

临床-影像融合模型在阿尔茨海默病中的研究进展

夏宇轩, 刘远康, 江翠萍, 吴樊宇, 曾雯雯, 张舒雅, 徐向阳

【摘要】 阿尔茨海默病(AD)是一种以认知功能进行性衰退为主要表现的神经退行性疾病,好发于老年人群。伴随人口老龄化趋势,AD发病率持续上升,其诊疗研究具有重要意义。临床-影像多模态融合技术的快速发展为AD研究提供了新思路。该技术通过融合影像学数据与电子病历、基因组学、实验室检查等多源临床信息,较单一数据模式更能全面刻画疾病特征,在AD早期识别、鉴别诊断、病情预测及风险分层等领域显示出良好前景。尽管如此,数据融合复杂性、模型性能优化等问题仍需深入探索。本文就临床-影像多模态融合模型在AD临床应用中的研究现状进行综述。

【关键词】 阿尔茨海默症;机器学习;人工智能;精准医学

【中图分类号】 R341;R977.6;R319;R-05;R459.9 **【文献标志码】** A

【文章编号】 1000-0313(2026)02-0217-06

DOI:10.13609/j.cnki.1000-0313.2026.02.018

开放科学(资源服务)标识码(OSID):



阿尔茨海默病(Alzheimer's disease, AD)是导致痴呆症最常见的疾病^[1],这种进行性神经退行性疾病不仅严重影响老年人的日常生活能力,还会给患者及其照护者造成巨大的精神和经济负担,随着病情发展,患者会出现记忆力减退、性格改变和思维能力下降等认知功能障碍,这些症状会逐步加重并最终导致生活完全不能自理^[2,3]。鉴于AD的早期诊断对于减缓疾病进程和维持认知功能具有重要意义,早期AD的准确诊断是当今医疗领域面临的挑战之一,它不仅与其他神经退行性疾病,如血管性痴呆(vascular dementia, VD)、路易体痴呆(Lewy body dementia, LBD)及额颞叶痴呆(frontotemporal dementia, FTD)在临床表现上存在交叉,而且很少独立存在,常与其他脑部病变共存^[4]。并且AD症状隐匿起病、进展方式多样,这些特点都增加了早期诊断的难度^[5]。

近年来,人工智能技术的快速发展为解决AD诊断的困境带来了新的方向。特别是在医学影像分析领域,许多研究者利用深度学习模型来分析患者的医学影像资料,在AD的早期诊断、进展预测等方面取得了进展。然而单一模态的影像分析往往难以全面反映疾病的复杂性。因此多模态模型逐渐成为研究热点^[6,7],包括多模态磁共振成像(MRI)、MRI及正电子发射断层扫描(positron emission tomography, PET)等不同模态影像间的整合分析^[8],还包括创新性地将

临床数据与影像特征相结合的研究,临床特征与影像融合模型可以同时整合影像学特征、电子健康记录(electronic health record, EHR)等多维度数据,这种分析方式与临床医生在诊疗过程中综合考虑多种信息后做出判断的思维模式相似,有望为医生提供更全面、更准确的辅助诊断依据,提高AD诊断的准确性。

在此就临床特征与影像融合模型在AD领域的最新研究进展予以综述,重点关注其在早期诊断、与其他痴呆病因的鉴别诊断及疾病进展预测等方面的应用价值。

融合模型概念及工作流程

深度学习作为机器学习的重要分支,通过多层神经网络实现了数据的自动特征学习。2015年,Ronneberger等^[9]提出的U-Net网络架构在医学图像分割任务中取得突破性进展,开创了深度学习在医学影像分析中的新时代。近年来,研究者发现将影像数据与EHR、基因组数据、临床检验等非影像数据结合分析能够在疾病分级、治疗方案选择及预后评估等方面为医生提供更全面、客观的决策支持,有望提高临床决策的准确性和效率。其工作流程包括采集并预处理相关的多模态数据,选择合适的融合策略与计算模型来整合信息,最后进行模型的训练、验证与性能评估。

1. 数据采集

阿尔茨海默病的全面评估需整合临床与影像两大类数据。临床信息主要包括量化认知功能的神经心理测试结果,以及指示病理状态的脑脊液生物标志物水平,例如关键的 β -淀粉样蛋白42、总tau和磷酸化tau蛋白。遗传学分析,特别是单核苷酸多态性(single

