Multi-Atlas Static–Dynamic Collaborative Diagnostic Network for Autism

Jingxia Chen [1], Yuhan Shi [1], Huiru Zheng[2*],Pengwei Zhang[1],Haifeng Chen[1]
[1] School of Electronic Information and Artificial Intelligence, Shaanxi University of Science and Technology
[2] School of Computing, Ulster University

***Abstract*—Functional brain networks derived from functional magnetic resonance imaging (fMRI) exhibit complex spatiotemporal patterns that are critical for accurate identification of autism spectrum disorder (ASD). Existing diagnostic methods primarily rely on single-atlas modelling and static functional connectivity, which limits their ability to capture complementary information across different brain parcellation schemes as well as the temporal variability of brain activity. To address these limitations, we propose a static–dynamic multi-atlas feature fusion network for automatic ASD diagnosis. The proposed framework incorporates static and dynamic functional connectivity features to comprehensively characterise the temporal diversity in brain functional activity. In addition, a collaborative multi-atlas feature fusion strategy is introduced to adaptively integrate feature representations from different brain atlases, thereby effectively exploiting complementary information under diverse spatial partitioning schemes and enhancing the discriminative representation of functional brain networks. Extensive experiments on the public ABIDE I dataset demonstrate that the proposed method achieves a classification accuracy of 81.64% and an AUC of 81.53%, outperforming existing baseline approaches. Furthermore, analysis of discriminative brain regions highlights the interpretability of the proposed multi-atlas fusion framework and provides new neurobiologically meaningful insights into ASD-related functional patterns.**

***Index Terms*—Multi-atlas; Feature Fusion; Functional Magnetic Resonance Imaging (fMRI); Functional Brain Network; Cross-attention; Autism Spectrum Disorder Classification**

I. Introduction

Neurodevelopmental disorders have become an increasingly prominent public health concern in modern society, exerting long-term impacts on individual development and family well-being. Autism Spectrum Disorder (ASD), as a representative neurodevelopmental condition, is characterised by early onset, pronounced clinical heterogeneity, and lifelong persistence. Individual with ASD often exhibit substantial impairments in social interaction, communication and cognitive functioning, underscoring the urgent need for objective, efficient, and automated diagnostic approaches to facilitate early identification and intervention [1].

Functional magnetic resonance imaging (fMRI) offers a non-invasive means to investigate whole-brain functional activity by measuring blood oxygen level–dependent (BOLD) signal fluctuations. Owing to its high spatial resolution and comprehensive brain coverage, fMRI has become a core neuroimaging modality for exploring functional abnormalities and disrupted connectivity patterns associated with ASD [2]. Early fMRI-based ASD identification studies primarily relied on traditional machine learning methods, in which handcrafted features were extracted from functional connectivity matrices or region-wise statistical descriptors and subsequently classified using models such as support vector machines (SVMs) and random forests (RFs) [3]. Although these approaches achieved moderate diagnostic performance, their strong dependence on manual feature engineering and prior domain knowledge limited their ability to capture the complex, nonlinear, and highly heterogeneous brain functional patterns characteristic of neurodevelopmental disorders.

With recent advances in deep learning, convolutional neural networks (CNNs), recurrent neural networks (RNNs), and Transformer-based architectures have been increasingly applied to fMRI analysis, enabling end-to-end representation learning and alleviating the reliance on handcrafted features. For example, Heinsfeld et al. [4] employed three-dimensional CNNs to model resting-state fMRI data, while Dvornek et al. [5] utilised long short-term memory (LSTM) networks to capture temporal dependencies in fMRI time series, both demonstrating improved classification performance on the ABIDE dataset. Nevertheless, these approaches typically treat functional connectivity matrices or fMRI time series as Euclidean-structured data and rely primarily on convolutional or sequential modelling paradigms, thereby limiting their capacity to explicitly encode the intrinsic topological organisation and inter-regional interactions of brain functional networks [6].

Graph neural networks (GNNs) provide a principled framework for modelling brain functional networks as graphs, enabling the explicit characterisation of inter-regional connectivity patterns and nonlinear topological interactions [7]. A growing body of research has demonstrated that GNN-based models can learn more discriminative and interpretable brain network representations for ASD identification [8]. However, most existing GNN approaches construct functional connectivity graphs using a single brain atlas, which constrains their ability to exploit complementary information across different parcellation schemes and limits comprehensive multi-scale modelling of brain functional organisation [9]. To mitigate this limitation, multi-atlas strategies have attracted increasing attention by integrating multiple brain parcellations to characterise functional organisation from diverse anatomical

*Corresponding author:Huiru Zheng (h.zheng@ulster.ac.uk).

This work was supported by the National Natural Science Foundation of China (61806118, 62306172) and the Key Research and Development Program of Shaanxi Province (2025CY-YBXM-191).

and functional perspectives. Previous studies have shown that multi-atlas fusion can significantly improve ASD classification performance and reproducibility compared with single-atlas methods [10]. Despite these advances, current multi-atlas GNN models predominantly rely on static functional connectivity representations and adopt relatively coarse cross-atlas fusion strategies, leaving complementary information embedded in temporal dynamics and fine-grained structural representations insufficiently exploited [11].

