arXiv:2602.09294v1 [cs.CE] 10 Feb 2026

BrainTAP：通过自适应蒸馏与选择性先验融合进行脑部疾病预测

雷振宇*，张爱英*，王松†，范涵*，李俊东*

*弗吉尼亚大学 †中佛罗里达大学

摘要

要在青少年脑认知发展（ABCD）研究等大规模神经影像学队列中，根据脑网络预测临床结局，需要有效整合功能连接（FC）与结构连接（SC），同时融入专家的神经生物学知识。然而，现有多模态融合方法要么融合较浅，要么过度同质化了功能连接和结构连接本身具有的异质性特征；同时，专家定义的解剖学先验未得到充分利用，且整合方式较为静态。为解决这些局限，我们提出了采用自适应相互蒸馏和选择性先验融合的脑Transformer（BrainTAP）。我们提出自适应互蒸馏（AMD），通过可学习的蒸馏—保留比例，实现模态间逐层信息交换，在保留模态特定信号的同时捕捉跨模态协同效应。我们进一步设计了选择性先验融合（SPF），以自适应方式整合专家定义的解剖学先验。在ABCD数据集上对注意力相关障碍进行预测的评估结果表明，BrainTAP的表现优于当前最先进的基线方法，证明了其在脑部疾病预测方面的有效性。

索引词— 脑部疾病预测，多模态融合，神经生物学先验

1. 引言

由神经影像学得出的脑连接网络提供了与认知和心理健康相关的脑组织结构的有意义表征 [1]。功能连接（FC）反映了静息态功能磁共振成像（fMRI）期间脑区之间血氧水平依赖（BOLD）信号波动的相关性，可捕捉动态的功能交互。结构连接（SC）表示根据弥散磁共振成像（dMRI）纤维束示踪估算的脑区间解剖白质通路强度，反映大脑的物理结构。近期研究日益广泛地利用这些网络预测脑部疾病。基于图的方法（如BrainGNN）将图卷积应用于功能连接网络 [2], ，而基于 Transformer 的模型（如脑网络变换器）则利用注意力机制来学习成对连接强度 [3]。然而，大多数现有方法依赖单一模态，忽略了功能连接和结构连接所提供的互补信息。

为发挥FC与SC的互补优势，研究者转向了多模态融合框架。早期方法采取了较为浅层的路径，仅简单地拼接或平均FC与SC特征 [4] ，或在多视图嵌入学习框架中将其视为相互独立的“视图” [5], ，因而未能捕捉这两种模态之间更深层次、更为复杂的协同作用。为增强两种模态之间的交互，人们提出了更为先进的方法。这些方法包括融合FC与SC的混合邻接矩阵 [6], 、使每种模态能够直接关注另一种模态的交叉注意力机制 [7], ，以及不仅应用于全脑图层面、也应用于区域尺度的多尺度融合策略 [8], ，从而实现更深层次的整合并提升预测性能。然而，大多数深度融合框架可能因过度强调二者的整合，而不经意地将本质上异质的FC与SC信息同质化，从而损害对预测至关重要的模态特异性特征。除方法学问题之外，大多数模型仅依赖影像数据，而忽视了专家推导的神经生物学先验的价值，例如已与特定认知功能或临床结局相关联的连接模式方面的既有知识。此类先验能够引导模型的关注重点、提升可解释性，并使预测建立在具有解剖学意义的基础之上。然而，在纳入这些先验时，它们通常被施加为静态二值掩码，将所有先验脑区视为同等信息量 [9]。这种粗粒度约束忽视了区域贡献的显著变异性，而有效整合需要细粒度的区域特异性重要性权重。

为应对这些挑战，我们提出了一种新颖的框架，称为**采用自适应互蒸馏与选择性先验融合的脑Transformer（BrainTAP）**。针对第一个挑战，我们引入**自适应互蒸馏（AMD）**，这是一种通过逐层相互蒸馏，使两种模态逐步交换信息的机制。在每一层中，仅蒸馏一部分特征，其余特征则保持模态完整，从而避免过度同质化。为此，AMD学习蒸馏–完整比，动态控制各层之间的信息共享程度。这种自适应