作者单位: 430077 武汉,华中科技大学同济医学院附属梨园医院放射科(夏宇轩、吴樊宇、曾雯雯、张舒雅、徐向阳); 430030 武汉,华中科技大学同济医学院附属同济医院放射科(刘远康、江翠萍)

作者简介: 夏宇轩(1999-),男,湖北武汉人,硕士研究生,主要从事人工智能在医学影像中的应用研究。

通讯作者: 徐向阳, E-mail: 1993ly0538@hust.edu.cn

nucleotide polymorphism, SNP), 可识别 APOE- ϵ 4 等高风险基因^[10]。影像学数据主要包括 MRI 以及 PET, MRI 提供了丰富信息:结构性 MRI 用于评估脑萎缩,功能性 MRI(functional magnetic resonance imaging, fMRI)分析脑网络动态,扩散张量成像(diffusion tensor imaging, DTI)则显示白质纤维束状况^[11]。PET 专注于特定靶点,淀粉样蛋白 PET 可显示淀粉样蛋白斑块沉积,或通过 FDG-PET 中的葡萄糖代谢水平评估神经元活性^[12]。

为保证分析的准确性与可靠性,这些多模态数据在用于研究前必须经过预处理。影像数据通常需要进行运动校正、颅骨去除、区域分割和空间平滑等操作,以标准化并减少噪声^[13]。相应的,其他临床及实验室数据也需经过数据清理、标准化和归一化等步骤,确保跨模态数据的一致性^[14]。

2. 数据融合策略

临床评估数据、脑脊液(cerebrospinal fluid, CSF)、遗传学测试以及患者的人口统计学数据等,这些可通过电子病历系统获取的数值型和分类型数据统称为“临床数据”。这些临床数据本质上属于结构化数据,具有预定义的格式和明确的变量类型,通常以表格形式存储和处理。与之相对的是以矩阵形式存储的高维医学影像数据,属于非结构化数据。

多模态数据融合策略主要包括早期融合、中期融合和晚期融合三种范式^[15]。早期融合通过在模型输入前直接拼接不同模态的原始数据或低级特征^[15],是概念上最直接的融合方式,但其主要局限在于难以有效处理低维、结构化的临床数据与高维、非结构化影像数据间的显著异质性。这种直接拼接方法存在缺陷,高维数据往往会压制低维信息、破坏各模态原有的内部结构,并且无法有效处理数据缺失情况,同时还容易引发特征数量远超样本数量导致模型性能崩溃的“维度灾难”现象,以及模型过拟合问题。因此中期与晚期融合策略在 AD 多模态研究中更为常用。

中期融合,又称特征级融合,在各不同的模态数据经过初步的独立特征提取后,于模型的中间层面对这些特征表示进行整合,再进行联合学习和预测。这种策略的核心优势在于能够在特征层面捕捉模态间的复杂交互与互补信息。常见的用于中期融合模型结构包括循环神经网络(recurrent neural network, RNN)^[16]适用于融合纵向数据特征与当前模态的时间动态;图神经网络(graph neural network, GNN)可将融合特征作为节点属性,用以学习个体或特征间的关系结构^[17];Transformer 模型可以利用自注意力和交叉注意力机制,动态地权衡不同模态特征的重要性并表示不同特征间的相互关系,尤其适合捕捉长距离依

赖关系^[18]。中期融合能更深入地探索模态协同效应并与 Transformer 等先进的深度学习架构搭配,是当前多模态 AD 研究中的热点与前沿方向。值得注意的是,传统的统计模型同样能在特征层面整合多模态信息,这从概念上可以视为一种类似中期融合的过程。深度学习模型更侧重于融合自动提取的深层特征表示,而这些传统模型融合的特征通常是经过预处理且定义明确的指标(如特定的影像测量值或基因风险评分),例如, Cox 回归模型及其变体适用于融合不同来源的特征指标,用以分析和预测疾病进展或转化等特定临床事件的发生风险^[19]。而结构方程模型(structural equation modeling, SEM)则提供了一个框架,用以检验和量化基于多个观测变量与潜在概念之间复杂的假设关系网络^[20]。

