·数智神经外科学·

基于宏观与微观影像学特征融合的胶质瘤 基因状态与分级预测研究

李臻 宋鹏斐 朱锐泽 江山 曹诗文 余锦华 史之峰

【摘要】目的 建立基于 MRI 与全视野数字切片(WSI)特征融合的双层特征蒸馏的多实例学习(DLFD-MIL)模型,实现对成人型弥漫性胶质瘤 *IDH1* 突变、1p/19q共缺失及世界卫生组织(WHO)分级的高精准性预测。方法 选择 2021年1月至 2024年6月复旦大学附属华山医院收治的 212例成人型弥漫性胶质瘤患者及美国癌症基因组图谱计划 42例成人型弥漫性胶质瘤病例,联合分析术前T₂-FLAIR影像与术后 WSI数据。构建 DLFD-MIL模型,采用伪包生成策略解决 WSI弱监督学习中的实例标签缺失问题,Concat 融合方式实现多模态融合;绘制受试者工作特征曲线,以曲线下面积比较单模态与多模态特征融合模型的曲线下面积大于单模态 WSI模型(Z = 2.752, P = 0.006)和单模态 T₂-FLAIR 模型(Z = 5.662, P = 0.000);在 1p/19q共缺失预测任务中,多模态特征融合模型的曲线下面积与单模态 T₂-FLAIR 模型(Z = 0.781, P = 0.435)差异均无统计学意义;在 WHO 分级预测任务中,多模态特征融合模型的曲线下面积大于单模态 T₂-FLAIR 模型(Z = 0.781, P = 0.435)差异均无统计学意义;在 WHO 分级预测任务中,多模态特征融合模型的曲线下面积大于单模态 T₂-FLAIR 模型(Z = 4.830, P = 0.000),而与单模态 WSI模型差异无统计学意义(Z = 1.739, P = 0.082)。结论 基于宏观与微观影像学特征融合模型可以提高胶质瘤 *IDH1* 基因分型和WHO 分级的预测精度,为临床制定个性化治疗方案提供可靠的人工智能决策支持工具。

【关键词】 神经胶质瘤; 磁共振成像; 病理学; 基因; 肿瘤分级; 深度学习; ROC曲线

Prediction of genetic status and grading in glioma based on fusion of macro- and micro-imaging features

LI Zhen¹, SONG Peng-fei², ZHU Rui-ze¹, JIANG Shan¹, CAO Shi-wen³, YU Jin-hua⁴, SHI Zhi-feng¹

¹Department of Neurosurgery, Huashan Hospital, Fudan University, Shanghai 201107, China

²Grade 2023 of Department of Biomedical Engineering, ³Grade 2021 of Department of Biomedical Engineering, ⁴Department of Biomedical Engineering, School of Information Science and Technology, Fudan University, Shanghai 200438, China

LI Zhen and SONG Peng-fei contributed equally to the article

Corresponding author: SHI Zhi-feng (Email: shizhifeng@fudan.edu.cn)

(Abstract) Objective To develop a dual-layer feature distillation multiple instance learning (DLFD-MIL) model integrating MRI and whole slide image (WSI) features for precise prediction of *IDH1* mutation, 1p/19q codeletion, and World Health Organization (WHO) grading in adult-type diffuse gliomas. **Methods** A retrospective cohort of 212 adult-type diffuse gliomas patients from Huashan Hospital, Fudan University (January 2021 to June 2024) and 42 cases from The Cancer Genome Atlas (TCGA) were included. Preoperative T₂-FLAIR and postoperative WSI data were jointly analyzed. The DLFD-MIL model addressed

doi:10.3969/j.issn.1672-6731.2025.03.002

基金项目:上海市卫生健康委员会优秀项目(项目编号:20234Z0009);国家重点研发计划项目(项目编号: 2022YFF1202804);国家自然科学基金资助项目(项目编号:82373018);国家自然科学基金资助项目(项目编号: 82072020);上海市科委医学创新研究项目(项目编号:23Y11906200)

作者单位:201107 上海,复旦大学附属华山医院神经外科(李臻,朱锐泽,江山,史之峰);200438 上海,复旦大学信息科学与工程学院生物医学工程系2023级(宋鹏斐),2021级(曹诗文),生物医学工程系(余锦华)

李臻与宋鹏斐对本文有同等贡献

通讯作者:史之峰,Email:shizhifeng@fudan.edu.cn

the lack of instance-level labels in weakly supervised WSI learning via a pseudo-bag generation strategy. Multimodal feature fusion was achieved through Concat. Diagnostic performance for molecular subtyping and WHO grading was evaluated by comparing area under the curve (AUC) of receiver operating characteristic (ROC) curve between single-mode (WSI or MRI) and multi-mode. **Results** In the *IDH1* mutation prediction task, AUC of the multi-mode feature fusion model surpassed single-mode WSI model (Z = 2.752, P = 0.006) and single-mode T₂-FLAIR model (Z = 5.662, P = 0.000). In the 1p/19q codeletion prediction task, no statistically significant differences in AUC were observed between the multi-mode feature fusion model and either single-mode WSI model (Z = -0.245, P = 0.806) or T₂-FLAIR model (Z = 0.781, P = 0.435). In the WHO grading prediction task, the multi-mode feature fusion model showed no significant differences in AUC compared to single -mode WSI model (Z = 1.739, P = 0.082), however its AUC was significantly higher than single-mode T₂-FLAIR model (Z = 4.830, P = 0.000). **Conclusions** Multi-mode fusion of macro- and micro-imaging features improves prediction accuracy for *IDH1* genotyping and WHO grading in gliomas, providing a reliable artificial intelligence (AI) decision - support tool for personalized clinical management.