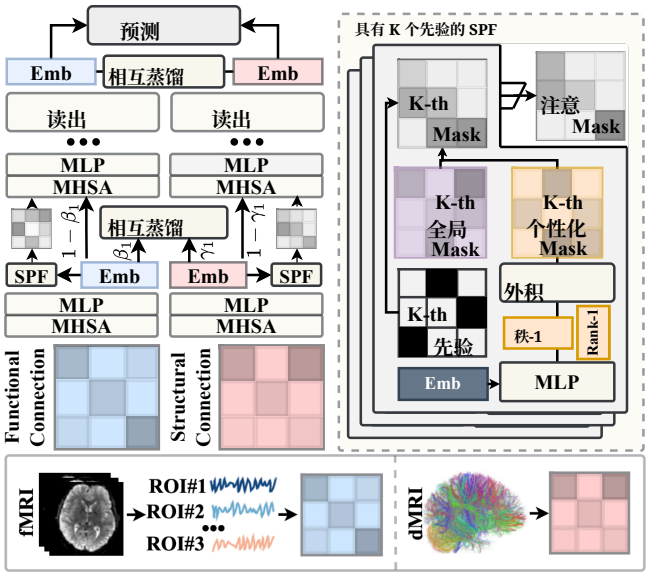

图 1. BrainTAP 框架概览。FC 和 SC 分别以蓝色和粉色邻接矩阵表示。

设计使模型能够在有益时蒸馏跨模态知识，同时保留对捕捉异质性功能连接（FC）–结构连接（SC）特征至关重要的模态特异性信号。针对第二个挑战，我们提出**选择性先验融合（SPF）**，这是一种以选择性方式整合专家定义的解剖学先验的框架。SPF并非在所有区域统一施加静态二值掩码，而是学习区域特异性掩码，以调节神经生物学先验的影响。这些掩码由两个互补组件构成：一个全局层级组件，用于编码跨受试者共享的群体层级神经生物学重要性；另一个个体层级组件，则以受试者特异性嵌入为条件。两者共同生成选择性注意力掩码，在Transformer架构中针对每位受试者自适应地调节连接的重要性。ABCD [10]数据集上的实证结果表明，BrainTAP在预测多种注意力相关障碍方面表现优异，超越了最先进的基线方法，并验证了所提框架的有效性。

2. 方法

图 1 展示了 BrainTAP 的整体概览。该框架基于Transformer架构构建，以捕捉细致的脑连接，由两个模块组成：(1) 自适应互蒸馏（AMD），通过逐层互蒸馏实现受控的跨模态信息交换，同时保留模态特定信号；(2) 选择性先验融合（SPF），选择性地将解剖学先验与可学习的全局及个性化先验掩码相结合。

个性化先验掩码。

问题定义。对于每个受试者，我们观测到两个由 N 感兴趣区域构成的加权脑网络，其中 FC 为 $A^F \in \mathbb{R}^{N\times N}$，SC 为 $A^S \in \mathbb{R}^{N\times N}$。我们的目标是根据这两个网络作为输入，预测疾病状态 $y \in \{0,1\}$。

2.1. Transformer架构

对于模态指示符 $M \in \{F, S\}$，ROI i 的连接指纹是 i 第 i 行 $\mathbf{a}_i^M$，通过模态特定映射 ϕ^M 嵌入为 token $X^{M,0}$ 其各行 $\mathbf{x}_i^{M\,0} =. \phi^M(\mathbf{a}_i^M)$ 用作 ROI i 的输入特征。每种模态均由一个 L层 Transformer 编码器进行编码，该编码器包含多头自注意力（MHSA）和一个多层感知机:

$$P_h^{M,\ell} = \mathrm{softmax}\big(Q_h K_h^\top/\sqrt{d}\big),$$
$$\mathrm{MHSA}(X) = \mathrm{Concat}_h\big(P_h^{M,\ell} V_h\big) W_O,$$
$$Z^{M,\ell} = X^{M,\ell-1} + \mathrm{MHSA}(X^{M,\ell-1}),$$
$$X^{M,\ell} = Z^{M,\ell} + \mathrm{MLP}(Z^{M,\ell}),$$

