	0.661±0.041	0.720±0.063	0.719±0.030	0.726±0.083	0.748±0.041	**0.801±0.039**
AUC	0.663±0.097	0.654±0.072	0.640±0.073	0.705±0.078	0.663±0.048	0.695±0.099	0.684±0.071	**0.783±0.076**
SEN	0.838±0.152	0.828±0.121	0.817±0.068	0.848±0.104	0.784±0.108	**0.888±0.074**	0.847±0.070	0.797±0.090
SPE	0.619±0.045	0.676±0.065	0.636±0.046	0.700±0.075	0.706±0.032	0.695±0.099	0.732±0.051	**0.804±0.051**
Metric	ADNI							
	SVM	BrainNetCNN	GAT	GCN	BrainGNN	MVS-GCN	KD-Transformer	Ours
ACC	0.622±0.042	0.713±0.051	0.707±0.066	0.693±0.105	0.723±0.106	0.743±0.069	0.750±0.061	**0.778±0.057**
AUC	0.661±0.105	0.682±0.050	0.686±0.064	0.662±0.131	0.699±0.119	0.715±0.084	0.738±0.095	**0.753±0.060**
SEN	0.721±0.099	0.799±0.125	**0.850±0.050**	0.799±0.116	0.800±0.100	0.836±0.073	0.743±0.151	0.811±0.080
SPE	0.635±0.043	0.654±0.050	0.689±0.043	0.646±0.065	0.694±0.072	0.715±0.084	0.721±0.055	**0.742±0.049**
Metric	Depression Subtype							
	SVM	BrainNetCNN	GAT	GCN	BrainGNN	MVS-GCN	KD-Transformer	Ours
ACC	0.567±0.054	0.681±0.059	0.714±0.046	0.717±0.094	0.738±0.034	0.738±0.093	0.741±0.058	**0.762±0.052**
AUC	0.486±0.055	0.668±0.094	0.693±0.071	0.669±0.129	0.683±0.098	0.669±0.117	**0.709±0.089**	0.701±0.050
SEN	0.517±0.148	0.631±0.268	0.708±0.167	0.647±0.254	0.636±0.163	0.580±0.306	0.689±0.184	**0.761±0.107**
SPE	0.564±0.054	0.675±0.065	0.714±0.049	0.717±0.095	0.724±0.041	0.669±0.117	0.740±0.051	**0.762±0.054**

TABLE III
COMPARISON WITH DIFFERENT PARCELLATIONS ON ABIDE DATASET (MEAN±STD)

Metric	AAL	Schaefer300	SLIC	SCSC	Ours
ACC	0.731±0.037	0.737±0.055	0.720±0.049	0.737±0.068	**0.801±0.039**
AUC	0.711±0.075	0.741±0.095	0.687±0.078	0.718±0.104	**0.783±0.076**
SEN	0.755±0.103	0.747±0.097	0.734±0.041	0.765±0.068	**0.797±0.090**
SPE	0.726±0.028	0.739±0.069	0.717±0.050	0.732±0.070	**0.804±0.051**

C. Classification Results on Different Parcellations

To demonstrate the effectiveness of our proposed random sampling strategy, we conduct a comparative analysis by replacing our sampled patches with atlas-based and data-driven ROI parcellations. For easy comparison, the experiments here are on ASD diagnosis only. Firstly, we consider several widely used atlases. The AAL atlas is a population-average brain atlas extensively utilized in neuroimage studies. Additionally, the frequently employed Schaefer atlases define functional regions. For fairness, we select the Schaefer300 atlas which consists of 300 ROIs, as the number closely matches the sampled patches in our method. Secondly, we compare an individualized brain atlas generated by simple linear iterative clustering (SLIC) [30]. Finally, we chose a data-driven brain atlas generated by spatially constrained spectral clustering (SCSC) [26]. To ensure a fair comparison upon different parcellations only, we combine our proposed position description and adaptive selection module into all experiments here.

The results are presented in Table III, where our method demonstrates superior performance compared to using brain atlases. This improvement can be attributed to the random sampling strategy, which provides ample brain patches to represent the brain while preserving individual-specific diagnostic information.