To address these limitations, we propose a Static–Dynamic Multi-Atlas Fusion Network (SD-MAFNet) for fMRI-based ASD identification. The proposed framework jointly models multi-atlas functional connectivity networks by integrating static and dynamic representations, enabling comprehensive characterisation of inter-regional spatiotemporal interaction patterns. Within a unified learning framework, multiple graph structures derived from different brain atlases are collaboratively optimised and fused, facilitating effective exploitation of intrinsic correlations across parcellation schemes. The resulting discriminative representations are subsequently used for downstream ASD classification, leading to improved diagnostic performance, robustness, and generalisation capability.

The main contributions of this work are summarised as follows:

1. We propose a static–dynamic multi-atlas feature fusion framework that jointly models static functional connectivity and dynamic fMRI time-series information, enabling effective capture of complementary spatiotemporal characteristics of brain functional interactions across multiple brain atlases.

2. We design a collaborative multi-atlas fusion strategy incorporating adaptive atlas weighting and cross-attention mechanisms to explicitly model inter-atlas complementarity, thereby enhancing multi-scale brain functional representation capacity and model robustness.

3. Extensive experiments on the ABIDE I dataset demonstrate that the proposed method consistently outperforms existing single-atlas and static multi-atlas approaches in ASD identification, validating the effectiveness of static–dynamic joint modelling and multi-atlas feature fusion.

II. Method

This section presents the proposed SD-MAFNet, including its architectural components and learning procedure. As illustrated in Fig. 1, the overall framework is organised into three stages. First, fMRI data are parcellated using three standard brain atlases—AAL [12], Harvard–Oxford (HO), and CC200 [13]—which exhibit complementary characteristics in terms of parcellation granularity and spatial organization, enabling the extraction of multi-scale, atlas-specific BOLD signals. Second, static spatial and dynamic temporal representations are extracted via a Static Spatial Feature Extraction Network (SSFE-Net) and a Dynamic Temporal Feature Extraction Network (DTFE-Net), respectively. Finally, a hierarchical feature fusion strategy is employed to integrate static and dynamic representations within each atlas as well as representations across multiple atlases, and the fused features are fed into an MLP classifier for ASD classification.

A. Static Feature Extraction

In this study, we propose a joint feature extraction strategy that integrates static and dynamic information across multiple brain atlases to comprehensively characterise both the stable

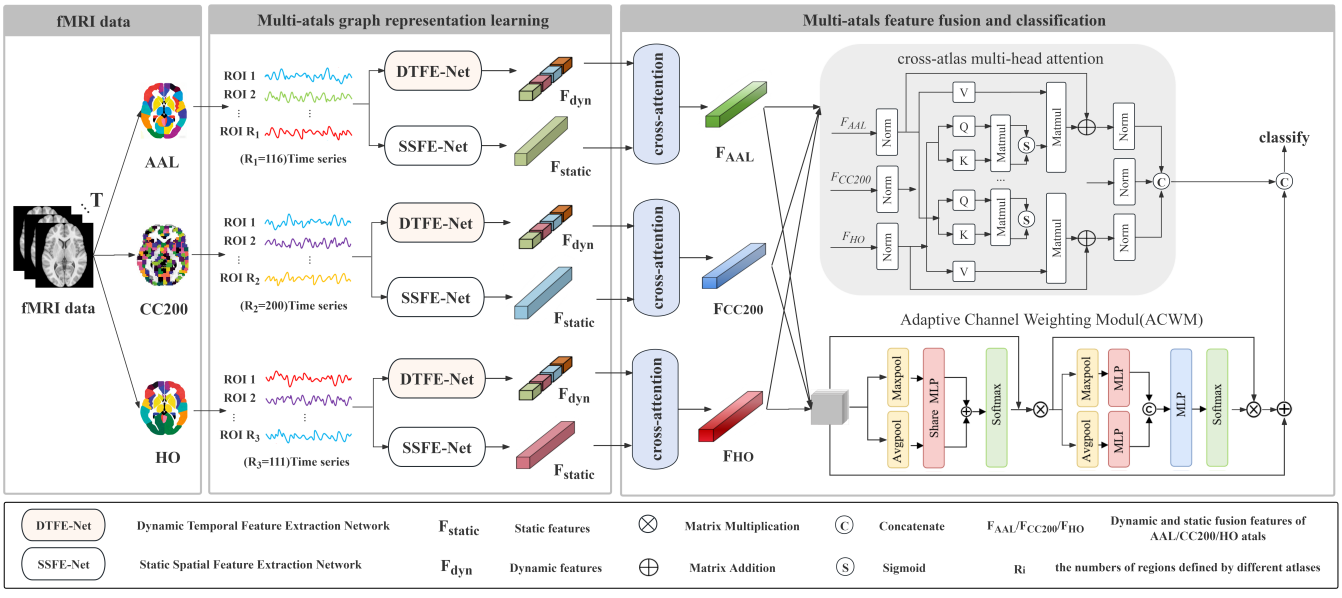

Fig. 1. SD-MAFNet framework. The brain is parcellated using three atlases, and region-level time series are extracted. These time series are fed into the multi-atlas representation learning module to obtain dynamic and static features for each atlas. Finally, all features are input into the Multi-atals feature fusion and classification module, where intra-atlas dynamic–static fusion and inter-atlas fusion are performed for final prediction.

and time-varying properties of brain function. The static feature extraction process based on SSFE-Net is illustrated in Fig. 2. For static feature extraction, each subject's brain is first parcellated using multiple atlases. For each atlas, regional fMRI time series are extracted and used to compute functional connectivity (FC) matrices based on Pearson correlation. Since FC matrices are symmetric, only the upper triangular elements are retained and vectorized as feature representations. To reduce redundancy and avoid overfitting caused by high-dimensional features, recursive feature elimination is further applied, and the feature dimension of each atlas is fixed to $d = 1500$.

