



· 论 著 ·

基于MRI的影像组学和深度学习模型构建：无创鉴别原发颅内弥漫大B细胞淋巴瘤分子亚型

曾延玮^{1, 2, 3}, 徐智坚⁴, 曹 鑫^{1, 2, 3}, 吕 锐^{1, 2, 3}, 李惠明⁵, 高 敏⁶, 居胜红⁵,

刘 军⁶, 耿道颖^{1, 2, 3}

1. 复旦大学附属华山医院放射科, 上海 200040;
2. 上海脑重大疾病智能影像工程技术研究中心, 上海 200040;
3. 复旦大学医学功能与分子影像研究所, 上海 200040;
4. 复旦大学工程与应用技术研究院, 上海 200433;
5. 东南大学附属中大医院放射科, 江苏 南京 210009;
6. 中南大学湘雅二医院放射科, 湖南 长沙 410011

[摘要] 背景与目的: 弥漫大B细胞淋巴瘤 (diffuse large B-cell lymphoma, DLBCL) 的生发中心B细胞样 (germinal center B-cell-like, GCB) 亚型和非GCB (non-GCB) 亚型在患者预后和治疗上存在差异, 但目前依赖有创病理学检查。本研究基于多参数MRI构建影像组学和深度学习模型, 旨在术前无创性区分这两种亚型。方法: 本研究回顾性分析2013年3月—2024年12月在复旦大学附属华山医院及外院经病理学检查确诊的DLBCL患者。使用多参数MRI扫描数据, 结合4种影像组学机器学习 [支持向量机 (support vector machine, SVM)、逻辑回归 (logistic regression, LR)、高斯过程 (Gaussian process, GP) 和朴素贝叶斯 (Naive Bayes, NB)] 和3种深度学习 [密集连接卷积网络121 (densely-connected convolutional networks 121, DenseNet121)、残差网络101 (residual network 101, ResNet101) 和高效网络B5 (EfficientNet-b5)] 建立DLBCL亚型分类模型。此外, 两名经验不同的放射科医师在盲法下基于MRI图像独立分类DLBCL。模型和医师的诊断性能均通过接收者操作特征曲线下面积 (area under the curve, AUC)、准确度 (accuracy, ACC) 和F1分数 (F1-score, F1) 等指标进行量化评估, 以衡量其区分GCB和non-GCB亚型的能力。本研究经复旦大学附属华山医院伦理委员会批准 (KY2024-663), 所有患者均知情同意。结果: 本研究共纳入173例患者 (GCB型55例, non-GCB型118例)。影像组学和深度学习方法能有效地区分DLBCL亚型。其中, GP影像组学模型 (基于T1-CE+T2-FLAIR+ADC序列) 和DenseNet121深度学习模型 (基于T1-CE+T2-FLAIR+ADC序列) 表现最佳, 在内部验证集上分别取得优异性能 (GP: AUC=0.900, ACC=0.896, F1=0.840; DenseNet121: AUC=0.846, ACC=0.854, F1=0.774), 并在外部验证集上保持稳健。并且, 最优AI模型的分类型效能优于经验丰富的放射科医师 (医师最高AUC=0.678)。结论: 基于多参数MRI特征的影像组学与深度学习模型可有效地鉴别DLBCL的GCB与non-GCB亚型。其中, GP与DenseNet121模型在处理复杂图像数据、特别是融合多序列特征组进行亚型分类时, 呈现出优异的性能。

[关键词] 弥漫大B细胞淋巴瘤; 生发中心B细胞样; 非GCB; 影像组学; 深度学习

中图分类号: R733.4 文献标志码: A
DOI: 10.19401/j.cnki.1007-3639.2025.08.001

基金项目: 国家自然科学基金 (82372048); 上海市科学技术委员会项目 (23S31904100)。

利益冲突: 作者声明无利益冲突。

伦理批件: KY2024-663。

知情同意: 已获得。

引用本文: 曾延玮, 徐智坚, 曹 鑫, 等. 基于MRI的影像组学和深度学习模型构建: 无创鉴别原发颅内弥漫大B细胞淋巴瘤分子亚型 [J]. 中国癌症杂志, 2025, 35(8): 735-742.

