



南京航空航天大学
Nanjing University of Aeronautics and Astronautics

基于认知和 MRI 对齐的语言-图像预训练 脑部疾病诊断模型

课 程 名 称 _____ 大模型原理与技术 _____

专 业 类 别 _____ 计算机科学与技术 _____

姓 名 _____ 赵春志 _____

学 号 _____ BX2524305 _____

二〇二六年一月

基于认知和 MRI 对齐的语言-图像预训练脑部疾病诊断模型

1. 背景

越来越多的神经影像学研究表明, 尽管阿尔茨海默病(Alzheimer's disease, AD)、自闭症谱系障碍(autism spectrum disorder, ASD)和注意力缺陷多动障碍(attention-deficit/hyperactivity disorder, ADHD)等主要神经疾病在临床表现与诊断标准上存在显著差异, 它们却存在共同的结构和功能异常[1-3]。这些异常通常表现为认知功能障碍的重叠现象, 例如执行功能和工作记忆缺陷等[4]。这种跨疾病神经生物学与认知表型趋同现象, 为构建能够诊断多种脑疾病并揭示其内在机制的通用模型奠定了理论基础[5]。然而, 目前大多数诊断模型仍局限于单一疾病类别, 且依赖于单模态数据, 这限制了其捕捉跨疾病病理机制以及整合多维生物信息的能力。标准化 NIH 工具箱提供了一个系统性的综合框架, 可用于评估包括执行功能和处理速度在内的核心认知能力[6], 这些测量指标与脑疾病中认知变化的特征密切相关。研究人员正致力于开发一种能够深度融合多源信息的通用诊断框架[7], 以提升对多种神经系统疾病的诊断能力及机制理解[8]。这一进展有望革新当前脑部疾病的诊断与治疗范式。

近年来, 视觉-语言预训练模型(如 CLIP[9]、BLIP[10, 11]和 ALBEF[12]等)在计算机视觉多个任务中取得了显著进展。通过将大规模数据集中的视觉与文本特征对齐到统一的表征空间, 这类模型展现出强大的零样本和少样本迁移能力。这一成功范式为神经科学领域带来了新的机遇。例如, MedCLIP[13]通过医学图像-文本对的对比学习, 获得具有泛化能力的多模态表征, 即使在标注数据有限的情况下仍能实现跨任务泛化。MedBLIP[14]则引入 MedQFormer 模块, 将 3D 医学影像数据与预训练的 2D 图像编码器和语言模型相融合, 在零样本分类和医学视觉问答任务中表现优异。此外, BrainSCK[8]模型采用两阶段训练策略: 首先通过知识注入提示将脑结构图像与认知功能文本进行对齐, 随后借助知识激活提示对诊断指令进行微调, 从而实现对多种脑部疾病的诊断。尽管取得了这些令人鼓舞的进展, 但将自然图像领域的视觉-语言预训练模型应用于神经影像学仍面临诸多挑战。

在神经科学研究中, 图像-文本配对数据的规模远小于自然图像领域的通用数据集规模, 这一局限性严重制约了神经科学视觉-语言预训练模型的发展。现有方法(如 ConVIRT[15]、GLORIA[16])高度依赖于成对的医学图像与诊断报告进行训练, 不仅受限于数据总量, 更存在对比学习中的“假阴性”问题(即描述相同病症的不同患者样本被错误归类为负样本)。此类固有噪声会干扰模型对细微医学语义的理解, 削弱其表征学习能力, 进而影响模型性能与数据利用效率。此外, 当前大多数医学视觉-语言预训练模型仅处理二维图像, 而如何将三维体积数据(如 structural MRI, sMRI)有效整合到为二维设计的大规模视觉编码器和语言模型中, 并实现跨模态语义对齐, 同时保持模型的轻量与实用性能, 仍是一个关键的问题。另一方面, 现有的预训练策略通常分为两类: 一是基于自然图像预训练的通域模型进行迁移学习, 二是利用健康人群脑数据通过代理任务(如年龄预测)预训练的域内模型。然而, 前者往往缺乏对脑疾病相关特征和个体差异的敏感度; 后者则依赖于代理任务与真实诊断任务之间微弱且隐性的关联, 未能显式地将脑结构信息与丰富的认知功能知识进行对齐, 导致其诊断效能与泛化能力有限。更重要的是, 当前仅依赖临床评分描述的方法往往难以充分捕捉与诊断密切相关的特定临床概念[13, 14](如“注意力不集中”、“抑制功能减退”等)。同时, 神经影像中的病理特征通常具备局部化与层次化特性, 从细微的局部结构变化到宏观层面的连接异常均可能存在。这种复杂性要求模型具备更精细的多层次跨模态对齐能力。为应对上述挑战, 我们提出了基于脑结构与认知对齐的引导式语言-图像预训练模型(BrainBLIP), 该模型旨在通过将健康人群的