To model inter-subject relationships, a population graph is constructed where each subject is treated as a node and node features correspond to the selected FC vectors. The edge weights are defined by combining phenotypic similarity and feature similarity. Specifically, we first construct a phenotypic affinity matrix $A_{pheno} \in \mathbb{R}^{N\times N}$ based on demographic information, where N is the number of subjects. Higher affinity is assigned to subject pairs with smaller age differences or sharing the same gender. Meanwhile, a feature similarity matrix $S_{feat} \in \mathbb{R}^{N\times N}$ is computed using a Gaussian kernel on node features. The final adjacency matrix is then obtained by element-wise multiplication of the two matrices, followed by thresholding to remove weak connections.

To further exploit complementary information across atlases, we construct multiple population graphs using a cross-atlas strategy. The key idea is that the edge weights of a graph are not computed from its own node features but from features of the same subjects under a different atlas. Assume that there are S atlases. For the graph corresponding to atlas s, node features are derived from atlas s, while edge weights are computed using features from another atlas t ($t \neq s$). In this process, node features from atlas t are first mapped into a latent space through a learnable transformation, and the edge weights are then calculated as the cosine similarity between the transformed node embeddings. This allows the graph structure to incorporate information from different atlases.

By iterating over different atlas pairs, each atlas can generate multiple graphs that share the same nodes but differ in their edge definitions. For example, when $S = 3$, each atlas produces two graphs, resulting in a total of six graphs. These graphs provide complementary structural views of the same set of subjects.

After graph construction, each graph is processed by a residual Chebyshev graph convolutional network to learn node representations. To integrate information from multiple graphs corresponding to the same atlas, a weighted fusion strategy is employed, where the contribution of each graph is determined by its classification performance. Compared with simple averaging or concatenation, this approach allows more informative graph features to dominate while reducing the impact of less discriminative ones.

Specifically, the static node features from all cross-

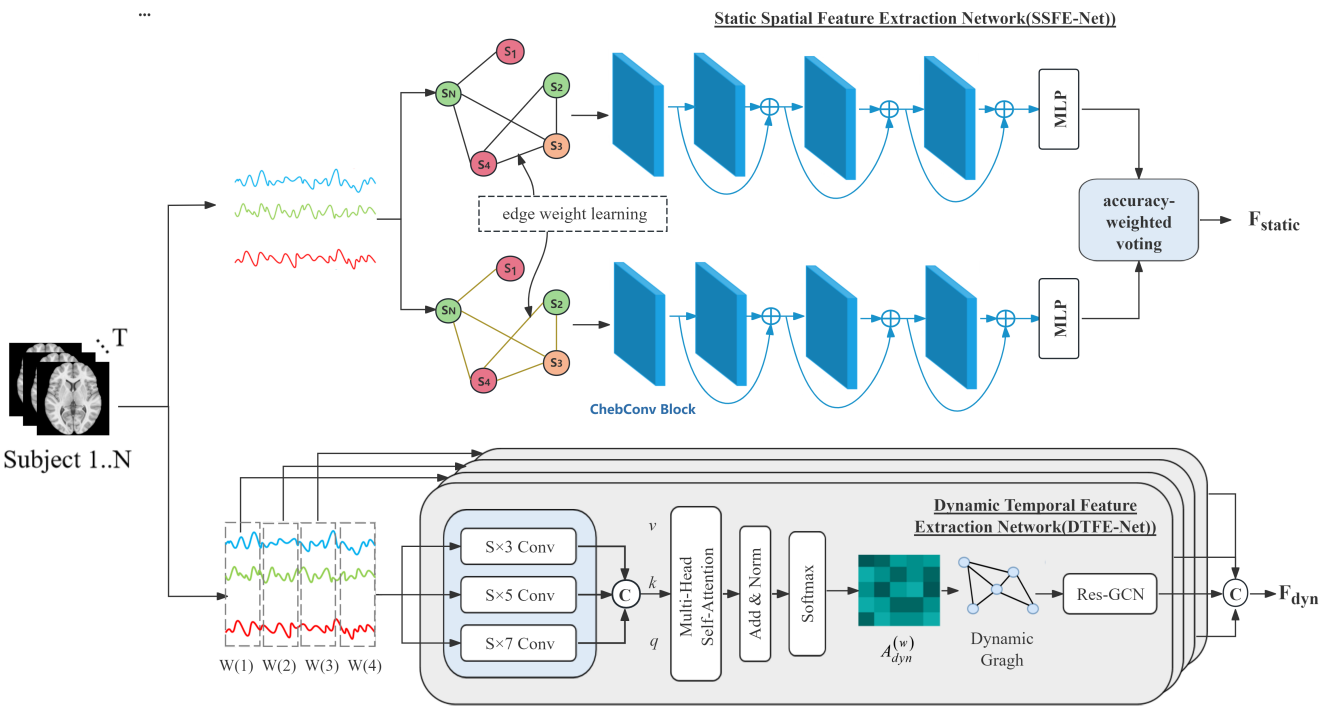

Fig. 2. Detailed framework diagrams of SSFE-Net and DTFE-Net.