其中$Q_h = XW_{Q_h}$、$K_h = XW_{K_h}$、$V_h = XW_{V_h}$、d 为每个注意力头的维度，且 $\ell = 1, \ldots, L$。最终表征 $X^{F,L}$ 和 $X^{S,L}$ 通过对 N 个 ROI 进行均值池化并取平均来聚合，然后输入多层感知机分类器，以生成疾病预测 $\hat{y}$。

2.2. 自适应相互蒸馏

为促使表征既互补又对齐，我们在每一层的模态之间蒸馏信息，并通过残差连接将其注入回去。设 g_F，g_S 表示映射至共享 d 维蒸馏空间的多层感知机投影头，h_F，h_S 表示其对应的逆投影。基于多层感知机处理后的表征，我们计算

$$Z^{F,\ell} = g_F(X^{F,\ell}), \qquad Z^{S,\ell} = g_S(X^{S,\ell}).$$

对于每个感兴趣区域 i，我们定义经过软化的逐通道分布

$$p_i^{F,\ell} = \mathrm{softmax}(Z_{i,:}^{F,\ell}/\tau), \qquad p_i^{S,\ell} = \mathrm{softmax}(Z_{i,:}^{S,\ell}/\tau),$$

其中 $\tau > 0$ 是温度参数。对称的互蒸馏损失用于对齐来自两种模态的分布:

$$\mathcal{L}_{\mathrm{distill}}^{\ell} = \frac{1}{N}\sum_{i=1}^{N}[\beta^\ell \,\mathrm{KL}(p_i^{F,\ell} \,\|\, p_i^{S,\ell}) + \gamma^\ell \,\mathrm{KL}(p_i^{S,\ell} \,\|\, p_i^{F,\ell})].$$

蒸馏得到的信息通过残差连接注入不同模态:

$$\Delta X^{F,\ell} = \gamma^\ell\, h_F(Z^{S,\ell}), \qquad \Delta X^{S,\ell} = \beta^\ell\, h_S(Z^{F,\ell}),$$

从而得到更新后的词元表示

$$\widetilde{X}^{F,\ell} = (1-\gamma^\ell)X^{F,\ell} + \Delta X^{F,\ell}, \widetilde{X}^{S,\ell} = (1-\beta^\ell)X^{S,\ell} + \Delta X^{S,\ell},$$

其中，可学习的 β^ℓ，$\gamma^\ell \in [0,1]$ 分别控制 SC 到 FC 和 FC 到 SC 传递的蒸馏强度。这些经过精炼的表征 $\widetilde{X}^{M,\ell}$ 作为第 $\ell+1$ 层的输入。

模型	ADHD	焦虑	OCD	ATT
MLP	55.45 ± 0.90	57.62 ± 1.22	54.38 ± 0.63	62.12 ± 0.15
MaskedGCN	58.15 ± 2.59	66.19 ± 2.54	57.17 ± 0.63	$\underline{67.40} \pm 2.27$
RH-BrainFS	$\underline{60.10} \pm 0.50$	61.44 ± 1.61	57.88 ± 0.11	64.40 ± 0.01
CrossGNN	57.95 ± 0.20	$\underline{67.25} \pm 0.62$	$\underline{60.15} \pm 0.13$	62.63 ± 0.10
BrainTAP	$\mathbf{61.25} \pm 2.03$	$\mathbf{73.08} \pm 2.02$	$\mathbf{61.67} \pm 2.58$	$\mathbf{73.77} \pm 0.59$

表1。主要结果（曲线下面积，3 个随机种子）。最佳和次佳分数以**粗体**和下划线标示。

2.3. 选择性先验融合

由专家知识推导出的解剖学先验可提供宝贵的结构约束，但将其简单地作为静态二值掩码统一施加并非最优。我们提出**选择性先验融合（SPF）**，通过学习共享的群体层面掩码和受试者特异性个体层面掩码，实现自适应性和可解释性。