sMRI 数据与认知描述进行对齐预训练，构建一个适用于多种脑部疾病的通用诊断框架。

提出的 BrainBLIP 模型包含两个阶段：(1) 在预训练阶段（图 1a），该模型该模型基于人类连接组计划[17]（Human Connectome Project, HCP）数据集中健康个体的 sMRI 及其临床数据（年龄、性别和认知功能的文本描述）进行训练。该阶段旨在捕捉认知功能与大脑结构之间的复杂关系。(2) 在微调阶段（图 1b），模型通过视觉问答（VQA）任务在多个脑部疾病数据集上进行领域自适应学习，使其能够根据输入的 sMRI 图像和问题提示生成相应的诊断答案。实验结果表明，BrainBLIP 在 6 种脑部疾病诊断任务中均优于其他先进方法，包括：AD、ASD、ADHD、精神分裂症（schizophrenia, SZ）、重度抑郁症（major depressive disorder, MDD）和轻度认知障碍（mild cognitive impairment, MCI）。总体而言，本研究的主要贡献可以概括如下：

- 提出了 BrainBLIP 这一新型引导式语言-图像预训练框架，通过将认知功能描述与 sMRI 特征进行对齐训练，全面捕捉与诊断密切相关的临床概念。
- 引入可学习的标签嵌入和多层对齐机制，实现了更细粒度、多层次跨模态语义对齐。
- BrainBLIP 在 6 种诊断任务中的表现优于其他最先进的方法，为多种脑疾病诊断提供了一个强大且可推广的通用框架。

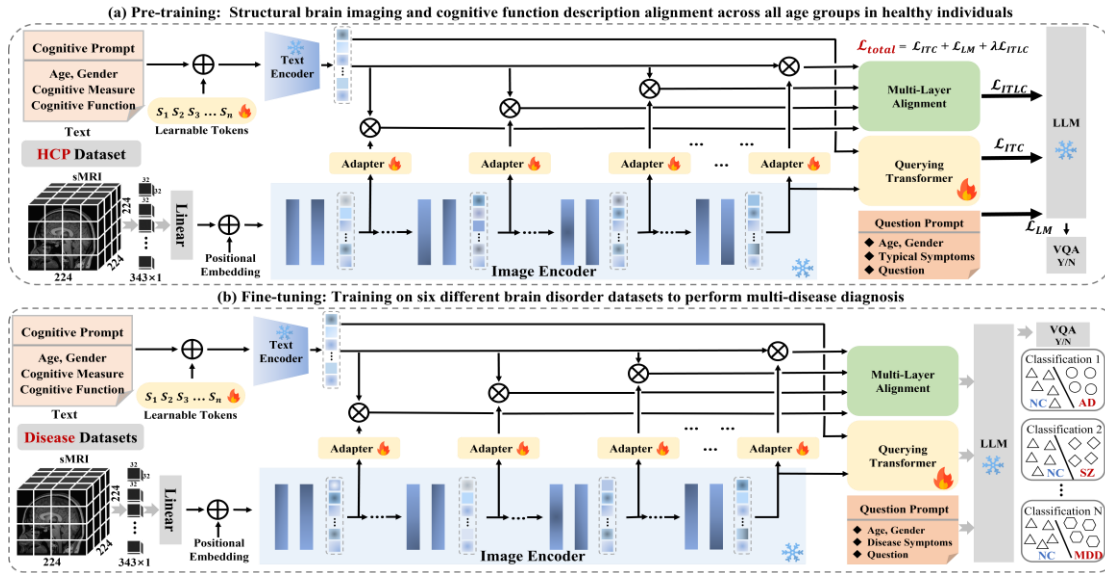


图 1. BrainBLIP 框架流程图，(a)预训练和(b)微调

2. 方法与材料

2.1. 网络架构

BrainBLIP 模型主要采用 BLIP-2[11]架构来处理图像和文本数据。如图 1 所示，该模型主要包括图像编码器、文本编码器以及作为解码器的大型语言模型（LLM）。文本编码器负责处理认知功能文本描述 T 和诊断问题 Q ，视觉编码器则从结构性脑影像 V 中提取特征。LLM 解码器有效地整合图像和文本特征信息，以自回归的方式生成用于诊断脑部疾病的答案 A 。此外，为进一步增强模型在神经影像领域的泛化能力，我们引入了可学习标记 S 和多层对齐模块。