D. Interpretability Exploration

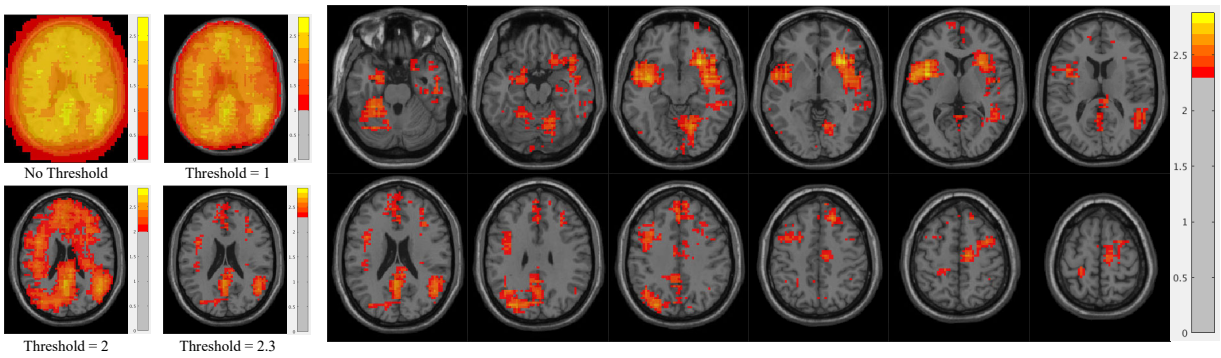
The aforementioned adaptive selection module can effectively screen all sampled patches to identify the most discriminative brain regions for disease diagnosis. These adaptively selected regions hold promise as potential disease biomarkers.

TABLE IV
THE INDICES AND NAMES OF THE MOST DISCRIMINATIVE REGIONS IDENTIFIED BY OUR METHOD ON ABIDE DATASET

Index	ROI Name in AAL	Index	ROI Name in AAL
1	Precentral_L	38	Hippocampus_R
2	Precentral_R	39	ParaHippocampal_L
4	Frontal_Sup_R	40	ParaHippocampal_R
7	Frontal_Mid_L	41	Amygdala_L
8	Frontal_Mid_R	45	Cuneus_L
11	Frontal_Inf_Oper_L	48	Lingual_R
16	Frontal_Inf_Orb_R	51	Occipital_Mid_L
17	Rolandic_Oper_L	56	Fusiform_R
20	Supp_Motor_Area_R	57	Postcentral_L
23	Frontal_Sup_Medial_L	65	Angular_L
24	Frontal_Sup_Medial_R	67	Precuneus_L
29	Insula_L	68	Precuneus_R
30	Insula_R	73	Putamen_L
31	Cingulum_Ant_L	74	Putamen_R
32	Cingulum_Ant_R	81	Temporal_Sup_L
34	Cingulum_Mid_R	82	Temporal_Sup_R
35	Cingulum_Post_L	84	Temporal_Pole_Sup_R
37	Hippocampus_L	86	Temporal_Mid_R

Specifically, we extract the brain patches identified by the adaptive selection module as detailed in Section III-D.1. Then, we aggregate the selected brain patches across the entire brain to generate a subject-specific heat map. Standardizing and averaging the heat maps of correctly predicted subjects yields the final set of regional biomarkers at the dataset level.

To interpret the identified biomarkers, we present the discriminative regions recognized by our models under different thresholds on the ABIDE dataset. As illustrated in Fig. 3, we can draw the following conclusions. Firstly, our proposed random sampling strategy effectively samples patches across the entire brain, as indicated by the 'No Threshold' setting. As the threshold increases, the discriminative regions gradually become more concentrated within the brain (Threshold = 1), eventually highlighting key brain regions (Threshold = 2, 2.3). Hence, we set the threshold at 2.3 to clearly visualize the most distinctive regions, which can potentially serve as biomarkers distinguishing ASD and control. To maintain consistency with previous studies, we utilize the brain region naming convention of the AAL atlas to establish connections with our detected



(a) Discriminative Regions Under Different Thresholds

(b) Discriminative Regions (Threshold = 2.3)

Fig. 3. Discriminative regions identified by our method on the ABIDE dataset: (a) under different thresholds, and (b) for threshold = 2.3.

