

Bi-BrainSSM: 一种基于双向 SSM 框架的 rs-fMRI 数据脑疾病诊断方法

张力¹, 顾恒¹, 黄硕¹, 马越², 任凯³, 张礼¹

(1. 南京林业大学信息科学技术学院、人工智能学院, 南京 210037; 2. 江苏健康卫生职业学院、中西医结合学院, 南京 210018; 3. 南京林业大学学院机械电子工程学院, 南京 210037)

摘要: 静息状态功能磁共振成像 (rs-fMRI) 技术被广泛应用于脑疾病预测和诊断。研究人员提出了大量基于 rs-fMRI 功能连接网络的诊断方法, 但这些方法存在未能充分考虑脑网络数据的复杂依赖关系、计算开销大等问题。为了解决上述问题, 本文提出了一种基于双向选择性状态空间 (State space model, SSM) 框架的脑疾病诊断方法 Bi-BrainSSM。该方法设计了 ABM (Attentive bi-Mamba) 模块, 将多头注意力机制与双向选择性状态空间模型相结合, 以增强脑功能连接网络的表征能力和提升模型的计算效率。同时引入正交聚类读出, 解决传统聚合方法信息丢失问题。本方法在公开的 ABIDE 自闭症数据集和 REST-MDD 抑郁症数据集上进行验证。实验结果表明, 相比其他基于 rs-fMRI 的对比方法, Bi-BrainSSM 获得了更精准的诊断性能, 同时拥有更高的计算效率和更低的 GPU 显存占有率。

关键词: 脑疾病诊断; rs-fMRI; 选择性状态空间; 注意力机制; 正交聚类
中图分类号: TP391.4; **文献标志码:** A

Bi-BrainSSM: A Brain Disease Diagnosis Method Based on

Bi-directional SSM Framework for rs-fMRI Data

ZHANG Li¹, GU Heng¹, HUANG Shuo¹, MA Yue², REN Kai³, ZHANG Li¹

(1. School of Information Science and Technology, School of Artificial Intelligence, Nanjing Forestry University, Nanjing 210037, China; 2. Jiangsu Health and Health Vocational College, College of Integrated Traditional Chinese and Western Medicine, Nanjing 210018, China; 3. College of Nanjing Forestry University, School of Mechanical and Electronic Engineering, Nanjing 210037, China)

Abstract: Resting state functional magnetic resonance imaging (rs-fMRI) technology is widely used for predicting and diagnosing brain diseases. Researchers have proposed a large number of diagnostic methods based on rs-fMRI functional connectivity networks, but these methods have problems such as not fully considering the complex dependency relationships of brain network data and high computational costs. To address the aforementioned issues, this paper proposes a brain disease diagnosis method based on the bidirectional selective state space model (SSM) framework, Bi-BrainSSM. This method designs an attentive bi-Mamba (ABM) module that combines multi head attention mechanism with a bidirectional selective state space model to enhance the representational ability of brain functional connectivity networks and improve the computational efficiency of the model. Simultaneously introducing orthogonal clustering readout

基金项目: 国家自然科学基金 (62506170;12404273)。

to solve the problem of information loss in traditional aggregation methods. This method was validated on the publicly available ABIDE autism dataset and REST-MDD depression dataset. The experimental results show that compared to other comparison methods based on rs-fMRI, Bi-BrainSSM achieves more accurate diagnostic performance, while also having higher computational efficiency and lower GPU memory occupancy.

Keywords: diagnosis of brain diseases; rs-fMRI; selective state space; attention mechanism; orthogonal clustering

引言

大脑不同区域（或神经元群体）的相互连接形成了复杂的神经网络，其信息传递与协同活动是实现高级认知、情感与行为过程的神经基础^[1]。静息态功能性磁共振成像（rs-fMRI）作为一种关键的无创神经成像技术，能够通过测量血氧水平依赖（BOLD）信号来间接探测神经元活动^[2]。研究者通常以感兴趣大脑区域（Regions of interest, ROIs）的 BOLD 信号时间序列为基础，计算其成对相关性，从而构建功能连接矩阵（Functional connectivity matrix, FC）^[3]。在临床研究中，FC 矩阵被广泛用于识别神经性或者精神性疾病的生物标志物。通过对比患者与健康对照组的功能连接模式，可以定位与特定疾病（如自闭症、抑郁症）相关的异常脑网络连接，为疾病的客观诊断、病情评估及疗效预测提供影像学依据^[4]。

在基于 rs-fMRI 的脑网络分析中，脑网络被抽象为一个由节点和边构成的无向图，图中节点（Node）代表不同的脑区，而边（Edge）则表示不同脑区之间 BOLD 信号时间序列的相关性，即功能连接的强度^[5]。早期的脑网络分析研究中，研究人员主要采用经典的图论指标来量化脑网络的拓扑特性，从而揭示其组织原则和信息处理模式。这些方法可分为两类，一类是提取图论拓扑特征，即将网络的宏观（如模块度、全局效率）与微观（如节点度、中心性）指标量化后，组合成一个特征向量输入分类器^[6]。另一类采用更直接的方法将 FC 矩阵的连接权重值直接作为特征，但这种方法容易产生高维稀疏数据。因此通常与主成分分析（PCA）等降维技术结合使用，以提取最具代表性的主成分作为最终特征^[7]。在获得特征向量后，研究人员通常利用，支持向量机（SVM）、随机森林等经典机器学习模型进行分类（如疾病诊断）或回归（如症状评分预测）^[8]。尽管这些方法为脑网络分析奠定了基础，但它们也存在明显的局限性。由于主要捕捉的是较为简单的、基于成对连接的拓扑属性，这些方法难以全面刻画脑区之间复杂且动态的高阶非线性依赖关系^[9]，同时不可避免地会丢失脑网络图结构信息^[10]。

随着深度学习的快速发展，基于卷积神经网络（CNN）的方法广泛运用于脑疾病诊断。如 Jeremy Kawahara 等提出的 BrainNetCNN，通过卷积操作，捕捉了脑网络中的局部拓扑结构^[11]。由于图神经网络（GNNs）强大的图结构数据分析能力，科研人员提出了大量基于图神经网络的脑疾病诊断模型^[12]。Li 等设计的 BrainGNN，通过为每个脑区分配独特的嵌入权重并引入多个正则化项，从而增强模型对脑图的理解^[13]。Cui 提出的 FBNetGen 通过生成任务感知脑网络，进一步提升了模型的准确性^[14]。BrainGB 是一个统一、模块化、可扩展且可重现的分析框架，为脑网络图模型的横向比较提供了统一的实验基准^[15]。