我们从 K 二值解剖先验掩码 $\{\Pi_k \in \{0,1\}^{N\times N}\}_{k=1}^{K}$ 开始，这些掩码编码了专家定义的关键区域。为确保完全覆盖，我们为未分配给任何先验的连接定义互补区域 $\Pi_{\text{free}} = \mathbf{1} - \max_k \Pi_k$，从而得到完整集合 $\mathcal{P} = \{\Pi_1, \ldots, \Pi_K, \Pi_{\text{free}}\}$

对于每个先验 Π_k，我们学习两个组件：一个在所有受试者之间共享的全局掩码矩阵 $W_k^{(g)} \in \mathbb{R}^{N\times N}$，以及一个适应个体差异的个性化秩一掩码矩阵。给定学习到的受试者嵌入$\mathbf{z}$，个性化组件生成如下

$$W_k^{(r)}(\mathbf{z}) = \frac{\alpha_k(\mathbf{z})}{2}\Big[u_{k,A}(\mathbf{z})u_{k,B}(\mathbf{z})^\top + u_{k,B}(\mathbf{z})u_{k,A}(\mathbf{z})^\top\Big],$$

其中，多层感知机从$\mathbf{z}$预测标量权重 $\alpha_k(\mathbf{z}) \in \mathbb{R}$ 和秩一因子$u_{k,A}(\mathbf{z})u_{k,B}(\mathbf{z}) \in \mathbb{R}^N$。这种低秩分解实现了高效的个性化掩码学习。

随后，我们聚合所有先验的贡献，形成统一的重要性分数矩阵：

$$S(\mathbf{z}) = \text{sym0}\Big(\sum_{k\in\mathcal{P}}(W_k^{(g)} + W_k^{(r)}(\mathbf{z})) \odot \Pi_k\Big),$$

其中，$\text{sym0}(M) = \frac{1}{2}(M + M^\top) - \text{diag}(\text{diag}(M))$用于确保对称性并移除自连接。由此得到的分数矩阵 $S(\mathbf{z})$反映了针对给定受试者，每个感兴趣区域对的预测重要性。这些分数通过Sigmoid函数转换为软性、可微的门控：

$$G(\mathbf{z}) = \sigma(S(\mathbf{z})/\tau) \in (0,1)^{N\times N},$$

其中 $\tau > 0$ 控制门控机制的锐度。学习得到的门控 $G(\mathbf{z})$ 以注意力偏置的形式融入模型，通过 $QK^\top/\sqrt{d} \mapsto QK^\top/\sqrt{d} + \eta G(\mathbf{z})$ 增强 Transformer 注意力逻辑值，其中 η用于缩放偏置项。通过这一设计，SPF 通过仅关注真正重要的区域，有选择地整合解剖学知识。

3. 实验

本节旨在回答五个研究问题：**研究问题1**：与现有基线方法相比，BrainTAP 的表现如何？**研究问题2**：BrainTAP 的各个组件分别有何贡献？**研究问题3**：AMD 如何自适应地调节模态交互强度？**研究问题4**：SPF 如何利用先验知识？**研究问题5**：BrainTAP 如何识别用于预测的关键感兴趣区域？

3.1. 实验设置

数据集。 本研究使用ABCD研究的静息态功能磁共振成像（rs-fMRI）数据。在基线时招募的11,099名9至11岁参与者中，排除缺少可用rs-fMRI扫描或未通过质量控制的参与者后，剩余7,844名。数据采用标准化的ABCD流程 [10]进行预处理。使用Glasser图谱对皮层区域进行分区，使用Aseg图谱对皮层下区域进行分区。功能连接通过感兴趣区域（ROI）时间序列之间的统计关联进行估计。我们的分析重点是预测四种儿童精神疾病：强迫症（OCD）、焦虑、注意缺陷多动障碍（ADHD）和一般注意障碍（ATT）。提供了四种先验掩码：更新与工作记忆修订相关；表现用于标示错误监控和表现调整；抑制支持反应抑制；以及注意介导自上而下分配和持续专注。

基线。 比较包括两种**单模态**基线方法：(1)多层感知机 和 (2) *MaskedGCN* ，以及两种**多模态**基线方法：(3)*RH-BrainFS* [11]和(4)*CrossGNN* [12]。对于单模态基线方法，我们将每个模型分别应用于功能连接和结构连接，并对最后一层的隐藏状态取平均。

