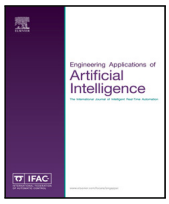
Engineering Applications of Artificial Intelligence 166 (2026) 113556

Contents lists available at ScienceDirect

Engineering Applications of Artificial Intelligence

journal homepage: www.elsevier.com/locate/engappai



Research paper

A consistency-driven pseudo-labeling framework for robust functional connectivity modeling in neuropsychiatric disorder diagnosis

Xin Wen [a], Shijie Guo [a], Li Dong [b,c,*], Xiaobo Liu [d], Wenbo Ning [a], Jie Shi [a], Songhua Liu [a], Cheng Luo [b,c], Rui Cao [a,*], Dezhong Yao [b,c]

[a] School of Software, Taiyuan University of Technology, Taiyuan, 030024, China
[b] The Clinical Hospital of Chengdu Brain Science Institute, MOE Key Lab for Neuroinformation, University of Electronic Science and Technology of China, Chengdu, 611731, China
[c] School of Life Science and Technology, University of Electronic Science and Technology of China, Chengdu, 611731, China
[d] Montreal Neurological Institute, Mcgill University, Quebec, H3A 2B1, Canada

ARTICLE INFO

Keywords:
Pseudo labeling
Computer-aided diagnosis
Functional magnetic resonance imaging
Functional connectivity
Deep learning

ABSTRACT

The incidence of neuropsychiatric disorders such as Autism Spectrum Disorder (ASD), Attention Deficit Hyperactivity Disorder (ADHD), and Major Depressive Disorder (MDD) continues to rise. Deep learning-based computer-aided diagnosis (CAD) has emerged as a promising approach to alleviate the increasing burden on neuroimaging-based clinical resources. However, neuroimaging modalities such as functional magnetic resonance imaging (fMRI) involve complex spatiotemporal characteristics, making their representations susceptible to various types of noise and interference, which in turn hampers the effectiveness of CAD. To address this challenge, we propose a pseudo-label consistency-driven framework for functional connectivity (FC) reconstruction and discriminative modeling (PL-FCDM), aiming to enhance both the representational quality and discriminative power of FC features. Specifically, two complementary pseudo-labeling models are developed to independently capture discriminative features from the temporal domain (time series) and spatial domain (dynamic functional connectivity), enabling pseudo label prediction from distinct modalities. Then a consistency-based filtering strategy is applied to construct high-confidence reconstructed functional connectivity. These graphs are subsequently fed into a classification model comprising a Feature Optimization Autoencoder and a Depthwise Separable Convolutional Neural Network for efficient identification of neuropsychiatric disorders. Extensive experiments conducted on four publicly available multi-site datasets—ABIDE I, ABIDE II, ADHD-200, and REST-meta-MDD demonstrate that the proposed method achieves classification accuracies of 76.14%, 74.37%, 72.89%, and 71.15%, respectively. These results consistently outperform several state-of-the-art approaches, validating the effectiveness and robustness of the proposed framework in feature refinement and multi-disorder recognition.

1. Introduction

As neuropsychiatric disorders become increasingly prevalent, illnesses such as Autism Spectrum Disorder (ASD), Attention Deficit Hyperactivity Disorder (ADHD), and Major Depressive Disorder (MDD) (Lauritsen, 2013; Tripp and Wickens, 2009; Otte et al., 2016) have turned into significant global public health issues. These conditions are frequently associated with notable emotional, behavioral, and cognitive difficulties (Thapar et al., 2017). Nonetheless, early diagnosis proves challenging due to the subtlety of symptoms and considerable clinical diversity, which greatly obstructs timely intervention and enhancement of patients' quality of life (World Health Organization, 2006). Consequently, it is crucial to develop effective and objective computer-aided diagnosis (CAD) methods for the early identification and precise treatment of these disorders.

In recent years, the rapid development of artificial intelligence (AI), particularly deep learning, has significantly advanced CAD in medical image analysis (Doi, 2007). Deep neural network based CAD methods have demonstrated strong potential in analyzing functional magnetic resonance imaging (fMRI) data by automatically learning functional connectivity (FC) representations, providing valuable support for the diagnosis of neuropsychiatric disorders (Canario et al., 2021; Ingalhalikar et al., 2021; Huang et al., 2020). FC reflects the

* Corresponding authors.
E-mail addresses: lidong@uestc.edu.cn (L. Dong), caorui@tyut.edu.cn (R. Cao).

https://doi.org/10.1016/j.engappai.2025.113556
Received 22 July 2025; Received in revised form 27 October 2025; Accepted 13 December 2025
Available online 22 December 2025
0952-1976/© 2025 Elsevier Ltd. All rights are reserved, including those for text and data mining, AI training, and similar technologies.

synchronized activity between brain regions, providing insights into neural mechanisms and functional abnormalities (Du et al., 2024). Static functional connectivity and dynamic functional connectivity (dFC) are the two widely utilized neuroimaging metrics in CAD. Static functional connectivity is derived through Pearson correlation of regional blood-oxygen-level-dependent (BOLD) time series over the entire fMRI acquisition period, yielding a stationary representation of inter-regional interactions (Ingalhalikar et al., 2021). However, it overlooks temporal variations and fails to capture transient brain states. In contrast, dFC (Hindriks et al., 2016) employs sliding window method to model time-varying connectivity patterns, enabling the detection of subtle dynamic changes associated with psychiatric conditions. Despite its advantages, dFC confronts several critical challenges, including high dimensionality, susceptibility to measurement noise, and temporal redundancy, all of which constrain its reliability and clinical applicability. Consequently, a central research objective in this domain is the development of effective methodologies that retain the dynamic sensitivity of dFC while simultaneously enhancing feature robustness, minimizing noise interference, and improving discriminative performance for accurate diagnostic inference.

Driven by sophisticated deep neural network architectures, deep learning techniques excel at the automated extraction of high-level latent features from complex and high-dimensional data, enabling robust and scalable pattern recognition (Shao et al., 2021; İlikhan et al., 2025). These attributes have positioned deep learning as a highly promising approach within CAD for neuropsychiatric disorders. In recent years, a range of deep learning models including Autoencoders (AE), Convolutional Neural Networks (CNNs), Long Short-Term Memory networks (LSTMs), and Graph Neural Networks (GNNs) have been increasingly applied to the study of neurodevelopmental and psychiatric conditions such as ASD, ADHD, and MDD (Yin et al., 2022; Alam et al., 2022; Liu et al., 2023; Jiang et al., 2020). Recent advances have further confirmed the efficacy of specific architectures, particularly CNNs, in diagnosing complex disorders like schizophrenia from neuroimaging data (Degirmenci et al., 2025). These models exhibit a strong ability to uncover high-level discriminative representations and improve diagnostic accuracy. Moreover, the successful application of advanced feature learning paradigms, such as manifold learning, in diagnosing other complex diseases like ovarian cancer (Yesilkaya et al., 2022) underscores the generalizability and versatility of data-driven approaches in the medical field. These models have demonstrated significant potential in uncovering high-level discriminative representations and enhancing diagnostic accuracy. Nevertheless, existing approaches continue to face notable limitations. A key challenge lies in their limited capacity to adequately capture the intrinsic spatial dependencies and topological structures embedded within FC. The learned representations may lack sufficient expressiveness to effectively characterize disease-related alterations in brain function.

To address the aforementioned challenges, this study proposes a novel pseudo-labeling framework for functional connectivity enhancement and disease classification—referred to as Pseudo-Label based Functional Connectivity Discrimination Model (PL-FCDM). By integrating two complementary modalities — time series (Ts) and functional connectivity — the framework enables cross-modal pseudo-label validation and reconstruction-based enhancement of pseudo-label quality. Specifically, we design two distinct pseudo-label generation models that independently extract features from the temporal domain (Ts matrix) and the spatial domain (dFC matrix) to generate corresponding pseudo-label predictions. A pseudo-label consistency filtering mechanism is then introduced to retain only high-confidence samples that receive consistent predictions from both models. These refined samples are subsequently utilized to construct an enhanced, more discriminative Reconstructed Functional Connectivity (R-FC) matrix. Finally, a Feature Optimization Autoencoder (FO-AE) is employed to extract robust and compact latent representations, which are further processed by a lightweight Depthwise Separable Convolutional Neural Network (DSCNN) to achieve efficient and accurate classification of neuropsychiatric disorders.

The main contributions of this study are summarized as follows:

- A Spatiotemporal-Guided Pseudo-Labeling Strategy: The proposed strategy leverages both temporal dynamics and spatial interactions to generate independent pseudo-label predictions. A cross-domain consistency constraint is then applied to filter low-confidence or inconsistent labels, thereby improving pseudo-label reliability and enhancing the representativeness and discriminative power of the unlabeled training set.
- Functional Connectivity Reconstruction: By leveraging pseudo-label-derived latent class information, the proposed approach refines the structure of dFC, significantly enhancing its signal stability and interpretability which reduces noise interference and improves resilience to mislabeled samples during classification.
- Optimized Encoder-Depthwise Separable Convolutional Neural Network: this network is developed by combining latent feature refinement through an autoencoder with lightweight depthwise separable convolution operations. The proposed architecture enhances model generalizability and efficiency while maintaining superior performance in the multi-diagnostic categorization of various neuropsychiatric conditions.

2. Related work

2.1. Semi-supervised domain adaptation

Semi-supervised domain adaptation (SSDA) effectively transfers knowledge from a labeled source domain to a target domain with scarce labels and abundant unlabeled data, reducing reliance on large-scale annotations and improving generalization in neuroimaging tasks (Kouw and Loog, 2019). Existing SSDA methods fall into three broad categories. Pseudo-labeling and consistency-based approaches assign pseudo-labels or enforce prediction stability under input or model perturbations; for example, Cheng et al. (2015) refined mild cognitive impairment classification via knowledge transfer from Alzheimer's disease data, and Dinsdale et al. (2020) harmonized multi-site MRI using consistency regularization, though such methods risk error propagation under large domain shifts. Adversarial alignment methods learn domain-invariant features by training with domain discriminators; for instance, Wang et al. (2020) aligned fMRI topics across sites using adversarial graph embedding, and Ali et al. (2020) applied Cycle-GAN for FLAIR-MRI mapping, yet these approaches often require careful trade-offs between domain alignment and class separability. A third category involves feature or parameter level adaptation, where knowledge is transferred through shared representations. For example, Hailong et al. (2018) fine-tuned a source-trained autoencoder on limited ASD labels with reconstruction consistency, and Sivaranjini and Sujatha (2020) adapted a pre-trained AlexNet for Parkinson's detection; recent work also demonstrates that freezing pre-trained encoders and applying semi-supervised objectives can effectively exploit unlabeled data (Kang et al., 2020). Overall, SSDA leverages both labeled and unlabeled structure, yielding more robust and data-efficient models than supervised transfer alone.

2.2. Robust and trustworthy AI in CAD

With the advancement of deep learning, developing robust and trustworthy AI systems has become a central goal in CAD (Wang and Qiao, 2024). Although data-driven SSDA methods have shown promising results in neuroimaging, the high dimensionality and complex dynamics of fMRI time series and dynamic functional connectivity pose significant challenges to building reliable CAD models (Kouw and Loog, 2019). Recent advances suggest that rethinking the modeling paradigm for complex time-series data can enhance robustness. For

Table 1
Demographic information of the subjects used in the experiments.