2.1.1 文本编码器

在文本编码器中，采用标准的分词方法将输入文本转换为嵌入表示[18]。具体而言，文本嵌入包括认

知文本描述序列 $T = \{t_1, t_2, \dots, t_{l_t}\} \in \mathbb{R}^{l_t \times e}$ ，问题序列 $Q = \{q_1, q_2, \dots, q_{l_q}\} \in \mathbb{R}^{l_q \times e}$ ，答案序列 $A = \{a_1, a_2, \dots, a_{l_a}\} \in \mathbb{R}^{l_a \times e}$ ，以及可学习的标记序列 $S = \{s_1, s_2, \dots, s_{l_s}\} \in \mathbb{R}^{l_s \times e}$ 。其中 l_t , l_q , l_a , l_s 表示嵌入序列的长度， e 表示嵌入维度。

2.1.2. 图像编码器

在图像编码过程中，sMRI 图像首先被分割为一系列尺寸为 $32 \times 32 \times 32$ 的三维图像块。每个图像块由一个可学习的线性投影层映射为嵌入向量，并通过添加位置嵌入信息进行增强。最终，使用参数冻结的预训练图像编码器对该序列进行处理，生成图像嵌入表示 $V = \{v_1, v_2, \dots, v_{l_v}\} \in \mathbb{R}^{l_v \times e}$ ，其中 l_v 表示序列长度。此外，从图像编码器提取的中间层特征用于后续与文本特征的对比学习。

2.1.3. 多层对齐

为提升模型在视觉问答（VQA）任务中的泛化能力，我们引入了基于多层感知器（MLP）并配备残差连接的适配器模块[19]。该模块将图像编码器的中间层特征 V_i 输入到适配器中，生成处理后的图像特征 P_i ，用于后续与文本特征之间的对比学习。输出 P_i 可以表示为：

$$P_i = V_i + \mathcal{F}(V_i) \quad (1)$$

其中 $\mathcal{F}()$ 表示 MLP 函数， i 表示图像编码器的层数。本研究提取图像编码器的第 8, 16, 24, 32 层和最终层的特征，在损失函数 L_{ITLC} 下进行对比学习：

$$L_{ITLC}^{(i)} = \frac{1}{2}CE(P_i, T) + \frac{1}{2}CE(T, P_i) \quad (2)$$

其中 $L_{ITLC}^{(i)}$ 表示第 i 层对比学习时的损失函数。 $CE()$ 表示交叉熵损失函数。经过多层对齐学习后的最终特征 P 可以表示为：

$$P = \sum_i^l \omega_i P_i \quad (3)$$

其中 ω_i 表示将第 i 层的图像特征与文本特征对齐时可学习的权重参数， l 表示图像编码器中间层的索引。

2.1.4. Q-Former

随后，经过处理的文本与图像特征序列被输入到可学习的查询转换器（Q-Former）中。在该模块中，图像嵌入通过交叉注意力机制与一组可学习的查询向量进行交互，提取出 k 个视觉特征 V' 。同时，Q-Former 将文本编码器的提示文本转换为文本嵌入表示 T' 。然后，通过图像-文本对比损失函数 L_{ITC} 在联合表示空间中对齐视觉特征 V' 和文本嵌入 T' 。我们选择与文本相似度最高的输出查询嵌入来计算图文交叉熵损失：

$$L_{ITC} = \frac{1}{2}CE(V', T') + \frac{1}{2}CE(T', V') \quad (4)$$

2.1.5. 语言模型

最终阶段，采用一个融合医学领域知识的大型语言模型（BioMedLM）作为解码器[20]。该模型接收来自 Q-Former 生成的自适应编码视觉特征 V' ，并结合问题提示与多尺度对齐特征 P 作为输入。在语言模型损失函数 L_{LM} 的监督下，以自回归方式进行训练并生成诊断答案。整体损失函数定义为如下加权形式：

$$L_{total} = L_{ITC} + L_{LM} + \lambda L_{ITLC} \quad (5)$$

其中 λ 表示可调参数。

2.2. 预训练

2.2.1. 数据准备

在模型预训练阶段，我们使用了来自人类连接组计划（HCP）数据集的 sMRI 图像。该数据集覆盖了从发育期、青年期到老年期的多个年龄阶段，并提供了包括年龄、性别和认知评分在内的人口统计学变量。我们排除了结构成像数据或认知评估不完整的受试者，最终纳入了 1597 名年龄介于 6 至 89 岁之间的参与者。认知功能的评估基于七项年龄校正后的 NIH Toolbox 心理测量测试。