[Key words] Glioma; Magnetic resonance imaging; Pathology; Genes; Neoplasm grading; Deep learning; ROC curve

This study was supported by the Excellent Project of Shanghai Municipal Health Commission (No. 20234Z0009), National Key Research and Development Program of China (No. 2022YFF1202804), the National Natural Science Foundation of China (No. 82373018, 82072020), and Medical Innovation Research Project of Shanghai Science and Technology Commission (No. 23Y11906200).

Conflicts of interest: none declared

脑胶质瘤是临床最为常见的原发性颅内肿瘤, 约占所有脑肿瘤的 22.9%,发病率为 5.80/10 万^[1-2]。 《2021年世界卫生组织(WHO)中枢神经系统肿瘤分 类(第五版)》将分子分型确立为胶质瘤诊断与治疗 的"金标准", IDH1突变、1p/19q共缺失等分子生物 学特征不仅决定其组织病理学分级,而且指导手术 策略的制定和治疗方案的选择^[34]。然而,当前临床 实践面临术前分子检测流程较慢、术中冰冻病理学 检测分子突变敏感性低、术后复发患者分子生物学 标志物动态监测缺乏有效手段等多重挑战[5-7]。近 年多模态机器学习(ML)技术为解决上述问题提供 了新的思路,如基于多参数MRI的影像组学模型和 基于组织病理切片的深度学习(DL)算法对术前分 子分型预测和识别以及辅助手术精准切除病灶,均 展现出巨大潜力;且其准确性、曲线下面积(AUC)等 预测效能均优于单一模态^[8-10]。基于此,本研究提 出一种基于双层特征蒸馏多实例学习(DLFD-MIL) 模型和包嵌入法的胶质瘤诊断方法[11],第一层在伪 包(pseudo-bag)内通过注意力权重提取具备高诊断 价值的伪包级特征;第二层则通过最大最小选择法 (MaxMinS)优化第二层特征输入,并输出跨伪包的 全局特征。同时构建基于 T,-FLAIR 成像与术后 HE 染色全视野数字切片(WSI)特征融合的多模态预测 体系,以期实现对胶质瘤 IDH1 突变、1p/19q 共缺失 及 WHO 分级的高准确性预测,通过分子-影像-病理 的宏观与微观特征融合,拟为胶质瘤诊断与外科治 疗提供术前分子预测-术中分子边缘界定-术后分子 随访的全周期应用。

对象与方法

一、研究对象

1.复旦大学附属华山医院病例纳入与排除标 准 (1)根据《2021年WHO中枢神经系统肿瘤分类 (第五版)》标准^[3],选择经术后组织病理学及分子病 理学检测首次确诊的成人型弥漫性胶质瘤患者,且 *IDH1*及1p/19q分子检测资料完整。(2)年龄>18岁。 (3)术前T₂-FLAIR成像及术后HE染色WSI质量均 符合诊断和评价要求,且分辨率≥0.25 μm/pixel。 (4)凡合并其他肿瘤或严重全身性疾病、影像学或 组织病理学数据不满足分析要求者均不属于本研 究观察范围。(5)本研究获得复旦大学附属华山医 院伦理委员会审核批准[审批号:(2023)临审第 (1012)号]。(6)遵循《赫尔辛基宣言》及相关伦理学 要求,所有纳入对象均对本研究内容知情并签署知 情同意书,研究数据严格保密,仅用于科研目的。

2.美国癌症基因组图谱计划公开数据库病例纳 入与排除标准 (1)经过详细的分子分型和标注, *IDH1*突变、1p/19q共缺失和WHO分级信息资料完

· 166 ·

整。(2)具备完整的术前 T_2 -FLAIR 影像数据(层厚 ≤ 5 mm)和术后 HE 染色组织病理学数字切片(分辨 率 > 0.50 µm²/pixel × 0.50 µm²/pixel)。(3)经 dbGaP 数 据库授权使用。(4)排除影像学或组织病理数据质 量不满足分析要求,以及术后复发或术前曾接受过 系统治疗的病例。

3. 一般资料 根据上述纳入与排除标准,选择 2021年1月至2024年6月在复旦大学附属华山医院 神经外科接受手术治疗的成人型弥漫性胶质瘤患 者 212 例(华山医院组),联合美国癌症基因组图谱 计划(TCGA)公开数据库中42例成人型弥漫性胶质 瘤病例(TCGA数据库组),共计254例,男性144例 (56.69%),女性110例(43.31%);年龄16~84岁,中 位值43(34,54)岁;组织学类型分别为星形细胞瘤 120 例(47.24%), 少突胶质细胞瘤 57 例(22.44%), 胶质母细胞瘤 77 例(30.31%)。经第二代测序技术 (NGS)获得 IDH1 突变状态分子分型,突变型 135 例 (53.15%)、野生型119例(46.85%);经荧光原位杂交 (FISH)获得 1p/19g 共缺失状态, 1p/19g 共缺失 72 例 (28.35%)、无共缺失182例(71.65%,含1例1p缺失 19q未缺失)。依据《2021年WHO中枢神经系统肿 瘤分类(第五版)》^[3],结合组织学和分子生物学特 征,最终判定为低级别胶质瘤(WHO 1~2级)121例 (47.64%)、高级别胶质瘤(WHO 3~4级)为133例 $(52.36\%)_{\odot}$