Funding: National Natural Science Foundation of China (82372048). Science and Technology Commission of Shanghai Municipality Fund (23S31904100).

Conflicts of interest: authors declare no conflicts of interest.

Ethical approval: KY2024-663.

Informed consent: acquired.

Cite this article: ZENG Y W, XU Z J, CAO X, et al. MRI-based radiomics and deep learning model construction: non-invasive differentiation of molecular subtypes in primary intracranial diffuse large B-cell lymphoma [J]. Chin Oncol, 2025, 35(8): 735-742.

MRI-based radiomics and deep learning model construction: non-invasive differentiation of molecular subtypes in primary intracranial diffuse large B-cell lymphoma ZENG Yanwei^{1,2,3}, XU Zhijian⁴, CAO Xin^{1,2,3}, LÜ Kun^{1,2,3}, LI Huiming⁵, GAO Min⁶, JU Shenghong⁵, LIU Jun⁶, GENG Daoying^{1,2,3} (1. Department of Radiology, Huashan

Hospital, Fudan University, Shanghai 200040, China; 2. Shanghai Engineering Research Center of Intelligent Imaging for Critical Brain Diseases, Shanghai 200040, China; 3. Institute of Functional and Molecular Medical Imaging, Fudan University, Shanghai 200040, China; 4. Academy for Engineering and Technology, Fudan University, Shanghai 200433, China; 5. Department of Radiology, Zhongda Hospital, Southeast University, Nanjing 210009, Jiangsu Province, China; 6. Department of Radiology, Second Xiangya Hospital, Central South University, Changsha 410011, Hunan Province, China)

Correspondence to: GENG Daoying E-mail: gdy_2019@163.com

[**Abstract**] **Background and purpose:** Diffuse large B-cell lymphoma (DLBCL) is subclassified into germinal center B-cell-like (GCB) and non-GCB subtypes, which differ in prognosis and treatment response. However, current distinction still relies on invasive pathological assays. This study developed radiomics and deep-learning models based on multiparametric magnetic resonance imaging (MRI) to non-invasively differentiate the two subtypes preoperatively, thereby reducing dependence on histopathological examination. **Methods:** This study retrospectively included patients with pathologically confirmed DLBCL diagnosed at Huashan Hospital, Fudan University, and other institutions between March 2013 and December 2024. Using multiparametric MRI data, we developed DLBCL-subtype classification models that combined 4 radiomics-based machine-learning algorithms: support vector machine (SVM), logistic regression (LR), Gaussian process (GP) and Naive Bayes (NB), with 3 deep-learning architectures [densely-connected convolutional networks 121 (DenseNet121), residual network 101 (ResNet101) and EfficientNet-b5]. Additionally, two radiologists with different experience levels independently classified DLBCL on MRI in a blinded fashion. Model and radiologist performance were quantified using the area under the receiver operating characteristic curve (AUC), accuracy (ACC), and F1-score to evaluate their ability to distinguish GCB from non-GCB subtypes. This study was approved by the Ethics Committee of Huashan Hospital of Fudan University (No. KY2024-663), and all patients signed informed consents. **Results:** A total of 173 patients were enrolled (55 with GCB subtype and 118 with non-GCB subtype). Radiomics and deep learning methods effectively distinguished DLBCL subtypes. Among these, the GP radiomics model (based on T1-CE+T2-FLAIR+ADC sequences) and DenseNet121 deep learning model (based on T1-CE+T2-FLAIR+ADC sequences) demonstrated optimal performance. Both achieved excellent results on the internal validation set (GP: AUC=0.900, ACC=0.896, F1=0.840; DenseNet121: AUC=0.846, ACC=0.854, F1=0.774) and maintained robustness on the external validation set. Furthermore, the classification efficacy of the optimal AI model surpassed that of experienced radiologists (highest physician AUC=0.678). **Conclusion:** Radiomics and deep-learning models based on multiparametric MRI features can effectively differentiate GCB from non-GCB subtypes of DLBCL. Among them, GP and DenseNet121 exhibit outstanding performance, especially when integrating multi-sequence feature sets for classifying DLBCL subtypes on complex imaging data.