Dataset	ABIDE I		ABIDE II		ADHD-200		REST-meta-MDD	
Information	ASD	NC	ASD	NC	ADHD	NC	MDD	NC
Subjects	505	530	264	322	353	429	741	673
Gender (F/M)	62/443	95/435	44/231	109/226	75/277	207/222	267/474	269/404
Age (mean ± std)	17.1 ± 8.6	16.8 ± 7.5	13.9 ± 8.0	13.5 ± 7.5	11.6 ± 3.0	11.7 ± 3.2	35.43 ± 12.42	35.69 ± 14.01

instance, treating images as sequences of tokens as in Vision Transformers (ViT) (Dosovitskiy et al., 2021), constructing hypergraphs (Zhu et al., 2025a), or discretizing continuous signals (Zhu et al., 2025b). Notably, spontaneous brain activity during rest exhibits rich temporal variability, characterized by transient interactions among brain regions that may engage different subregions of a network at different times (Karahanoğlu and Van De Ville, 2015; Wen et al., 2020). Inspired by ViT's success in modeling spatial data as sequential inputs, we propose to treat both fMRI time series and dFC as structured "images" and model them within an SSDA framework. This perspective leverages the inherent spatiotemporal structure of resting-state fMRI and, by enforcing cross-modal consistency, offers a promising pathway toward more robust and trustworthy CAD systems.

3. Methods and materials

3.1. Datasets and preprocessing

ABIDE I, ABIDE II (Di Martino et al., 2014), ADHD-200 (Consortium, 2012), and REST-meta-MDD (Chen et al., 2022) datasets were employed to evaluate the proposed method. In this study, we used 1035 participants from the ABIDE I dataset (505 ASD and 530 normal control), 586 participants from the ABIDE II dataset (264 ASD and 322 normal control), and 782 participants from the ADHD-200 dataset (353 ADHD and 429 normal control). Additionally, 1414 participants from the REST-meta-MDD dataset were included (741 MDD and 673 normal control). The number of subjects included in this study and their demographics are summarized in Table 1.

For data preprocessing, the ABIDE dataset was processed using the Preprocessed Connectomes Project (PCP) pipeline (Craddock et al., 2013), which applies the Configurable Pipeline for the Analysis of Connectomes (CPAC). The Athena pipeline (Bellec et al., 2017) was employed for the preprocessing of the ADHD-200 dataset. For the REST-meta-MDD dataset, the standard preprocessing procedures provided by DPARSF (Data Processing Assistant for Resting-State fMRI) (https://rfmri.org/DPARSF) were applied. All the preprocessed datasets are publicly available from their respective official websites.

3.2. Method

The proposed PL-FCDM framework is illustrated in Fig. 1. As shown in Fig. 1A, the sliding window technique is employed to segment the time series and construct dynamic functional connectivity (dFC) matrices. Fig. 1B presents the dual-perspective pseudo-label generation and consistency filtering strategy, which reconstructs high-quality feature-level functional connectivity by integrating information from both temporal and spatial domains. Fig. 1C shows the proposed OE-DSCNN classification model, which is designed to deeply extract latent features from the reconstructed FCs and leverage them to improve classification performance.

3.2.1. Time-series segment matrix and dFCs construction

As shown in Fig. 1A, the Time series-Segment (Ts-S) matrix and dFCs of each subject were constructed using a sliding window method (Hindriks et al., 2016). The fMRI data were processed using the AAL atlas (Rolls et al., 2020) to extract time series from 116 predefined brain regions. The sliding window approach was applied to segment the time series into a series of overlapping windows, each capturing local temporal features. Let the time series be denoted as $\mathbf{X} = [x_1, x_2, \ldots, x_T]$, where T is the total length of the time series, W is the window length ($W = 50$), and S is the step size ($S = 1$). For each time window starting at index i, the window can be represented as: $\mathbf{X}_{w_i} = [x_i, x_{i+1}, \ldots, x_{i+W-1}]$, As the window slides along the time series, the starting index i increases from 0 by a step of S, until the condition $i + W \leq T$ is satisfied. This procedure generates multiple windows systematically, with each window forming one Ts-S segment. Finally, for each Ts-S, a symmetric FC matrix is computed using the Pearson correlation, as shown in Eq. (1):

$$FC = \frac{\sum_{i=1}^{n}(X_i - \bar{X})(Y_i - \bar{Y})}{\sqrt{\sum_{i=1}^{n}(X_i - \bar{X})^2 \sum_{i=1}^{n}(Y_i - \bar{Y})^2}} \tag{1}$$

where n is the window length, and X, Y are two time series within the window. $\bar{X}$ and $\bar{Y}$ represent the mean values of X and Y, respectively.

Assuming a subject's time series length is L, the total number of generated Ts-S and FC matrices, denoted as N, can be computed using Eq. (2):

$$N = \left\lfloor \frac{L - W}{S} \right\rfloor + 1 \tag{2}$$

where W is the window length and S is the step size.

Thus, for a given subject, the complete set of matrices can be represented as: Ts-S matrix = $[\text{Ts-S}_0, \text{Ts-S}_1, \ldots, \text{Ts-S}_N]$, dFC = $[\text{FC}_0, \text{FC}_1, \ldots, \text{FC}_N]$. In summary, both the Ts-S matrices and corresponding dFCs can be obtained from the fMRI data.

3.2.2. Time-series segment matrix and dFCs (T–F) Pseudo Label Generative

To generate pseudo labels for data with different feature characteristics, we design two models: the Ts-S Pseudo Label Generative Model and the dFCs Pseudo Label Generative Model, which are responsible for learning temporal and spatial representations, respectively.

The Ts-S Pseudo Label Generative Model employs a five-layer Transformer encoder to learn representations from the Ts-S matrix and predict pseudo labels. Leveraging the ability to capture long-range dependencies, the model extracts more informative and discriminative temporal features.

The dFCs Pseudo Label Generative Model is implemented based on a Depthwise Separable Convolutional Network, which consists of two specialized convolutional kernels — ColumnFeatureConv and RowFeatureConv — as illustrated in Fig. 1B. The ColumnFeatureConv operation (see Eq. (3)) aggregates information from the edges connected to each node, transforming edge information into node-level feature representations. This enables each node to not only retain its own attributes but also integrate features from neighboring nodes, thereby capturing more comprehensive and robust spatial representations.

$$Y_1[i, j] = \sum_{m=0}^{c_1-1} \sum_{n=0}^{\dim-1} X[i, j+n] \cdot W_1[m, n] + b_1 \tag{3}$$

where X is the input feature matrix, Y_1 is the output matrix, W_1 (1*116@32) is the convolution kernel, b_1 is the bias term, i and j are the indices of the output matrix, dim represents the dimension of the feature matrix, and c_1 is the number of output channels in the first convolution layer.

The RowFeatureConv operation, as defined in Eq. (4), aggregates information from all nodes into a global representation. This operation

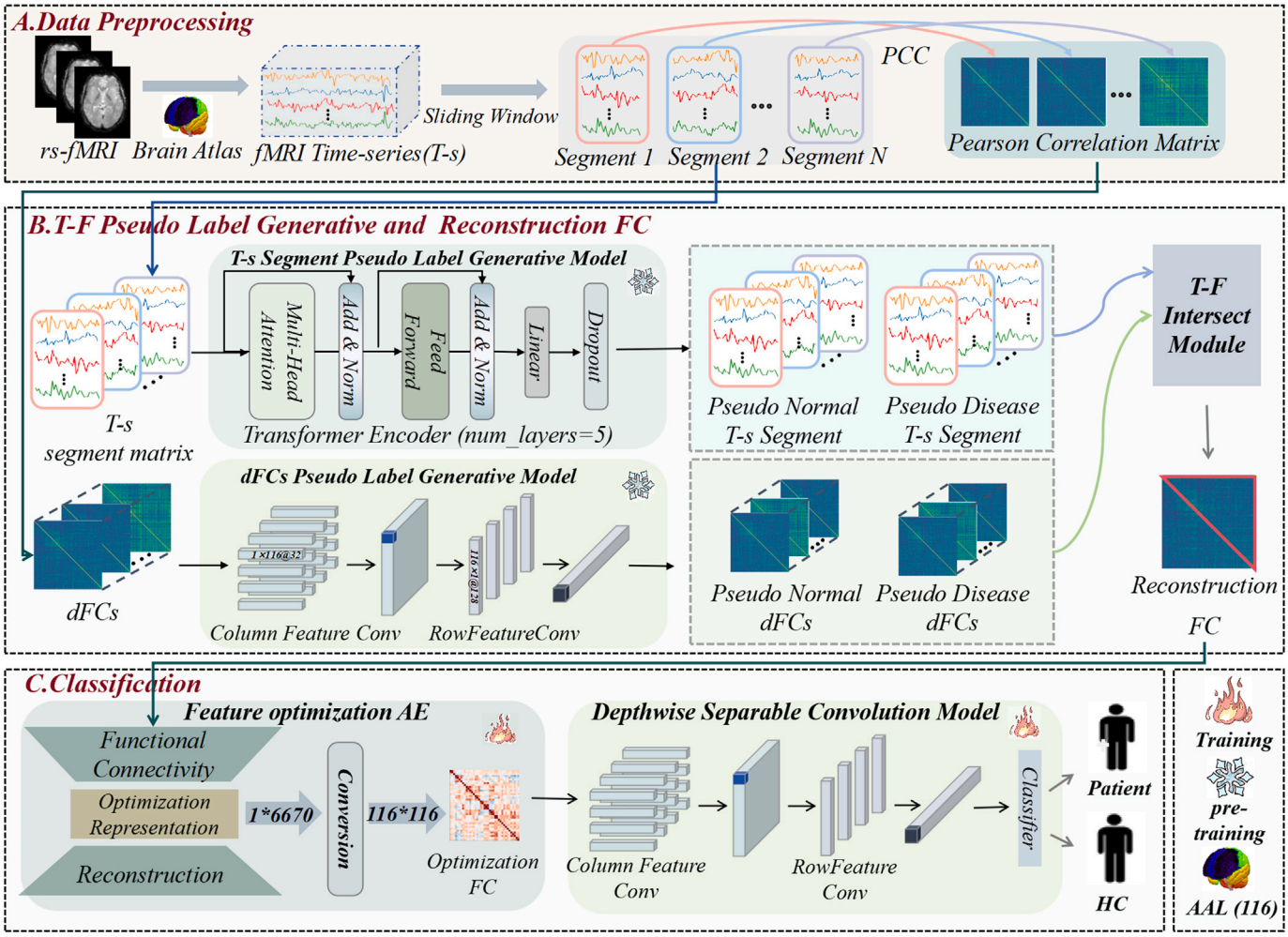


Fig. 1. The proposed PL-FCDM framework consists of three main components: (a) The sliding window technique is employed to segment the time series and construct dFC matrices. (b) a spatiotemporal-guided pseudo-label generation and consistency filtering strategy for reconstructing high-confidence functional connectivity representations. (c) a classification model, OE-DSCNN that integrates feature optimization AE with depthwise separable convolution, which is designed to deeply extract latent features from the reconstructed FCs and leverage them to improve classification performance.

enables the extraction of holistic features that characterize the entire functional connectivity (FC) matrix.

$$Y_2[i,j] = \sum_{m=0}^{c_2-1} \sum_{n=0}^{\dim-1} \tanh\left(Y_1[i+n,j]\right) \cdot W_2[m,n] + b_2 \tag{4}$$

where W_2 (116*1@128) is the convolution kernel, Y_2 is the convolution output, b_2 is the bias term, i and j are the indices of the output matrix, and C_2 is the number of output channels in the second convolution layer.