晚期融合,也称为决策级融合,采用先分别训练再整合的策略。这种方法首先为各数据模态建立独立模型,结构化临床数据通常经过特征选择技术优化后,输入随机森林(random forest, RF)、支持向量机(support vector machine, SVM)等传统机器学习模型,或使用自编码器(autoencoder, AE)等神经网络提取特征,然后将这些特征输入到多层感知机(multilayer perceptron, MLP)等分类器中进行预测,医学影像则先通过卷积神经网络(convolutional neural network, CNN)、视觉 Transformer 等深度学习架构提取关键特征。最终阶段在决策层整合各模型预测结果,融合机制包括简单的投票或平均法,以及使用 RF、MLP 等元学习器等。晚期融合的优势在于实现较为简单、模块化程度高且允许各模型独立优化,其主要局限是无法捕捉不同模态间早期特征层面的交互信息^[21,22]。

3. 模型的可解释性方法

深度学习模型在处理复杂数据时的表现优异,但同时也具有一定的“黑箱”特性,理解其内部决策对于增强临床信任和促进科学发现至关重要。为此,可解释性人工智能技术随之发展起来。可解释性技术的选择通常取决于模型的具体结构、数据的类型以及解释的目标。例如,类激活图(class activation mapping, CAM)及其变体如 Grad-CAM 等方法,是针对 CNN 设计的,常被用于可视化分析,通过生成显著性图来高亮模型在进行预测时最关注的输入区域,多用于医学影像分析中。对于本身包含注意力机制的 Transformer 模型,其注意力权重可作为一种内在的可解释性手段^[23]。此外,还有如 SHAP(SHapley Additive exPlanations)等模型无关的方法,它们系统性地扰动模型的输入数据并观察模型输出的变化,从而量化不同特征对模型最终预测的贡献度,适用于多种不同模型。运用这些技术有助于验证模型是否聚焦于符合医

学知识的生物标志物或脑区,不仅提升了模型透明度和可靠性,也可能为探索新的疾病相关特征提供线索^[24]。

4. 模型验证与评估

模型评估与验证包括在独立验证集上评估模型性能、稳定性、泛化能力等。准确率、敏感度、特异度、受试者操作特征曲线下面积(area under the curve, AUC)等性能评估指标能够反映模型在不同决策阈值下的表现。对于多模态融合模型,还需评估单一或多个模态缺失以及数据质量下降时的鲁棒性。比较验证时应将融合模型与单模态模型、不同的融合策略、现有的临床指南或评分系统及临床专家诊断结果对比。

多模态融合模型在 AD 中的应用

1. AD 的诊断

AD 的早期诊断具有难度,疾病具有很长的临床前期,神经退行性改变可能先于临床症状出现长达 10~15 年^[25]。淀粉样 β 蛋白积累和 tau 蛋白病理改变在此期间已发生,但认知功能尚未明显下降,早期诊断更多依赖于生物标志物检测和影像学技术^[26]。早期仅可表现为轻度认知障碍(mild cognitive impairment, MCI),症状隐匿^[5,27]。Venugopalan 等^[28]探索了多模态模型用于 AD 的早期识别的潜力。他们比较了单一模态和多模态融合模型的性能,评估了 MRI、电子病历数据(人口统计学信息、临床评估、生化指标、神经影像学测量及用药史)和 SNP 这几种不同类型组合的效能。在区分正常认知(normal cognition, NC)、MCI 和 AD 患者的三分类任务中,多模态模型的准确率在内部交叉验证和外部测试集分别达到了 79% 和 78%。研究发现多模态数据的预测效果优于单模态数据的预测效果,并且仅使用电子病历和 MRI 数据的模型也获得相近准确率,这也表明这说明 AD 的早期筛查在数据有限的情况下仍然可行。

当拥有更完整的临床和影像数据时,如何有效地整合这些异构信息并深入挖掘其内在关联成为当前研究的重点。近年来,图神经网络凭借其在处理非欧几里得数据和建模复杂关系方面的天然优势,为多模态数据整合提供了新的研究方向。Xu 等^[29]提出了一种可扩展的分层图卷积网络,能够整合结构 MRI、DTI 和 fMRI 三种模态的数据。模型可以全面分析大脑灰质萎缩、白质纤维损伤和功能连接异常,建模不同脑区之间的结构和功能连接,同时可以整合人口统计学数据,捕捉 SVM 和 3D-CNN 可能会忽略的重要关系。该模型在 AD、MCI 和 NC 三组人群的鉴别诊断中表现出色:AD 与 NC 的鉴别准确率达 88.71%,AD 与 MCI 的鉴别准确率为 82.71%,MCI 与 NC 的鉴别准