[**Key words**] Diffuse large B-cell lymphoma; Germinal center B-cell-like; Non-GCB; Radiomics; Deep learning

弥漫大B细胞淋巴瘤 (diffuse large B-cell lymphoma, DLBCL) 占原发性中枢神经系统淋巴瘤 (primary central nervous system lymphoma, PCNSL) 的95%以上^[1-2]。虽然基因表达谱 (gene expression profiling, GEP) 是精确的分型方法, 但因技术复杂、成本高而造成临床应用受限。目前临床上主要采用Hans分类法^[3], 根据细胞起源 (cell of origin, COO) 的差异将DLBCL分为两种亚型: 生发中心B细胞样 (germinal center B-cell-like, GCB) 型和非生发中心B细胞样 (non-GCB) 型。研究^[4-6]表明, non-GCB型更具侵袭性, 对于标准化疗方案的反应较差, 患者预后显著劣于GCB型。

免疫组织化学检测因需侵入性获取脑部肿瘤组织样本, 会增加患者痛苦与风险, 还可能引发采样不准确及损伤健康脑组织等问题。在我们既往的研究^[7]中发现, 多参数MRI中GCB和non-GCB亚型的部分图像特征存在显著差异, 但传统影像学诊断易受医师经验影响。随着影像组

学和深度学习技术的快速发展, 研究人员开始探索利用这些技术从MRI中提取定量特征, 以实现DLBCL的非侵入性分型。影像组学通过从MRI中提取高通量图像特征^[8], 已在区分PCNSL与其他脑肿瘤方面展现出潜力^[9], 并被用于评估PCNSL的Ki-67增殖指数和双重表达状态等^[10-11]。深度学习则无需手动选取特征, 在医学图像分析与处理领域广泛应用, 已在提升PCNSL术前诊断的准确性和效率方面展现出显著的应用潜力^[12]。

本研究基于DLBCL的GCB型和non-GCB型的多参数MRI特征构建影像组学和深度学习模型, 并评估其在术前鉴别DLBCL亚型中的应用价值, 以提高影像分型DLBCL的准确率和效率。

1 资料和方法

1.1 研究对象

本研究回顾性纳入2013年3月—2024年12月在复旦大学附属华山医院经病理学检查确诊的

DLBCL患者。患者纳入标准为：① 年龄 ≥ 18 岁；② 病理学检查确诊为DLBCL；③ 具有完整的CD10、BCL-6和MUM1免疫组织化学染色结果，可明确分为GCB型或non-GCB型；④ 无先天性或获得性免疫缺陷。排除标准包括：① MRI影像资料不完整，包括T1加权对比增强成像（T1-weighted contrast-enhanced imaging, T1-CE）、T2加权液体衰减反转恢复序列（T2-weighted fluid-attenuated inversion recovery, T2-FLAIR）和表观弥散系数（apparent diffusion coefficient, ADC）图；② 在手术、化疗、放疗或使用类固醇后进行的MRI扫描；③ MRI图像质量差；④ 临床资料不完善。外部验证患者来自同时期内东南大学附属中大医院和中南大学附属湘雅二医院，纳入和排除标准相同。本研究经复旦大学附属华山医院伦理委员会批准（KY2024-663），所有患者均知情同意。

1.2 仪器和材料

在本研究中，MRI图像由5台配备8通道头部线圈的3T MR扫描仪（DISCOVERY MR750w、Verio、Vantage Titan、uMR790和Ingenia）采集，扫描范围覆盖整个大脑。扫描序列包括轴位T1-CE、T2-FLAIR和DWI，基于DWI扫描自动生成ADC图。T1-CE扫描前，通过肘静脉注射钆双胺作为对比剂，剂量为0.1 mmol/kg，注射后立即启动T1-CE扫描，并用20 mL生理盐水冲洗以确保对比剂完全注入并减少残留。