After the construction of the two models, pseudo-label prediction is carried out. To prevent data leakage, separate datasets are used for model pre-training. Specifically, the REST-meta-MDD dataset is employed to pre-train the model for pseudo-label prediction on the ABIDE-I, ABIDE-II, and ADHD-200 datasets. For the REST-meta-MDD dataset, the model pre-trained on ABIDE-I is adopted. During pre-training, each dataset is split into a training set (80%) and a testing set (20%) in a 4:1 ratio. Both the *Ts-S Pseudo Label Generative Model* and the *dFCs Pseudo Label Generative Model* are trained via the cross-entropy loss function, as defined in Eq. (5).

$$\text{CrossEntropyLoss} = -\frac{1}{N}\sum_{i=1}^{N} \log\left(\frac{\exp(z_{i,y_i})}{\sum_{k=1}^{C}\exp(z_{i,k})}\right) \tag{5}$$

where z_{i,y_i} denotes the predicted logit (score) for the correct class y_i of the ith sample, $z_{i,k}$ is the logit for class k, N is the total number of samples, and C is the number of classes.

To facilitate understanding of the pseudo-label prediction process, the pseudocode is provided in Algorithm 1. As shown in Fig. 1B, a subject obtains multiple Ts-S segments and dFCs after the sliding window computation. When pseudo labels are predicted based on the subject's Ts-S matrix and dFCs, two types of pseudo labels can be obtained: Normal Pseudo Label and Disease Pseudo Label, as defined in Eqs. (6), (7) and (8). The total number of Normal and Disease Pseudo Labels predicted from the Ts-S matrix and dFCs equals the number of dFCs for that subject.

$$y_{i,j}^{\text{dFC}} = \text{PseudoLabel}(\text{Pre-trained dFCs Pseudo Label Generative Model} \times (\mathbf{X}_{i,j}^{\text{dFC}})) \tag{6}$$

$$y_{i,j}^{\text{Ts-S}} = \text{PseudoLabel}(\text{Pre-trained Ts-S Pseudo Label Generative Model} \times (\mathbf{X}_{i,j}^{\text{Ts-S}})) \tag{7}$$

$$y_{i,j}^{\text{dFC}} / y_{i,j}^{\text{Ts-S}} = \begin{cases} \text{Normal Pseudo Label}, & \text{if } y_{i,j} = 0 \\ \text{Disease Pseudo Label}, & \text{if } y_{i,j} = 1 \end{cases} \tag{8}$$

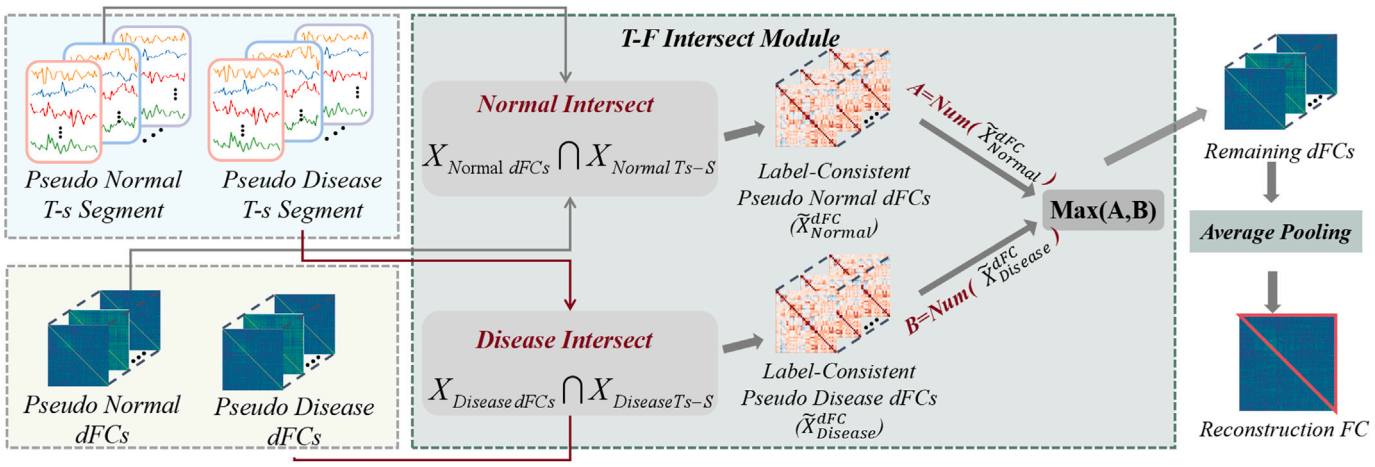


Fig. 2. Illustration of the proposed T–F Intersect Module. The module integrates pseudo-labels from both temporal (Ts-S matrices) and spatial (dFCs) representations to improve the quality of functional connectivity data. First, pseudo labels are predicted for each Ts-S segment and corresponding dFC using the pre-trained generative models. Then, only samples whose pseudo labels are consistent across both temporal and spatial domains (i.e., both labeled as *Normal* or both labeled as *Disease*) are retained, while inconsistent samples are removed. To mitigate the impact of class imbalance within each subject, samples from the minority pseudo-label class are discarded. Finally, an average pooling operation is applied to the remaining dFC samples to construct a high-quality static functional connectivity matrix (Reconstruction FC) for downstream classification.

where $j = 1, 2, \ldots, N$, $X_{i,j}$ denotes the jth Ts-S matrix or dFC obtained from the ith subject after applying the sliding window.

Algorithm 1: Pseudo-Label Generation

Input: $X_1(X_{Ts-S}, X_{dFCs})$: pretraining dataset
$X_2(X_{Ts-S}, X_{dFCs})$: pseudo-label prediction dataset
Output: Optimized Ts-S and dFCs pseudo-label generation models

Split X_1 into train/validation sets (80%/20%);
Initialize Ts-S model and dFCs model;
while *not converged* **do**
  Temporal Feature Learning:
  $\hat{Y}_{Ts-S} \leftarrow$ Ts-S Model($X_1^{train}[Ts-S]$);
  $\mathcal{L}_{Ts-S} \leftarrow$ CrossEntropy($\hat{Y}_{Ts-S}, Y$);
  Update: $\min_{W_{Ts-S}} \mathcal{L}_{Ts-S} + \lambda \|W_{Ts-S}\|^2$;
  Spatial Feature Learning:
  $\hat{Y}_{dFCs} \leftarrow$ dFCs Model($X_1^{train}[dFCs]$);
  $\mathcal{L}_{dFCs} \leftarrow$ CrossEntropy($\hat{Y}_{dFCs}, Y$);
  Update: $\min_{W_{dFCs}} \mathcal{L}_{dFCs} + \lambda \|W_{dFCs}\|^2$;
  Save intermediate models;
Model Selection:
$Acc^{Ts-S} \leftarrow$ Evaluate($X_1^{val}[Ts-S]$);
$Acc^{dFCs} \leftarrow$ Evaluate($X_1^{val}[dFCs]$);
Save best models based on validation accuracy;
Pseudo Label Prediction:
Predict pseudo labels for $X_2[Ts-S]$ using best Ts-S model;
Predict pseudo labels for $X_2[dFCs]$ using best dFCs model;

3.2.3. Pseudo-label consistency-driven reconstruction-based functional connectivity

To improve the quality of dFC data obtained through the sliding window method and reduce the impact of low-quality samples on downstream tasks, we propose a pseudo-label consistency-driven reconstruction method (T–F Intersect Module). This module integrates pseudo-label from both temporal (Ts-S matrix) and spatial (dFCs) representations to filter and reconstruct samples, aiming to construct a more representative and discriminative static functional connectivity (Reconstruction FC), as illustrated in Fig. 2.

Specifically, four pseudo-label sets are defined as follows: $\mathcal{N}_N^{\text{dFC}}$ represents the set of dFC samples labeled as *Normal*, D_D^{dFC} denotes the set labeled as *Disease*, $\mathcal{N}_N^{\text{Ts-S}}$ and $D_D^{\text{Ts-S}}$ denote the corresponding *Normal* and *Disease* Ts-Segment sample sets. For each dFC sample and its corresponding Ts-Segment (i.e., from the same subject and same time window), we retain only those samples whose pseudo-label predictions are consistent across both models (i.e., dFC and Ts-Segment are both labeled as *Normal* or both as *Disease*). Only these consistent samples are used for downstream tasks. The index set of consistent samples for the ith subject is defined as:

$$\mathcal{I}_i = \left\{ j \in \{1, 2, \ldots, N_i\} \,\middle|\, y_{i,j}^{\text{dFC}} = y_{i,j}^{\text{Ts-S}} \right\} \tag{9}$$

where, $y_{i,j}$ represents the pseudo label of the jth Ts-Segment or dFC matrix of the ith subject. $\mathcal{I}_i$ denotes the set of indices where the pseudo labels of dFC and Ts-Segment match for the ith subject.

Based on the index set $\mathcal{I}_i$, we construct two new filtered matrices containing only the consistent samples, as defined in Eq. (10):

$$\tilde{\mathbf{X}}_i^{\text{dFC}} = \left\{ \mathbf{X}_{i,j}^{\text{dFC}} \,\middle|\, j \in \mathcal{I}_i \right\}, \quad \tilde{\mathbf{X}}_i^{\text{Ts-S}} = \left\{ \mathbf{X}_{i,j}^{\text{Ts-S}} \,\middle|\, j \in \mathcal{I}_i \right\} \tag{10}$$

where $\mathbf{X}_{i,j}^{\text{dFC}}$ denotes the dFC of the ith subject in the jth time window, and $\mathbf{X}_{i,j}^{\text{Ts-S}}$ denotes the corresponding Ts-Segment in the same time window.

Subsequently, we reconstructed the static FC using the retained high-quality dFC samples. Specifically, for each subject, there often exists an imbalance between the Normal and Disease pseudo-labeled dFC samples. To mitigate the potential bias introduced by such imbalance, we discard the minority class and retain only the dFC samples belonging to the majority pseudo-label category. Thereafter, an average pooling operation is applied to the Remaining dFC samples to aggregate their features and generate the final Reconstruction FC for each subject, as defined in Eq. (11).

$$\mathbf{X}_{\text{Reconstruction, i}} = \begin{cases} \frac{1}{N_1} \sum_{i=1}^{N_1} \tilde{\mathbf{X}}_{1,i}^{\text{dFC}}, & \text{if } N_{1,i} > N_{0,i} \\ \frac{1}{N_2} \sum_{i=1}^{N_2} \tilde{\mathbf{X}}_{2,i}^{\text{dFC}}, & \text{if } N_{0,i} > N_{1,i} \end{cases} \tag{11}$$

where, $\tilde{\mathbf{X}}_{1,i}^{\text{dFC}}$ denotes the set of dFC samples for the ith subject with pseudo-label $y = 1$ (Disease), containing $N_{1,i}$ samples, while $\tilde{\mathbf{X}}_{0,i}^{\text{dFC}}$ denotes the set of dFC samples for the ith subject with pseudo-label $y = 0$ (Normal), containing $N_{0,i}$ samples.

3.3. Classification

In this study, we propose a novel classification model termed OE-DSCNN, designed to enhance the representational capability of the Reconstruction FC constructed after pseudo-label consistency filtering. Specifically, the Reconstruction FC is first fed into a Feature Optimization Autoencoder (FO-AE) to extract high-quality features, which are then input to a Depthwise Separable Convolutional Neural Network to perform the final classification and discrimination.