constructed population graphs are first aggregated. A normalized weight w, reflecting the discriminative power of each graph, is then applied to emphasize the most informative features. The fused static node features are thus computed as:

$$F_{static} = w \sum_{z=1}^{Z} F_{static}^{z} \tag{1}$$

where F_{static}^{z} denotes the node features extracted from the z-th cross-constructed population graph, and F_{static} represents the final fused static features across all graphs. This weighted aggregation ensures that the resulting node representations capture complementary information from multiple graph views while prioritizing the most informative structures.

B. Dynamic Feature Extraction

To complement static connectivity modeling, dynamic feature extraction is introduced to capture temporal variability in brain function. An illustrative example of the dynamic feature extraction process using DTFE-Net is shown in Fig. 2. All fMRI time series are first normalized to a fixed length of 120 time points using truncation or zero-padding to ensure consistency across subjects. The resulting time-series matrix is denoted as $X \in \mathbb{R}^{R \times T}$, where R represents the number of brain regions and T denotes the number of time points. The normalized time series is then divided into $W = 4$ non-overlapping windows, each containing 30 time points. This window-based strategy enables the model to capture short-term temporal variations in brain activity while reducing noise and improving the robustness of dynamic connectivity estimation.

Within each window, multi-scale one-dimensional convolutions with kernel sizes $k = 3, 5, 7$ are applied to capture temporal patterns at different resolutions. The outputs of these convolutional branches are batch-normalized, followed by ReLU activation, and then concatenated to form window-level features $F^{(w)} \in \mathbb{R}^{R \times h}$, where h denotes the feature dimension for each window.

To further model dependencies across temporal features, a multi-head self-attention mechanism with four heads is applied to the window-level representations. This operation enhances informative temporal patterns while preserving global contextual relationships.

Based on the enhanced temporal features $F_{atm}^{(w)}$, dynamic functional connectivity matrices are constructed and used as adjacency matrices:

$$A_{dyn}^{(w)} = \text{softmax}\left(F_{atm}^{(w)} F_{atm}^{(w)\top}\right) \tag{2}$$

where the softmax function is applied to normalize connectivity strengths. Subsequently, two layers of residual graph convolution are employed to extract dynamic graph features $H^{(w)}$.

Finally, the dynamic features from the four windows are concatenated and projected to match the dimensionality d of the static features, yielding atlas-specific dynamic representations F_{dyn}.

C. feature fusion

In multi-atlas functional brain modeling, feature heterogeneity arises from both static–dynamic discrepancies within each atlas and structural inconsistencies across atlases. We adopt a hierarchical feature fusion strategy that integrates intra-atlas collaboration and inter-atlas complementary fusion to construct discriminative representations.

Within a single atlas, we introduce an intra-atlas feature interaction mechanism based on cross-attention, which explicitly models the dependencies between static and dynamic features to enable guided information exchange and adaptive feature enhancement. Specifically, static and dynamic features are first projected into a shared latent space, with static features serving as queries (Q) and dynamic features as keys (K) and values (V). The cross-attention operation is formulated as follows:

$$\text{Attn}(F_{\text{static}}, F_{\text{dyn}}) = \text{Softmax}\left(\frac{QK^{\top}}{\sqrt{d}}\right) V \tag{3}$$

where

$$Q = F_{\text{static}} W^{Q}, \quad K = F_{\text{dyn}} W^{K}, \quad V = F_{\text{dyn}} W^{V} \tag{4}$$

and W^{Q}, W^{K}, W^{V} are learnable projection matrices. A symmetric operation is further applied to enhance dynamic features under static structural constraints, and the resulting interaction features are integrated via residual connections to preserve the original semantic structure. This process effectively incorporates dynamic temporal discriminative information while maintaining the stability of static functional organization. As a result, unified static-dynamic representations are obtained for each atlas, denoted as F_{AAL}, F_{CC200}, and F_{HO}, which serve as consistent high-level inputs for subsequent multi-atlas fusion.

To explicitly model dependencies among atlases, a cross-atlas multi-head attention mechanism is applied to obtain a joint representation. This mechanism follows the same principle as the fusion of static and dynamic features within each atlas, highlighting complementary information while suppressing redundant features, with the joint representation denoted as F_{cross}. However, due to individual differences in brain structure and function, standardized brain atlases may exhibit regional mismatches or uneven information reliability for specific subjects. To address this issue, we introduce an Adaptive Channel Weighting Module (ACWM) following cross-atlas interaction. Unlike fixed-weight fusion, ACWM can adaptively learn each subject's personalized dependence on different brain atlases and dynamically adjust their weights in the classification decision. At the same time, it performs fine-grained two-stage adaptive weighting of channels, enabling the dynamic computation of the contribution of each channel and atlas. Before the first stage, the graph-level feature vectors from the three atlases are concatenated along the channel dimension to form a combined feature vector, denoted as F, which serves as the input to the ACWM.