为了更全面地建模脑区之间的全局依赖关系，研究者们将目光投向了在序列建模领域取得巨大成功的 Transformer 架构 Transformer^[16]。例如 GraphAttentionNetworks（GATs）通过自注意力机制为每个节点的邻居分配不同的权重，从而实现了节点间复杂关系的建模^[17]。Graph Transformer 将 Transformer 架构推广到图数据，利用局部邻域的连通性来计算注意力机制，进一步提升了模型对图数据稀疏性和拓扑结构的利用能力^[18]。SAN 提出了一种基于谱图理论的学习位置编码（LPE）方法，增强了模型对图结构的表示能力^[19]。Graphormer 通过引入结构编码改进了 Transformer 在图数据上的表现，提升了其性能^[20]。Brain Network

Transformer (BrainNetTF) 结合了注意力机制和图神经网络方法, 使模型专注于脑网络分析, 以改进疾病诊断和脑功能预测的能力^[21]。李建彤等人通过引入 Hub 约束的注意力机制, 使模型能够专注于网络中作为信息枢纽的“Hub 节点”及其异常连接模式, 从而提取与疾病更相关的判别性特征^[22]。刘子凡等人提出了一种自适应聚合功能网络模型 (AAFN), 该模型利用自注意力机制在端到端的学习过程中动态构建功能脑网络, 使其能够根据具体任务自适应地调整, 从而增强了特征的有效性^[23]。BrainDGT 通过自适应地捕获和分析功能脑模块内的时序动态脑连接来改进脑疾病的诊断准确性, 从而克服了传统滑动窗口方法中固定时间长度和全局空间范围的限制^[27]。但是基于 Transformer 的脑疾病诊断方法仍面临挑战, 主要在于其自注意力机制的二次方计算复杂度会导致巨大的计算开销, 同时模型对脑网络中长序列特征信息的捕捉能力仍有待提升^[28]。

为解决上述问题, 本文提出了一种基于双向选择性状态空间框架^[29]的脑疾病诊断方法, Bi-BrainSSM。该方法的核心 ABM 模块将多头注意力机制与双向选择性状态空间模型 (Selective SSM) 相结合。在 ABM 模块中, 多头自注意力机制首先被用于显式地建模脑网络中任意两个脑区节点之间的空间交互关系。同时, 该模块引入了选择性机制, 使其核心参数能够根据输入数据动态调整, 从而有效地“记住”关键信息并“遗忘”无关信息。Bi-BrainSSM 采用了双向 (Bidirectional) 结构, 通过一个前向和一个后向的 Mamba 编码器并行处理节点序列, 通过融合正反两个方向的输出, 每个节点的最终表示能够同时包含来自序列前后文的完整信息, 这不仅有效地反映了整个脑网络的相互依赖结构, 也显著减少了节点在序列中固定排列顺序可能对模型性能造成的偏差。进一步优化其在脑网络功能模块分类和全局特征提取方面的性能。最后通过公开 ABIDE 自闭症数据集和 REST-MDD 抑郁症数据集来验证 Bi-BrainSSM 的有效性。

1. 模型方法

Bi-BrainSSM 的整体框架如图 1 所示。其核心创新在于两个关键模块的设计: ABM (注意力双向 Mamba) 模块, 它创新性地融合了多头注意力机制与双向选择性状态空间模型, 多头注意力为节点间的全局交互动态分配权重, 而双向 Mamba 则同时刻画正、反向依赖, 用于在较低的计算开销下高效捕捉脑网络复杂的空间交互与长程依赖关系; (2) OCFC (正交约束特征聚类) 模块, 解决了传统聚合方法导致的信息丢失问题, 将局部节点嵌入聚合为全局的图级表示, 为后续的疾病诊断任务提供更具代表性的特征。

$$h_i^m = \sum_{j=1}^N \alpha_{ij}^m W_V^m h_j \quad (3)$$

式中： $W_V^m \in \mathbb{R}^{d \times d/M}$ 表示第 m 个头的投影矩阵， M 为注意力头的数量。将 M 个注意力头的输出拼接并映射生成最终的节点特征

$$h_i^{\text{multi}} = W_O[h_i^1 \parallel h_i^2 \parallel \dots \parallel h_i^M] \quad (4)$$

式中： $W_O \in \mathbb{R}^{d \times d}$ 为输出映射矩阵，“ $\parallel$ ”为拼接操作。

通常在计算注意力权重时需结合节点坐标与显式边关系。但是脑功能网络往往呈现近似全连接的高密度拓扑，特征分解在此场景下计算开销高昂，而边的存在本身几乎不提供判别信息。对于脑网络这样的全连接图，脑网络上的 ROI 节点特征本身就包含了足够的位置信息，因此 Bi-BrainSSM 模型无需额外的位置编码，从而进一步简化了计算复杂度。

1.1.2 双向 Mamba 序列编码

为了捕捉脑网络中的全局长程依赖关系，ABM 模块引入了双向 Mamba 编码器。该编码器基于选择性状态空间模型（SSM）。首先将脑网络表示为节点序列 $[Z_1, Z_2, \dots, Z_N]$ （其中 Z_t 为节点 t 的特征向量， $Z \in \mathbb{R}^{N \times d}$ 为节点特征矩阵），然后构建两个并行的 Mamba 序列编码器：一个沿节点序列的正向（从第1到第 N ）进行编码，另一个沿反向（从第 N 到第1）进行编码。通过双向处理，使得每个节点的表示同时包含来自序列的前向和后向信息，从而有效地反映整个脑网络的依赖结构，减少节点顺序排列对模型的影响。