确率为 79.68%,多模态模型的各种性能指标均优于单模态模型。该模型还能够灵活地添加其他模态的图像特征和非图像信息,具有很好的可扩展性。Lei 等^[30]在传统图神经网络的基础上,提出了一种改进的双多级图神经网络框架。该框架通过特征归纳学习模块,为不同模态在相同特征空间下学习特征表示,减轻“模态竞争”的问题,从而更好地捕捉这些数据之间的复杂关系,模型结合结构 MRI、人口统计学信息、SNP、脑脊液蛋白浓度诊断 AD。在 AD 患者与 NC 对照分类任务中,多模态模型在 NC 和 AD 的分类任务中准确率达到 93.2%,AUC 达到了 93.1%,超过了单模态模型以及 RF、SVM 等传统机器学习模型。

2. AD 与其他痴呆病因的鉴别诊断

对于疑似 AD 的痴呆患者,进行全面的病因评估和鉴别诊断对于制定个性化治疗方案和临床管理至关重要。AD 的临床表现常与其他类型痴呆如 VD、LBD、FTD 存在重叠,并且这些疾病可能与 AD 共存,形成混合型痴呆,增加了诊断的复杂性。缺乏准确的鉴别诊断,可能使患者接受不当治疗,影响患者预后。

在个性化治疗方案方面,不同类型痴呆的治疗策略存在差异。例如,早期 AD 患者可能适合接受抗淀粉样蛋白治疗^[31,32],但若同时存在 VD 等其他病因时,可能会增加淀粉样蛋白相关成像异常的风险^[33]。同样,胆碱酯酶抑制剂虽然对 AD 患者具有一定疗效,但对 FTD 患者可能无效,甚至可能加剧认知功能障碍和行为症状^[34,35]。在临床管理方面,不同类型痴呆存在特定的并发症风险谱系。例如,VD 患者具有更高的卒中风险,需要加强脑血管风险因素的管理^[36]。因此,基于影像学技术的多模态模型在 AD 的鉴别诊断中具有重要的临床应用价值,可以提高诊断准确性,指导临床决策。

Xue 等^[37]开发了一种基于 Transformer 架构的模型,每个特征通过特定模态的嵌入策略被处理成固定长度的向量,作为 Transformer 模型的输入,利用多模态数据进行痴呆症的鉴别诊断。该研究分析了来自 9 个独立、地理位置多样的数据集的 51,269 名参与者,包括 T₁WI、T₂WI、FLAIR、SWI 在内的多种 MRI 序列的图像、人口统计学特征、个人及家族病史、用药情况、神经心理学评估和功能评估数据。该模型能够识别包括 AD、LBD、VD、FTD、朊病毒病、正常压力脑积水、系统性和环境因素、精神疾病、创伤性脑损伤及其他痴呆病症共 10 种不同的痴呆病因,在区分不同痴呆病因方面的微平均 AUC 达到 0.96。对于混合性痴呆病例,当同时存在两种病理改变时,该模型的平均 AUC 达到 0.78。在随机选择的 100 例样本中,神经科医生使用多模态模型评估时较单独医生评估的 AUC

提高了 26.25%。此外,模型预测与生物标志物证据一致,并通过尸检结果验证了其不同蛋白病变的关联。研究表明,基于多模态的辅助鉴别诊断模型在临床实践和药物试验中,有潜力成为痴呆症病因筛查的重要手段。

3. AD 的病程预测与风险评估

MCI 是 AD 的早期阶段,流行病学研究表明,每年约有 10%~12% 的 MCI 患者会进展为 AD^[38,39]。多模态模型能够预测 MCI 患者向 AD 转化的可能性,也能评估认知正常人群发展为 AD 的风险。通过分析纵向随访的影像数据,这些模型可以捕捉到疾病进展的动态特征,这对于制定有效的临床早期干预策略至关重要。