First stage: global average pooling (GAP) and global max pooling (GMP) are applied to extract channel-wise descriptors:

$$s_1 = \text{GAP}(F) + \text{GMP}(F) \tag{5}$$

GAP captures the overall activation trend of each channel across the input features, while GMP emphasizes locally prominent activations. Together, they provide complementary information that reflects the relative importance of each channel. These descriptors are passed through a shared non-linear mapping f_1 and combined via element-wise addition followed by a Sigmoid activation to generate the first-stage channel weights:

$$W_1 = \sigma\big(f_1(\text{GAP}(F)) + f_1(\text{GMP}(F))\big) \tag{6}$$

The features are then scaled channel-wise:

$$F' = W_1 \odot F \tag{7}$$

This weighting effectively enhances channels that carry strong discriminative information while suppressing weaker or redundant channels, enabling ACWM to dynamically quantify each channel's contribution.

Second stage: To further capture fine-grained discriminative information, F' undergoes GAP and GMP again, followed by independent mappings f_2 and f_3 to extract higher-order channel descriptors:

$$s_2 = f_2\big(\text{GAP}(F')\big) + f_3\big(\text{GMP}(F')\big) \tag{8}$$

These descriptors are concatenated and passed through a non-linear mapping f_4 to produce the second-stage channel weights:

$$W_2 = f_4(s_2) \tag{9}$$

The features are reweighted channel-wise:

$$F'' = W_2 \odot F' \tag{10}$$

A residual connection is added to preserve the original cross-atlas information:

$$F_{ACWM} = F'' + F \tag{11}$$

Finally, F_{ACWM} is concatenated with the cross-atlas interaction features to form the final multi-atlas fused representation:

$$F_{\text{final}} = \big[F_{\text{cross}}, F_{ACWM}\big] \tag{12}$$

By extracting channel-wise statistics and applying two-stage adaptive weighting, ACWM can evaluate the contribution of each channel and brain atlas for each subject, enhancing discriminative channels, suppressing redundant ones, and ensuring that the fused representation reflects both individual-specific dependencies and atlas-level complementary information.

D. Loss Function

To jointly enforce discriminability, complementarity, and structural consistency of multi-atlas features, we design a three-term loss function.

Firstly, during the static feature extraction stage, two different graph structures are constructed for each brain atlas. To ensure diversity between the representations derived from these graph structures, a Hilbert–Schmidt Independence Criterion (HSIC) loss is introduced at the GCN output layer. Minimizing this loss enhances the independence between the two graph-based representations, thereby preserving intra-atlas diversity:

$$L_{HSIC} = \text{tr}(K_i R K_j) \tag{13}$$

where K_i and K_j denote the Gram matrices computed from the two graph representations, R is the centering matrix, and $\text{tr}(\cdot)$ denotes the matrix trace.

Secondly, after fusing the static and dynamic features within each atlas, to ensure that the fused features from different brain atlases are consistent in distribution and can be reliably integrated, we introduce the cross-atlas consistency loss:

$$L_{\text{cons}} = \sum_{i<j} \|\text{Norm}(F_i) - \text{Norm}(F_j)\|_F^2 \tag{14}$$

Here, $F_i, F_j \in \{F_{\text{AAL}}, F_{\text{CC200}}, F_{\text{HO}}\}$, $\|\cdot\|_F$ denotes the Frobenius norm, which measures the difference between matrices.

Finally, for the classification task, we use the standard cross-entropy loss:

$$L_{CE} = -\frac{1}{N}\sum_{n=1}^{N} y_n \log \hat{y}_n \tag{15}$$

where N denotes the number of samples, y_n is the ground-truth label of the n-th sample, and $\hat{y}_n$ is the predicted probability by the model. The total loss is thus:

$$L_{\text{total}} = \lambda_1 L_{\text{HSIC}} + \lambda_2 L_{\text{cons}} + \lambda_3 L_{CE} \tag{16}$$

where $\lambda_1, \lambda_2, \lambda_3$ balance each term.

III. Experiments

A. Datasets and Preprocessing

This study utilised the ABIDE I dataset [14] from the multi-centre Autism Brain Imaging Data Exchange , which compiles resting-state functional MRI (rs-fMRI) data and associated clinical information from multiple independent sites worldwide.

TABLE I
PARAMETER CONFIGURATION

parameter	Value
Seed	42
Input dim	1500
Chebyshev layers	4
Hidden layers	24
Learning rate	0.001
Weight decay	5e-4
Dropout rate	0.3
patience	20

To ensure data quality and analytical consistency, participants with fully preprocessed data that passed strict quality control were selected, yielding a total of 1,035 subjects, including 505 individuals with autism spectrum disorder (ASD) and 530 typically developing controls (HC). All rs-fMRI data were standardised using the CPAC pipeline [15], including motion correction, slice-timing correction, spatial normalisation to MNI space, Friston 24-parameter motion regression, and band-pass filtering (0.01–0.1 Hz). In addition to imaging features, demographic variables—age, sex, and full-scale IQ (FIQ)—were included as covariates to control for potential confounding effects on the classification models.