该编码器基于高效的选择性状态空间模型（State space model, SSM）构建。SSM 的核心优势在于其选择性机制，该机制使其核心参数（ A 、 B 、 C ）能够根据输入特征 Z^t 动态调整^[24]： $A_d = f_A(Z^t)$, $B_d = f_B(Z^t)$, $C_d = f_C(Z^t)$ 。这使得模型能有效“记住”关键信息并“遗忘”无关信息，在处理长序列时对关键特征保持高度敏感^[25]。结合 SSM 的状态更新与选择性机制^[26]，考虑到脑网络的数据特性，基于 Mamba 架构的特征编码状态更新方程为

$$h_{t+1} = f_A(Z^t)h_t + f_B(Z^t)Z^t \quad (5)$$

输出特征计算方程为

$$y_t = f_C(Z^t)h_t \quad (6)$$

通过隐状态 h_t 和输入特征 Z^t 的结合更新隐状态，同时通过动态调整 f_A 、 f_B 和 f_C 来选择性记忆或忽略特定的输入，从而提高长序列建模能力，最后隐状态通过输出矩阵 C_d 转换为节点的输出特征。通过状态转移和输入矩阵的递归更新，模型能够编码节点间的长程依赖。由于状态转移矩阵和输入矩阵的参数是数据自适应的，因此 Mamba 编码器动态调整信息的选择性记忆或遗忘，以捕捉最为关键的信息。

由于 Mamba 的单向性质使其无法像基于 Transformer 的方法一样捕捉全局依赖。为了克服限制，模型的 Mamba 层采用双向结构分别对正向和反向序列进行计算。具体来说，经过多头注意力机制增强后的节点特征序列，被并行送入两个独立的 Mamba 编码器。第一个编码器按原始顺序处理序列，得到前向隐藏状态序列 h_{forward} ，第二个编码器则处理逆序序列，得到反向隐藏状态序列 h_{backward} 。具体公式为

$$h_t^{\text{forward}} \leftarrow \text{Mamba}(h_i^{\text{multi}}) \quad (7)$$

$$h_t^{\text{backward}} \leftarrow \text{Mamba}(\text{reverse}(h_i^{\text{multi}})) \quad (8)$$

通过这种双向并行处理，每个节点的表示都同时蕴含了来自序列前向和后向的完整信息，有效消除了节点顺序排列对模型性能造成的影响。最终，通过融合正反两个方向的输出特征，得到每个节点的最最终双向特征表示

$$h_t = h_t^{\text{forward}} + h_t^{\text{backward}} \quad (9)$$

得益于 Mamba 架构所采用的硬件感知优化算法及并行计算机制^[30]，ABM 在维持高性

能的同时显著降低了计算开销，并在序列处理任务中实现线性时间复杂度。通过融合正向与反向的信息流编码，该模型能够有效整合脑网络的全局特征，增强长程依赖的建模能力，从而提升脑网络分类任务中的特征表示能力。

1.2 OCFC（正交约束特征聚类）模块

为解决传统聚合方法（如求和或平均操作）未能充分利用脑功能模块关系而导致的信息丢失问题，Bi-BrainSSM 引入了正交约束特征聚类读出（OCFC）模块。该模块结合了软聚类和正交化约束机制，将节点嵌入分配到不同的聚类中心^[31]。通过在嵌入空间中对聚类中心进行正交化，它确保了功能模块间的独立性，从而提高了模型的区分能力和对功能模块的捕捉能力。首先，通过 Gram-Schmidt 正交化过程对聚类中心 E 进行正交化处理，保证聚类中心在嵌入空间中互相正交，最大化它们之间的区分度。该过程首先通过 Xavier 初始化随机生成 K 个聚类中心 $C \in \mathbb{R}^{K \times d}$ ，使用 Gram-Schmidt 正交化过程生成正交的聚类中心 E

$$u_k = C_k - \sum_{j=1}^{k-1} \frac{\langle u_j, C_k \rangle}{\langle u_j, u_j \rangle} u_j, E_k = \frac{u_k}{\|u_k\|} \quad (10)$$

式中： u_j 表示在正交化过程中的中间结果， E_k 是第 K 个正交化聚类中心^[32]。正交化后的聚类中心确保每个聚类在嵌入空间中彼此独立，这样可以提高模型在脑网络数据中的聚类效果和分类性能。将节点分配到聚类中心进行软聚类，给定节点的嵌入矩阵 $Z^L \in \mathbb{R}^{N \times d}$ ，其中 N 表示节点数量， d 表示节点嵌入的维度，正交聚类读出首先通过软聚类操作，将每个节点分配到 K 个聚类中心之一。每个聚类中心 $E \in \mathbb{R}^{K \times d}$ 也是一个可学习的参数。节点 i 被分配到聚类 k 的概率 P_{ik} 通过以下公式计算

$$P_{ik} = \frac{\exp(\langle Z_i^L, E_k \rangle)}{\sum_{k'=1}^K \exp(\langle Z_i^L, E_{k'} \rangle)} \quad (11)$$

式中： $\langle \cdot, \cdot \rangle$ 表示内积操作， Z_i^L 是节点 i 的嵌入， E_k 是聚类中心 k 的嵌入向量。该公式通过 Softmax 操作计算每个节点属于不同聚类的概率，允许每个节点根据其特征与多个功能模块相关联，而不是强制性地归属到某个单一模块。这种柔性分配方式更符合脑网络中不同区域的复杂关联特性。在得到节点的聚类概率矩阵 $P \in \mathbb{R}^{N \times K}$ 后，节点的最终嵌入 Z_G 将通过聚合所有节点的嵌入得到图级表示

$$Z_G = P^T Z^L \quad (12)$$

式中 $Z_G \in \mathbb{R}^{K \times d}$ 是图级嵌入，代表了整个脑网络的特征。这一聚合过程充分利用了节点和功能模块之间的相似性，将局部节点嵌入转换为全局图嵌入。全局表征 Z_G 被进一步通过全连接层（FCN）映射为任务相关的输出表示。

1.3 模型损失函数

为优化自闭症脑网络分类性能，本文采用了结合任务损失与正交约束损失的多重损失机制。任务损失通过交叉熵函数定义，旨在最小化预测结果与真实标签之间的差异。正交约束损失则用于提升模型的可解释性和结构化特性。给定真实标签 y 和预测结果 $\hat{y}$ ，交叉熵损失定义为

$$\mathcal{L}_{\text{task}} = - \sum_{i=1}^n [y_i \log(\hat{y}_i) + (1 - y_i) \log(1 - \hat{y}_i)] \quad (13)$$