2.2.2. 认知提示

如图 2a 所示，对于人口统计学特征，我们将实际数值填入预设的文本模板中以生成规范化描述。参照 Wang 等人[6, 8]所设定的标准，认知评分低于 85 分被解释为相应认知功能受损，高于该阈值则被视为功能正常。因此，除了生成如“.....分数.....低于/高于.....”等数值型描述外，我们还进一步提供了诸如“.....表示表现不佳.....”或“.....表明功能正常.....”等语义解释。预训练过程使用了 HCP 数据集中全部七种心

理测量的数据。

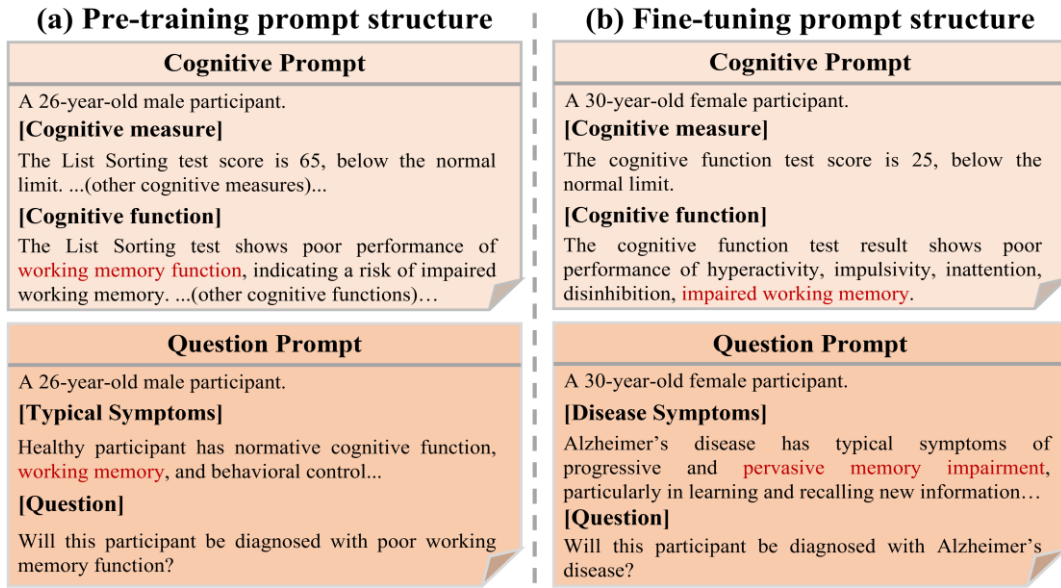


图 2. 预训练(a)和微调(b)阶段提示结构

2.3. 微调

2.3.1. 数据准备

在微调阶段，模型使用了六种脑部疾病的 sMRI 图像及对应人口统计数据。具体而言，自闭症数据使用自闭症脑成像数据交换（Autism Brain Imaging Data Exchange, ABIDE）[21, 22]，包括 519 名患者和 591 名健康对照（healthy controls, HCs）；ADHD 数据来自 ADHD-200 数据集[23]，包含 344 名患者和 577 名 HCs；SZ 数据取自双相与精神分裂症中间表型网络（Bipolar and Schizophrenia Intermediate Phenotype Network, BSNIP）[24]，包含 320 名患者和 431 名 HCs；MDD 数据来自[25]，包括 261 名患者和 286 名 HCs；MCI 和 AD 数据来自阿尔茨海默病神经影像学倡议（Alzheimer's Disease Neuroimaging Initiative, ADNI）[26]，包含 1006 名 MCIs 患者、362 名 ADs 患者及 68 名 HCs。每个数据集均按 7:1:2 的比例随机划分为训练集、验证集和测试集，并进行了 5 折交叉验证。