二、研究方法

1.数据采集 (1)MRI数据采集。华山医院组: 采用 MAGNETOM Trio 3.0T MRI 扫描仪(德国 Siemens Healthineers公司)获取患者术前T,-FLAIR 影像。TCGA数据库组:通过TCGA官方数据门户 v32.0版本(https://portal.gdc.cancer.gov)检索,筛选 "Project"为"TCGA-LGG"和"TCGA-GBM"的病例。 随后通过癌症影像存档库(TCIA)的API接口交叉 匹配 TCGA-LGG 和 TCGA-GBM 中含 T,-FLAIR 影像 的病例。两组病例的影像数据均通过DICOM格式 导出,采用 dcm2niix 开源工具(https://github.com/ rordenlab/dcm2niix)将 DICOM 格式转换为 NIfTI 格 式,再通过ANTs开源工具(https://github.com/ ANTsX/ANTs)配准至标准加拿大蒙特利尔神经病 学研究所(MNI)152空间。基于三维U-Net网络 (3D U-Net)模型(预训练于 BraTS 2020数据集)自动 勾画并分割肿瘤感兴趣区(ROI),分割结果经两位 神经放射科医师盲法复核,以影像组学特征的组内

相关系数(ICC) > 0.80 判定为高度稳定,若 ICC < 0.80,则由第三位资深医师采用 ITK-SNAP 3.8.0 软件重新勾画感兴趣区,以人工勾画边界重新计算 ICC,经两轮迭代后仍 ICC < 0.80 者予以排除,无排除病例。结果以 NIfTI 格式存储,用于后续多模态特征融合。(2)组织学图像采集。华山医院组:从 Digital Microscopic Imaging System Ocus[®](芬兰Grundium公司)获取 200 倍放大的 HE染色 WSI 图像,图像分辨率 0.4793 μ m²/pixel × 0.4793 μ m²/pixel。 TCGA数据库组:获得上述 TCGA数据库与 T₂-FLAIR 成像匹配的组织病理学数据,并提取相匹配的 HE 染色 WSI 图像。两组 WSI 数据均以虚拟切片图像 (SVS)格式进行标准化存储,用于后续 DLFD-MIL模型构建。

2. 双层特征蒸馏多实例学习模型构建 采用 DLFD-MIL模型的双层特征蒸馏和包嵌入法对WSI 数据进行特殊处理,以实现胶质瘤WSI基因状态和 WHO分级的精准预测。(1)数据准备及结构预处理: 按随机策略将单个WSI切片划分为50个伪包^[11], 每个伪包包含约50个图像块(patch), 伪包标签继 承其父包(原始包)的整体标签,并通过分层随机抽 样保证伪包内肿瘤细胞密度分布与父包一致。采 用深度残差网络(ResNet)提取伪包内图像块深度特 征并进行归一化,构建维度为50伪包×50实例× 512特征的张量。(2)第一层基于注意力机制的多实 例学习(AB-MIL)模型^[12]:其输入为结构预处理后提 取的图像块级别特征,输出为每个伪包的蒸馏特 征。每个伪包通过 AB-MIL 模型(图1)聚合实例特 征生成伪包级表示。为量化每个图像块对包级预 测的贡献,利用梯度加权类激活映射(Grad-CAM)技 术^[13],根据AB-MIL模型提供的梯度信息定位关键 图像块,生成类激活图(CAM),验证模型关注区域 与组织病理图像标注的一致性。(3)第二层 AB-MIL 模型:其输入为第一层输出的伪包蒸馏特征,输出 为最终用于分类或多模态特征融合的跨伪包全局 WSI特征。为克服伪包划分引入的噪声,分别采用 最大选择(MaxS)、最大最小选择、最大注意力分数 选择(MAS)和聚合特征选择(AFS)共4种特征蒸馏 方法,筛选关键实例优化第二层AB-MIL模型输入。 经筛选试验,以最大最小选择法可在特征蒸馏过程 中同时选择最大和最小特征值,使模型对关键区域 的敏感性增强,同时减少噪声干扰,能更好地平衡 正负样本的识别,其受试者工作特征(ROC)曲线下



面积最大,诊断特征明显,因此最终选择最大最小选择法策略构建最终的DLFD-MIL模型(图2)。将经过DLFD-MIL模型处理后的WSI特征输入多模态特征融合模型,完成基因及WHO分级的联合预测任务。

3. 多模态特征融合 T₂-FLAIR 数据经配准后, 通过 3D U-Net 模型分割 T₂-FLAIR 影像的肿瘤感兴 趣区并提取图像深度特征(1024 维),同时基于上述 DLFD-MIL 模型从 WSI 图像中生成全局组织病理特 征(768 维,图 3)。以 Add、Concat 和 TFN 张量 3 种融 合策略整合跨模态信息,经筛选试验,以 Concat 融 合方式的 ROC 曲线下面积最大,表明其诊断特征显 著、预测效能最佳,因此选择 Concat 融合方式的多 模态特征融合获得最终的基因及 WHO 分级的预测 信息。

4.多模态特征融合模型训练及验证 为减少数 据来源差异引起的系统性偏倚,采用分层随机抽样 策略将 254 例患者数据划分为训练集(170 例, 66.93%)和测试集(84 例,33.07%),其中测试集包含 内部测试集(华山医院组,42 例)和外部测试集 (TCGA数据库组,42 例);根据*IDH1*突变型/野生型、 1p/19q共缺失及WHO分级,于华山医院组进行组内 分层抽样以保证肿瘤突变类型均衡。在训练集内 部,通过分层五折交叉验证优化模型参数将170 例 数据分为5个子集(各34 例),每次迭代使用4个子 集(136 例,80%)进行训练,余1个子集(34 例,20%)