式中： n 是训练样本的数量， y_i 是第 i 个样本的真实标签， $\hat{y}_i$ 是模型的预测概率。正交约束损失用于约束模型中的聚类矩阵 P 使其保持正交性，从而提升模型的解释能力和稳定性。该损失函数旨在鼓励脑网络中不同脑区之间的连接特征保持独立性，避免不同脑区间信息过于冗余。

$$\mathcal{L}_{\text{orth}} = \|P^T P - I\|_F^2 \quad (14)$$

式中： $\mathbf{P}$ 是正交聚类矩阵， $\mathbf{I}$ 是单位矩阵， $\|\cdot\|_F$ 是 Frobenius 范数。这个损失项通过惩罚 $\mathbf{P}^\top \mathbf{P}$ 与单位矩阵 $\mathbf{I}$ 的偏差，确保了聚类矩阵的正交性，进而保证模型在空间特征建模中的独立性。最终的总损失函数为

$$\mathcal{L} = \mathcal{L}_{\text{task}} + \lambda \mathcal{L}_{\text{orth}} \quad (15)$$

式中 λ 为正交约束损失的权重超参数，控制任务损失和正交约束损失之间的平衡。

Bi-BrainSSM 算法如算法 1 所示。

算法 1 Bi-BrainSSM 算法

输入： fMRI 数据时间序列 X ，正交投影矩阵 P ，批量大小 B ，注意力头数 M ，聚类中心数量

- 1: 初始化参数 A, B, C 和聚类矩阵 P 。
- 2: **特征预处理：**对 X 归一化，生成脑网络特征 Z^t 。
- 3: **for** 每批次样本 Z^t **do**
- 4: ABM (Attentive Bi-Mamba) 模块:
- 5: -多头注意力机制

$$h_i^{\text{attn}} = \sum_{j=1}^N \alpha_{ij} W_V h_j$$

$$h_i^{\text{multi}} = W_O [h_i^1 \parallel h_i^2 \parallel \dots \parallel h_i^M]$$

- 6: -双向 Mamba 序列编码

$$h_t^{\text{forward}} \leftarrow \text{Mamba}(h_i^{\text{multi}})$$

$$h_t^{\text{backward}} \leftarrow \text{Mamba}(\text{reverse}(h_i^{\text{multi}}))$$

$$h_t = h_t^{\text{forward}} + h_t^{\text{backward}}$$

- 7: **end for**
 - 8: OCFC (Orthogonal constraint feature clustering) 模块
 - 9: - 投影 Z^L 到正交聚类中心，生成模块特征： Z_G 。
 - 10: -映射到全局表征 $Z_G = f_{\text{proj}}(Z_G)$ 。
 - 11: **分类任务：**
 - 12: -使用全连接网络 (FCN) 预测 $Y = \text{FCN}(Z_G)$ 。
 - 13: **输出** Y
-

2. 实验设计与结果分析

2.1 数据集

本研究中的 rs-fMRI 影像来自两个公开数据库：ABIDE (Autism brain imaging data exchange) [33] 自闭症数据集和 REST-MDD[34] 重度抑郁症数据集。ABIDE 数据集包含来自 19 个国际站点的静息态 fMRI 数据，涉及 1 084 名受试者，其中 526 名为自闭症谱系障碍 (ASD) 患者，558 名为健康对照组。REST-MDD 数据集则包含来自中国 17 家医院 23 个站点的 2 356 个数据，涉及 1 263 名抑郁症 (MDD) 患者和 1 093 名健康对照组。两个数据集的统计信息如表 1 所示。

表 1 数据集的人口学统计信息

Table 1 Demographic and statistical information of the dataset

ABIDE				REST-MDD			
ASD		HC		MDD		HC	
年龄/岁	性别(男/女)	年龄/岁	性别(男/女)	年龄/岁	性别(男/女)	年龄/岁	性别(男/女)
35.2±19	461/65	31.8±25	459/99	47.4±35	455/808	47.7±33	457/636

数据预处理阶段，首先采用 DPABI (Data processing & analysis for brain imaging) [35] 工具箱对 rs-fMRI 图像数据进行处理，具体步骤包括：丢弃每个受试者的前 5 个时间点来消除

扫描初期的信号不稳定性；校正切片间的时间偏差，以确保各切片同步；对功能图像进行刚体配准，以校正头动引起的图像位移；对结构性 T1 加权图像进行分割，提取灰质、白质及脑脊液等组织的概率图；将结构像配准到标准脑模板空间（如 MNI 空间），并将该配准参数应用于功能图像，以实现功能图像的标准化；去除无关变量，如头动参数、白质和脑脊液信号的影响；使用高斯滤波器对功能图像进行空间平滑；去除功能图像时间序列中的低频漂移，确保信号稳定性；应用带通滤波器（0.01–0.08 Hz）以保留与静息态功能相关的低频信号。本文使用图谱 CC200 (Craddock 200)^[36]对大脑进行区域划分，提取各 ROI 脑区的 BOLD 时间序列。利用皮尔森相关系数 (Pearson correlation coefficient, PCC) 计算了不同 ROI 的 BOLD 时间序列之间的成对相关性，最终得到了相应受试者的脑功能连接网络^[37]。

2.2 评估标准

为了评估模型诊断性能，本文选择了准确率 (Accuracy)，特异性 (Specificity)、敏感性 (Sensitivity) 及 F_1 分数 (F_1 -score)，以及 ROC 曲线下面积 (AUC) 五个指标。其计算公式分别为

$$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{TN} + \text{FP} + \text{FN}} \quad (16)$$

$$\text{Specificity} = \frac{\text{TN}}{\text{TN} + \text{FP}} \quad (17)$$

$$\text{Sensitivity} = \frac{\text{TP}}{\text{TP} + \text{FN}} \quad (18)$$

$$F1 = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \quad (19)$$

式中：TP (True positives) 表示真阳性，即被正确识别为正类的样本数量；FN (False negatives) 表示假阴性，即实际为正类但被错误预测为负类的样本数量；TN (True negatives) 表示真负性，即被正确识别为负类的样本数量；FP (False positives) 表示假阳性，即实际为负类但被错误预测为正类的样本数量。