First, the ReconstructionFC of each subject is transformed into a one-dimensional feature vector. Since the functional connectivity matrix is symmetric, we extract its upper triangular part and flatten it to construct the one-dimensional vector $\mathbf{V}$, as expressed in Eq. (12):

$$\mathbf{V} = \{\text{Reconstruction FC}_{ij} \mid 1 \le i < j \le n\}, \tag{12}$$

where the length of $\mathbf{V}$ is $L = \frac{n(n-1)}{2}$, and n (116) is the dimension of the matrix.

The one-dimensional vector $\mathbf{V}$ is then input into the FO-AE for feature extraction. Structurally, autoencoder is a kind of feedforward neural network, which is composed of two parts: encoder and decoder. The encoder encodes the input data into a low-dimensional representation, as shown in Eq. (13):

$$\mathbf{Z}_{\text{encoder}} = f_{\text{encoder}}(\mathbf{V}) = \sigma(\mathbf{V}\mathbf{W}_{\text{encoder}} + \mathbf{b}_{\text{encoder}}), \tag{13}$$

where σ is the activation function, $\mathbf{W}_{\text{encoder}}$ and $\mathbf{b}_{\text{encoder}}$ are the weight matrix and bias vector of the encoder, respectively.

The decoder reconstructs the encoder's output back to the original input data, as shown in Eq. (14):

$$\mathbf{X}' = f_{\text{decoder}}(\mathbf{Z}_{\text{encoder}}) = \mathbf{Z}_{\text{encoder}}\mathbf{W}_{\text{decoder}} + \mathbf{b}_{\text{decoder}}. \tag{14}$$

The hidden layer's output is then transformed into a matrix. This matrix is transposed and added to the original matrix, with the diagonal elements set to 1, resulting in the final optimized functional connectivity matrix OptimizationFC.

The extracted Optimization FC is classified using the Depthwise Separable Convolution Model. The classifier is trained jointly with the Optimization FC, two loss functions are set for training. The loss function for training the Depthwise Separable Convolution Model is Cross-Entropy loss (Eq. (5)), while the loss function for training the Feature Optimization AE is the Cosine Embedding Loss (Eq. (15)).

$$L\left(x_1, x_2, y\right) = \begin{cases} 1 - \cos\left(x_1, x_2\right) & \text{if } y = 1 \\ \max\left(0, \cos\left(x_1, x_2\right) + 1\right) & \text{if } y = -1 \end{cases} \tag{15}$$

The value of y is 1, indicating that a high similarity is desired between x_1 and x_2, and -1, indicating that a low similarity is desired. In this paper, we aim for a high similarity between the Reconstruction FC and the Optimization FC. $\cos\left(x_1, x_2\right)$ is the cosine similarity between x_1 and x_2.

3.4. Implementation details

Our model is implemented in PyTorch (Imambi et al., 2021). The experiments are accelerated on an NVIDIA GeForce RTX 3090 24G. Hyperparameter tuning is performed using the NNI automated tuning tool (Microsoft, 2021), which helps optimize the model's performance by automatically adjusting parameters. The final implementation details are as follows. For the sliding window technique, a window size of 50 and a step size of 1 were adopted. Each Ts-S has a dimensionality of 50 × 116. Each FC matrix has a dimensionality of 116 × 116. The ColumnFeatureConv and RowFeatureConv layers are configured with kernel sizes of 1 × 116 (output channels: 32) and 116 × 1 (output channels: 128), respectively. The code is available at https://github.com/ggguo666/PL-FCDM.

4. Experiments and results

The proposed PL-FCDM was evaluated against a range of baseline methods, which can be divided into two categories for clarity: (1) For classical machine learning methods (SVM (Jakkula, 2006), MLP (Taud and Mas, 2017)) and re-implemented deep learning models (BrainNetCNN (Kawahara et al., 2017), HI-GCN (Jiang et al., 2020)), the hyperparameters were optimized using the NNI toolkit. (2) For other state-of-the-art models (BrainGNN (Li et al., 2021), ASD-HNet (Luo et al., 2025), S3TL (Hu et al., 2023)), we directly reported the best performance from their respective original papers.

4.1. Classification performance

The performance of the proposed framework is evaluated using four metrics: accuracy, sensitivity, specificity, and AUC. In addition, the average value of these four metrics is reported. To mitigate the potential bias caused by a single split of the dataset, 10-fold cross-validation is employed. As shown in Tables 2, 3, 4, 5, we compare the classification performance of various machine learning and deep learning methods across four datasets: ABIDE I, ABIDE II, ADHD-200, and REST-meta- MDD. The proposed method achieves high accuracy on these datasets (76.14%, 74.37%, 72.89%, and 71.15%, respectively). Furthermore, the model attains balanced sensitivity (77.04%, 72.27%, 71.67%, and 70.38%) and specificity (75.97%, 76.24%, 74.61%, and 71.51%) across the four datasets, indicating that our approach maintains a well-balanced predictive capability in distinguishing between patients and healthy controls.

4.2. Ablation study

The proposed method comprises three key components: the Pseudo Label Generative module, the T–F Intersect module, and the OE-DSCNN classification model, which collectively constitute the core of its performance. Specifically, the Pseudo Label Generative module is responsible for generating pseudo-labels for dFC and constructing Reconstruction FC via max-average pooling for subsequent classification tasks. The T–F Intersect module incorporates temporal information into the original dFC by leveraging time-series based features, facilitating the selection of high-quality data segments and enhancing feature representation. The OE-DSCNN model serves as the classification architecture proposed in this study, which integrates a Feature Optimization Autoencoder with a Depthwise Separable Convolutional Neural Network to improve classification performance.

To further validate the contribution of each component, we conducted five groups of ablation experiments. In each experiment, one or more modules were removed or replaced, and 10-fold cross-validation was employed. Classification performance was assessed using accuracy, sensitivity, specificity, and AUC. The average results across the 10 folds are summarized in Table 6, where the best results are highlighted in bold. The results across the 10 folds are summarized in Fig. 3. The ablation study is designed as follows:

(1) Classification performance of CNN: In the first experiment, a CNN was directly applied to classify the raw dFCs. As shown in Table 6, the accuracy on the ABIDE I, ABIDE II, ADHD, and MDD datasets was 63.3%, 64.68%, 60.85%, and 60.9%, respectively, indicating that the raw dFCs alone are insufficient for extracting discriminative features.
(2) Effect of the Pseudo Label Generative module: In the second experimental phase, the Pseudo Label Generative module was incorporated to refine dFC using the generated pseudo-labels, with the reconstructed FC being derived through the application of max-average pooling. This was followed by classification using the same CNN. Compared with the first experiment, the accuracy on the ABIDE I, ABIDE II, ADHD, and MDD datasets improved

Table 2
Performance comparison of different methods on ABIDE-I dataset. Results (in %) are presented as mean ± standard deviation. The best results in each column are highlighted in bold. Missing values are indicated by "–".

Method	ACC	SEN	SPE	AUC
SVM	65.56 ± 5.78	70.00 ± 7.86	60.99 ± 11.28	71.42 ± 7.86
MLP	68.40 ± 5.30	64.00 ± 18.20	72.20 ± 14.70	69.30 ± 4.20
HI-GCN (Jiang et al., 2020)	69.31 ± 5.14	73.00 ± 10.63	70.26 ± 4.57	69.21 ± 5.14
BrainGNN (Li et al., 2021)	69.31 ± 3.03	68.18 ± 11.05	68.64 ± 13.77	69.52 ± 3.52
BrainNetCNN (Kawahara et al., 2017)	69.73 ± 2.76	66.73 ± 9.93	72.54 ± 8.26	71.99 ± 3.06
ASD-HNet (Luo et al., 2025)	73.57	72.63	74.90	–
BrainSCK (Wang et al., 2024)	63.60	–	–	–
FMTLJD (Huang et al., 2022)	69.48 ± 1.60	73.38 ± 4.80	69.65 ± 3.80	–
S3TL (Hu et al., 2023)	69.14 ± 3.87	67.27 ± 4.32	70.88 ± 3.21	72.52 ± 3.56
STCAL (Liu et al., 2024a)	73.00	**79.80**	65.90	78.20
Pre-Trained Transformer (Li et al., 2024)	69.09 ± 2.23	60.16 ± 10.50	68.92 ± 3.91	73.23 ± 2.85
Transformer (Li et al., 2022)	74.18	–	–	–
STIGR(Dynamic Graph) (Liu et al., 2025)	72.7	76.1	68.8	**79.0**
Ours	**76.14 ± 2.95**	77.04 ± 5.86	**75.97 ± 5.40**	75.94 ± 3.07

Table 3
Performance comparison of different methods on ABIDE-II dataset. Results (in %) are presented as mean ± standard deviation where available. The best results in each column are highlighted in bold. Missing values are indicated by "–".

Method	ACC	SEN	SPE	AUC
SVM	64.48 ± 4.29	68.57 ± 4.36	60.39 ± 5.28	62.81 ± 4.80
MLP	61.18 ± 5.16	60.12 ± 6.10	62.24 ± 2.30	62.85 ± 5.15
HI-GCN (Jiang et al., 2020)	64.17	64.17	56.58	–
BrainGNN (Li et al., 2021)	66.25 ± 6.72	–	–	67.48 ± 5.05
BrainNetCNN (Kawahara et al., 2017)	65.16 ± 1.60	57.73 ± 6.74	71.52 ± 6.74	66.75 ± 1.93
MCG-Net	65.51	61.55	64.06	–
ASD-HNet (Luo et al., 2025)	72.55	65.23	78.04	–
S3TL (Hu et al., 2023)	69.65	64.98	71.31	69.27
STCAL (Liu et al., 2024a)	72.00	**74.40**	69.30	76.80
Pre-Trained Transformer (Li et al., 2024)	69.43 ± 1.53	57.69 ± 4.47	68.38 ± 6.23	69.43 ± 1.45
STIGR(Dynamic Graph) (Liu et al., 2025)	72.5	73.0	71.6	**77.6**
Ours	**74.37 ± 3.62**	72.27 ± 8.03	**76.24 ± 5.88**	72.75 ± 4.01

Table 4
Performance comparison of different methods on ADHD-200 dataset. Results (in %) are presented as mean ± standard deviation where available. The best results in each column are highlighted in bold. Missing values are indicated by "–".

Method	ACC	SEN	SPE	AUC
SVM	54.02 ± 6.98	36.08 ± 24.64	69.71 ± 15.61	49.64 ± 15.57
MLP	60.10 ± 2.90	59.90 ± 6.10	61.20 ± 3.30	62.30 ± 4.50
HI-GCN (Jiang et al., 2020)	64.23 ± 4.39	63.13 ± 11.94	66.19 ± 8.99	58.97 ± 4.45
BrainGNN (Li et al., 2021)	66.63 ± 5.41	70.10 ± 12.77	63.97 ± 10.79	64.03 ± 3.88
BrainNetCNN (Kawahara et al., 2017)	63.77 ± 3.84	69.87 ± 12.76	58.37 ± 13.57	62.97 ± 5.10
ASD-HNet (Luo et al., 2025)	71.33	58.49	80.30	–
BrainSCK (Wang et al., 2024)	70.10	–	–	–
FMTLJD (Huang et al., 2022)	71.44 ± 3.20	65.55 ± 6.30	75.68 ± 2.80	–
S3TL (Hu et al., 2023)	72.62 ± 3.96	62.71 ± 3.65	**77.43 ± 4.32**	73.66 ± 2.85
STCAL (Liu et al., 2024a)	72.50	79.20	64.60	**78.00**
Pre-Trained Transformer (Li et al., 2024)	70.11 ± 2.07	51.38 ± 6.49	73.75 ± 5.04	70.78 ± 4.63
Ours	**72.89 ± 3.91**	**71.67 ± 9.16**	74.61 ± 5.66	71.84 ± 2.81

Table 5
Performance comparison of different methods on REST-meta- MDD. dataset. Results (in %) are presented as mean ± standard deviation where available. The best results in each column are highlighted in bold. Missing or invalid values are indicated by "–".