1.3 肿瘤分割

将原始DICOM图像数据转换为NIfTI格式后，运用ANTs软件包对ADC图与T2-FLAIR、T1-CE进行空间配准，并对每幅图像进行N4偏场校正以校正强度不均匀性。随后计算平均值和标准差，采用z分数方法对所有序列的MRI图像进行归一化。两名经验丰富的放射科医师（分别有7年和11年的神经影像学诊断经验）使用ITK-SNAP软件（v3.8.0）以盲法手动勾画肿瘤及其周围水肿，创建感兴趣区域（region of interest, ROI）。其中，在T1-CE图像上勾画明显强化的肿瘤病变作为ROI，T2-FLAIR图像上勾画最大异常区域（涵盖肿瘤病变和瘤周水肿）作为ROI。两名医师分割结果的重叠部分作为最终ROI（图1）。

1.4 特征提取

使用Pyradiomics（<https://pyradiomics.readthedocs.io/en/latest/index.html>）开源Python包进行特征提取。从原始图像中提取了14个形状特征、18个一阶直方图特征和68个纹理特征，以及13个额外滤波图像（包括平方、平方根、指数、对数、高斯拉普拉斯和8个小波滤波器）上的相同数量的一阶和纹理特征。

1.5 特征选择

使用scikit-learn包（<https://scikit-learn.org/stable/>）进行特征选择，步骤如下：① 排除类内相关系数（intraclass correlation coefficient, ICC）低于0.8的特征，确保高可重复性；② 应

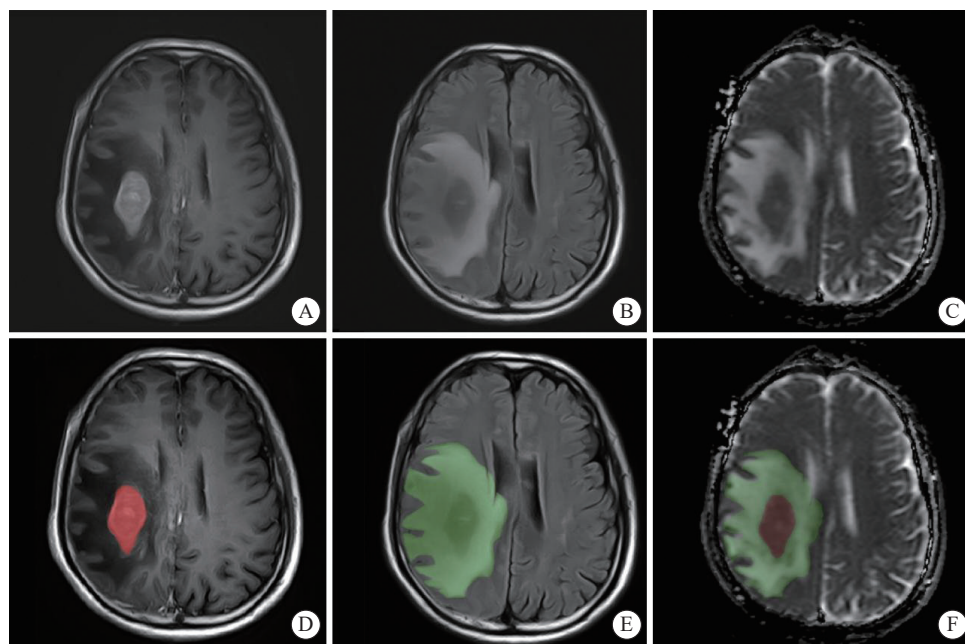


图1 ROI示意图

Fig. 1 Schematic diagram of ROI

A: T1-CE image; B: T2-FLAIR image; C: ADC map; D: Red area, tumor ROI; E: Green area, ROI within edema region; F: Tumor ROI and edema ROI mapped onto the ADC map.