B. Experimental Setup

The performance of the proposed model was evaluated on the ABIDE I dataset using 10-fold cross-validation. All subjects were randomly and evenly divided into 10 subsets, ensuring that the distribution of ASD and typically developing controls was balanced within each fold. In each iteration, 9 folds were used for training and the remaining fold for testing; this process was repeated 10 times, and the average of the 10 test results was taken as the final performance metric. To ensure reproducibility, all sources of randomness—including data splitting, model parameter initialisation, feature selection, and Dropout—were controlled by fixed random seeds. Experiments were conducted using the PyTorch deep learning framework on a CentOS 7 system equipped with an NVIDIA GeForce RTX 3090 GPU (24 GB VRAM). Detailed experimental parameter settings are listed in Table I.

C. Comparative Experiments

To validate the effectiveness of the proposed method, we compare it with traditional machine learning methods (Random Forest and SVM), single-atlas GNN methods (DHG-Former [19], Cheb-GCN [16], MDGL [17], and PLSNet [18]), and representative multi-atlas GNN methods (MADE-for-ASD [22], CcSi-MHAHGEL [23], AIDFusion [20], MCGC-SPFFN [24], and MGCARAFFNet [21]). All methods were reimplemented for fair comparison, and original results were used when reproduced values differed by less than 2%.

The results show that our method achieves the highest performance across most metrics. It outperforms traditional machine learning approaches, consistently surpasses single-atlas GNNs, and provides strong and balanced results compared with state-of-the-art multi-atlas GNNs. These findings demonstrate the effectiveness of multi-atlas feature fusion in improving the stability and reliability of ASD identification.

D. Ablation experiment

To evaluate the contributions of each module, we conducted ablation experiments on ACWM and DTFE-Net (Table III). Using Model A, which excludes both modules, as the baseline, Model B with only ACWM improves accuracy by 1.91% and AUC by 1.95%, showing its effectiveness in capturing the global structure of brain functional networks. Model C, with only DTFE-Net, increases accuracy by 1.80% and AUC by 2.02%, confirming its role in enhancing class-discriminative capability. The full model, combining both ACWM and DTTE-Net, achieves the best performance, with accuracy and AUC gains of 3.19% and 3.22% over the baseline. These results highlight the complementary contributions of the two modules and their ability to improve identification accuracy.

TABLE II
THE COMPARATIVE RESULTS ON ABIDE-I DATASET

Model	Acc	AUC	Pre	SEN	F1	SPE
Random Forest	57.42	59.12	74.96	66.38	70.46	49.53
SVM	61.46	60.82	64.13	66.70	65.48	55.26
Cheb-GCN [16]	63.00	62.06	65.37	69.41	67.31	60.73
MDGL [17]	72.00	74.00	71.00	66.00	66.00	76.19
PLSNet [18]	72.40	72.00	65.48	71.60	68.14	71.30
DHGFormer [19]	73.21	76.46	73.69	72.78	73.23	74.02
AIDEfusion [20]	74.58	81.02	72.41	77.78	75.00	71.30
MGCARAFFNet [21]	75.07	77.00	76.04	76.04	76.03	74.06
MADE-for-ASD [22]	75.20	78.38	71.08	82.90	76.51	69.70
CcSi-MHAHGEL [23]	76.18	79.51	74.12	70.45	71.41	74.90
MCGC-SPFFN [24]	78.45	78.31	77.08	83.96	79.83	76.34
Ours	**81.64**	**81.53**	**80.83**	**84.90**	**82.63**	**80.92**

TABLE III
ABLATION TEST RESULTS

Model	ACWM	DTFE-Net	ACC	AUC	SEN	SPE
Model A	×	×	78.45	78.31	83.96	76.34
Model B	√	×	80.36	80.26	84.28	78.31
Model C	×	√	80.25	80.33	80.88	80.29
Ours	√	√	**81.64**	**81.53**	**84.90**	**80.92**

IV. DISCUSSION

A. Results of SD-MAFNet for FCN based on different atlases

To study the impact of different brain atlases and their combinations on ASD classification, we conducted multiple multi-atlas experiments to evaluate their performance in multi-scale feature fusion. EZ [25] and DOS160 were also included to form three-atlas comparison groups with existing atlases. As shown in Table IV, multi-atlas combinations generally outperformed two-atlas ones, with the AAL, CC200, and HO combination achieving the best results, demonstrating advantages in accuracy and discriminative power.Some two-atlas combinations outperformed certain three-atlas sets on specific metrics, likely due to differences in region granularity, anatomical priors, or functional coverage. While larger atlas sets include more information, they may introduce redundancy or noise, slightly reducing performance on some measures. These results indicate that multi-atlas fusion effectively enhances discriminative ability, but atlas selection and fusion strategy significantly affect performance.

B. Feature Visualisation

To further visually assess the discriminative capability of the features learned by the proposed model, we performed a visualisation analysis of the high-dimensional fused features. As shown in Figure 3, we selected the fold with the best performance from cross-validation and, with model parameters fixed, fed all subject samples through the network via forward propagation to extract the final fused feature representations before the classifier. The high-dimensional features were then mapped to a two-dimensional space using t-distributed Stochastic Neighbor Embedding (t-SNE) to preserve the local neighborhood structure and examine the distribution of different classes in feature space. The visualization shows that HC and individuals with ASD form two relatively compact and well-separated clusters, with clear inter-class boundaries and consistent intra-class distributions, and only slight overlap in a few regions.