实现细节。 我们使用曲线下面积作为评估指标。学习率设为 1×10^{-4}，权重衰减设为1×10^{-6}，批量大小设为32。温度参数 τ 固定为 1，隐藏维度 d 设为 64，层数设为 3。

3.2. 主要结果

为回答**研究问题1**，我们在表1中给出了结果，并观察到：(1) BrainTAP始终优于所有基线方法，证实了其在预测不同脑部疾病方面的有效性。(2) 多模态基线方法优于单模态方法，凸显了功能连接和结构连接表征之间交互的价值。(3) BrainTAP在焦虑和ATT上的表现更好，这表明这些疾病可能对脑结构和功能产生更大影响，且更容易捕捉到异常连接。为探究**研究问题2**，我们进行了消融研究，以量化各组件的贡献。我们使用“去除”表示移除；“*G-SPF*”和“*P-SPF*”分别表示全局SPF变体和个性化SPF变体。如表2所示，移除任何组件都会降低表现，证实了所有组件的有效性。

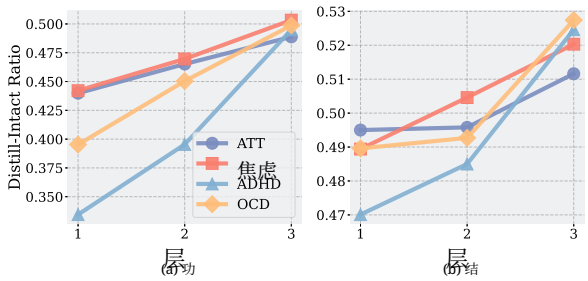

图2。功能连接与结构连接各AMD层的蒸馏–完整比。数值越大，表明跨模态交互越强。

Task	无AMD	无G-SPF	无P-SPF	BrainTAP
ADHD	59.74 ± 0.50	59.03 ± 1.26	59.77 ± 0.36	**61.25** ± 2.03
焦虑	69.19 ± 2.57	69.59 ± 0.18	71.15 ± 1.24	**73.08** ± 2.02
OCD	59.21 ± 0.65	58.34 ± 2.24	60.19 ± 0.27	**61.67** ± 2.58
ATT	70.40 ± 1.07	67.65 ± 2.12	71.40 ± 1.86	**73.77** ± 0.59

表 2. 消融研究结果（曲线下面积，3 个随机种子）。最佳和次佳分数以**粗体**和下划线标示。

3.3. 进一步分析

为回答**研究问题3**，我们计算了AMD模块中的蒸馏–完整比率。如图2所示，功能连接和结构连接的比率均随深度增加而上升，这表明浅层主要保留模态特异性信息，而更深层则逐步整合跨模态信息。这些结果表明，AMD能够有效地实现对模态间信息交换的自适应控制。

为回答**研究问题4**，我们计算学习得到的注意力掩码下先验脑区的平均倒数排名（MRR）。我们首先根据注意力掩码值对先验脑区排序，计算受试者级平均倒数排名，再跨受试者取平均，得到先验—疾病关联得分。如图3所示，认知控制环路内部的连接——尤其是支持目标更新和反应抑制的连接——在所有疾病中均表现出强关联，提示执行过程的破坏可能是这些疾病共有易感性的基础。这与已有证据一致：额顶控制网络（FPN）和扣带-岛盖网络/显著性网络（CON/SN）是认知控制的核心，且在这些障碍中均出现失调 [13]。具体而言，注意缺陷多动障碍研究发现，额顶控制网络及额叶-纹状体连接的改变与抑制和注意控制受损有关 [14]；强迫症研究显示，在抑制控制和错误处理过程中，前额叶、眶额及扣带-岛盖活动异常 [15]；比较研究则强调，支持干扰控制和表现监控的网络存在重叠缺陷 [16]。注意力构念在我们的功能连接结果中并不突出，这很可能是因为我们使用了静息态功能磁共振成像，此时注意力网络的活跃程度低于外部导向任务期间 [17]。