用最小冗余最大相关性 (minimum redundancy maximum relevance, mRMR) 方法, 将剩余特征数量减少到病变数量的20%以下, 最小化过拟合风险; ③ 使用最小绝对收缩和选择算子 (least absolute shrinkage and selection operator, LASSO) 方法进一步优化特征, 提高分类准确性和可解释性。

1.6 分类模型构建

本研究整合T1-CE、T2-FLAIR和ADC共3种MR序列, 构建7种组合模式, 包括单一序列、双序列组合及三序列联合。基于筛选后的影像组学特征, 使用scikit-learn工具包中的4种分类器支持向量机 (support vector machine, SVM)、逻辑回归 (logistic regression, LR)、高斯过程 (Gaussian process, GP) 和朴素贝叶斯 (Naive bayes, NB) 进行训练。同时构建密集连接卷积网络121 (densely-connected convolutional networks 121, DenseNet121)、残差网络101 (residual network 101, ResNet101) 和高效网络B5 (EfficientNet-b5) 3种深度学习网络架构。训练在配备Intel Core i9-10900X和Nvidia Geforce 3090 GPU的Linux工作站上完成, 采用Adam优化器, 初始学习率为0.001。训练至少100个epoch, 当训练集loss值低于 $1e^{-3}$ 且连续10个epoch波动小于 $1e^{-5}$ 时停止训练, 以确保模型稳定优化^[13]。

1.7 放射科医师诊断

两名放射科医师 (分别具备7年及11年中枢神经系统影像学诊断经验) 在未知病理学诊断结果的前提下, 对纳入内部验证集的所有病例MRI图像进行了独立判读, 并将其分类为GCB型与非-GCB型。

1.8 评价指标

本研究采用以下指标评估模型和放射科医师的诊断性能: 受试者工作特征曲线下面积 (area under the receiver operating characteristic curve, AUC)、准确度 (accuracy, ACC)、敏感性 (sensitivity, SEN)、特异性 (specificity, SPE)、阳性预测值 (positive predictive value, PPV)、阴性预测值 (negative predictive value, NPV) 和F1分数 (F1-score)。AUC衡量模型的整体区分能力, ACC反映分类正确的总体比例, SEN和SPE分别揭示模型对正类和负类的识别能力, PPV和NPV表示预测结果的可信度, 而F1-score则在样本不均衡时综合评估模型对某一亚型的精确性与召回率的平衡。所有指标均针对二分类任务计算。

1.9 统计学处理

连续变量经Shapiro-Wilk检验评估正态性, 符合正态分布且方差齐性者采用独立样本 t 检验, 否则采用Mann-Whitney U 检验。分类变量组间比较, 对于独立样本, 采用 χ^2 检验 (当总样本量 $n \geq 40$ 且所有理论频数 $T \geq 5$ 时) 或Fisher精确概率法 (当总样本量 $n \geq 40$ 且任一理论频数 $1 \leq T < 5$ 时; 当 $n < 40$ 或 $T < 1$ 时)。所有统计检验均为双侧检验, $P < 0.05$ 为差异有统计学意义。采用SPSS 29.0软件进行统计学分析。

2 结果

2.1 临床资料

根据纳排标准, 本研究最终纳入了173例DLBCL患者 (图2)。其中, 160例来自复旦大学附属华山医院的患者, 按7:3比例随机分配至训练集 ($n=112$) 和内部验证集 ($n=48$); 另外13例来自东南大学附属中大医院和中南大学湘雅二医院的患者, 作为外部验证集。患者的临床信息见表1。组间比较显示, 训练集与内部验证集在患者年龄、性别及病理学分型上差异均无统计学意义 ($P > 0.05$); 而训练集与外部验证集在性别分布上差异有统计学意义 ($P < 0.05$), 但在年龄和病理学分型上差异无统计学意义 ($P > 0.05$)。

2.2 影像组学特征筛选

采用影像组学方法从每例患者的影像资料中提取了8 526个特征 (7个MR序列组 \times 1 218个/组)。各序列组经特征选择后保留数量为: T1-CE ($n=3$)、T2-FLAIR ($n=2$)、ADC ($n=10$)、T1-CE+T2-FLAIR ($n=5$)、T1-CE+ADC ($n=12$)、T2-FLAIR+ADC ($n=12$)、T1-CE+T2-FLAIR+ADC ($n=13$)。所有分类器均通过五折交叉验证并计算指标平均值。