Velazquez 等^[40]结合 DTI 和电子病历中的脑室和海马体体积、APOE4、神经心理量表分数及人口统计学数据,创建了评估早期 MCI 患者转化为 AD 的风险预测模型,在 95 名测试患者中准确率达 98.81%。该研究采用随机森林进行特征重要性排序,使用 Grad-CAM 技术生成热力图显示关键脑区,发现年龄、海马体积和脑室体积是预测 AD 转化的重要因素。而 Rahim 等^[41]指出 Grad-CAMs 在医疗领域应用存在局限性,无法提供精确的定位受损组织。他们使用双向 RNN 处理从 3D CNN 提取的 MRI 时间序列特征,利用基线、6 个月和 12 个月的 MRI 数据,集合临床评分,评估疾病进展的动态特征。该模型可预测认知正常个体在基线的 4 年后发展为 AD 的风险,准确率为 96%,AUC 为 96.9%。此外,研究采用时间序列可解释性技术,通过 2D MRI 的体素级激活图和 3D 大脑表面渲染,直观展示 AD 影响的脑区及其时间演变特征。基于类似的改进可视化思路,Eslami 等^[42]提出了一种创新的色彩编码可视化系统。与传统分类方法不同,该研究不仅关注分类准确率,着重强调利用可视化技术来辅助临床决策过程。该系统通过融合 MRI 和 PET 两种模态的神经影像学数据、人口统计学数据、临床评估量表及脑脊液生物标志物,使用深度学习生成色彩编码图像,为临床提供更为直观的疾病发展轨迹。

在临床中,患者的检查数据时常会出现不连续的情况,影响决策过程,当某些检查数据缺失时,Martí-Juan 等^[43]开发的模型可以通过其他现有的临床指标对缺失信息进行预测,为无法进行某些检查的患者提供替代性评估方案。他们所设计的模型通过 RNN 处理时间序列数据,通过变分自编码器将输入数据编码为概率分布,这种模型能更好地处理数据的不确定性、捕捉不同类型数据间的关系并生成新数据,从而进行多模态 AD 进展建模,整合分析 MRI 中的脑部体积、

皮层厚度测量值及认知评分等数据。与传统方法相比,该模型在预测缺失模态数据时平均误差降低了 5%。研究发现,模型可基于认知评分等临床指标预估脑结构改变,为无法完成 MRI 检查的患者提供替代评估。还可以分析皮质厚度、脑体积、认知评分等生物标志物的动态关联,有助于理解疾病进展规律。该模型还有望应用于新药临床试验中,进行患者分组和疗效评估,为精准医疗提供数据支持。

除了上述基于深度学习的探索,传统的统计模型框架也能够融合多模态信息为 AD 病程预测和风险评估提供不同的视角。它们能有效处理时间依赖性事件、深入检验变量间的复杂关系,并且能提供具有临床意义的可解释性。

Reas 等^[44]运用 Cox 回归模型,整合了年龄、基于 31 个 SNP 和两个 APOE 变体计算的多基因风险评估、基于多区域 MRI 量化测量的脑萎缩评分以及代表认知分数的 Rey 听觉词语学习测试得分。他们构建的多模态风险评估有效地预测了从 MCI 向 AD 痴呆转化的时间依赖性风险,风险评分处于最高 20% 的高风险组相比评分处于最低 20% 的低风险组,其 5 年内转化为 AD 的风险比高达 27.03,研究还表明,该评分能将模拟临床试验所需的样本量减少约 67%。该研究还发现,即使仅使用多基因风险评估这一项遗传指标,就可预测个体出现淀粉样蛋白和 tau 病理特征的大致年龄。这项研究展示了 Cox 模型在融合多模态生物标志物进行精准风险预测和辅助临床试验设计方面的潜力。

而 SEM 及其变体,偏最小二乘结构方程模型 (partial least squares structural equation modeling, PLS-SEM),为探索多模态生物标志物、风险因素与认知衰退之间复杂的相互作用网络提供了定量分析框架。Tosun 等^[45]应用 PLS-SEM 方法,研究了 β 淀粉样蛋白阳性的老年人群,包括 NC、MCI 和痴呆患者。该研究整合了多种影像生物标志物,包括全局 $A\beta$ -PET 负荷、区域性 tau-PET 摄取、区域性 MRI 结构萎缩,以及年龄、性别、教育年限、APOE 基因型和脑白质病变负荷等风险因素。该研究量化了不同因素对认知的影响程度。结果表明,tau 病理与基线认知水平的关联最为密切,可解释约 46%~47% 的认知分数差异,其次是脑萎缩和 β 淀粉样蛋白。并且,研究进一步构建了一个融合模型,同时考虑了年龄、基因、血管等风险因素,以及它们如何通过 $A\beta$ -tau-萎缩这一系列病理变化影响认知。这个更完整的模型能够解释 53%~58% 的基线认知水平差异,以及 50%~56% 的后续认知衰退速度差异。研究表明,要全面理解患者的认知状况及其变化,不能仅看单个生物标志物,而要整合多