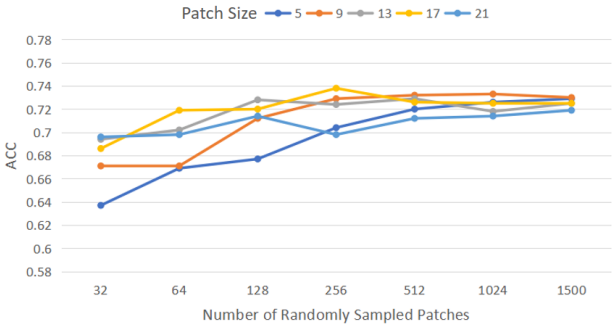


Fig. 4. ACC of ABIDE dataset with different randomly sampled patch sizes and numbers under multiple runs. The ACC is the mean of five-fold cross-validation for ten times.

TABLE V
ABLATION STUDY RESULTS OF POSITION DESCRIPTION (PD) AND ADAPTIVE SELECTION (AS) MODULE ON ABIDE DATASET (MEAN±STD). FD: FUNCTION DESCRIPTION

FD	PD	AS	ACC	AUC	SEN	SPE
✓			0.643±0.040	0.588±0.074	0.806±0.039	0.616±0.038
✓	✓		0.760±0.036	0.761±0.058	0.836±0.099	0.747±0.054
✓		✓	0.655±0.040	0.608±0.046	0.745±0.053	0.641±0.047
✓	✓	✓	0.801±0.039	0.783±0.076	0.797±0.090	0.804±0.051

biomarkers. A total of 36 regions are identified, with their respective AAL indices and names summarized in Table IV.

The majority of our identified biomarkers are consistent with previously reported ASD studies. Specifically, most of the identified regions (Table IV, Index 4, 16, 23, 24, 31, 32, 35, 65, 81, 82, 86) are implicated in the default mode network (DMN). The DMN is a crucial brain system involved in processing information about the self and others, and it has emerged as a key system associated with social dysfunction in ASD [44]. Moreover, other identified regions have also proven to be related to ASD. For instance, the insula has been widely utilized in FC analyses for ASD diagnosis [45]. The putamen has been shown to play a fundamental role in predicting ASD [46]. Previous studies have also reported brain abnormalities in various regions among individuals with ASD, such as the frontal lobe, temporal lobe, hippocampus, and precuneus [47], [48].

V. DISCUSSION

A. Ablation Studies

To demonstrate the effectiveness of the position description and adaptive selection module, we conduct ablation studies. As presented in Table V, the results highlight the impact of these components on the performance of our method. Specifically, compared to the absence of position description, our method achieves a significant improvement of 0.117 in terms of ACC (Table V, Row 1-2). This enhancement highlights the importance of position description in enabling the recognition and comparability of sampled patches within the randomizing framework, ultimately resulting in improved performance. Furthermore, the introduction of the adaptive selection module further enhances diagnostic performance, with an additional ACC improvement of 0.041 (Table V, Row 2, Row 4).

B. Sampled Patch Size and Number

In the random sampling strategy, the size of the sampled patch is a critical hyperparameter. We investigate the impact of patch size on the results using five different scales: 5, 9, 13, 17, and 21. We conduct the five-fold cross-validation for ten times by randomly dividing the dataset. As illustrated in Fig. 4 and Table VI, we can draw the conclusion that the larger sampled patches tend to yield improved results within a specific range. This observation can be attributed to the fact that larger brain patches exhibit a higher tolerance for BOLD noise. However, when the patch size exceeds a certain threshold, diagnostic performance will be compromised. This suggests that excessive brain parcellation leads to increased aliasing of brain function, which can be detrimental to disease diagnosis. Consequently, we adopt a patch size of $9\times9\times9$ in our study.