Method	ACC	SEN	SPE	AUC
SVM	60.04 ± 2.42	54.83 ± 3.97	64.79 ± 7.35	63.73 ± 1.45
MLP	60.25 ± 2.57	59.21 ± 7.21	61.11 ± 4.42	59.47 ± 3.30
BrainGNN (Li et al., 2021)	67.60	66.70	68.70	68.20
DHGNN (Liu et al., 2024b)	68.60 ± 0.90	65.80 ± 1.10	71.00 ± 1.30	71.80 ± 0.01
FDAA-Net (Zheng et al., 2024)	66.10 ± 0.15	–	–	64.20 ± 0.16
MSSTAN (Kong et al., 2025)	68.65 ± 0.90	**74.60 ± 73.26**	59.45 ± 4.84	67.06 ± 1.39
Transformer–Encoder (Dai et al., 2024)	68.61	69.78	67.93	–
APHGNN (Shen et al., 2024)	69.09 ± 5.97	54.08 ± 26.23	**77.83 ± 15.78**	66.64 ± 8.12
Ours	**71.15 ± 4.49**	70.38 ± 5.86	71.51 ± 4.36	**70.38 ± 4.50**

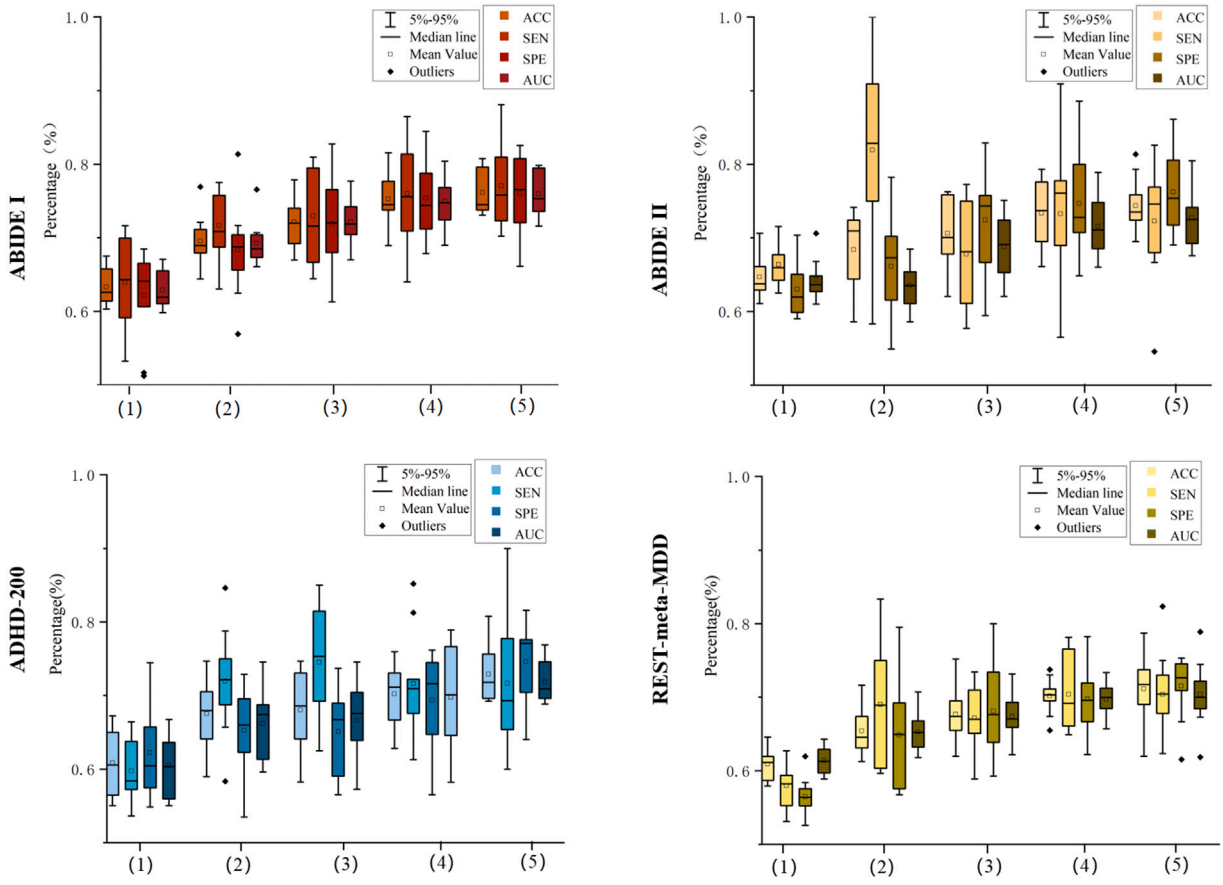


Fig. 3. Results of the ablation experiments conducted on four datasets. Each bar represents the mean ± standard deviation over 10-fold cross-validation for ACC, SEN, SPE and AUC. The x-axis denotes different ablation configurations corresponding to Table 6: (1). Baseline CNN: classification using raw dFCs without any proposed modules; (2). (1)+ Pseudo Label Generative module: refinement of dFCs using pseudo-label-guided Reconstruction FC via max-average pooling; (3). (2)+ T–F Intersect module: incorporation of temporal (Ts-S) and spatial (dFC) pseudo-label consistency for constructing Reconstruction FC; (4). (2)+ OE-DSCNN classifier: replacing the standard CNN with the proposed Feature Optimization Autoencoder and Depthwise Separable CNN architecture; (5). Full model (PL-FCDM): integration of all three modules: Pseudo Label Generative, T–F Intersect, and OE-DSCNN.

Table 6
Ablation study results of the proposed method across four datasets (ABIDE I, ABIDE II, ADHD-200 and REST-meta-MDD). The best performance for each dataset is highlighted in bold.

Dataset	Pseudo label	T–F intersect	OE-DSCNN	ACC	SEN	SPE	AUC
ABIDE I	x	x	x	63.30 ± 2.65	63.93 ± 6.84	62.14 ± 6.11	64.88 ± 2.68
	✓	x	x	69.56 ± 3.38	71.68 ± 4.52	68.35 ± 6.35	69.29 ± 3.01
	✓	✓	x	72.17 ± 3.73	72.94 ± 6.24	71.92 ± 6.45	72.19 ± 3.49
	✓	x	✓	75.27 ± 3.69	76.00 ± 7.17	75.40 ± 5.16	74.96 ± 3.54
	✓	✓	✓	**76.14 ± 2.95**	**77.04 ± 5.86**	**75.97 ± 5.40**	**75.94 ± 3.07**
ABIDE II	x	x	x	64.68 ± 2.78	66.39 ± 3.00	63.03 ± 3.82	64.31 ± 2.67
	✓	x	x	68.39 ± 5.11	81.93 ± 11.61	66.14 ± 6.62	63.56 ± 3.16
	✓	✓	x	70.61 ± 4.85	67.78 ± 6.92	72.45 ± 7.13	68.76 ± 4.50
	✓	x	✓	73.36 ± 4.36	**73.28** ± 9.65	74.64 ± 7.45	71.57 ± 4.30
	✓	✓	✓	**74.37 ± 3.62**	72.27 ± 8.03	**76.24 ± 5.88**	**72.75 ± 4.01**
ADHD-200	x	x	x	60.85 ± 4.45	58.99 ± 3.87	62.27 ± 6.98	61.36 ± 4.28
	✓	x	x	67.52 ± 4.98	71.95 ± 7.18	65.25 ± 6.07	66.22 ± 5.00
	✓	✓	x	68.03 ± 5.34	74.45 ± 8.01	65.13 ± 6.12	66.60 ± 5.38
	✓	x	✓	70.20 ± 4.34	71.59 ± 7.00	69.36 ± 6.48	69.74 ± 6.58
	✓	✓	✓	**72.89 ± 3.91**	**71.67 ± 9.16**	**74.61 ± 5.66**	**71.84 ± 2.81**
REST-meta-MDD	x	x	x	60.90 ± 2.34	57.96 ± 3.25	56.47 ± 2.64	61.41 ± 1.81
	✓	x	x	65.42 ± 3.09	69.04 ± 8.32	64.83 ± 7.42	65.29 ± 2.70
	✓	✓	x	67.69 ± 3.50	67.19 ± 4.53	68.13 ± 6.29	67.39 ± 3.02
	✓	x	✓	70.16 ± 2.44	**70.42** ± 5.11	69.77 ± 4.63	69.70 ± 2.28
	✓	✓	✓	**71.15 ± 4.49**	70.38 ± 5.86	**71.51 ± 4.36**	**70.38 ± 4.50**

by 6.26%, 3.71%, 6.67%, and 4.52%, respectively. These results demonstrate the effectiveness of this module in extracting high-quality features and enhancing classification performance.

(3) Effect of the T–F Intersect module: Building on the previous setup, the third experiment incorporated the T–F Intersect module, which combines temporal (Ts-S matrix) and spatial (dFCs) pseudo-label predictions for Reconstruction FC construction. Compared with the second experiment, the accuracy increased by 2.61%, 2.22%, 0.51%, and 2.05% on the ABIDE I, ABIDE II, ADHD, and MDD datasets, respectively. The results indicate that incorporating pseudo-label information from both temporal and spatial domains enhances the discriminability and robustness of the reconstructed FC representations.

(4) Effect of the OE-DSCNN: The influence of the proposed classification architecture was assessed by substituting the CNN employed in prior experiments with the OE-DSCNN model, with the remaining components held constant. A comparison between Experiments 2 and 4, as well as Experiments 3 and 5 (results shown in Table 6), reveals that the OE-DSCNN model consistently improves classification accuracy and achieves a better balance between sensitivity and specificity. This demonstrates its critical role in learning complex patterns in brain functional connectivity and enhancing overall performance.

4.3. Sliding window parameter optimal setting

To evaluate the impact of sliding window parameters on model performance, we conducted a systematic comparative analysis of the window length W and the step size S in our experiments. Specifically, nine parameter combinations were designed with $W \in \{40, 50, 60\}$ and $S \in \{1, 2, 3\}$, forming a 3×3 experimental matrix. For each parameter configuration, all other hyperparameters were kept constant, and the entire training and evaluation process was performed, with classification accuracy and other performance metrics recorded on the validation set. As shown in Fig. 4, the model exhibited significant differences in performance under different sliding window settings. Among them, the configuration with $W = 50$ and $S = 1$ achieved the best performance across multiple evaluation metrics, demonstrating superior stability and discriminative ability. This setting effectively captures dynamic changes in neural activity while balancing temporal resolution and data redundancy. Therefore, we adopted $W = 50$ and $S = 1$ as the final sliding window configuration in all subsequent experiments.

5. Discussion

5.1. Statistical analysis of reconstruction FC

In this study, we processed the Reconstruction FCs derived from the ABIDE I, ABIDE II, ADHD-200, and REST-meta-MDD datasets. Specifically, the upper triangular portion of each Reconstruction FC matrix was extracted and flattened into a one-dimensional vector to facilitate subsequent statistical analysis and pattern recognition. A two-sample t-test (Cressie and Whitford, 1986) was then conducted to evaluate the differences in functional connectivity between groups. Connections with p-values less than 0.05 were considered statistically significant. Based on the identified significant connections, a SVM classifier (Jakkula, 2006) was employed to perform classification analysis. To further validate the effectiveness of the Reconstruction FC, we used the original FC as a baseline for comparison, applying the same processing and classification pipeline. As illustrated in Fig. 5, the Reconstruction FC consistently outperformed the original FC when classified using a simple SVM, thereby demonstrating the superiority and effectiveness of the proposed feature extraction approach in functional connectivity modeling.