2.3.2. 问题提示

为辅助脑部疾病的诊断，我们设计了一个包含人口统计信息、症状描述及二元答案的问题提示模板，如图 2b 所示。为加强预训练任务与下游诊断任务之间的语义一致性，我们从多种脑部疾病的症状列表中提取关键认知症状，并以与健康受试者认知提示中“功能较弱”相似的表述方式进行描述。脑部疾病提示中涵盖的典型症状包括注意力不集中、行为脱抑制和工作记忆受损等，这些症状均与注意力和工作记忆等认知功能密切相关。。

2.3.3. 用于脑疾病诊断的 VQA

在微调阶段，BrainBLIP 模块使用预训练参数进行初始化。随后，随后对脑图像与相应的认知功能障碍提示进行图像-文本对齐学习（图 1b）。在此基础上，我们引入问题提示，并融合经对齐的视觉特征，以引导大型语言模型（LLM）对脑功能障碍进行诊断。在推理阶段，该模型能够基于配对的图像与提示输入，自动生成相应的诊断答案。

3. 实验及结果

3.1. 实现细节

在预训练阶段，我们使用与 BLIP-2 相同的冻结图像编码器 ViT-g/14 [30] 和 Q-Former 模块进行初始化，并将 FLAN-T5 替换为富含医学领域知识的大型语言模型 BioMedLM。在微调阶段，Q-Former 继承了预训练版本的权重参数。我们将批大小设置为 16，可学习标记的大小设置为 256，并使用学习率为 1×10^{-4} 的 AdamW 优化器进行了 200 个 epoch 的训练。所有实验均在八块 NVIDIA RTX A6000 GPU 上进行。

3.2. 脑部疾病的诊断性能

我们将 BrainBLIP 在 6 种脑部疾病中的诊断性能与几种先进的预训练方法进行了比较，包括 MedCLIP、MedBLIP、BrainSCK 和 BLIP-2。模型的性能通过计算每种疾病数据集在 5 折交叉验证方案下的诊断准确率来评估。表 1 展示了 6 项诊断任务的结果，其中性能最佳的以粗体标出，性能第二好的以加下划线标出。结果表明，我们的方法在 6 种脑部疾病中始终优于其他方法，并实现了出色的泛化性能，证明了其有效性。重要的是，与 BLIP-2（使用 BioMedLM 进行 VQA 辅助）相比，BrainBLIP 在诊断大多数脑部疾病方面取得了显著改进，且无需特定领域的数据或认知提示预训练。

表 1. BrainBLIP 与其他方法在 5 个疾病数据集上的诊断性能 (%) 比较。

Method	ABIDE	ADHD-200	BSNIP	MDD	ADNI	
	ASD vs. HC	ADHD vs. HC	SZ vs. HC	MDD vs. HC	MCI vs. HC	AD vs. HC
BLIP-2	63.51±1.03	65.65±2.01	73.20±1.59	63.42±1.36	66.49±2.47	88.00±1.19
MedBLIP	68.47±4.93	67.39±2.38	<u>83.20±3.81</u>	74.78±5.34	<u>79.65±4.67</u>	87.14±1.76
MedCLIP	63.06±1.59	66.96±2.65	76.67±1.79	73.87±3.89	69.32±1.36	87.05±2.22
BrainSCK	<u>69.10±4.23</u>	<u>69.23±5.13</u>	82.71±3.16	<u>79.32±3.22</u>	74.86±4.57	87.65±2.67
BrainBLIP	79.10±6.35	69.89±3.52	86.40±4.10	81.18±3.25	80.95±5.11	93.72±4.84

3.3 文本信息的贡献

此外，我们对提示模块的输入内容进行了对比实验。首先，我们从认知提示和问题提示中移除年龄和性别信息，以评估它们对诊断准确性的影响。其次，我们从认知提示模块中排除认知评分和功能描述，以评估认知与脑结构匹配的重要性。最后，我们从问题提示中省略疾病症状描述，以评估其对模型泛化能力的影响。如表 2 所示，所有 6 项疾病诊断任务的性能均有所下降，这证实了我们提示设计的每个组成部分对于提高诊断的精确度和泛化能力都至关重要。值得注意的是，结果表明，认知语义描述和症状相关描述对诊断准确性的贡献比年龄和性别信息更大。