Regarding random sampling, another critical hyperparameter to consider is the number of sampled patches. We assess the influence of the patch number on the results at: 32, 64, 128, 256, 512, 1024, and 1500. We conduct the five-fold cross-validation for ten times by randomly dividing the dataset. As illustrated in Fig. 4 and Table VI, we can draw the following conclusions. Initially, as the number of patches increases from a small value, the model performance improves. However,

TABLE VI
ACC OF ABIDE DATASET AND ITS 95% CONFIDENCE INTERVAL (CI) FOR RANDOMLY SAMPLED PATCH SIZES AND NUMBERS UNDER FIVE-FOLD CROSS-VALIDATION FOR TEN TIMES

ACC (CI)		Patch Size				
		5	9	13	17	21
Patch Number	32	0.637 (0.624, 0.650)	0.671 (0.656, 0.685)	0.694 (0.675, 0.712)	0.686 (0.672, 0.700)	0.696 (0.680, 0.712)
	64	0.669 (0.653, 0.684)	0.671 (0.657, 0.686)	0.702 (0.688, 0.717)	0.719 (0.705, 0.732)	0.698 (0.684, 0.713)
	128	0.677 (0.661, 0.692)	0.712 (0.695, 0.729)	0.728 (0.710, 0.747)	0.720 (0.706, 0.734)	0.714 (0.697, 0.730)
	256	0.704 (0.685, 0.724)	0.729 (0.716, 0.742)	0.724 (0.706, 0.741)	0.738 (0.721, 0.754)	0.698 (0.684, 0.711)
	512	0.720 (0.703, 0.737)	0.732 (0.715, 0.748)	0.729 (0.713, 0.746)	0.726 (0.709, 0.742)	0.712 (0.699, 0.726)
	1024	0.726 (0.708, 0.745)	0.733 (0.717, 0.749)	0.718 (0.700, 0.735)	0.725 (0.707, 0.744)	0.714 (0.698, 0.730)
	1500	0.729 (0.711, 0.747)	0.730 (0.714, 0.746)	0.725 (0.710, 0.740)	0.725 (0.708, 0.743)	0.719 (0.704, 0.734)

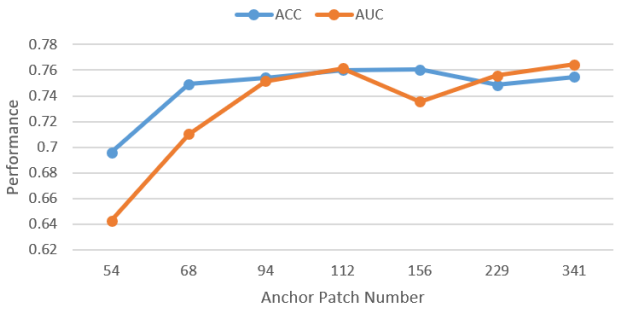


Fig. 5. ACC and AUC of ABIDE dataset with different anchor patch numbers.

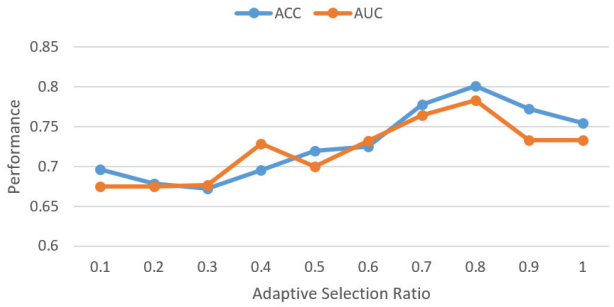


Fig. 6. ACC and AUC of ABIDE dataset with different adaptive selection ratios.

once the number of patches surpasses a certain threshold, the model's performance tends to stabilize and reach saturation. This observation can be attributed to the fact that when the number of patches is small, there is insufficient information for disease diagnosis, so increasing the number of patches enhances the model's performance accordingly. Nonetheless, once the number of patches is sufficiently large to adequately represent the brain, further increasing the number does not yield significant improvements in model performance. Thus, to strike a balance between efficiency and performance, we choose 256 as the number of sampled patches in our study.