2.3 实验设置

本文实验均在单块 NVIDIA RTX 4090D GPU (24 GB 显存) 上运行，采用 Python 3.8 和 PyTorch 2.2.1 深度学习框架实现。参数优化采用 Adam 算法，初始学习率与权重衰减系数均设置为 1×10^{-4} 。批量大小固定为 64，训练轮次为 200 次。实验数据集采用 5 次随机划分，每次划分 70% 数据作为训练集、10% 数据为验证集，剩余 20% 为测试集。

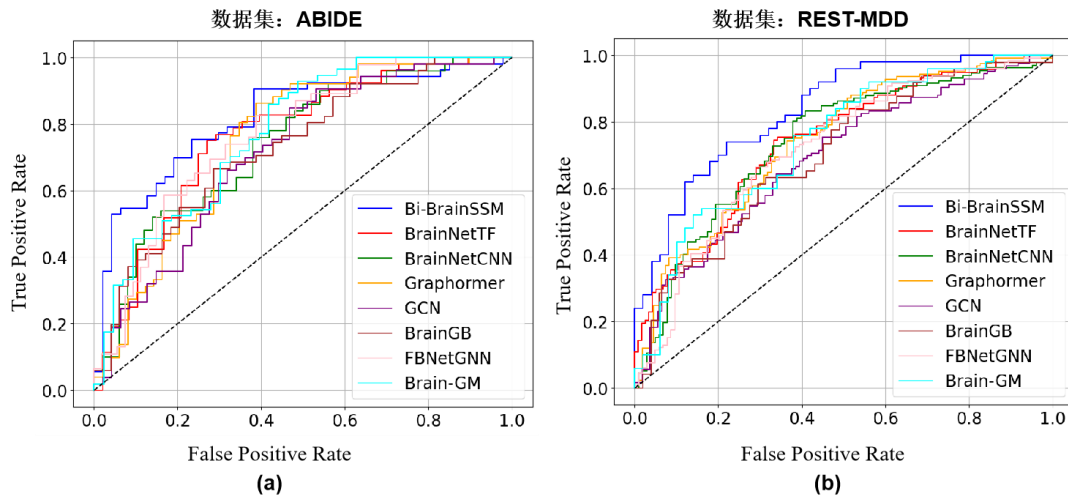
2.4 与当前流行方法比较

为了评估 Bi-BrainSSM 方法，本文选择 7 种不同的对比方法，包括经典的图卷积网络 GCN^[38]；基于 CNN 框架的 BrainNetCNN 和 BrainGB 方法；基于图神经网络的 FBNetGNN 方法；基于 Transformer 的 Graphormer 和 BrainNetTF 方法；以及基于 Mamba 的 BrainGM^[39] 方法。表 2 中展示了不同方法在 ABIDE 和 REST-MDD 数据集上的性能。从表 2 可知，Bi-BrainSSM 方法在各项指标上均优于其他对比方法，尤其在 ACC、AUC 和 F_1 三项全局指标。这表明 Bi-BrainSSM 通过双向 Mamba 状态空间模块能有效捕捉脑功能连接网络中的依赖关系，从而显著提升诊断性能。值得注意的是，同样基于 Mamba 模型的 Brain-GM 方法，其性能却与 FBNetGNN 相近。原因在于，Brain-GM 方法的主干网络是 GCN，仅在网络最末端采用了单向 Mamba 层。图 2 展示了 Bi-BrainSSM 模型与其他对比方法在两个数据集上的 ROC 曲线，可以看出 Bi-BrainSSM 曲线明显优于其他对比方法，尤其在低 FPR 区间。

表 2 不同方法在 ABIDE 和 REST-MDD 数据集上的性能

数据集	方法	ACC	AUC	SEN	SPE	F_1
-----	----	-----	-----	-----	-----	-------

ABIDE	GCN	63.6±2.3	69.7±3.2	67.1±3.2	67.5±4.7	65.8±2.1
	BrainNetCNN	66.8±3.1	72.1±2.8	64.2±4.5	69.7±6.6	66.2±2.2
	FBNetGNN	68.1±1.8	76.7±2.2	66.4±5.8	63.8±5.7	68.3±1.3
	BrainGB	64.5 ± 2.7	69.2 ± 1.6	65.5±6.4	64.9±7.5	64.6±1.9
	Graphormer	65.7±3.7	77.6±3.1	66.6±3.9	70.1±4.4	66.7±1.9
	BrainNetTF	71.1±2.2	78.8±4.2	73.0±8.8	68.9±7.2	70.9±1.1
	Brain-GM	69.2±3.0	77.4±2.7	72.3±4.1	69.8±4.1	71.2±1.4
	Bi-BrainSSM	75.8±2.9	81.2±0.8	70.3±3.8	81.5±2.6	75.1±2.8
REST-MDD	GCN	65.1±2.4	70.4±3.6	69.0±4.1	71.8±3.3	67.9±2.4
	BrainNetCNN	68.4±1.9	73.7±3.1	66.2±3.8	70.5±5.6	67.8±2.3
	FBNetGNN	66.7±3.0	74.2±1.9	65.6±4.6	62.4±4.5	65.8±1.8
	BrainGB	67.2±2.7	70.7±2.4	68.8±3.0	67.9±5.1	66.1±2.2
	Graphormer	70.1±2.9	78.6±3.3	70.4±3.5	73.2±4.2	69.7±2.6
	BrainNetTF	75.4±1.8	81.5±3.6	74.4±5.1	70.8±4.8	73.4±2.8
	Brain-GM	68.8±1.2	76.3±2.8	69.5±3.2	68.5±3.2	67.4±1.5
	Bi-BrainSSM	77.3±2.6	83.3±1.5	72.1±3.4	82.4±3.1	76.5±2.7



(a) ABIDE 数据集

(b) REST-MDD 数据集

(a) ABIDE dataset

(b) REST-MDD dataset

图 2 不同方法在 ABIDE 和 REST-MDD 数据集上的 ROC 曲线

Fig.2 ROC curves of different methods on ABIDE and REST-MDD datasets

2.5 与当前流行方法的计算效率比较

为了验证 Bi-BrainSSM 方法在计算效率的优势, 本文通过比较不同方法在 ABIDE 和 REST-MDD 数据集上的训练时间和 GPU 显存占用, 结果如图 3 所示。