5.2. Visualization of the important brain regions

The FWE-corrected results not only enhance statistical robustness but also enable a more clinically grounded interpretation of functional dysconnectivity across neuropsychiatric disorders. Critically, several subcortical and limbic regions (most notably the caudate, pallidum, and insula) emerged as shared hubs of abnormal connectivity across ASD, ADHD, and MDD. This transdiagnostic pattern aligns with growing evidence supporting a shared neurobiological substrate underlying diverse psychiatric conditions. For instance, the caudate and pallidum, key components of the cortico-striato-thalamo-cortical circuits are consistently implicated in cognitive control, reward processing, and habit formation, functions that are broadly disrupted across diagnostic categories (Doss et al., 2022). Similarly, insular dysfunction has been linked to impaired interoception, emotional awareness, and salience detection in multiple disorders (Robbins et al., 2024). The recurrence of these regions in our FWE-corrected maps lends empirical support to dimensional models of psychopathology, which posit that core neural circuits underlie symptom dimensions such as cognitive inflexibility, emotional dysregulation, and impaired social functioning (see Fig. 6). At the same time, disorder specific patterns provide important clinical nuance. In ASD, abnormalities extended prominently into orbitofrontal, superior temporal pole, and thalamic regions, areas tied to social cognition and sensory integration, consistent with the disorder's hallmark deficits (Watanabe and Rees, 2017; Holiga et al., 2019). ADHD was characterized by dysconnectivity involving the paracentral lobule and orbitofrontal-insular-striatal circuits, reflecting disruptions in motor inhibition and attentional control (Konrad and Eickhoff, 2010). In MDD, widespread alterations spanned the mid-cingulate, amygdala, middle temporal cortex, and cerebellar Crus I/II, regions central to affective processing, self-referential thought, and cognitive-emotional integration (Kong et al., 2024; Dai et al., 2024). Notably, while the angular gyrus was not among the top regions in our current analysis, its frequent appearance in meta-analyses of transdiagnostic structural and functional alterations (Uddin, 2015). Future work with higher spatial resolution or alternative parcellations may further clarify its role. Together, these findings demonstrate that our data recapitulate both shared and distinct neural signatures predicted by contemporary transdiagnostic theories, thereby strengthening the clinical validity of the identified biomarkers and offering a circuit-level framework for understanding symptom overlap and differentiation across disorders.

Furthermore, overlapping regional activations, particularly in the Caudate, Pallidum, and Angular gyri, were identified across various neuropsychiatric disorders, indicating their potential role as common dysfunctional hubs. Such findings highlight both shared and distinct neurobiological mechanisms and contribute to the establishment of cross-disorder functional connectivity based biomarkers.

6. Limitations

Although this study leverages large-scale, publicly available datasets (ABIDE I/II, ADHD-200 and REST-meta-MDD) to enhance reproducibility, several methodological constraints warrant consideration. First, the estimation of dFC is highly sensitive to preprocessing choices and parcellation schemes. As demonstrated in Section 4.3 and Fig. 4, model performance varies across different dFC configurations; although we evaluated nine parameter combinations per dataset and selected a window length of 50 TRs with a step size of 1 as the optimal setting, the model's dependence on specific analytical pipelines may still limit its generalizability to data processed with alternative protocols. Second, our approach relies on a data-driven pseudo-labeling strategy that aligns temporal (time-segment similarity) and spatial (dFC) features to generate training labels. The quality of these pseudo-labels directly affects sample selection in the T–F Intersect Module and the fidelity of the reconstructed functional connectivity, which influence the training and predictive performance of the OE-DSCNN classifier. Despite efforts

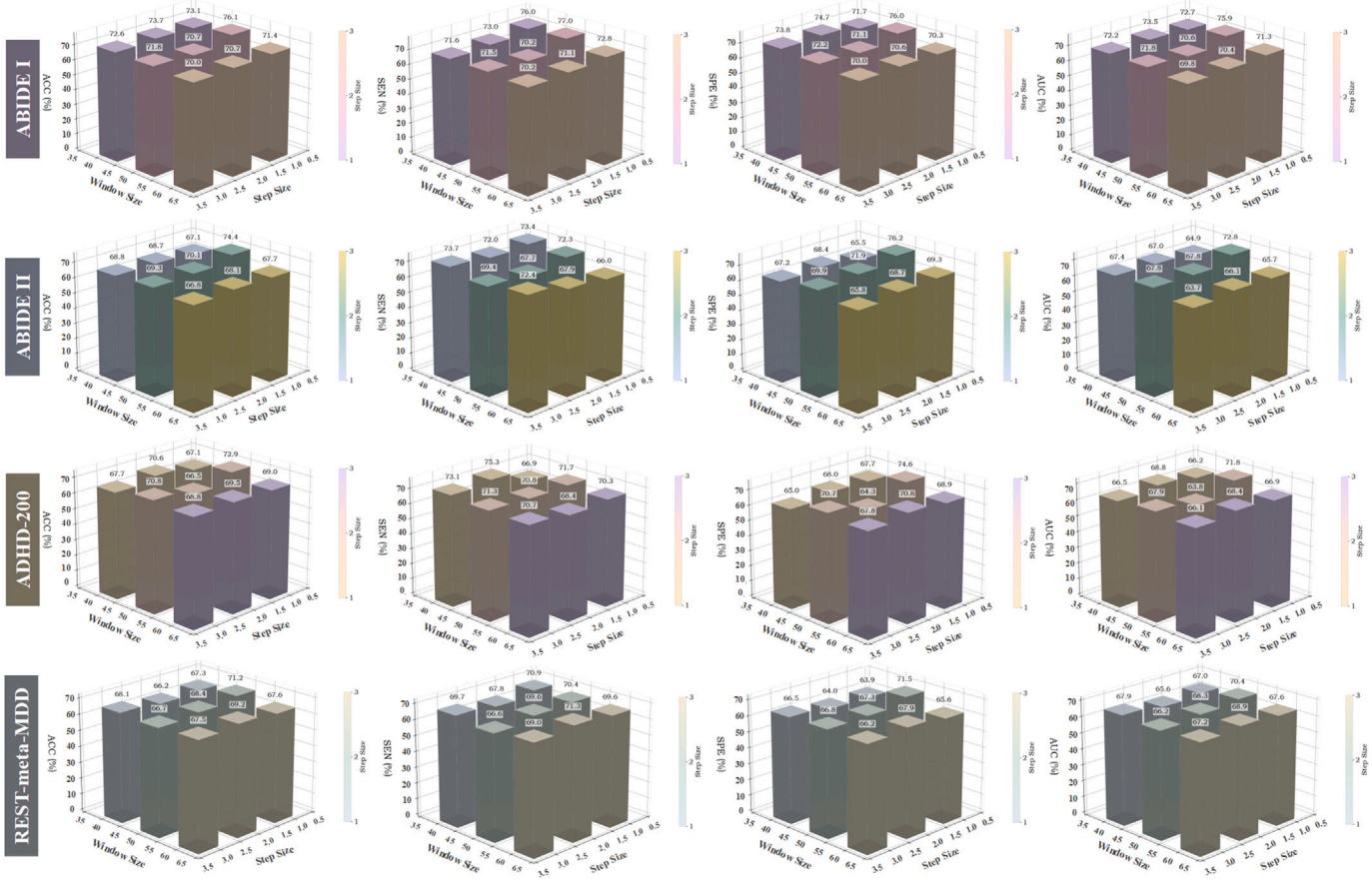


Fig. 4. Classification results of ACC, SEN, SPE, and AUC at different sliding window parameters on the ABIDE I, ABIDE II, ADHD-200, and REST-meta-MDD datasets.

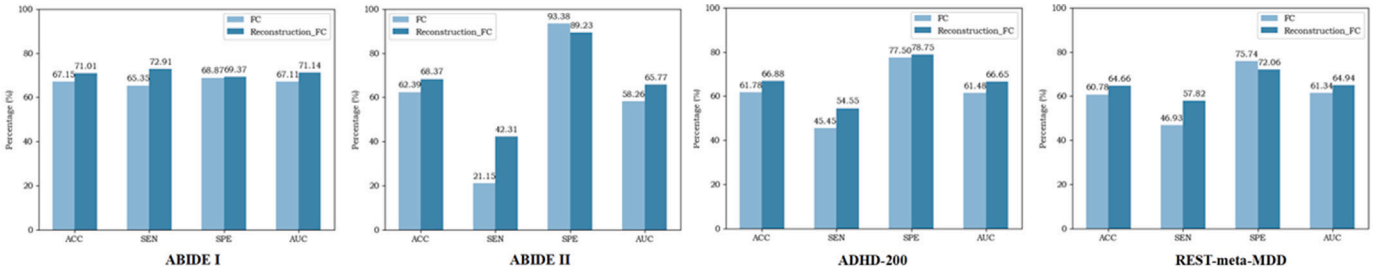


Fig. 5. SVM classification results for the original FC and Reconstruction FC from the ABIDE I, ABIDE II, ADHD-200, and REST-meta-MDD datasets.

to ensure temporal–spatial consistency through pre-trained generative models, residual noise or mislabeling may introduce bias and constrain discriminative power. Moreover, heterogeneity in sample sizes across diagnostic categories and age groups could further compromise model calibration and cross-dataset generalizability, especially in low-data or external validation scenarios.

7. Conclusion

This study presents a pseudo-label consistency-driven framework that integrates temporal and spatial features for sample purification and classification, effectively addressing the challenges of noise and sample quality in dFC. By generating pseudo-labels from both time-series-based features (Ts-S) and dFC representations, and applying consistency filtering, the reliability of pseudo-labels and the robustness of training samples are significantly improved. The reconstruction of high-quality static functional connectivity maps, in conjunction with the OE-DSCNN, enables accurate identification of multiple neuropsychiatric disorders. Extensive experiments conducted on four large-scale datasets demonstrate the framework's superior performance and strong generalization ability. Visualization of discriminative brain regions confirms the model's neurobiological plausibility and highlights common abnormal areas across disorders, offering valuable insights into shared and distinct pathological mechanisms. However, the underlying functional mechanisms associated with these regions warrant further investigation through dedicated neuroimaging and neuroscience studies. The proposed framework offers a promising tool for automated diagnosis and functional connectivity modeling in neuropsychiatry, with potential extensions to multimodal data fusion and cross-disorder analysis.