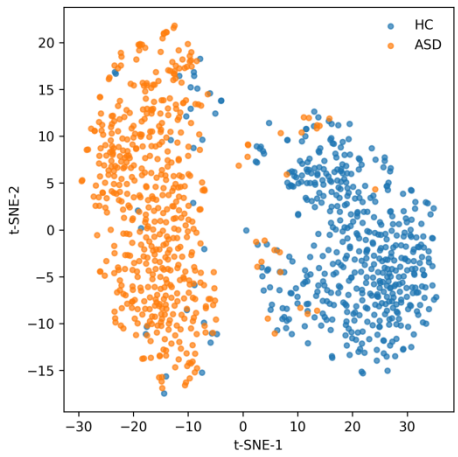

Fig. 3. Feature distribution visualisation based on T-SNE.

C. Discriminative Brain Region Detection

During feature extraction and fusion, we performed a visualization analysis of brain regions associated with autism spectrum disorder (ASD) to identify potential biologically meaningful biomarkers. Using the BrainNet Viewer [26] toolbox, we visualized the top 10 brain regions in the AAL atlas that contributed most to ASD classification. As shown in Figure 4, these regions include the left medial superior frontal gyrus (SFGmed.L), right middle frontal gyrus (MFG.R), left triangular part of the inferior frontal gyrus (IFGtriang.L), right opercular part of the inferior frontal gyrus (IFGoperc.R), right insula (INS.R), left hippocampus (HIP.L), bilateral middle temporal gyri (MTG.L, MTG.R), left angular gyrus (ANG.L), and right precuneus (PCUN.R). These regions are mainly located in the frontal, temporal, and parietal lobes, as well as the hippocampus, and are closely associated with key neural functions such as social interaction, language comprehension, emotion regulation, memory, and executive cognition. This analysis not only provides an intuitive visualization of potential ASD biomarkers but also further validates the effectiveness of the proposed model in multi-atlas feature extraction and fusion, offering a biologically plausible interpretation.

TABLE IV
ASD/HC CLASSIFICATION PERFORMANCE BY SD-MAFNET USING DIFFERENT ATLASES

Model	Atlas	Acc	AUC	Pre	SEN	F1	SPE
SD-MAFNet	CC200+DOS160	76.88	76.78	75.73	80.72	78.08	49.53
SD-MAFNet	AAL+CC200	78.04	77.93	77.05	82.03	79.29	55.26
SD-MAFNet	CC200+HO+DOS160	77.07	76.97	76.21	80.48	77.92	60.73
SD-MAFNet	AAL+CC200+EZ	79.30	79.28	79.89	80.34	79.89	76.19
SD-MAFNet	**AAL+CC200+HO**	**81.64**	**81.53**	**80.83**	**84.90**	**82.63**	**80.92**

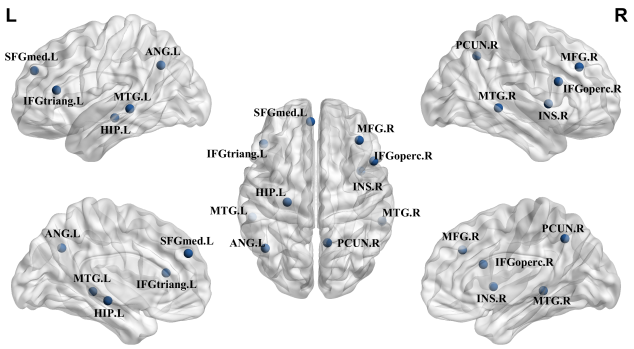

Fig. 4. The most distinctive brain region in the classification of ASD.

V. Conclusion

This study proposes a multi-atlas feature fusion graph neural network for fMRI-based ASD recognition. The model separately encodes static functional connectivity and dynamic features, and employs a cross-attention mechanism to fuse them in a shared semantic space to capture complementary information between stable structural patterns and temporal dynamics. Multi-atlas fusion is further applied to collaboratively model representations across different brain atlases, enhancing overall representation. Experimental results demonstrate that the proposed method outperforms approaches using only static features or a single atlas and provides interpretable insights into discriminative brain regions. Although incorporating dynamic and multi-atlas features increases model complexity, future work will aim to improve efficiency and validate the model on larger multi-site datasets to enhance generalizability.

References

[1] R. A. Hoekstra, M. Bartels, C. J. Verweij, and D. I. Boomsma, "Heritability of autistic traits in the general population," *Archives of pediatrics & adolescent medicine*, vol. 161, no. 4, pp. 372–377, 2007.

[2] S. L. Simpson, F. D. Bowman, and P. J. Laurienti, "Analyzing complex functional brain networks: fusing statistics and network science to understand the brain," *Statistics surveys*, vol. 7, p. 1, 2013.

[3] C. P. Santana, E. A. de Carvalho, I. D. Rodrigues, G. S. Bastos, A. D. de Souza, and L. L. de Brito, "rs-fmri and machine learning for asd diagnosis: a systematic review and meta-analysis," *Scientific reports*, vol. 12, no. 1, p. 6030, 2022.