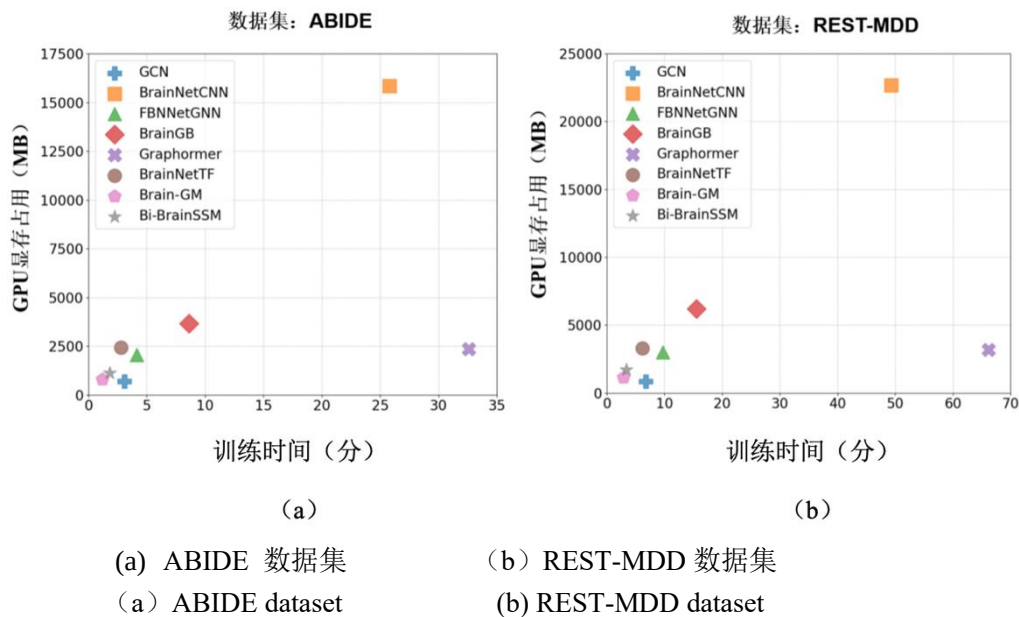


图 3 不同模型的训练时间和 GPU 显存占用对比
Fig.3 Comparison of training time and GPU memory usage for different models

如图 3 所示, Bi-BrainSSM 的训练时间和显存占用明显低于其他基于 CNN 和 Transformer 的诊断模型。具体而言, 与基于 Transformer 架构的 BrainNetTF 相比, Bi-BrainSSM 的计算速度提升了接近 50%; 与基于 CNN 的 BrainNetCNN 相比, 速度提升高达约 94%。Graphormer 模型的训练时间最长, 主要是由于其没有针对大脑网络的特定属性进行优化, 而且其生成边缘特征的计算进一步增加了计算负担。BrainNetCNN 显存占用最高的原因在于其“边到边”卷积层的实现方式非常消耗资源。而 Brain-GM 计算开销较低的原因在于采用了基于稀疏邻接矩阵的 GNN 作为主干网络, 并将 Mamba 用于处理短得多的图级时间序列, 而非长节点级空间序列, 所以其计算开销自然远低于我们的模型。实验结果表明 Bi-BrainSSM 能够在保证高性能的同时, 显著降低对硬件资源的需求。

2.6 消融实验

为量化各组成模块对 Bi-BrainSSM 性能的独立贡献, 并验证整体架构的设计合理性, 本文在 ABIDE 与 REST-MDD 数据集上实施了系统性消融实验。具体做法为: 依次移除或替换关键模块 (ABM 模块中的双向 Mamba、多头注意力机制, 以及 OCFC 模块), 比较所得模型版本在分类准确率 (ACC) 与曲线下面积 (AUC) 上的差异。从表 3 的实验结果可以看出, Bi-BrainSSM 方法中不同模块对最终性能都具有重要作用。移除注意力机制后模型性能出现最显著的下降, 这凸显了该模块在建模脑区间空间交互中的重要作用。移除 OCFC 模块后, 模型性能下降, 因为该模块对于生成高质量的图级表征至关重要, 证明了其在抑制特征冗余、强化全局判别性方面的必要性。移除双向 Mamba 模块后, 模型性能亦有明显下滑, 这表明双向序列建模对于完整捕捉脑功能网络中的长程依赖关系具有不可或缺的价值。

表 3 在 ABIDE 和 RSET-MDD 数据集上消融实验结果

Table 3 Ablation experimental results on ABIDE and RSET-MDD datasets						%
注意力机制	双向 Mamba	OCFC模块	ABIDE		REST-MDD	
			Acc	AUC	Acc	AUC
√		√	74.5±2.7	79.3±1.1	75.1±3.2	80.6±1.6

√	√		70.1± 3.2	74.7±2.2	73.3±1.9	76.8±2.5
	√	√	60.2±1.8	63.6 ± 1.3	61.1±1.6	65.7±2.7
√	√	√	75.8±2.9	81.2±0.8	77.3±2.6	83.3±1.5

2.7 脑功能连接的可解释性分析

为了进一步证明 Bi-BrainSSM 所识别的生物标志物的可解释性,本文采用了一种基于梯度的事后可解释性分析 (Post-hoc explainability analysis) 方法,将学习到的特征权重映射到原始脑空间中进行可视化分析。当模型训练收敛后,可通过计算分类输出(即特定疾病的预测得分)关于输入特征(即功能连接矩阵中的每一个连接强度)的偏导数(Partial derivatives),该梯度值的大小直接反映了每个输入特征对模型决策的敏感度与贡献度。梯度绝对值越高的功能连接,即被认为是模型进行分类时所依赖的关键特征。最终,筛选出贡献度最高的前十脑区连接,并利用 BrainNet Viewer^[40]工具进行三维可视化,其结果如图 4 所示。