表 2. 提示结构在 5 个疾病数据集上的诊断性能比较。

Age, Gender	Cognitive Function	Disease Symptoms	ABIDE	ADHD-200	BSNIP	MDD	ADNI	
			ASD vs. HC	ADHD vs. HC	SZ vs. HC	MDD vs. HC	MCI vs. HC	AD vs. HC
	✓	✓	76.13±3.90	66.25±6.95	81.33±3.90	79.37±2.65	76.49±6.17	88.05±1.88
✓		✓	69.37±5.02	64.67±3.23	78.67±3.76	72.97±3.75	77.46±2.56	81.90±2.94
✓	✓		68.02±5.73	66.09±4.41	75.33±3.76	75.77±3.45	70.41±3.26	85.62±6.12
✓	✓	✓	79.10±6.35	69.89±3.52	86.40±4.10	81.18±3.25	80.95±5.11	93.72±4.84

3.4. 消融研究

我们进行了一系列消融实验，以验证 BrainBLIP 中可学习词元和多层对齐的有效性，相应的评估结果列于表 3。首先，在预训练和微调阶段移除可学习词元嵌入模块会导致 6 项诊断任务的准确率下降。其次，从大型语言模型的输入中移除多层对齐特征也会导致所有 6 项诊断任务的准确率下降。这些结果证实，可

学习词元和多层对齐模块都是维持模型性能和鲁棒性的必要组成部分。

表 3. BrainBLIP 对 5 种疾病数据集的消融研究。

Learnable Tokens	Multi-Layer Alignment	ABIDE	ADHD-200	BSNIP	MDD	ADNI	
		ASD vs. HC	ADHD vs. HC	SZ vs. HC	MDD vs. HC	MCI vs. HC	AD vs. HC
✓		74.05±2.33	69.42±8.47	83.33±5.89	78.65±1.41	77.58±6.87	85.71±1.79
	✓	75.10±5.59	68.18±5.32	82.72±6.81	74.50±2.39	79.65±4.67	87.20±5.44
✓	✓	79.10±6.35	69.89±3.52	86.40±4.10	81.18±3.25	80.95±5.11	93.72±4.84

4. 结论

本研究提出了一种基于脑结构和认知对齐的引导式语言-图像预训练模型（BrainBLIP）。BrainBLIP 通过将认知功能描述与 sMRI 特征对齐，并结合可学习标签和多层对齐模块，旨在建立一个适用于多种脑部疾病的通用诊断框架。结果表明，BrainBLIP 在广泛的诊断任务中均取得了优异的性能，显著优于其他方法，凸显了其强大的泛化能力和有效性。值得注意的是，提示结构实验表明，认知语义和症状相关描述对诊断准确性的贡献远大于年龄和性别信息。此外，消融实验证实了所提出的核心模块在提升模型整体性能方面发挥的关键作用。总而言之，本研究提供了一种强大且具有普适性的新范式，利用健康人群预训练的 sMRI 和认知对齐信息，为诊断各种脑部疾病提供了可能。