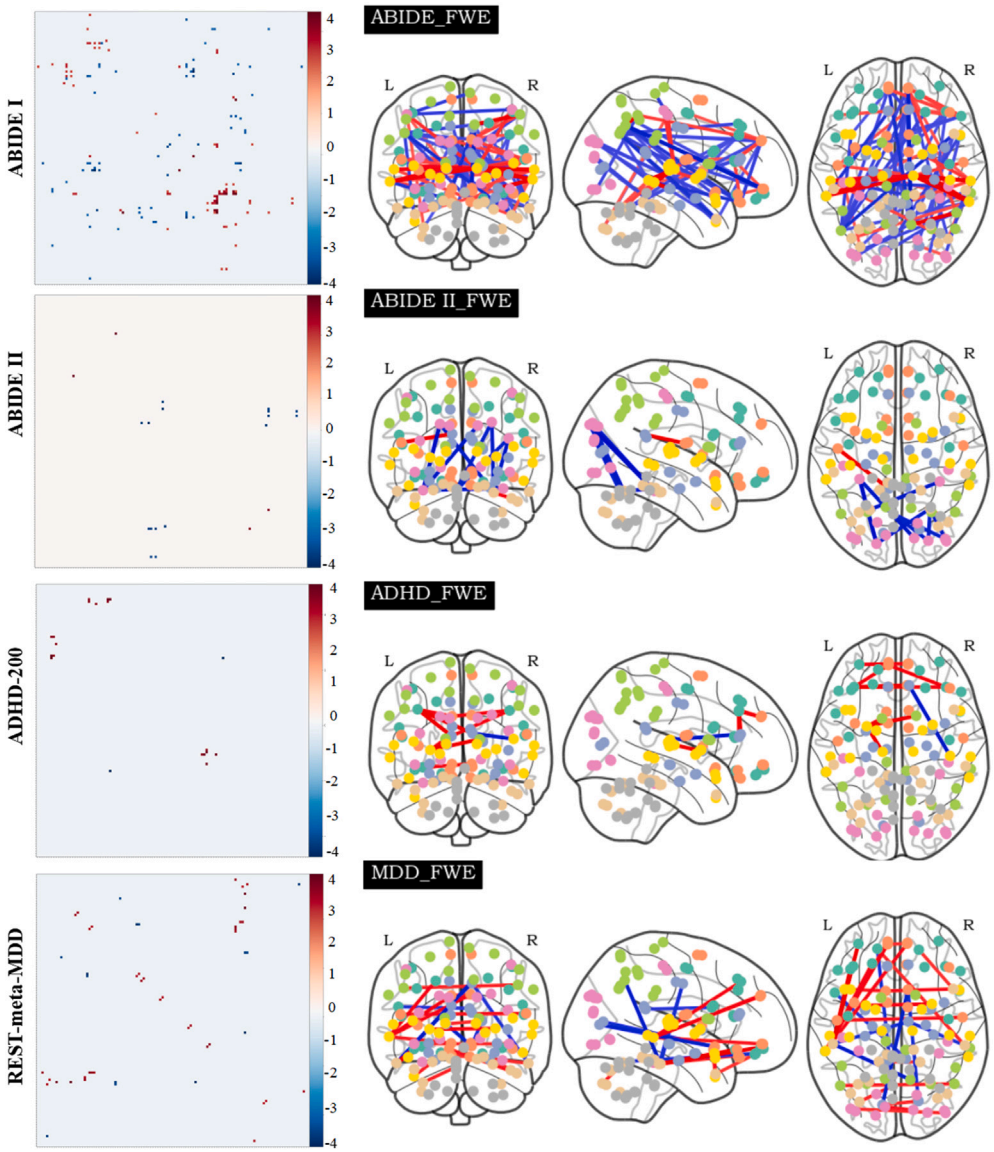


Fig. 6. Visualization of brain regions that remained statistically significant after FWE correction.

CRediT authorship contribution statement

Xin Wen: Writing – review & editing, Writing – original draft, Funding acquisition. **Shijie Guo:** Writing – original draft, Methodology, Data curation. **Li Dong:** Writing – review & editing, Supervision, Conceptualization. **Xiaobo Liu:** Methodology, Formal analysis. **Wenbo Ning:** Visualization, Resources, Data curation. **Jie Shi:** Visualization, Formal analysis, Data curation. **Songhua Liu:** Resources, Funding acquisition. **Cheng Luo:** Validation, Conceptualization. **Rui Cao:** Writing – review & editing, Supervision, Conceptualization. **Dezhong Yao:** Supervision.

Declaration of competing interest

The authors declare no competing interests.

Acknowledgments

This work was supported by the National Natural Science Foundation of China (62206196) and Shanxi Province Natural Science Foundation Program, China (202203021211144, 202403021222057).

Data availability

All datasets used in this work are publicly available, and their access URLs are provided in the main text.

References

Alam, Sumbul, Raja, Pravinth, Gulzar, Yonis, 2022. Investigation of machine learning methods for early prediction of neurodevelopmental disorders in children. Wirel. Commun. Mob. Comput. 2022 (1), 5766386.

Ali, Muhaddisa Barat, Gu, Irene Yu-Hua, Berger, Mitchel S., Pallud, Johan, Southwell, Derek, Widhalm, Georg, Roux, Alexandre, Vecchio, Tomás Gomez, Jakola, Asgeir Store, 2020. Domain mapping and deep learning from multiple MRI clinical datasets for prediction of molecular subtypes in low grade gliomas. Brain Sciences 10 (7).

Bellec, Pierre, Chu, Carlton, Chouinard-Decorte, Francois, Benhajali, Yassine, Margulies, Daniel S, Craddock, R Cameron, 2017. The neuro bureau ADHD-200 preprocessed repository. Neuroimage 144, 275–286.

Canario, Edgar, Chen, Donna, Biswal, Bharat, 2021. A review of resting-state fMRI and its use to examine psychiatric disorders. Psychoradiology 1 (1), 42–53.

Chen, Xiao, Lu, Bin, Li, Hui-Xian, Li, Xue-Ying, Wang, Yu-Wei, Castellanos, Francisco Xavier, Cao, Li-Ping, Chen, Ning-Xuan, Chen, Wei, Cheng, Yu-Qi, Cui, Shi-Xian, Deng, Zhao-Yu, Fang, Yi-Ru, Gong, Qi-Yong, Guo, Wen-Bin, Hu, Zheng-Jia-Yi, Kuang, Li, Li, Bao-Juan, Li, Le, Li, Tao, Lian, Tao, Liao, Yi-Fan, Liu, Yan-Song,

Liu, Zhe-Ning, Lu, Jian-Ping, Luo, Qing-Hua, Meng, Hua-Qing, Peng, Dai-Hui, Qiu, Jiang, Shen, Yue-Di, Si, Tian-Mei, Tang, Yan-Qing, Wang, Chuan-Yue, Wang, Fei, Wang, Hua-Ning, Wang, Kai, Wang, Xiang, Wang, Ying, Wang, Zi-Han, Wu, Xiao-Ping, Xie, Chun-Ming, Xie, Guang-Rong, Xie, Peng, Xu, Xiu-Feng, Yang, Hong, Yang, Jian, Yao, Shu-Qiao, Yu, Yong-Qiang, Yuan, Yong-Gui, Zhang, Ke-Rang, Zhang, Wei, Zhang, Zhi-Jun, Zhu, Jun-Juan, Zuo, Xi-Nian, Zhao, Jing-Ping, Zang, Yu-Feng, the DIRECT consortium, Yan, Chao-Gan, 2022. The DIRECT consortium and the REST-meta-MDD project: towards neuroimaging biomarkers of major depressive disorder. Psychoradiol. 2 (1), 32–42.

Cheng, Bo, Liu, Mingxia, Suk, Heung Il, Shen, Dinggang, Zhang, Daoqiang, 2015. Multimodal manifold-regularized transfer learning for MCI conversion prediction. Brain Imaging Behav. 9 (4), 913–926.

Consortium, The Adhd, 2012. The ADHD-200 consortium: A model to advance the translational potential of neuroimaging in clinical neuroscience. Front. Syst. Neurosci. 6, 62.

Craddock, Cameron, Benhajali, Yassine, Chu, Carlton, Chouinard, Francois, Evans, Alan, Jakab, András, Khundrakpam, Budhachandra Singh, Lewis, John David, Li, Qingyang, Milham, Michael, et al., 2013. The neuro bureau preprocessing initiative: open sharing of preprocessed neuroimaging data and derivatives. Front. Neuroinform. 7 (27), 5.

Cressie, N.A.C., Whitford, H.J., 1986. How to use the two sample t-test. Biom. J. 28 (2), 131–148.

Dai, Peishan, Zhou, Ying, Shi, Yun, Lu, Da, Chen, Zailiang, Zou, Beiji, Liu, Kun, Liao, Shenghui, Consortium, Rest Meta Mdd, 2024. Classification of MDD using a transformer classifier with large-scale multisite resting-state fMRI data. Hum. Brain Mapp. 45 (1), 14.

Degirmenci, Murside, Surucu, Murat, Perc, Matjaž, Isler, Yalcin, 2025. Convolutional neural networks can diagnose schizophrenia. J. Comput. Sci. 90, 102634.

Di Martino, Adriana, Yan, Chao-Gan, Li, Qingyang, Denio, Ellen, Castellanos, F Xavier, Alaerts, Kaat, Anderson, Jeffrey S, Assaf, Michal, Bookheimer, Susan Y, Dapretto, Mirella, Deen, Ben, Delmonte, Sara, Dinstein, Ilan, Ertl-Wagner, Birgit, Fair, Damien A, Gallagher, Louise, Kennedy, Daniel P, Keown, Christopher L, Keysers, Christian, Lainhart, Janet E, Lord, Catherine, Luna, Beatriz, Menon, Vinod, Minshew, Nancy J, Monk, Christopher S, Mueller, Susanne, Muller, Ralph-Axel, Nebel, Mary Beth, Nigg, Joel T, O'Hearn, Kirsten, Pelphrey, Kevin A, Peltier, Scott J, Rudie, Jeffrey D, Sunaert, Stefan, Thioux, Marc, Tyszka, J Michael, Uddin, Lucina Q, Verhoeven, Judith S, Wigg, Karen, Wiggins, Jillian L, Mostofsky, Stewart H, Milham, Michael P, 2014. The autism brain imaging data exchange: towards a large-scale evaluation of the intrinsic brain architecture in autism. Molecular Psychiatry 19 (6), 659–667.

Dinsdale, Nicola K., Jenkinson, Mark, Namburete, Ana I.L., 2020. Unlearning scanner bias for MRI harmonisation in medical image segmentation. In: Medical Image Understanding and Analysis: 24th Annual Conference, MIUA 2020, Oxford, UK, July 15-17, 2020, Proceedings 24. Springer, pp. 15–25.

Doi, K., 2007. Computer-aided diagnosis in medical imaging: Historical review, current status and future potential.. Comput. Med. Imaging Graph.: Off. Jounal Med. Imaging Soc. (4/5), 31.

Dosovitskiy, Alexey, Beyer, Lucas, Kolesnikov, Alexander, Weissenborn, Dirk, Zhai, Xiaohua, Unterthiner, Thomas, Dehghani, Mostafa, Minderer, Matthias, Heigold, Georg, Gelly, Sylvain, Uszkoreit, Jakob, Houlsby, Neil, 2021. An image is worth 16x16 words: Transformers for image recognition at scale. In: International Conference on Learning Representations. ICLR.

Doss, Manoj K, Madden, Maxwell B, Gaddis, Andrew, Nebel, Mary Beth, Griffiths, Roland R, Mathur, Brian N, Barrett, Frederick S, 2022. Models of psychedelic drug action: modulation of cortical-subcortical circuits. Brain 145 (2), 441–456.

Du, Yuhui, Fang, Songke, He, Xingyu, Calhoun, Vince D., 2024. A survey of brain functional network extraction methods using fmri data. Trends Neurosci..

Hailong, Li, Parikh, Nehal A., Lili, He, 2018. A novel transfer learning approach to enhance deep neural network classification of brain functional connectomes. Front. Neurosci. 12.

Hindriks, Rikkert, Adhikari, Mohit H, Murayama, Yusuke, Ganzetti, Marco, Mantini, Dante, Logothetis, Nikos K, Deco, Gustavo, 2016. Can sliding-window correlations reveal dynamic functional connectivity in resting-state fmri? Neuroimage 127, 242–256.