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

[5] N. C. Dvornek, P. Ventola, K. A. Pelphrey, and J. S. Duncan, "Identifying autism from resting-state fmri using long short-term memory networks," in *International Workshop on Machine Learning in Medical Imaging*. Springer, 2017, pp. 362–370.

[6] L. Wang, K. Li, and X. Hu, "Graph convolutional network for fmri analysis based on connectivity neighborhood. netw. neurosci. 5 (1), 83–95 (2021)."

[7] X. Luo, J. Wu, J. Yang, S. Xue, A. Beheshti, Q. Z. Sheng, D. McAlpine, P. Sowman, A. Giral, and P. S. Yu, "Graph neural networks for brain graph learning: A survey," *arXiv preprint arXiv:2406.02594*, 2024.

[8] J. Pan, Y. Dong, and H. Chen, "Review of research on auxiliary diagnosis of autism based on graph neural networks," *Comput. Eng*, vol. 48, pp. 1–11, 2022.

[9] M. Liu, H. Zhang, F. Shi, and D. Shen, "Hierarchical graph convolutional network built by multiscale atlases for brain disorder diagnosis using functional connectivity," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 35, no. 11, pp. 15 182–15 194, 2023.

[10] H. Mohammadi and W. Karwowski, "Graph neural networks in brain connectivity studies: Methods, challenges, and future directions," *Brain Sciences*, vol. 15, no. 1, p. 17, 2024.

[11] 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.

[12] 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.

[13] 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.

[14] A. Di Martino, C.-G. Yan, Q. Li, E. Denio, F. X. Castellanos, K. Alaerts, J. S. Anderson, M. Assaf, S. Y. Bookheimer, M. Dapretto *et al.*, "The autism brain imaging data exchange: towards a large-scale evaluation of the intrinsic brain architecture in autism," *Molecular psychiatry*, vol. 19, no. 6, pp. 659–667, 2014.

[15] C. Craddock, S. Sikka, B. Cheung, R. Khanuja, S. S. Ghosh, C. Yan, Q. Li, D. Lurie, J. Vogelstein, R. Burns *et al.*, "Towards automated analysis of connectomes: The configurable pipeline for the analysis of connectomes (c-pac)," *Front Neuroinform*, vol. 42, no. 10.3389, 2013.

[16] M. Defferrard, X. Bresson, and P. Vandergheynst, "Convolutional neural networks on graphs with fast localized spectral filtering," *Advances in neural information processing systems*, vol. 29, 2016.

[17] Y. Ma, Q. Wang, L. Cao, L. Li, C. Zhang, L. Qiao, and M. Liu, "Multi-scale dynamic graph learning for brain disorder detection with functional mri," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 31, pp. 3501–3512, 2023.

[18] Y. Wang, H. Long, Q. Zhou, T. Bo, and J. Zheng, "Plsnet: Position-aware gcn-based autism spectrum disorder diagnosis via fc learning and rois sifting," *Computers in Biology and Medicine*, vol. 163, p. 107184, 2023.

[19] R. Xue, H. Hu, Z. Zhang, X. Han, J. Wang, Y. Gao, and S. Du, "Dhgformer: Dynamic hierarchical graph transformer for disorder brain disease diagnosis," in *International Conference on Medical Image Computing and Computer-Assisted Intervention*. Springer, 2025, pp. 268–278.

[20] J. Xu, M. Lan, X. Dong, K. He, W. Zhang, Q. Bian, and Y. Ke, "Multi-atlas brain network classification through consistency distillation and complementary information fusion," *IEEE Journal of Biomedical and Health Informatics*, 2025.

[21] Y. Ma, W. Cui, J. Liu, Y. Guo, H. Chen, and Y. Li, "A multi-graph cross-attention-based region-aware feature fusion network using multi-template for brain disorder diagnosis," *IEEE Transactions on Medical Imaging*, vol. 43, no. 3, pp. 1045–1059, 2023.

[22] X. Liu, M. R. Hasan, T. Gedeon, and M. Z. Hossain, "Made-for-asd: A multi-atlas deep ensemble network for diagnosing autism spectrum disorder," *Computers in Biology and Medicine*, vol. 182, p. 109083, 2024.

[23] W. Wang, L. Xiao, G. Qu, V. D. Calhoun, Y.-P. Wang, and X. Sun, "Multiview hyperedge-aware hypergraph embedding learning for multisite, multiatlas fmri based functional connectivity network analysis," *Medical Image Analysis*, vol. 94, p. 103144, 2024.

[24] X. Jiang, X. Ding, Z. Xia, H. Wang, and Z. Jiao, "Self-perceptive feature fusion network with multi-channel graph convolution for brain disorder diagnosis," *Expert Systems with Applications*, p. 127984, 2025.

[25] S. B. Eickhoff, K. E. Stephan, H. Mohlberg, C. Grefkes, G. R. Fink, K. Amunts, and K. Zilles, "A new spm toolbox for combining probabilistic cytoarchitectonic maps and functional imaging data," *Neuroimage*, vol. 25, no. 4, pp. 1325–1335, 2005.

[26] M. Xia, J. Wang, and Y. He, "Brainnet viewer: a network visualization tool for human brain connectomics," *PloS one*, vol. 8, no. 7, p. e68910, 2013.