作为验证集,重复5次以评估模型稳定性,最终模型 参数选择基于五折交叉验证的平均曲线下面积。 最终生成的模型在测试集中评估模型预测效能,取 ROC曲线上垂直距离最大的点即最大Youden指数 点(Youden指数=灵敏度+特异度-1)作为诊断阈 值,其中灵敏度(%)=真阳性例数/(真阳性例数+ 假阴性例数)×100%,特异度(%)=真阴性例数/ (真阴性例数+假阳性例数)×100%(真阳性例数为 "金标准"检测阳性且模型正确预测亦呈阳性者,真 阴性例数为"金标准"检测阴性且模型正确预测亦 呈阴性者,假阳性例数为"金标准"检测为阴性但模 型误判为阳性者,假阴性例数为"金标准"检测阳性 但模型误判为阴性者)。

5.统计分析方法 本研究模型训练和所有数据 分析处理均通过 Python 3.9及其开源科学计算库实 现。计数资料以相对数构成比(%)或率(%)表示; 采用 Shapiro-Wilk 检验评估连续变量的正态性,呈 非正态分布的计量资料以中位数和四分位数间距 $[M(P_{25}, P_{75})]$ 表示。采用诊断试验评价单模态和多 模态特征融合对胶质瘤基因状态和WHO分级的预 测价值,通过绘制 ROC 曲线并计算曲线下面积评估 其预测效能,两两比较采用 Delong 检验。为系统评 估不同预测模型的临床实用性,进一步通过决策曲 线分析(DCA)量化多模态特征融合模型与单模态模 型的临床净收益差异。以 $P \leq 0.05$ 为差异具有统计 学意义。

结 果

以 NGS 测序检测 *IDH1* 突变状态作为"金标 准",测试集84 例患者中野生型44 例、突变型40 例。 对 3 种诊断方法效能的评价显示,单模态 T₂-FLAIR 诊断灵敏度为 27.50%(11/40)、特异度97.73%(43/ 44);单模态 WSI诊断灵敏度为90%(36/40)、特异度 75%(33/44);多模态特征融合诊断灵敏度为97.50% (39/40)、特异度81.82%(36/44,表1)。

以荧光原位杂交为检测 1p/19q共缺失状态的 "金标准",测试集 84 例患者中 1p/19q共缺失 19 例、 无共缺失 65 例。对 3 种诊断方法的效能评价显示, 单模态 T₂-FLAIR 诊断灵敏度为 42.11%(8/19)、特异 度 87.69%(57/65);单模态 WSI 诊断灵敏度为 57.89%(11/19)、特异度 83.08%(54/65);多模态特征 融合诊断灵敏度为 78.94%(15/19)、特异度 61.54% (40/65,表2)。

以《2021年WHO中枢神经系统肿瘤分类(第五版)》整合诊断标准作为"金标准"进行胶质瘤分级,测试集84例患者中高级别为46例、低级别38例。对3种诊断方法的效能评价显示,单模态T₂-FLAIR诊断灵敏度为34.78%(16/46)、特异度94.74%(36/38);单模态WSI诊断灵敏度为76.09%(35/46)、特异度86.84%(33/38);多模态特征融合诊断灵敏度

· 169 ·

Table 1. Comparison of predictive performance of different diagnostic models for IDH1 mutation status (case)											
诊断试验	金标准		一会计	いた地にいたった	金标准		스러		金标准		스러
	突变型	野生型	百月	区的风驰	突变型	野生型	百月	区的风驰	突变型	野生型	百月
单模态 T ₂ -FLAIR				单模态WSI				多模态特征融合			
预测突变型	11	1	12	预测突变型	36	11	47	预测突变型	39	8	47
预测野生型	29	43	72	预测野生型	4	33	37	预测野生型	1	36	37
合计	40	44	84	合计	40	44	84	合计	40	44	84

表1 不同诊断模型对 IDH1 突变状态预测效能的比较(例)

WSI, whole slide image, 全视野数字切片。The same for Table 1-3

表2 不同诊断模型对1p/19q共缺失状态预测效能的比较(例)

Table 2. Comparison of predictive performance of different diagnostic models for 1p/19q codeletion status (case)

公底计应	金标准		<u>л</u> н	いん 城市 いーヤ 市人	金标准		A 21	公库计办	金标准		A 21
诊断试验	共缺失	无共缺失	- 合り	区时风驰	共缺失	无共缺失	ΠИ		共缺失	无共缺失	百月
单模态 T ₂ -FLAIR				单模态WSI				多模态特征融合			
预测共缺失	8	8	16	预测共缺失	11	11	22	预测共缺失	15	25	40
预测无共缺失	11	57	68	预测无共缺失	8	54	62	预测无共缺失	4	40	44
合计	19	65	84	合计	19	65	84	合计	19	65	84

表3 不同诊断模型对 WHO 分级预测效能的比较(例)

Table 3. Comparison of predictive performance of different diagnostic models for WHO grading (case)

诊断试验	金标准		스러		金标准		へけ	→△ w⊆ →→ 元→	金标准		会社
	高级别	低级别	合订	区时风驰	高级别	低级别	百月	区的风驰	高级别	低级别	百月
单模态T ₂ -FLAIR				单模态WSI				多模态特征融合			
预测高级别	16	2	18	预测高级别	35	5	40	预测高级别	38	4	42
预测低级别	30	36	66	预测低级别	11	33	44	预测低级别	8	34	42
合计	46	38	84	合计	46	38	84	合计	46	38	84