5. REFERENCES

- [1] M. Liu *et al.*, "A common spectrum underlying brain disorders across lifespan revealed by deep learning on brain networks," *Iscience*, vol. 26, no. 11, 2023.
- [2] N. Opel, J. Goltermann, M. Hermesdorf, K. Berger, B. T. Baune, and U. Dannlowski, "Cross-disorder analysis of brain structural abnormalities in six major psychiatric disorders: a secondary analysis of mega-and meta-analytical findings from the ENIGMA consortium," *Biological Psychiatry*, vol. 88, no. 9, pp. 678-686, 2020.
- [3] M. T. M. Park, A. Raznahan, P. Shaw, N. Gogtay, J. P. Lerch, and M. M. Chakravarty, "Neuroanatomical phenotypes in mental illness: identifying convergent and divergent cortical phenotypes across autism, ADHD and schizophrenia," *Journal of Psychiatry and Neuroscience*, vol. 43, no. 3, pp. 201-212, 2018.
- [4] V.-R. Bourque *et al.*, "Genetic and phenotypic similarity across major psychiatric disorders: a systematic review and quantitative assessment," *Translational Psychiatry*, vol. 14, no. 1, p. 171, 2024.
- [5] A. S. Greene and R. T. Constable, "Clinical promise of brain-phenotype modeling: a review," *JAMA psychiatry*, vol. 80, no. 8, pp. 848-854, 2023.
- [6] S. Weintraub *et al.*, "Cognition assessment using the NIH Toolbox," (in eng), *Neurology*, vol. 80, no. 11 Suppl 3, pp. S54-64, Mar 12 2013, doi: 10.1212/WNL.0b013e3182872ded.
- [7] S. Qi *et al.*, "Multimodal fusion with reference: searching for joint neuromarkers of working memory deficits in schizophrenia," *IEEE transactions on medical imaging*, vol. 37, no. 1, pp. 93-105, 2017.
- [8] L. Wang, M. Liu, S. Zhang, and X. Wang, "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*, 2024: Springer, pp. 45-55.
- [9] A. Radford *et al.*, "Learning transferable visual models from natural language supervision," in *International conference on machine learning*, 2021: PmLR, pp. 8748-8763.
- [10] J. Li, D. Li, C. Xiong, and S. Hoi, "Blip: Bootstrapping language-image pre-training for unified vision-language understanding and generation," in *International conference on machine learning*, 2022: PMLR, pp. 12888-12900.
- [11] J. Li, D. Li, S. Savarese, and S. Hoi, "Blip-2: Bootstrapping language-image pre-training with frozen image encoders and large language models," in *International conference on machine learning*, 2023: PMLR, pp. 19730-19742.
- [12] J. Li, R. Selvaraju, A. Gotmare, S. Joty, C. Xiong, and S. C. H. Hoi, "Align before fuse: Vision and language representation learning with momentum distillation," *Advances in neural information processing systems*, vol. 34, pp. 9694-9705, 2021.
- [13] Z. Wang, Z. Wu, D. Agarwal, and J. Sun, "Medclip: Contrastive learning from unpaired medical images and text," in *Proceedings of the Conference on Empirical Methods in Natural Language Processing. Conference on Empirical Methods in Natural Language Processing*, 2022, vol. 2022, p. 3876.
- [14] Q. Chen and Y. Hong, "Medblip: Bootstrapping language-image pre-training from 3d medical images and texts," in *Proceedings of the Asian conference on computer vision*, 2024, pp. 2404-2420.
- [15] Y. Zhang, H. Jiang, Y. Miura, C. D. Manning, and C. P. Langlotz, "Contrastive learning of medical visual representations from paired images and text," in *Machine learning for healthcare conference*,

- 2022: PMLR, pp. 2-25.
- [16] S.-C. Huang, L. Shen, M. P. Lungren, and S. Yeung, "Gloria: A multimodal global-local representation learning framework for label-efficient medical image recognition," in *Proceedings of the IEEE/CVF international conference on computer vision*, 2021, pp. 3942-3951.
 - [17] D. C. Van Essen *et al.*, "The WU-Minn human connectome project: an overview," *Neuroimage*, vol. 80, pp. 62-79, 2013.
 - [18] S. M. Jain, "Introduction to transformers for NLP," *With the hugging face library and models to solve problems*, 2022.
 - [19] S.-A. Rebuffi, F. Croce, and S. Gowal, "Revisiting adapters with adversarial training," *arXiv preprint arXiv:2210.04886*, 2022.
 - [20] E. Bolton *et al.*, "Biomedlm: A 2.7 b parameter language model trained on biomedical text," *arXiv preprint arXiv:2403.18421*, 2024.
 - [21] S. Qi *et al.*, "Common and unique multimodal covarying patterns in autism spectrum disorder subtypes," *Molecular autism*, vol. 11, no. 1, p. 90, 2020.
 - [22] A. Di Martino *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.
 - [23] S. Qi *et al.*, "Reward processing in novelty seekers: a transdiagnostic psychiatric imaging biomarker," *Biological psychiatry*, vol. 90, no. 8, pp. 529-539, 2021.
 - [24] C. A. Tamminga, G. Pearlson, M. Keshavan, J. Sweeney, B. Clementz, and G. Thaker, "Bipolar and schizophrenia network for intermediate phenotypes: outcomes across the psychosis continuum," *Schizophrenia bulletin*, vol. 40, no. Suppl_2, pp. S131-S137, 2014.
 - [25] S. Qi *et al.*, "MicroRNA132 associated multimodal neuroimaging patterns in unmedicated major depressive disorder," *Brain*, vol. 141, no. 3, pp. 916-926, 2018.
 - [26] K. Li *et al.*, "Polygenic Hazard Score Associated Multimodal Brain Networks Along the Alzheimer's Disease Continuum," *Frontiers in Aging Neuroscience*, vol. 13, p. 725246, 2021.