种风险因素和病理通路的全貌。

需要指出的是,这些基于统计模型的融合方法与前述的许多深度学习模型在目标和评估指标上存在差异。Cox模型侧重于预测事件发生时间的风险比,PLS-SEM关注路径系数和解释方差,而深度学习分类模型则常用AUC或准确率评估。它们的性能通常不能直接比较,它们在AD多模态研究中更多的是互补的作用,深度学习更擅长从高维原始数据中自动提取复杂非线性特征进行分类或分割,而Cox回归是分析生存数据的经典方法,其在提供可解释的风险比例方面的优势突出,SEM则能够检验预设的理论模型和变量间的复杂关系。根据具体的科学问题和数据特性选择合适的融合策略与模型,是未来研究的关键。

面临的挑战与未来发展方向

多模态融合模型在AD研究中显示出良好应用前景。这种技术有望提高诊断准确率,有助于揭示疾病机制。但目前该领域仍面临诸多挑战^[46],临床实践中,多模态数据类别多样,这些数据在采集设备、成像参数和数据格式上存在显著差异,且普遍存在缺失和类别不平衡问题,尤其对于样本量有限的情况,需要探索和应用针对小样本学习的策略,例如利用迁移学习将在大规模数据集上预训练的模型知识迁移到目标任务中进行微调^[47],或结合图像变换、生成对抗网络合成数据等数据增强技术来扩充训练样本,从而更有效地训练模型并提升其泛化能力^[48]。不同中心的数据质量参差不齐,并且多模态数据维度高、模态多导致模型结构复杂,模型的可解释性仍需进一步提升,例如Grad-CAM这类基于梯度的方法可能对噪声敏感或定位不够精确,有时会产生误导性的解释^[49],SHAP值虽然理论上更完备,但计算量大,同时可能无法完全捕捉特征间的复杂依赖^[50],Transformer的注意力权重的大小与输入特征对于模型预测的实际重要性之间并不总是有很强的相关性,并非总能直接等同于符合临床直觉的特征重要性^[51],这些可解释性技术在提供可靠且易于理解的决策依据方面仍有局限。研究中表现良好的模型在实际临床环境中难以复现,复杂模型的计算成本高且难以实时应用,而数据隐私保护要求又限制了模型推广。未来需要建立标准化的数据采集和处理流程,开发计算效率更高的算法,加强多中心验证研究,同时探索符合隐私保护要求的联邦学习等新技术^[52],推进多模态模型向临床实践转化。

参考文献:

[1] 2023 Alzheimer's disease facts and figures[J]. *Alzheimers Dement*, 2023, 19(4):1598-1695.
 [2] Rao RV, Subramaniam KG, Gregory J, et al. Rationale for a multi-

factorial approach for the reversal of cognitive decline in Alzheimer's disease and MCI: A review[J]. *Int J Mol Sci*, 2023, 24(2):1659.