为82.61%(38/46)、特异度89.47%(34/38,表3)。

基于模型预测概率与"金标准"标签的对比,绘 制ROC曲线并计算曲线下面积。在IDH1突变预测 任务中,单模态T₂-FLAIR模型、单模态WSI模型及 多模态特征融合模型的曲线下面积分别为0.589 (95%CI: 0.463 ~ 0.714) \ 0.867 (95%CI: 0.788 ~ 0.947)及0.955(95%CI:0.918~0.991);在1p/19g共 缺失预测任务中,3种模型曲线下面积分别为0.643 (95%CI: 0.492 ~ 0.794) \ 0.731 (95%CI: 0.600 ~ 0.862)及0.708(95%CI:0.583~0.832);在WHO分级 预测任务中,3种模型曲线下面积分别为0.627 (95%CI: 0.509 ~ 0.745) 及 0.868 (95%CI: 0.791 ~ 0.946)、0.918(95%CI: 0.862 ~ 0.973, 图 4)。 Delong 检验结果显示,多模态特征融合模型在IDH1突变预 测任务中表现优异,其预测的曲线下面积大于单模 态 T₂-FLAIR 模型(P=0.000)和单模态 WSI模型(P= 0.006),单模态 WSI 模型预测的曲线下面积大于单 模态 T₂-FLAIR 模型(P = 0.000);在 1p/19q 预测任务中,上述 3 种模型预测能力一般(曲线下面积 < 0.75),且在 3 个预测任务中差异均无统计学意义(P > 0.05);在 WHO 分级预测任务中,多模态特征融合模型预测的曲线下面积大于单模态 T₂-FLAIR 模型(P = 0.000),而与单模态 WSI 模型预测差异无统计学意义(P = 0.082),单模态 WSI 模型预测的曲线 下面积大于单模态 T₂-FLAIR 模型(P = 0.000,表4)。

进一步通过决策曲线分析量化多模态特征融 合模型与单模态模型的临床净收益差异。在高风 险阈值(模型预测下对患者采取干预措施的最低患 病率,图5中红线与蓝线交点)<0.77时,多模态特 征融合模型预测*IDH1*突变的临床净收益高于单模 态模型;而在1p/19q预测任务中上述3种模型的决 策曲线均在多个高风险阈值点与全治疗(treat all) 或全不治疗(treat none)曲线交叉,表明3种模型并 未显示出显著的临床收益;在WHO分级预测中,在



表 4	多模态特征融合模型与单模态模型预测效能的比
较	

 Table 4.
 Comparison of predictive performance between multi-mode feature fusion model and single-mode model

41 间 西 西 中	IDH1突变						
坦回四四日	Z 值	P 值					
多模态特征融合:单模态 WSI	2.752	0.006					
多模态特征融合:单模态T ₂ -FLAIR	5.662	0.000					
单模态WSI:单模态T ₂ -FLAIR	4.064	0.000					
41 问 西 西 中	1p/19q共缺失						
组间网内比	Z 值	P值					
多模态特征融合:单模态 WSI	- 0.245	0.806					
多模态特征融合:单模态T ₂ -FLAIR	0.781	0.435					
单模态WSI:单模态T ₂ -FLAIR	0.712	0.476					
41 问 更 更 以	WHO 分级						
组间网网比	Z 值	P 值					
多模态特征融合:单模态WSI	1.739	0.082					
多模态特征融合:单模态T ₂ -FLAIR	4.830	0.000					
单模态WSI:单模态T ₂ -FLAIR	3.819	0.000					
WSI, whole slide image, 全视野数字切片; IDH1, isocitrate dehydrogenase 1, 异柠檬酸脱氢酶 1; WHO, World Health Organization,世界卫生组织							

0.06~0.39及0.47~0.89的高风险阈值范围内多模态特征融合模型展现出稳定优势(图5)。



AUC,曲线下面积;WSI,全视野数字切片

图 4 基于宏观与微观影像学特征融合预测基因状态 和 WHO 分级的 ROC 曲线 4a *IDH1* 突变状态预测 4b 1p/19q共缺失状态预测 4c WHO 分级预测 Figure 4 ROC curves for predicting genetic status and WHO grading based on macro - and micro - imaging features fusion *IDH1* mutation status prediction (Panel 4a). 1p/19q codeletion status prediction (Panel 4b). WHO grading prediction (Panel 4c).

讨 论

本研究采用基于宏观与微观影像学特征融合 预测胶质瘤基因状态和WHO分级:宏观MRI影像 学数据经3DU-Net模型提取特征后直接纳入多模 态特征融合分析,而考虑到HE染色切片中肿瘤细 胞可能仅占一小部分,故将WSI数据经DLFD-MIL 模型进行双层蒸馏提取特征后再融合,即通过双层 蒸馏和伪包策略将HE染色切片中真正显示肿瘤细 胞特征的图像块准确提取,然后再进行多模态特征 融合。

在基于WSI分析的多实例学习模型的分类架构中,通常分为两个主要类别:第一类基于从单个 实例的预测结果直接导出包级预测^[14-17],即传统的 多实例学习模型;第二类通过整合单个实例的特征 形成更高级别的包表示,然后基于这种表示构建分 类器以进行包级别预测^[12,18-20],通常称为包嵌入法。 传统的多实例学习模型将WSI视为由局部图像块 (实例)所构成的包,通过注意力机制筛选关键区域 完成图像分类任务,但其单层结构难以应对*IDH1*突 变灶稀疏分布、局部坏死等胶质瘤组织病理学的复 杂性;与之相比,包嵌入法通常具有更高的分类性

• 171 •

· 172 ·





WHO 分级的决策曲线分析图 5a *IDH1* 突变状态预测 5b 1p/19q共缺失状态预测 5c WHO 分级预测 **Figure 5** DCA for predicting genetic status and WHO grading based on macro- and micro-imaging features fusion *IDH1* mutation status prediction (Panel 5a). 1p/19q codeletion status prediction (Panel 5b). WHO grading prediction (Panel 5c).