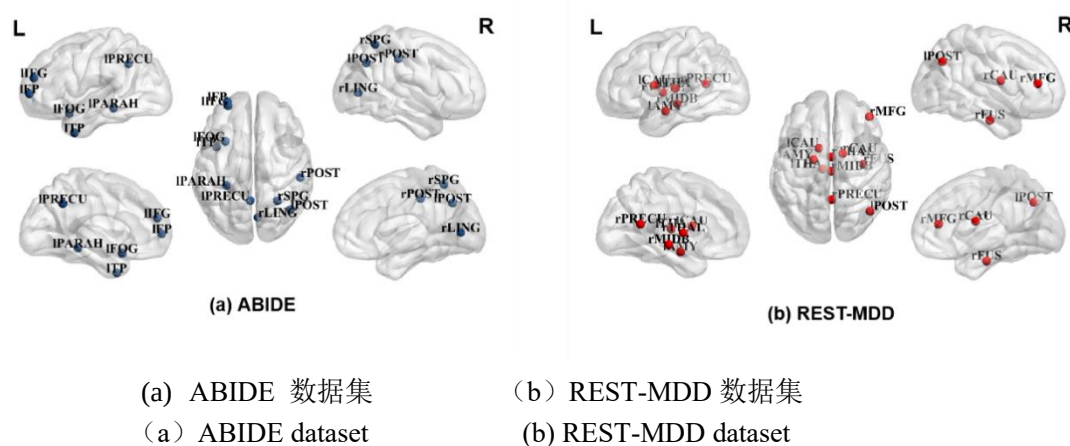


图 4 Top 10 最具辨别力的脑区可视化

Fig.4 Top 10 most discriminative brain area visualizations

图 4 的(a)中展示了与自闭症最相关的前 10 个脑区,主要分布在控制感觉与运动功能(IPOST, IPRECU, rPOST),高级认知与执行功能(IFP, IIFG, IFPG, rSPG),情绪与记忆功能(IPARAH, IFOG, ITP),视觉和多感官整合(rLING, rSPG)等区域。现有研究已证实 ASD 与上述部分脑区有直接关系,Padmanabhan 等人^[41]的研究表明,ASD 患者的 IPARAH (左侧海马旁回) 异常可能导致记忆和环境适应能力受损,而 RLING (右侧舌回) 异常则会导致自闭症患者在面部表情和视觉场景处理上有障碍。图 4 的(b)中展示了与抑郁症最相关的前 10 个脑区。主要分布在情绪处理与调控中心(IAMY, ITHA, rTHAL),奖赏、动机与学习环路(ICAU, rCAU, rMIDB),高级认知与执行功能(rMFG),自我认知与感觉整合(rPRECU, IPOST, rFUS)等区域。现有研究^[42]表明 IAMY (杏仁核) 与 MDD 存在密切联系,抑郁症患者常表现为杏仁核的过度激活,导致情绪反应过于敏感或对负面情绪的反应被放大。Dam 等^[43]的研究发现,抑郁症患者的 ICAU, rCAU (尾状核) 可能表现出功能连接异常,导致抑郁症的典型症状,如对活动缺乏兴趣。上述结果验证了 Bi-BrainSSM 能够有效识别疾病相关生物标志物,提升模型可解释性。

3. 结束语

本文提出了一种基于选择性状态空间模型 (SSM) 的脑网络分析方法 Bi-BrainSSM。该方法的 ABM 模块,将多头注意力机制和双向 Mamba 架构相结合,能够有效挖掘脑网络中

节点间的复杂依赖关系，并通过聚类操作提取出关键的图级特征。实验结果表明，Bi-BrainSSM 在 ABIDE 和 REST-MDD 数据集上的分类性能明显优于其他对比模型。同时，Bi-BrainSSM 在计算资源消耗方面也表现出显著优势。未来的研究将进一步考虑多站点的影响，探索更高效的特征提取模块，进一步提升模型的适用性和扩展性，有望为更多脑疾病分析提供有力支持，并在实际应用中发挥更大的应用价值。

参考文献：

- [1] Wang Y, Guo Y. A hierarchical independent component analysis model for longitudinal neuroimaging studies[J]. *NeuroImage*, 2019, 189: 380-400.
- [2] Smith S M, Miller K L, Salimi-Khorshidi G, et al. Network modelling methods for FMRI[J]. *Neuroimage*, 2011, 54(2): 875-891.
- [3] Wang Y, Kang J, Kemmer P B, et al. An efficient and reliable statistical method for estimating functional connectivity in large scale brain networks using partial correlation[J]. *Frontiers in Neuroscience*, 2016, 10: 123.
- [4] Simpson S L, Bowman F D B, Laurienti P J. Analyzing complex functional brain networks: fusing statistics and network science to understand the brain[J]. *Statistics Surveys*, 2013, 7: 1.
- [5] 张道强, 接标. 基于机器学习的脑网络分析方法及应用 [J]. *数据采集与处理*, 2015, 30(1): 68-76.
- [6] Stam C J, Reijneveld J C. Graph theoretical analysis of complex networks in the brain[J]. *Nonlinear Biomedical Physics*, 2007, 1(1): 3.
- [7] Honey C J, Sporns O, Cammoun L, et al. Predicting human resting-state functional connectivity from structural connectivity[J]. *Proceedings of the National Academy of Sciences*, 2009, 106(6): 2035-2040.
- [8] Ivanoska I, Trivodaliev K, Kalajdziski S, et al. Statistical and machine learning link selection methods for brain functional networks: Review and comparison[J]. *Brain Sciences*, 2021, 11(6): 735.
- [9] Herzog R, Rosas F E, Whelan R, et al. Genuine high-order interactions in brain networks and neurodegeneration[J]. *Neurobiology of Disease*, 2022, 175: 105918.
- [10] Waikhom L, Patgiri R. A survey of graph neural networks in various learning paradigms: Methods, applications, and challenges[J]. *Artificial Intelligence Review*, 2023, 56(7): 6295-6364.
- [11] Kawahara J, Brown C J, Miller S P, et al. BrainNetCNN: Convolutional neural networks for brain networks; towards predicting neurodevelopment[J]. *NeuroImage*, 2017, 146: 1038-1049.
- [12] 李文杰, 韩秋蕾, 孙大文, 等. 图神经网络在脑网络中的应用研究综述[J]. *计算机应用*, 2024, 44(S1): 124-132.
- [13] Li X, Zhou Y, Dvornek N, et al. BrainGNN: Interpretable brain graph neural network for fMRI analysis[J]. *Medical Image Analysis*, 2021, 74: 102233.
- [14] Kan X, Cui H, Lukemire J, et al. FBNetGen: Task-aware GNN-based fMRI analysis via functional brain network generation[C]//*International Conference on Medical Imaging with Deep Learning*. [S.l.]: PMLR, 2022: 618-637.
- [15] Cui H, Dai W, Zhu Y, et al. BrainGB: A benchmark for brain network analysis with graph neural networks[J]. *IEEE Transactions on Medical Imaging*, 2022, 42(2): 493-506.
- [16] Vaswani A. Attention is all you need[J]. *Advances in Neural Information Processing Systems*, 2017.
- [17] Veličković P, Cucurull G, Casanova A, et al. Graph attention networks[J]. *arXiv preprint arXiv:1710.10903*, 2017.