Holiga, Štefan, Hipp, Joerg F, Chatham, Christopher H, Garces, Pilar, Spooren, Will, D'Ardhuy, Xavier Liogier, Bertolino, Alessandro, Bouquet, Céline, Buitelaar, Jan K, Bours, Carsten, et al., 2019. Patients with autism spectrum disorders display reproducible functional connectivity alterations. Sci. Transl. Med. 11 (481), eaat9223.

Hu, Yao, Huang, Zhi-An, Liu, Rui, Xue, Xiaoming, Sun, Xiaoyan, Song, Linqi, Tan, Kay Chen, 2023. Source free semi-supervised transfer learning for diagnosis of mental disorders on fmri scans. IEEE Trans. Pattern Anal. Mach. Intell..

Huang, Zhi-An, Hu, Yao, Liu, Rui, Xue, Xiaoming, Zhu, Zexuan, Song, Linqi, Tan, Kay Chen, 2022. Federated multi-task learning for joint diagnosis of multiple mental disorders on MRI scans. IEEE Trans. Biomed. Eng. 70 (4), 1137–1149.

Huang, Zhi-An, Liu, Rui, Tan, Kay Chen, 2020. Multi-task learning for efficient diagnosis of ASD and ADHD using resting-state fMRI data. In: 2020 International Joint Conference on Neural Networks. IJCNN, IEEE, pp. 1–7.

İlikhan, Sevil Uygun, Özer, Mahmut, Perc, Matjaz, Tanberkan, Hande, Ayhan, Yavuz, Complementary use of artificial intelligence in healthcare. Medical Journal of Western Black Sea, 9(1) 7–17.

Imambi, Sagar, Prakash, Kolla Bhanu, Kanagachidambaresan, G.R., 2021. Pytorch. Program. TensorFlow: Solut. Edge Comput. Appl. 87–104.

Ingalhalikar, Madhura, Shinde, Sumeet, Karmarkar, Arnav, Rajan, Archith, Rangaprakash, D, Deshpande, Gopikrishna, 2021. Functional connectivity-based prediction of autism on site harmonized ABIDE dataset. IEEE Trans. Biomed. Eng. 68 (12), 3628–3637.

Jakkula, Vikramaditya, 2006. Tutorial on Support Vector Machine (Svm). Vol. 37, School of EECS, Washington State University, p. 3.

Jiang, Hao, Cao, Peng, Xu, MingYi, Yang, Jinzhu, Zaiane, Osmar, 2020. Hi-GCN: A hierarchical graph convolution network for graph embedding learning of brain network and brain disorders prediction. Comput. Biol. Med. 127, 104096.

Kang, Li, Jiang, Jingwan, Huang, Jianjun, Zhang, Tijiang, 2020. Identifying early mild cognitive impairment by multi-modality MRI-based deep learning. Front. Aging Neurosci. 12, 206.

Karahanoğlu, Fikret Işik, Van De Ville, Dimitri, 2015. Transient brain activity disentangles fMRI resting-state dynamics in terms of spatially and temporally overlapping networks. Nat. Commun. 6 (1), 7751.

Kawahara, Jeremy, Brown, Colin J, Miller, Steven P, Booth, Brian G, Chau, Vann, Grunau, Ruth E, Zwicker, Jill G, Hamarneh, Ghassan, 2017. BrainNetCNN: Convolutional neural networks for brain networks; towards predicting neurodevelopment. NeuroImage 146, 1038–1049.

Kong, Youyong, Zhang, Xiaotong, Wang, Wenhan, Zhou, Yue, Li, Yueying, Yuan, Yonggui, 2024. Multi-scale spatial-temporal attention networks for functional connectome classification. IEEE Trans. Med. Imaging.

Kong, Youyong, Zhang, Xiaotong, Wang, Wenhan, Zhou, Yue, Li, Yueying, Yuan, Yonggui, 2025. Multi-scale spatial-temporal attention networks for functional connectome classification. IEEE Trans. Med. Imaging 44 (1), 475–488.

Konrad, Kerstin, Eickhoff, Simon B., 2010. Is the ADHD brain wired differently? A review on structural and functional connectivity in attention deficit hyperactivity disorder. Hum. Brain Mapp. 31 (6).

Kouw, Wouter Marco, Loog, Marco, 2019. A review of domain adaptation without target labels. IEEE Trans. Pattern Anal. Mach. Intell. PP (99).

Lauritsen, Marlene Briciet, 2013. Autism spectrum disorders. Eur. Child Adolesc. Psychiatry 22, 37–42.

Li, Jia, Liu, Gan, Ji, Junzhong, 2024. Pre-training and fine-tuning transformer for brain network classification. In: 2024 IEEE International Conference on Medical Artificial Intelligence (MedAI). pp. 136–144.

Li, Wenhui, Wang, Shiyuan, Liu, Guangyuan, 2022. Transformer-based model for fMRI data: ABIDE results. In: 2022 7th International Conference on Computer and Communication Systems. ICCCS, pp. 162–167.

Li, Xiaoxiao, Zhou, Yuan, Dvornek, Nicha, Zhang, Muhan, Gao, Siyuan, Zhuang, Juntang, Scheinost, Dustin, Staib, Lawrence H, Ventola, Pamela, Duncan, James S, 2021. Braingnn: Interpretable brain graph neural network for fmri analysis. Med. Image Anal. 74, 102233.

Liu, Rui, Hu, Yao, Wu, Jibin, Wong, Ka Chun, Huang, Zhi An, Huang, Yu An, Tan, Kay Chen, 2025. Dynamic graph representation learning for spatio-temporal neuroimaging analysis. IEEE Trans. Cybern. 55 (3), 1121–1134.

Liu, Rui, Huang, Zhi-An, Hu, Yao, Zhu, Zexuan, Wong, Ka-Chun, Tan, Kay Chen, 2024a. Spatial–temporal co-attention learning for diagnosis of mental disorders from resting-state fMRI data. IEEE Trans. Neural Netw. Learn. Syst. 35 (8), 10591–10605.

Liu, Lingwen, Wen, Guangqi, Cao, Peng, Hong, Tianshun, Yang, Jinzhu, Zhang, Xizhe, Zaiane, Osmar R, 2023. BrainTGL: A dynamic graph representation learning model for brain network analysis. Comput. Biol. Med. 153, 106521.

Liu, Jingyu, Yang, Wenxin, Ma, Yulan, Dong, Qunxi, Li, Yang, Hu, Bin, Consortium, DIRECT, et al., 2024b. Effective hyper-connectivity network construction and learning: Application to major depressive disorder identification. Comput. Biol. Med. 171, 108069.

Luo, Yiqian, Chen, Qiurong, Li, Fali, Yi, Liang, Xu, Peng, Zhang, Yangsong, 2025. Hierarchical feature extraction on functional brain networks for autism spectrum disorder identification with resting-state fMRI data. Neural Netw. 188.

Microsoft, 2021. Neural Network Intelligence. https://github.com/microsoft/nni.

Otte, Christian, Gold, Stefan M, Penninx, Brenda W, Pariante, Carmine M, Etkin, Amit, Fava, Maurizio, Mohr, David C, Schatzberg, Alan F, 2016. Major depressive disorder. Nat. Rev. Dis. Prim. 2 (1), 1–20.

Robbins, Trevor W., Banca, Paula, Belin, David, 2024. From compulsivity to compulsion: the neural basis of compulsive disorders. Nature Reviews Neuroscience 25 (5), 313–333.

Rolls, Edmund T, Huang, Chu-Chung, Lin, Ching-Po, Feng, Jianfeng, Joliot, Marc, 2020. Automated anatomical labelling atlas 3. Neuroimage 206, 116189.

Shao, Lizhen, Fu, Cong, You, Yang, Fu, Dongmei, 2021. Classification of ASD based on fMRI data with deep learning. Cogn. Neurodynamics 15 (6), 961–974.

Shen, Guangfeng, Zeng, Weiming, Yang, Jiajun, 2024. Research on migraine classification model based on hypergraph neural network. J. Supercomput. 80 (17), 25403–25423.

Sivaranjini, S., Sujatha, C.M., 2020. Deep Learning Based Diagnosis of Parkinson's Disease Using Convolutional Neural Network. Springer US.

Taud, Hind, Mas, Jean-Franccois, 2017. Multilayer perceptron (MLP). In: Geomatic Approaches for Modeling Land Change Scenarios. Springer, pp. 451–455.

Thapar, Anita, Cooper, Miriam, Rutter, Michael, 2017. Neurodevelopmental disorders. Lancet Psychiatry 4 (4), 339–346.

Tripp, Gail, Wickens, Jeffery R., 2009. Neurobiology of ADHD. Neuropharmacology 57 (7–8), 579–589.
Uddin, Lucina Q., 2015. Salience processing and insular cortical function and dysfunction. Nature Rev. Neurosci. 16 (1), 55–61.
Wang, Lilong, Liu, Mianxin, Zhang, Shaoting, Wang, Xiaosong, 2024. Brainsck: Brain structure and cognition alignment via knowledge injection and reactivation for diagnosing brain disorders. In: International Conference on Medical Image Computing and Computer-Assisted Intervention. Springer, pp. 45–55.
Wang, Weizheng, Qiao, Hong, 2024. How Natural Language Processing Enables AIGC Recognition? –Latest Trends and Future Prospects. ICSIM '24, Association for Computing Machinery, New York, NY, USA, pp. 103–109.
Wang, Jun, Zhang, Lichi, Wang, Qian, Chen, Lei, Shen, Dinggang, 2020. Multi-class ASD classification based on functional connectivity and functional correlation tensor via multi-source domain adaptation and multi-view sparse representation. IEEE Trans. Med. Imaging PP (99).
Watanabe, Takamitsu, Rees, Geraint, 2017. Brain network dynamics in high-functioning individuals with autism. Nat. Commun. 8, 16048.
Wen, Xin, Dong, Li, Chen, Junjie, Xiang, Jie, Yao, Dezhong, 2020. Detecting the information of functional connectivity networks in normal aging using deep learning from a big data perspective. Front. Neurosci. 13, 1435.
World Health Organization, 2006. Neurological Disorders: Public Health Challenges. World Health Organization.
Yesilkaya, Bartu, Perc, Matjaž, Isler, Yalcin, 2022. Manifold learning methods for the diagnosis of ovarian cancer. J. Comput. Sci. 63, 101775.
Yin, Wutao, Li, Longhai, Wu, Fang-Xiang, 2022. A semi-supervised autoencoder for autism disease diagnosis. Neurocomputing 483, 140–147.
Zheng, Weihao, Bao, Cong, Mu, Renhui, Wang, Jun, Li, Tongtong, Zhao, Ziyang, Yao, Zhijun, Hu, Bin, 2024. Frequency-specific dual-attention based adversarial network for blood oxygen level-dependent time series prediction. Hum. Brain Mapp. 45 (14).
Zhu, Peng, Chen, Xiaotian, Zhang, Ziyi, Li, Peishan, Cheng, Xi, Dai, Yucheng, 2025a. AI-driven hypergraph neural network for predicting gasoline price trends. Energy Econ. 108895.
Zhu, Peng, Zhang, Han, Shi, Yunsheng, Xie, Wanli, Pang, Mingyong, Shi, Yuhui, Hens, Luc, 2025b. A novel discrete conformable fractional grey system model for forecasting carbon dioxide emissions. Environ. Dev. Sustain. 27 (6).