能^[19]。本研究提出一种基于分层注意力蒸馏的改 进AB-MIL模型,其核心创新在于利用包嵌入法与 最大最小选择法策略,解决WSI实例标签缺失与噪 声干扰问题。具体而言,模型将单张WSI划分为多 个伪包,并通过第一层 AB-MIL 模型提取伪包级特 征:第二层则采用最大最小选择法策略筛选关键特 征,从而优化输入至最终分类器的特征表示。这种 双层特征蒸馏过程使得第一层 AB-MIL 模型为第二 层模型提供了初始征候选集,从而在整个训练过程 中逐步优化模型参数。本模型突破传统弱监督学 习下存在的如单模态 WSI 对微浸润灶漏检等局限 性^[21-22],首次在WSI分类任务中实现实例-伪包-父 包的三级特征提取,为组织病理学数据分析提供更 精细的弱监督学习路径。本研究结果显示,与单模 态模型相比,融合MRI与WSI特征的多模态模型在 IDH1突变预测任务中表现最优异,曲线下面积达到 0.955(95%CI:0.918~0.991),其敏感性和特异性在 所有诊断试验中最为均衡,验证了 DLFD-MIL 模型 在特征提炼与信息互补中的有效性。

当前基于人工智能(AI)的胶质瘤多模态数据融合研究多聚焦于宏观影像组学模态(如MRI、PET与超声)的模态内多模态整合分析^[23-27],而宏观与微观的融合研究仍处于探索阶段。尽管面临宏观与微观模态数据存在对齐复杂、组织病理学图像块级标注缺失等问题,但MRI与WSI的多模态特征融合

仍展现出独特的互补优势:T₂-FLAIR 成像提供肿瘤 富水组织的空间定位信息,但其对微环境分子表达 异质性的解析能力有限;而WSI虽然缺乏整体空间 关联性,但通过200倍高分辨率成像揭示细胞核形 态、染色质分布等微观特征。本研究结果发现, Concat融合方式在保留双模态原始信息的同时,可 通过简单而高效的维度叠加显著提高预测效能;本 研究也纳入TFN 张量融合法,该方法基于高阶张量 交互将两种特征映射至1024维,融合特征输入线性 分类器输出多基因分类及分级结果,但本研究所得 结果提示,尽管理论上高阶融合方法能够捕捉复杂 模态交互^[28],但其计算复杂性与硬件需求限制其临 床实用性,而Concat融合方式在多模态融合场景中 可能更具普适性。

本研究构建的多模态特征融合模型在胶质瘤 的诊断与治疗全周期中有一定的临床转化潜力,其 应用场景可系统划分为3个阶段:第一阶段,术前基 于穿刺活检获取WSI数据,模型可在30min内完成 *IDH1*突变状态与1p/19q共缺失预测,较传统分子检 测(5~7d)显著缩短决策时间,使靶向治疗方案或 新辅助化疗得以在确诊后快速制定或启动。第2阶 段,术中通过配准术前MRI与术中冰冻切片快速扫 描WSI,模型可动态界定肿瘤分子活跃区,指导术者 在神经导航下实现分子全切。第三阶段,术后整合 已有的基线WSI与随访MRI数据,模型可捕捉*IDH1* 野生型向突变型转化或1p/19q二次缺失等分子事件,可在单纯影像学复发诊断的基础上提前发出分子分型转化预警。

本研究存在一定的局限性:(1)样本规模与多 样性受限。尽管本研究初步验证了模型的内部及 外部适用性,但受限于两组样本的胶质瘤分子分型 特征差异及TCGA数据库组西方病例占比较高的种 族构成特征,当前样本量在深度学习模型训练中仍 显不足,且存在潜在的非高加索人群泛化风险。此 外,因数据来源为多机构采集,TCGA数据库组MRI 数据存在包括场强差异、层厚波动及扫描序列参数 不统一等异质性,WSI数据存在空间分辨率、HE染 色试剂批次及切片厚度等多样性,虽然经过一定预 处理,仍将引入一定的多中心效应,影响模型的预 测效能。(2)多模态数据空间配准的生物学限制。 由于胶质瘤周围水肿带与IDH1/2突变区域并不完 全重合,故T₂-FLAIR成像不能完全反应胶质瘤的分 子边界[5],水肿区内混杂的非突变组织干扰关键特 征的提取,或分子活跃区信号被周围水肿高信号稀 释,降低多模态特征融合的判别权重。(3)多模态性 能的灵敏度与特异度失衡。如单模态 T,-FLAIR 模 型具有较低的灵敏度与较高的特异度(表1~3),而 单模态WSI模型在IDH1突变预测任务中表现出相 反趋势,以IDH1突变预测任务为例,这一现象可能 源于T,-FLAIR 成像对微环境水肿信号的高敏感性, 但水肿信号本质上是非特异性炎症反应,与IDH1突 变缺乏关联,而IDH1野生型一般更具侵袭性生长模 式,局部炎症和水肿范围更大,导致模型易遗漏真 阳性病例而识别真阴性病例:WSI可识别 IDH1 突变 相关细胞形态学特征,但模型可能将一部分反应性 胶质增生或坏死周围细胞等误判为突变相关表型, 从而表现为高灵敏度和低特异度。未来可通过单 模态知识迁移等方法进一步优化灵敏度与特异度 的均衡性^[29-30],抑制多模态噪声;通过上下文感知生 成对抗网络^[31]等方法合成缺失的MRI模态,以改善 单模态 T,-FLAIR 不能完全反映分子边界的限制;通 过矫正病理学图像不均匀照明^[32],精细化WSI拼接 预处理等。