- [3] Ruiz-Fernández MD, Hernández-Padilla JM, Ortiz-Amo R, et al. Predictor factors of perceived health in family caregivers of people diagnosed with mild or moderate Alzheimer's disease[J]. *Int J Environ Res Public Health*, 2019, 16(19):3762.
- [4] Shim YS, Roe CM, Buckles VD, et al. Clinicopathologic study of Alzheimer's disease: Alzheimer mimics[J]. *J Alzheimers Dis*, 2013, 35(4):799-811.
- [5] Galvin JE, Sadowsky CH. Practical guidelines for the recognition and diagnosis of dementia[J]. *J Am Board Fam Med*, 2012, 25(3):367-382.
- [6] Sharma R, Goel T, Tanveer M, et al. Deep-learning-based diagnosis and prognosis of Alzheimer's disease: a comprehensive review[J]. *IEEE T Cogn Dev Syst*, 2023, 15(3):1123-1138.
- [7] Behrad F, Saniee Abadeh M. An overview of deep learning methods for multimodal medical data mining[J]. *Expert Syst Appl*, 2022, 200:117006.
- [8] 董俊伊, 苗延巍. 阿尔茨海默病多模态MRI形态和功能研究进展[J]. *放射学实践*, 2019, 34(8):930-934.
- [9] Ronneberger O, Fischer P, Brox T. U-Net: Convolutional networks for biomedical image segmentation[C]. Springer, 2015. DOI:10.1007/978-3-319-24574-4_28.
- [10] Krishnamurthy HK, Jayaraman V, Krishna K, et al. An overview of the genes and biomarkers in Alzheimer's disease[J]. *Ageing Res Rev*, 2025, 104:102599.
- [11] 詹秀萍, 祁维民, 史琴, 等. 磁共振扩散成像技术在阿尔茨海默病的研究进展[J]. *放射学实践*, 2024, 39(6):815-819.
- [12] Moonis G, Subramaniam RM, Trofimova A, et al. ACR appropriateness criteria dementia[J]. *J Am Coll Radiol*, 2020, 17(5):S100-S12.
- [13] Alfaro-Almagro F, Jenkinson M, Bangerter NK, et al. Image processing and quality control for the first 10,000 brain imaging datasets from UK Biobank[J]. *NeuroImage*, 2018, 166:400-424.
- [14] Abhyankar S, Demner-Fushman D, McDonald CJ. Standardizing clinical laboratory data for secondary use[J]. *J Biomed Inform*, 2012, 45(4):642-650.
- [15] Ramachandram D, Taylor GW. Deep multimodal learning: a survey on recent advances and trends[J]. *IEEE Signal Proc Mag*, 2017, 34(6):96-108.
- [16] Mao S, Sejdi? E. A review of recurrent neural network-based methods in computational physiology[J]. *IEEE Trans Neural Netw Learn Syst*, 2023, 34(10):6983-7003.
- [17] Li M M, Huang K, Zitnik M. Graph representation learning in biomedicine and healthcare[J]. *Nat Biomed Eng*, 2022, 6(12):1353-1369.
- [18] Nerella S, Bandyopadhyay S, Zhang J, et al. Transformers and large language models in healthcare: a review[J]. *Artifi Intell Med*, 2024, 154:102900.
- [19] Cox DR. Regression models and life-tables[J]. *J R Stat Soc Ser B Methodol*, 1972, 34(2):187-202.
- [20] Bollen KA. Structural equations with latent variables[M]. Wiley-Interscience, 1989:1-9.
- [21] Stahlschmidt SR, Ulfenborg B, Synnergren J. Multimodal deep