- [18] Dwivedi V P, Bresson X. A generalization of transformer networks to graphs[J]. arXiv preprint arXiv: 2012.09699, 2020.
- [19] Kreuzer D, Beaini D, Hamilton W, et al. Rethinking graph transformers with spectral attention[J]. Advances in Neural Information Processing Systems, 2021, 34: 21618-21629.
- [20] Ying C, Cai T, Luo S, et al. Do transformers really perform badly for graph representation?[J]. Advances in Neural Information Processing Systems, 2021, 34: 28877-28888.
- [21] Kan X, Dai W, Cui H, et al. Brain network transformer[J]. Advances in Neural Information Processing Systems, 2022, 35: 25586-25599.
- [22] 李建彤, 姚垚, 高俊涛, 等. 基于 HubGLasso 注意力机制的脑网络分类研究[J]. 计算机技术与发展, 2024, 34(9): 131-137.
- [23] 刘子凡, 孙道清, 赵善辉, 等. 基于功能脑网络和图特征学习的 ADHD 分类模型[J]. 自动化学报, 2025, 51(7): 1651-1661.
- [24] Shehzad A, Zhang D, Yu S, et al. Dynamic graph transformer for brain disorder diagnosis[J]. IEEE Journal of Biomedical and Health Informatics, 2025.
- [25] Zhang M, Saab K K, Poli M, et al. Effectively modeling time series with simple discrete state spaces[J]. arXiv preprint arXiv: 2303.09489, 2023.
- [26] Gu A, Dao T. Mamba: Linear-time sequence modeling with selective state spaces[J]. arXiv preprint arXiv: 2312.00752, 2023.
- [27] Gu A, Dao T, Ermon S, et al. HiPPO: Recurrent memory with optimal polynomial projections[J]. Advances in neural information processing systems, 2020, 33: 1474-1487.
- [28] Aoki M. State space modeling of time series[M]. Springer Science & Business Media, 2013.
- [29] Gu A, Johnson I, Goel K, et al. Combining recurrent, convolutional, and continuous-time models with linear state space layers[J]. Advances in Neural Information Processing Systems, 2021, 34: 572-585.
- [30] Fu D Y, Dao T, Saab K K, et al. Hungry hungry hippos: Towards language modeling with state space models[J]. arXiv preprint arXiv: 2212.14052, 2022.
- [31] 张东风, 余道衡, 张里千. 一种全局优化新方法——正交聚类法[C]//中国电子学会电子线路与系统学会 1987 年第七届年会. 1987.
- [32] Glorot X, Bengio Y. Understanding the difficulty of training deep feedforward neural networks[C]//Proceedings of the Thirteenth International Conference on Artificial Intelligence and Statistics. [S.l.]: JMLR Workshop and Conference Proceedings, 2010: 249-256.
- [33] Craddock C, Benhajali Y, Chu C, et al. The neuro bureau preprocessing initiative: Open sharing of preprocessed neuroimaging data and derivatives[J]. Frontiers in Neuroinformatics, 2013, 7(27): 5.
- [34] Yan C G, Chen X, Li L, et al. Reduced default mode network functional connectivity in patients with recurrent major depressive disorder[J]. Proceedings of the National Academy of Sciences, 2019, 116(18): 9078-9083.
- [35] Yan C G, Wang X D, Lu B. DPABISurf: Data processing & analysis for brain imaging on surface[J]. Science Bulletin, 2021, 66(24): 2453-2455.
- [36] Craddock R C, James G A, Holtzheimer III P E, et al. A whole brain fMRI atlas generated via spatially constrained spectral clustering[J]. Human Brain Mapping, 2012, 33(8): 1914-1928.
- [37] 黄剑辉, 马迪, 张礼. 多视图低秩子空间的图结构学习多站点自闭症诊断方法[J]. 数据采集与处理, 2024, 39(4): 984-995.
- [38] Kipf T N, Welling M. Semi-supervised classification with graph convolutional networks[J].

arXiv preprint arXiv: 1609.02907, 2016.

[39] Wang J, Lang J, Yang L Z, et al. Learning dynamic brain network representation based on graph Mamba architecture[C]//2024 IEEE International Conference on Bioinformatics and Biomedicine (BIBM). IEEE, 2024: 5143-5150.

[40] Xia M, Wang J, He Y. BrainNet viewer: A network visualization tool for human brain connectomics[J]. PloS One, 2013, 8(7): e68910.

[41] Padmanabhan A, Lynch C J, Schaer M, et al. The default mode network in autism[J]. Biological Psychiatry: Cognitive Neuroscience and Neuroimaging, 2017, 2(6): 476-486.

[42] Murray E A, Wise S P, Drevets W C. Localization of dysfunction in major depressive disorder: Prefrontal cortex and amygdala[J]. Biological Psychiatry, 2011, 69(12): e43-e54.

[43] Dam S, Batail J M, Robert G H, et al. Structural brain connectivity and treatment improvement in mood disorder[J]. Brain Connectivity, 2024, 14(4): 239-251.

作者简介:



张力 (2000-), 男,
硕士研究生, 研究方
向: 医学影像学分析,
E-mail:
zhanglibeiyoung@njfu.edu.cn.



顾恒 (1999-),
男, 硕士研究
生, 研究方向:
医学影像学分
析。



黄硕 (1992-), 男,
讲师, 研究方向: 模式
识别、神经计算与认知
解码等。



马越 (1986-), 女,
讲师, 研究方向: 神
经内科、精神病学等。



任凯 (1992-),
男, 副教授, 研
究方向: 深度学
习、力学。



张礼 (1985 -), **通信
作者**, 男, 副教授, 硕
士导师, 研究方向: 机
器学习、医学影像、生
物信息学等, E-mail:
lizhang@njfu.edu.cn.