为回答**研究问题5**，我们展示平均学习得到的注意力掩码中权重最高的连接——

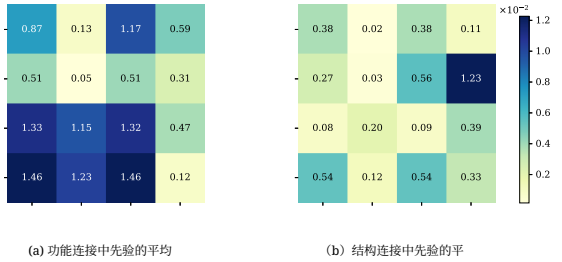

图 3。学习得到的注意掩码下先验的平均倒数排名。平均倒数排名越高，既往疾病关联越强。

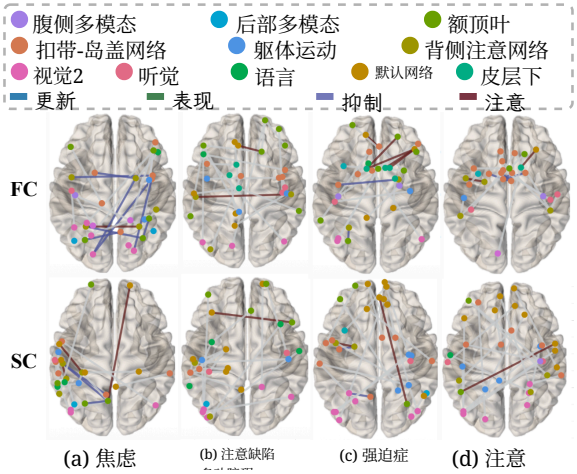

图 4。学习得到的注意力掩码中排名前 0.01% 的连接。浅灰色边不属于任何先验。

（图4）。Brain-TAP 主要从抑制和注意力先验中识别关键连接。与图3相比，此处注意力更为突出，因为少量具有高度判别力的注意力相关连接推动了预测，尽管注意力先验的整体排名较低。此外，功能连接识别出的关键区域多于结构连接，表明其捕获了更多与精神疾病预测相关的诊断信息。

3.4. 结论

本文提出 BrainTAP，这是一种 Transformer框架，包含自适应互蒸馏（AMD）和选择性先验融合（SPF），可根据功能连接（FC）和结构连接（SC）预测脑部疾病，同时保留模态特异性信号，并通过全局和个性化注意力掩码有选择地整合专家先验。ABCD 数据集上的实验表明，BrainTAP 的表现优于强基线方法。进一步分析发现，AMD 能有效调节模态间交互强度，而 SPF 则突出显示了神经生物学环路——尤其是涉及认知控制和抑制的环路——是关键的判别特征，这与已知临床证据相符。

符合伦理标准。本研究使用了公开可用的人类数据 ABCD。根据许可条款，无需额外的机构伦理审批。**致谢**。本研究获得了 UVA TYDE 种子基金和 UVA 脑研究所总统跨学科神经科学研究员项目的支持。

4. 参考文献

[1] Hae-Jeong Park 与 Karl Friston，“结构与功能脑网络：从连接到认知”，*Science*，第 342 卷，第 6158 期，第 1238411 页，2013。

[2]Xiaoxiao Li、Yuan Zhou、Nicha Dvornek、Muhan Zhang、Siyuan Gao、Juntang Zhuang、Dustin Scheinost、Lawrence H Staib、Pamela Ventola 与 James S Duncan，“BrainGNN：用于功能性磁共振成像分析的可解释脑图神经网络”，*Medical ImageAnalysis*，第 74 卷，第 102233 页，2021。

[3] Xuan Kan、Wei Dai、Hejie Cui、Zilong Zhang、Ying Guo 与 Carl Yang，“脑网络变换器”，神经信息处理系统进展，第 35 卷，第 25586–25599 页，2022。

[4]Anees Abrol、Zening Fu、Yuhui Du 和 Vince D Calhoun，“融合深度学习与动态功能连接特征的多模态数据，以预测阿尔茨海默病的进展”，载于*2019*年第*41*届 *IEEE* 医学与生物学工程学会国际年会（*EMBC*）。IEEE，2019年，第4409–4413页。