It is worth noting that the size and number of sampled patches are not independent hyperparameters, but rather interconnected. The experimental results depicted in Fig. 4 indirectly support this notion. For example, when the number of sampled patches is insufficient (e.g., number = 32), the performance improves as the size of the patches increases. This can be attributed to the fact that a combination of a small number of large-sized patches captures a substantial amount of information necessary for effective disease diagnosis.

C. Anchor Patch Number

We conduct experimental investigations to assess the impact of anchor patch granularity (i.e., the number of anchor patches and their corresponding sizes) on brain disease diagnosis. The experimental results are depicted in Fig. 5. As the number of anchor patches increases, the diagnostic performance initially improves significantly, followed by stabilization. Based on these findings, we select a total of 112 anchor patches for subsequent experiments. This choice is justified by the results presented in Fig. 5, which confirm that this level of granularity is sufficient to calculate the function representation of the randomly sampled patches.

D. Adaptive Selection Ratio

We design an adaptive selection module to identify the most salient patches for diagnosis among the sampled patches. This module operates by adjusting the adaptive selection ratio r. To determine the optimal value for r, we conducted a series of experiments within the range of [0, 1]. The results, as presented in Fig. 6, reveal a distinct pattern. The classification performance initially improves and then decreases as the adaptive selection ratio increases. Notably, the best performance is achieved when $r = 0.8$. Extremely small values of r lead to an insufficient number of patches for effective diagnosis, while excessively large values allow redundant and noisy patches to interfere with the diagnostic process. Consequently, we adopt an adaptive selection ratio of $r = 0.8$ in our approach.

E. Limitations and Future Work

In our study, we have showcased the effectiveness of using anchor patches to calculate function representation. However, our current implementation employs a unified parcellation of the gray matter as the anchor patch. To enhance flexibility and adaptability, alternative approaches for defining the anchor patch should be explored, such as a random sampling strategy.

In fact, our random sampling strategy offers additional potential for data augmentation of brain functions by iteratively sampling patches and constructing corresponding FCs within the same subject, which will be our future work. Currently, we simply add function embeddings and position embeddings to combine them, and optimizing the way to effectively fuse function and position representations will be the future work.

VI. Conclusion

In this paper, we present a novel random sampling strategy to generate brain function representation for neural disease diagnosis. Our approach leverages random sampling to derive brain patches, thereby circumventing the limitations associated with atlas-based parcellation used for defining ROIs. Additionally, we introduce a new framework for brain function representation, which combines function description with reference to anchor patches and position description. Furthermore, the adaptive selection module designed by us optimizes the computational cost of the Transformer network and realizes interpretability exploration. Our method has demonstrated competitive diagnostic performance across three distinct brain diseases, showcasing its efficacy and versatility in the study of neural disorders.

References

[1] K. Sakai and K. Yamada, "Machine learning studies on major brain diseases: 5-year trends of 2014–2018," *Japanese journal of radiology*, vol. 37, pp. 34–72, 2019.

[2] M. Solmi, M. Song, D. K. Yon, S. W. Lee, E. Fombonne, M. S. Kim, S. Park, M. H. Lee, J. Hwang, R. Keller *et al.*, "Incidence, prevalence, and global burden of autism spectrum disorder from 1990 to 2019 across 204 countries," *Molecular Psychiatry*, pp. 1–9, 2022.

[3] H. Herrman, C. Kieling, P. McGorry, R. Horton, J. Sargent, and V. Patel, "Reducing the global burden of depression: a lancet–world psychiatric association commission," *The Lancet*, vol. 393, no. 10189, pp. e42–e43, 2019.

[4] J. Zhang, A. Kucyi, J. Raya, A. N. Nielsen, J. S. Nomi, J. S. Damoiseaux, D. J. Greene, S. G. Horovitz, L. Q. Uddin, and S. Whitfield-Gabrieli, "What have we really learned from functional connectivity in clinical populations?" *NeuroImage*, vol. 242, p. 118466, 2021.

[5] E. Canario, D. Chen, and B. Biswal, "A review of resting-state fmri and its use to examine psychiatric disorders," *Psychoradiology*, vol. 1, no. 1, pp. 42–53, 2021.