综上所述,本研究构建 DLFD-MIL模型,通过伪 包生成和最大最小选择法策略,有效解决 WSI 实例 标签缺失与多模态噪声干扰问题,实现 MRI 与病理 切片宏观与微观影像学特征的高效融合,提高胶质 瘤 *IDH1* 基因分型和 WHO 分级的预测精度。通过术 前快速预测、术中精准导航及术后动态监测的闭环 式应用路径,有望为胶质瘤全周期精准诊疗体系的 建立提供可落地的人工智能解决方案。 利益冲突 无

参考文献

- [1] Price M, Ballard C, Benedetti J, Neff C, Cioffi G, Waite KA, Kruchko C, Barnholtz-Sloan JS, Ostrom QT. CBTRUS statistical report: primary brain and other central nervous system tumors diagnosed in the United States in 2017-2021[J]. Neuro Oncol, 2024, 26(Supplement_6): vi 1-vi 85.
- [2] Schaff LR, Mellinghoff IK. Glioblastoma and other primary brain malignancies in adults: a review[J]. JAMA, 2023, 329: 574-587.
- [3] Louis DN, Perry A, Wesseling P, Brat DJ, Cree IA, Figarella-Branger D, Hawkins C, Ng HK, Pfister SM, Reifenberger G, Soffietti R, von Deimling A, Ellison DW. The 2021 WHO Classification of Tumors of the Central Nervous System: a summary[J]. Neuro Oncol, 2021, 23:1231-1251.
- [4] Horbinski C, Berger T, Packer RJ, Wen PY. Clinical implications of the 2021 edition of the WHO classification of central nervous system tumours [J]. Nat Rev Neurol, 2022, 18: 515-529.
- [5] Weller M, van den Bent M, Preusser M, Le Rhun E, Tonn JC, Minniti G, Bendszus M, Balana C, Chinot O, Dirven L, French P, Hegi ME, Jakola AS, Platten M, Roth P, Rudà R, Short S, Smits M, Taphoorn MJB, von Deimling A, Westphal M, Soffietti R, Reifenberger G, Wick W. EANO guidelines on the diagnosis and treatment of diffuse gliomas of adulthood [J]. Nat Rev Clin Oncol, 2021, 18:170-186.
- [6] Makino K, Nakamura H, Hide T, Yano S, Kuroda J, Takahashi Y, Iyama K, Kuratsu J. Usefulness of immunohistochemical expression analysis of metabolic - related molecules in differentiating between intracranial neoplastic and non neoplastic lesions[J]. Brain Tumor Pathol, 2013, 30:144-150.
- [7] Brandes AA, Tosoni A, Franceschi E, Sotti G, Frezza G, Amistà P, Morandi L, Spagnolli F, Ermani M. Recurrence pattern after temozolomide concomitant with and adjuvant to radiotherapy in newly diagnosed patients with glioblastoma: correlation with MGMT promoter methylation status[J]. J Clin Oncol, 2009, 27: 1275-1279.
- [8] Hu ZY, Wei W, Hu WZ, Ma MH, Li Y, Wu XS, Yin H, Xi YB. Prediction of IDH mutations in glioma based on MRI multiparametric image fusion and DenseNet network [J]. Ci Gong Zhen Cheng Xiang, 2023, 14:10-17.[胡振远,魏炜,胡文 鐘,马梦航,李艳,吴旭莎,印弘,席一斌.基于磁共振多参数 图像融合及 DenseNet 网络的脑胶质瘤 IDH 突变预测研究[J]. 磁共振成像, 2023, 14:10-17.]
- [9] Liu ZP, Jiang JX, Wu QW, Zhou Y, Bian XF, Zhu YX. Prediction of glioma grading based on multi-sequence MRI radiomics and deep transfer learning features [J]. Nanjing Yi Ke Da Xue Xue Bao (Zi Ran Ke Xue Ban), 2024, 44:372-379. [刘志鹏, 降建新, 吴琪炜, 周炎, 卞雪峰, 朱银杏. 基于多序列 MRI影像组学与深度迁移学习特征的脑胶质瘤分级预测研究 [J]. 南京医科大学学报(自然科学版), 2024, 44:372-379.]
- [10] Xue H, Han Z, Li H, Li X, Jia D, Qi M, Zhang H, Zhang K, Gong J, Wang H, Feng Z, Ni S, Han B, Li G. Application of intraoperative rapid molecular diagnosis in precision surgery for glioma: mimic the world health organization CNS5 integrated diagnosis[J]. Neurosurgery, 2023, 92:762-771.
- [11] Zhang H, Meng Y, Zhao Y, Qiao Y, Yang X, Coupland SE,

Zheng Y. DTFD-MIL: double-tier feature distillation multiple instance learning for histopathology whole slide image classification [C]//2022 IEEE/CVF conference on computer vision and pattern recognition (CVPR), New Orleans: IEEE, 2022: 18780-18790.