2.3 影像组学模型诊断效能评估

影像组学模型在训练集上均展现出良好的分类能力, 4种模型平均AUC: 0.858~0.916。在内部验证集中, 单序列特征组中ADC特征组表现较为出色, GP模型AUC达0.836, 但NB模型的ACC和F1-score更高 (0.833和0.734)。多序列特征组中, T1-CE+T2-FLAIR+ADC特征组表现最优, GP模型AUC为0.900, ACC为0.896, F1-score为0.840。在外部验证集中, 单序列特征组中ADC特征组再次表现突出, NB模型AUC为0.725, 但LR模型的ACC和F1-score更高 (0.833和0.769)。多序列特征组中, T1-CE+T2-

FLAIR+ADC特征组仍表现最佳，NB模型AUC为0.852，ACC为0.833，F1-score为0.800；GP模型的ACC和F1-score则更高，分别为0.889和0.833（表2）。

2.4 深度学习模型诊断效能评估

基于优化后的深度学习模型。内部验证显示：单序列特征组中，基于ADC序列的EfficientNet-b5模型表现最佳，AUC为0.834，ACC为0.792，F1-score为0.706；多序列特征组中，T1-CE+T2-FLAIR+ADC特征组的DenseNet121模型性能综合较佳，AUC为0.846，ACC为0.854，F1-score为0.774。外部验证显示，单序列特征组中，基于ADC序列的DenseNet121模型表现出色，AUC为0.723，ACC为0.833，

F1-score为0.800；多序列特征组中，T1-CE+T2-FLAIR+ADC特征组的DenseNet121模型继续领先，AUC为0.807，ACC为0.889，F1-score为0.875（表2）。

2.5 放射科医师的诊断结果

在内部验证集中，两位放射科医师的诊断结果对比如下：较少工作经验的医师诊断结果对应的AUC为0.589，ACC为0.583，SEN为0.611，SPE为0.567，PPV为0.458，NPV为0.798，F1-score为0.524；较多工作经验的医师诊断结果对应的AUC为0.678，ACC为0.667，SEN为0.722，SPE为0.633，PPV为0.542，NPV为0.792，F1-score为0.619。较多经验的医师在所有诊断指标上均优于较少经验的医师。

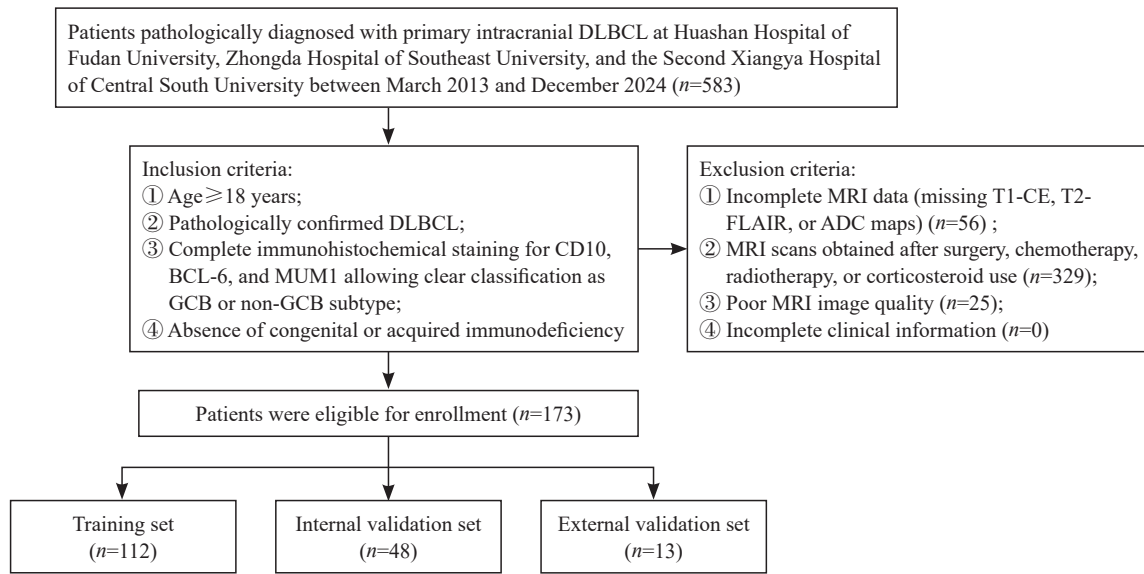


图2 患者纳入和排除标准的流程图

Fig. 2 Flow diagram of patient inclusion and exclusion criteria

DLBCL: Diffuse large B-cell lymphoma; GCB: Germinal center B-cell-like; non-GCB: Non-germinal center B-cell-like; T1-CE: Contrast-enhanced T1-weighted imaging; T2-FLAIR: T2-weighted fluid-attenuated inversion recovery; ADC: Apparent diffusion coefficient.