[6] Z. Wang, X. Zhou, Y. Gui, M. Liu, and H. Lu, "Multiple measurement analysis of resting-state fmri for adhd classification in adolescent brain from the abcd study," *Translational Psychiatry*, vol. 13, no. 1, p. 45, 2023.

[7] H. Jiang, P. Cao, M. Xu, J. Yang, and O. Zaiane, "Hi-gcn: A hierarchical graph convolution network for graph embedding learning of brain network and brain disorders prediction," *Computers in Biology and Medicine*, vol. 127, p. 104096, 2020.

[8] E. Bullmore and O. Sporns, "Complex brain networks: graph theoretical analysis of structural and functional systems," *Nature reviews neuroscience*, vol. 10, no. 3, pp. 186–198, 2009.

[9] X. Li, Y. Zhou, N. Dvornek, M. Zhang, S. Gao, J. Zhuang, D. Scheinost, L. H. Staib, P. Ventola, and J. S. Duncan, "Braingnn: Interpretable brain graph neural network for fmri analysis," *Medical Image Analysis*, vol. 74, p. 102233, 2021.

[10] J. Huang, L. Zhou, L. Wang, and D. Zhang, "Attention-diffusion-bilinear neural network for brain network analysis," *IEEE transactions on medical imaging*, vol. 39, no. 7, pp. 2541–2552, 2020.

[11] M. D. Greicius, B. H. Flores, V. Menon, G. H. Glover, H. B. Solvason, H. Kenna, A. L. Reiss, and A. F. Schatzberg, "Resting-state functional connectivity in major depression: abnormally increased contributions from subgenual cingulate cortex and thalamus," *Biological psychiatry*, vol. 62, no. 5, pp. 429–437, 2007.

[12] D. Yao, J. Sui, M. Wang, E. Yang, Y. Jiaerken, N. Luo, P.-T. Yap, M. Liu, and D. Shen, "A mutual multi-scale triplet graph convolutional network for classification of brain disorders using functional or structural connectivity," *IEEE transactions on medical imaging*, vol. 40, no. 4, pp. 1279–1289, 2021.

[13] M. Cole, K. Murray, E. St-Onge, B. Risk, J. Zhong, G. Schifitto, M. Descoteaux, and Z. Zhang, "Surface-based connectivity integration: An atlas-free approach to jointly study functional and structural connectivity," *Human Brain Mapping*, vol. 42, no. 11, pp. 3481–3499, 2021.

[14] K. Ota, N. Oishi, K. Ito, H. Fukuyama, S.-J. S. Group *et al.*, "A comparison of three brain atlases for mci prediction," *Journal of neuroscience methods*, vol. 221, pp. 139–150, 2014.

[15] N. Tzourio-Mazoyer, B. Landeau, D. Papathanassiou, F. Crivello, O. Etard, N. Delcroix, B. Mazoyer, and M. Joliot, "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.

[16] A. Schaefer, R. Kong, E. M. Gordon, T. O. Laumann, X.-N. Zuo, A. J. Holmes, S. B. Eickhoff, and B. T. Yeo, "Local-global parcellation of the human cerebral cortex from intrinsic functional connectivity mri," *Cerebral cortex*, vol. 28, no. 9, pp. 3095–3114, 2018.

[17] R. S. Desikan, F. Ségonne, B. Fischl, B. T. Quinn, B. C. Dickerson, D. Blacker, R. L. Buckner, A. M. Dale, R. P. Maguire, B. T. Hyman *et al.*, "An automated labeling system for subdividing the human cerebral cortex on mri scans into gyral based regions of interest," *Neuroimage*, vol. 31, no. 3, pp. 968–980, 2006.

[18] R. M. Lawrence, E. W. Bridgeford, P. E. Myers, G. C. Arvapalli, S. C. Ramachandran, D. A. Pisner, P. F. Frank, A. D. Lemmer, A. Nikolaidis, and J. T. Vogelstein, "Standardizing human brain parcellations," *Scientific data*, vol. 8, no. 1, p. 78, 2021.

[19] C. H. Gomez, L. Dodero, A. Gozzi, V. Murino, and D. Sona, "Atlas-free connectivity analysis driven by white matter structure," in *2017 IEEE 14th International Symposium on Biomedical Imaging (ISBI 2017)*. Ieee, 2017, pp. 89–92.