- [12] Ilse M, Tomczak JM, Welling M. Attention-based deep multiple instance learning [J]. Proc Mach Learn Res, 2018, 80:2127 -2136.
- [13] Zhou B, Khosla A, Lapedriza A, Oliva A, Torralba A. Learning deep features for discriminative localization [C]//2016 IEEE conference on computer vision and pattern recognition (CVPR), Las Vegas: IEEE, 2016: 2921-2929.
- [14] Hou L, Samaras D, Kurc TM, Gao Y, Davis JE, Saltz JH. Patchbased convolutional neural network for whole slide tissue image classification [J]. Proc IEEE Comput Soc Conf Comput Vis Pattern Recognit, 2016:2424-2433.
- [15] Xu G, Song Z, Sun Z, Ku C, Yang Z, Liu C, Wang S, Ma J, Xu W. CAMEL: aweakly supervised learning framework for histopathology image segmentation [C]//2019 IEEE/CVF international conference on computer vision (ICCV), Seoul: IEEE, 2019: 10681-10690.
- [16] Campanella G, Hanna MG, Geneslaw L, Miraflor A, Werneck Krauss Silva V, Busam KJ, Brogi E, Reuter VE, Klimstra DS, Fuchs TJ. Clinical-grade computational pathology using weakly supervised deep learning on whole slide images [J]. Nat Med, 2019, 25:1301-1309.
- [17] Kanavati F, Toyokawa G, Momosaki S, Rambeau M, Kozuma Y, Shoji F, Yamazaki K, Takeo S, Iizuka O, Tsuneki M. Weaklysupervised learning for lung carcinoma classification using deep learning[J]. Sci Rep, 2020, 10:9297.
- [18] Sharma Y, Shrivastava A, Ehsan L, Moskaluk CA, Syed S, Brown DE. Cluster-to-conquer: a framework for end-to-end multi-instance learning for whole slide image classification [J]. Proc Mach Learn Res, 2021, 143:682-698.
- [19] Shao Z, Bian H, Chen Y, Wang Y, Zhang J, Ji X, Zhang Y. TransMIL: transformer based correlated multiple instance learning for whole slide image classification [J]. Adv Neural Inform Process Syst, 2021, 34:2136-2147.
- [20] Li B, Li Y, Eliceiri KW. Dual-stream multiple instance learning network for whole slide image classification with self-supervised contrastive learning [J]. Conf Comput Vis Pattern Recognit Workshops, 2021:14318-14328.
- [21] Hsu WW, Guo JM, Pei L, Chiang LA, Li YF, Hsiao JC, Colen R, Liu P. A weakly supervised deep learning-based method for glioma subtype classification using WSI and mpMRIs [J]. Sci Rep, 2022, 12:6111.
- [22] Despotovic V, Kim SY, Hau AC, Kakoichankava A, Klamminger GG, Borgmann FBK, Frauenknecht KBM, Mittelbronn M, Nazarov PV. Glioma subtype classification from histopathological images using in - domain and out - of - domain

transfer learning: an experimental study[J]. Heliyon, 2024, 10: e27515.

- [23] Chen Y, Yang Z, Zhao J, Adamson J, Sheng Y, Yin FF, Wang C. A radiomics-incorporated deep ensemble learning model for multi-parametric MRI-based glioma segmentation [J]. Phys Med Biol, 2023:ID37586382.
- [24] Xie X, Shen C, Zhang X, Wu G, Yang B, Qi Z, Tang Q, Wang Y, Ding H, Shi Z, Yu J. Rapid intraoperative multi-molecular diagnosis of glioma with ultrasound radio frequency signals and deep learning[J]. EBioMedicine, 2023, 98:104899.
- [25] Alongi P, Arnone A, Vultaggio V, Fraternali A, Versari A, Casali C, Arnone G, DiMeco F, Vetrano IG. Artificial intelligence analysis using MRI and PET imaging in gliomas: a narrative review[J]. Cancers (Basel), 2024, 16:407.
- [26] Paprottka KJ, Kleiner S, Preibisch C, Kofler F, Schmidt-Graf F, Delbridge C, Bernhardt D, Combs SE, Gempt J, Meyer B, Zimmer C, Menze BH, Yakushev I, Kirschke JS, Wiestler B. Fully automated analysis combining [¹⁸F] - FET - PET and multiparametric MRI including DSC perfusion and APTw imaging: a promising tool for objective evaluation of glioma progression [J]. Eur J Nucl Med Mol Imaging, 2021, 48:4445-4455.
- [27] Yuan Y, Li G, Mei S, Hu M, Chu YH, Hsu YC, Li C, Song J, Hu J, Feng D, Xie F, Guan Y, Yue Q, Liu M, Mao Y. Deep mutual learning on hybrid amino acid PET predicts H3K27M mutations in midline gliomas [J]. NPJ Precis Oncol, 2024, 8: 274.
- [28] Zadeh A, Chen M, Poria S, Cambria E, Morency LP. Tensor fusion network for multimodal sentiment analysis [C]//Palmer M, Hwa R, Riedel S. Proceedings of the 2017 conference on empirical methods in natural language processing, Copenhagen: Association for Computational, 2017: 1103-1114.
- [29] Xing X, Chen Z, Hou Y, Yuan Y. Gradient modulated contrastive distillation of low-rank multi-modal knowledge for disease diagnosis[J]. Med Image Anal, 2023, 88:102874.
- [30] Xing X, Zhu M, Chen Z, Yuan Y. Comprehensive learning and adaptive teaching: distilling multi - modal knowledge for pathological glioma grading [J]. Med Image Anal, 2024, 91: 102990.
- [31] Huang P, Li D, Jiao Z, Wei D, Cao B, Mo Z, Wang Q, Zhang H, Shen D. Common feature learning for brain tumor MRI synthesis by context-aware generative adversarial network [J]. Med Image Anal, 2022, 79:102472.
- [32] Wang J, Wang X, Zhang P, Xie S, Fu S, Li Y, Han H. Correction of uneven illumination in color microscopic image based on fully convolutional network[J]. Opt Express, 2021, 29: 28503-28520.

(收稿日期:2025-02-13) (本文编辑:袁云)

下期内容预告 本刊2025年第4期报道专题为内镜颅底外科,重点内容包括:颅底沟通骨性病变诊断与治疗进展;经鼻内镜极 内侧入路切除下斜坡侵犯枕髁及咽旁间隙脊索瘤;内镜下经腹侧自然腔道切除侧颅底病变;经鼻内镜手术治疗自发性蝶窦外 侧隐窝脑脊液鼻漏;内镜下经鼻蝶入路脑脊液鼻漏修补术在基层医院的应用价值探讨;梗死性垂体腺瘤卒中三例并文献复习