[5] Ye Liu、Lifang He、Bokai Cao、Philip Yu、Ann Ragin 和 Alex Leow，“用于脑网络聚类分析的多视图多图嵌入”，载于*AAAI* 人工智能大会论文集，2018年，第32卷。

[6] Lu Zhang、Li Wang、Jean Gao、Shannon L Risacher、Jingwen Yan、Gang Li、Tianming Liu、Dajiang Zhu、阿尔茨海默病神经影像学计划等，“轻度认知障碍中的脑结构–功能深度融合”，*Medical Image Analysis*，第72卷，第102082页，2021年。

[7] Moein Khajehnejad、Forough Habibollahi 和 Adeel Razi，“BrainSymphony：Transformer 驱动的功能性磁共振成像时间序列与结构连接融合”，*arXiv* 预印本 *arXiv:2506.18314*，2025年。

[8] Rui Liu、Ximan Hou、Shuyu Liu、Yuan Zhou、Jingjing Zhou、Kaini Qiao、HanQi、Ruinan Li、Zhi Yang、Ling Zhang等，“通过局部-全局图神经网络和神经影像生物标志物预测抗抑郁药反应”，*npjDigitalMedicine*，第8卷，第1期，第515页，2025年。

[9]Colin J Brown、Jeremy Kawahara和Ghassan Hamarneh，“深度神经网络中用于预测自闭症的连接组先验”，收录于*2018*年*IEEE*第*15*届生物医学成像国际研讨会（*ISBI 2018*年）。IEEE，2018年，第110–113页。[10] Donald J Hagler Jr、SeanN Hatton、M Daniela Cornejo、Carolina Makowski、Damien A Fair、Anthony Steven Dick、Matthew T Sutherland、BJ Casey、Deanna M Barch、Michael P Harms等，“青少年大脑认知发展研究的图像处理和分析方法”，*NeuroImage*，第202卷，第116091页，2019。

[11] Hongting Ye、Yalu Zheng、Yueying Li、Ke Zhang、Youy- ong Kong 和 Yonggui Yuan，“RH-BrainFS：区域异质多模态脑网络融合策略”，神经信息处理系统进展，第36卷，第59286–59303页，2023年。

[12] Yanwu Yang、Chenfei Ye、Xutao Guo、Tao Wu、YangXi- ang 和 Ting Ma，“通过跨模态互学习绘制用于脑疾病诊断的多模态脑连接组”，*IEEE*医学影像汇刊，第43卷，第1期，第108–121页，2023年。

[13] Michael W. Cole、Grega Repovs、Alex Anticevic 和 Todd S. Braver，《额顶控制系统：在心理健康中的核心作用》，*Neuroscientist*，第 20 卷，第 6 期，第 652–664 页，2014年。

[14] Kerstin Konrad 和 Simon B Eickhoff，《ADHD 大脑的神经连接方式是否有所不同？注意缺陷多动障碍结构和功能连接的综述》，*Human brain mapping*，第 31 卷，第 6 期，第 904–916 页，2010年。

[15] Luke J. Norman、Stephan F. Taylor、Yanni Liu、Joaquim Radua、Yann Chye、Stella J. De Wit、Chaim Huyser、F. Isik Karahanoglu、Tracy Luks、Dara Manoach、Carol Mathews、Katya Rubia、Chao Suo、Odile A. van den Heuvel、Murat Y¨ucel 和 Kate Fitzgerald，《强迫症中的错误处理与抑制控制：基于统计参数图的荟萃分析》，*Biological Psychiatry*，第 85 卷，第 9 期，第 713–725 页，2018年。

[16] Silvia Brem、Edna Gr¨unblatt、Renate Drechsler、Peter Riederer和Susanne Walitza，“强迫症与注意缺陷多动障碍之间的神经生物学联系”，*ADHD AttentionDeficitand Hyperactivity Disorders*，第6卷，第3期，第175–202页，2014年。

[17] B.A. Seitzman等，“静息态网络的现状”，*NeuroImage*，第200卷，第1–9页，2019。