[20] M. A. de Reus and M. P. Van den Heuvel, "The parcellation-based connectome: limitations and extensions," *Neuroimage*, vol. 80, pp. 397–404, 2013.

[21] D. Wang, R. L. Buckner, M. D. Fox, D. J. Holt, A. J. Holmes, S. Stoecklein, G. Langs, R. Pan, T. Qian, K. Li *et al.*, "Parcellating cortical functional networks in individuals," *Nature neuroscience*, vol. 18, no. 12, pp. 1853–1860, 2015.

[22] Y. Zhang, H. Zhang, E. Adeli, X. Chen, M. Liu, and D. Shen, "Multiview feature learning with multiatlas-based functional connectivity networks for mci diagnosis," *IEEE Transactions on Cybernetics*, vol. 52, no. 7, pp. 6822–6833, 2020.

[23] A. Abraham, M. P. Milham, A. Di Martino, R. C. Craddock, D. Samaras, B. Thirion, and G. Varoquaux, "Deriving reproducible biomarkers from multi-site resting-state data: An autism-based example," *NeuroImage*, vol. 147, pp. 736–745, 2017.

[24] K. Qin, D. Lei, W. H. Pinaya, N. Pan, W. Li, Z. Zhu, J. A. Sweeney, A. Mechelli, and Q. Gong, "Using graph convolutional network to characterize individuals with major depressive disorder across multiple imaging sites," *EBioMedicine*, vol. 78, p. 103977, 2022.

[25] N. U. Dosenbach, B. Nardos, A. L. Cohen, D. A. Fair, J. D. Power, J. A. Church, S. M. Nelson, G. S. Wig, A. C. Vogel, C. N. Lessov-Schlaggar *et al.*, "Prediction of individual brain maturity using fmri," *Science*, vol. 329, no. 5997, pp. 1358–1361, 2010.

[26] R. C. Craddock, G. A. James, P. E. Holtzheimer III, X. P. Hu, and H. S. Mayberg, "A whole brain fmri atlas generated via spatially constrained spectral clustering," *Human brain mapping*, vol. 33, no. 8, pp. 1914–1928, 2012.

[27] L. Fan, H. Li, J. Zhuo, Y. Zhang, J. Wang, L. Chen, Z. Yang, C. Chu, S. Xie, A. R. Laird *et al.*, "The human brainnetome atlas: a new brain atlas based on connectional architecture," *Cerebral cortex*, vol. 26, no. 8, pp. 3508–3526, 2016.

[28] P. Bellec, P. Rosa-Neto, O. C. Lyttelton, H. Benali, and A. C. Evans, "Multi-level bootstrap analysis of stable clusters in resting-state fmri," *Neuroimage*, vol. 51, no. 3, pp. 1126–1139, 2010.

[29] W. Cui, Y. Ma, J. Ren, J. Liu, G. Ma, H. Liu, and Y. Li, "Personalized functional connectivity based spatio-temporal aggregated attention network for mci identification," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 2023.

[30] J. Wang, Z. Hu, and H. Wang, "Parcellating whole brain for individuals by simple linear iterative clustering," in *Neural Information Processing: 23rd International Conference, ICONIP 2016, Kyoto, Japan, October 16–21, 2016, Proceedings, Part III 23*. Springer, 2016, pp. 131–139.

[31] M. Liu, J. Zhang, E. Adeli, and D. Shen, "Landmark-based deep multi-instance learning for brain disease diagnosis," *Medical image analysis*, vol. 43, pp. 157–168, 2018.

[32] C. Lian, M. Liu, Y. Pan, and D. Shen, "Attention-guided hybrid network for dementia diagnosis with structural mr images," *IEEE transactions on cybernetics*, vol. 52, no. 4, pp. 1992–2003, 2020.

[33] A. Dosovitskiy, L. Beyer, A. Kolesnikov, D. Weissenborn, X. Zhai, T. Unterthiner, M. Dehghani, M. Minderer, G. Heigold, S. Gelly *et al.*, "An image is worth 16x16 words: Transformers for image recognition at scale," *arXiv preprint arXiv:2010.11929*, 2020.