- learning for biomedical data fusion: a review[J]. *Brief Bioinform*, 2022, 23(2): bbab569.
- [22] Kline A, Wang H, Li Y, et al. Multimodal machine learning in precision health: a scoping review[J]. *NPJ Digit Med*, 2022, 5(1): 171.
- [23] Wiegrefe S, Pinter Y. Attention is not not Explanation[C]// Inui K, Jiang J, Ng V, et al. Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing (EMNLP-IJCNLP). Association for Computational Linguistics, 2019: 11-20. 10.18653/v1/D19-1002
- [24] Allgaier J, Mulansky L, Draelos RL, et al. How does the model make predictions? A systematic literature review on the explainability power of machine learning in healthcare[J]. *Artif Intell Med*, 2023, 143: 102616.
- [25] A El-Gamal FE, Elmogy MM, Ghazal M, et al. Medical imaging diagnosis of early Alzheimer's disease[J]. *Front Biosci (Landmark Ed)*, 2018, 23(4): 671-725.
- [26] Rani S, Dhar SB, Khajuria A, et al. Advanced overview of biomarkers and techniques for early diagnosis of Alzheimer's disease[J]. *Cell Mol Neurobiol*, 2023, 43(6): 2491-2523.
- [27] Li P, Quan W, Wang Z, et al. Early-stage differentiation between Alzheimer's disease and frontotemporal lobe degeneration: clinical, neuropsychology, and neuroimaging features[J]. *Front Aging Neurosci*, 2022, 14: 981451.
- [28] Venugopalan J, Tong L, Hassanzadeh HR, et al. Multimodal deep learning models for early detection of Alzheimer's disease stage[J]. *Sci Rep*, 2021, 11(1): 3254.
- [29] Tian X, Liu Y, Wang L, et al. An extensible hierarchical graph convolutional network for early Alzheimer's disease identification[J]. *Comput Methods Programs Biomed*, 2023, 238: 107597.
- [30] Lei B, Li Y, Fu W, et al. Alzheimer's disease diagnosis from multimodal data via feature inductive learning and dual multilevel graph neural network[J]. *Med Image Anal*, 2024, 97: 103213.
- [31] Sevigny J, Chiao P, Bussière T, et al. The antibody aducanumab reduces A β plaques in Alzheimer's disease[J]. *Nature*, 2016, 537(7618): 50-56.
- [32] Van Dyck Christopher H, Swanson Chad J, Aisen P, et al. Lecanemab in early Alzheimer's disease[J]. *New Engl J Med*, 2023, 388(1): 9-21.
- [33] Hampel H, Elhage A, Cho M, et al. Amyloid-related imaging abnormalities (ARIA): radiological, biological and clinical characteristics[J]. *Brain*, 2023, 146(11): 4414-4424.
- [34] Noufi P, Khoury R, Jeyakumar S, et al. Use of Cholinesterase inhibitors in non-Alzheimer's dementias[J]. *Drugs Aging*, 2019, 36(8): 719-731.
- [35] Li Y, Hai S, Zhou Y, et al. Cholinesterase inhibitors for rarer dementias associated with neurological conditions[J]. *Cochrane Database Syst Rev*, 2015, 3(2015): CD009444.
- [36] Imfeld P, Bodmer M, Schuerch M, et al. Risk of incident stroke in patients with Alzheimer disease or vascular dementia[J]. *Neurology*, 2013, 81(10): 910-919.
- [37] Xue C, Kowshik SS, Lteif D, et al. AI-based differential diagnosis of dementia etiologies on multimodal data[J]. *Nat Med*, 2024, 30(10): 2977-2989.
- [38] Petersen RC, Roberts RO, Knopman DS, et al. Mild cognitive impairment: ten years later[J]. *Arch Neurol*, 2009, 66(12): 1447-1455.
- [39] Schneider JA, Arvanitakis Z, Leurgans SE, et al. The neuropathology of probable Alzheimer disease and mild cognitive impairment[J]. *Ann Neurol*, 2009, 66(2): 200-208.
- [40] Velazquez M, Lee Y. Multimodal ensemble model for Alzheimer's disease conversion prediction from early mild cognitive impairment subjects[J]. *Comput Biol Med*, 2022, 151: 106201.
- [41] Rahim N, El-Sappagh S, Ali S, et al. Prediction of Alzheimer's progression based on multimodal deep-learning-based fusion and visual explainability of time-series data[J]. *Inform Fusion*, 2023, 92: 363-388.
- [42] Eslami M, Tabarestani S, Adjouadi M. A unique color-coded visualization system with multimodal information fusion and deep learning in a longitudinal study of Alzheimer's disease[J]. *Artif Intell Med*, 2023, 140: 102543.
- [43] Marti-Juan G, Lorenzi M, Piella G. MC-RVAE: multi-channel recurrent variational autoencoder for multimodal Alzheimer's disease progression modelling[J]. *NeuroImage*, 2023, 268: 119892.
- [44] Reas ET, Shadrin A, Frei O, et al. Improved multimodal prediction of progression from MCI to Alzheimer's disease combining genetics with quantitative brain MRI and cognitive measures[J]. *Alzheimers Dement*, 2023, 19(11): 5151-5158.
- [45] Tosun D, Demir Z, Veitch DP, et al. Contribution of Alzheimer's biomarkers and risk factors to cognitive impairment and decline across the Alzheimer's disease continuum[J]. *Alzheimers Dement*, 2022, 18(7): 1370-1382.
- [46] Carini C, Seyhan AA. Tribulations and future opportunities for artificial intelligence in precision medicine[J]. *J Transl Med*, 2024, 22(1): 411.
- [47] Alzubaidi L, Al-Amidie M, Al-Asadi A, et al. Novel transfer learning approach for medical imaging with limited labeled data[J]. *Cancers (Basel)*, 2021, 13(7): 1590.
- [48] Goceri E. Medical image data augmentation: techniques, comparisons and interpretations[J]. *Artif Intell Rev*, 2023, 56(11): 12561-12605.
- [49] Draelos RL, Carin L. Use HiResCAM instead of Grad-CAM for faithful explanations of convolutional neural networks[J]. *arXiv*, 2021. <https://doi.org/10.48550/arXiv.2011.08891>.
- [50] Salih AM, Galazzo IB, Raisi-Estabragh Z, et al. Characterizing the contribution of dependent features in XAI methods[J]. *IEEE J Biomed Health Inform*, 2024, 28(11): 6466-6473.
- [51] Jain S, Wallace BC. Attention is not Explanation[C]// Burstein J, Doran C, Solorio T. Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long and Short Papers). Association for Computational Linguistics, 2019: 3543-3556. 10.18653/v1/N19-1357.
- [52] Aouedi O, Sacco A, Piamrat K, et al. Handling privacy-sensitive medical data with federated learning: challenges and future directions[J]. *IEEE J Biomed Health Inform*, 2023, 27(2): 790-803.