[34] M. Graña and M. Silva, "Impact of machine learning pipeline choices in autism prediction from functional connectivity data," *International journal of neural systems*, vol. 31, no. 04, p. 2150009, 2021.

[35] C. Yan and Y. Zang, "Dparsf: a matlab toolbox for" pipeline" data analysis of resting-state fmri," *Frontiers in systems neuroscience*, p. 13, 2010.
[36] R. W. Cox, "Afni: software for analysis and visualization of functional magnetic resonance neuroimages," *Computers and Biomedical research*, vol. 29, no. 3, pp. 162–173, 1996.
[37] M. P. Van Den Heuvel and H. E. H. Pol, "Exploring the brain network: a review on resting-state fmri functional connectivity," *European neuropsychopharmacology*, vol. 20, no. 8, pp. 519–534, 2010.
[38] H. Gao and S. Ji, "Graph u-nets," in *international conference on machine learning*. PMLR, 2019, pp. 2083–2092.
[39] J. Zhang, L. Zhou, L. Wang, M. Liu, and D. Shen, "Diffusion kernel attention network for brain disorder classification," *IEEE Transactions on Medical Imaging*, vol. 41, no. 10, pp. 2814–2827, 2022.
[40] J. Kawahara, C. J. Brown, S. P. Miller, B. G. Booth, V. Chau, R. E. Grunau, J. G. Zwicker, and G. Hamarneh, "Brainnetcnn: Convolutional neural networks for brain networks; towards predicting neurodevelopment," *NeuroImage*, vol. 146, pp. 1038–1049, 2017.
[41] P. Veličković, G. Cucurull, A. Casanova, A. Romero, P. Lio, and Y. Bengio, "Graph attention networks," *arXiv preprint arXiv:1710.10903*, 2017.
[42] G. Wen, P. Cao, H. Bao, W. Yang, T. Zheng, and O. Zaiane, "Mvs-gcn: A prior brain structure learning-guided multi-view graph convolution network for autism spectrum disorder diagnosis," *Computers in Biology and Medicine*, vol. 142, p. 105239, 2022.
[43] R. Musil, F. Seemüller, S. Meyer, I. Spellmann, M. Adli, M. Bauer, K.-T. Kronmüller, P. Brieger, G. Laux, W. Bender *et al.*, "Subtypes of depression and their overlap in a naturalistic inpatient sample of major depressive disorder," *International Journal of Methods in Psychiatric Research*, vol. 27, no. 1, p. e1569, 2018.
[44] A. Padmanabhan, C. J. Lynch, M. Schaer, and V. Menon, "The default mode network in autism," *Biological Psychiatry: Cognitive Neuroscience and Neuroimaging*, vol. 2, no. 6, pp. 476–486, 2017.
[45] M. Plitt, K. A. Barnes, and A. Martin, "Functional connectivity classification of autism identifies highly predictive brain features but falls short of biomarker standards," *NeuroImage: Clinical*, vol. 7, pp. 359–366, 2015.
[46] L.-A. Sapey-Triomphe, L. Pattyn, V. Weilnhammer, P. Sterzer, and J. Wagemans, "Neural correlates of hierarchical predictive processes in autistic adults," *Nature Communications*, vol. 14, no. 1, p. 3640, 2023.
[47] D. Sussman, R. Leung, V. Vogan, W. Lee, S. Trelle, S. Lin, D. Cassel, M. Chakravarty, J. Lerch, E. Anagnostou *et al.*, "The autism puzzle: Diffuse but not pervasive neuroanatomical abnormalities in children with asd," *NeuroImage: Clinical*, vol. 8, pp. 170–179, 2015.
[48] Y. Xiao, T. H. Wen, L. Kupis, L. T. Eyler, V. Taluja, J. Troxel, D. Goel, M. V. Lombardo, K. Pierce, and E. Courchesne, "Atypical functional connectivity of temporal cortex with precuneus and visual regions may be an early-age signature of asd," *Molecular Autism*, vol. 14, no. 1, p. 11, 2023.