表1 原发颅内DLBCL患者临床信息

Tab. 1 Clinical characteristics of patients with primary intracranial DLBCL

Baseline characteristics	Training set (n=112)	Internal validation set (n=48)	P value	External validation set (n=13)	P value*
Age/year	59.06 ± 13.43	60.46 ± 10.79	0.174	64.62 ± 8.81	0.148
Gender			1.000		0.049
Male	63 (56.2)	27 (56.3)		11 (84.6)	
Female	49 (43.8)	21 (43.7)		2 (15.4)	
Pathology			0.971		0.194
GCB	37 (33.0)	16 (33.3)		2 (15.4)	
Non-GCB	75 (67.0)	32 (66.7)		11 (84.6)	

*: Training set compared with external validation set; Continuous variables were presented as $\bar{x} \pm s$; Categorical variables were expressed as n(%); DLBCL: Diffuse large B-cell lymphoma; GCB: Germinal center B-cell-like; Non-GCB: Non-germinal center B-cell-like.

表2 影像组学和机器学习模型在内外验证集中的表现

Tab. 2 Performance of radiomics models in the internal and external validation set

Sequence	Model	Internal validation set [AUC (95% CI)/ACC/F1]	External validation set [AUC (95% CI)/ACC/F1]
T1-CE+T2-FLAIR+ADC	SVM	0.887 (0.789-0.984)/0.833/0.778	0.822 (0.700-0.944)/0.889/0.833
	LR	0.878 (0.771-0.987)/0.854/0.800	0.840 (0.718-0.960)/0.833/0.769
	GP	0.900 (0.805-0.906)/0.896/0.840	0.834 (0.710-0.958)/0.889/0.833
	NB	0.805 (0.678-0.932)/0.813/0.743	0.852 (0.738-0.965)/0.833/0.800
	DenseNet121	0.846 (0.710-0.982)/0.854/0.774	0.807 (0.672-0.941)/0.889/0.875
	ResNet101	0.842 (0.701-0.983)/0.833/0.750	0.826 (0.706-0.947)/0.833/0.800
	EfficientNet-b5	0.802 (0.674-0.929)/0.771/0.703	0.802 (0.673-0.937)/0.889/0.833
T1-CE+T2-FLAIR	SVM	0.709 (0.562-0.856)/0.646/0.622	0.703 (0.552-0.854)/0.722/0.615
	LR	0.719 (0.572-0.866)/0.667/0.636	0.703 (0.552-0.854)/0.778/0.714
	GP	0.719 (0.574-0.863)/0.667/0.652	0.678 (0.517-0.839)/0.611/0.588
	NB	0.688 (0.529-0.846)/0.708/0.632	0.717 (0.571-0.863)/0.556/0.556
	DenseNet121	0.713 (0.565-0.860)/0.646/0.622	0.691 (0.526-0.856)/0.667/0.666
	ResNet101	0.723 (0.576-0.869)/0.729/0.629	0.738 (0.585-0.892)/0.611/0.632
	EfficientNet-b5	0.725 (0.582-0.867)/0.688/0.651	0.732 (0.578-0.886)/0.611/0.364
T1-CE+ADC	SVM	0.793 (0.660-0.926)/0.750/0.700	0.562 (0.375-0.750)/0.778/0.666
	LR	0.840 (0.718-0.962)/0.771/0.718	0.615 (0.440-0.790)/0.778/0.714
	GP	0.844 (0.726-0.961)/0.792/0.750	0.547 (0.359-0.734)/0.611/0.588
	NB	0.820 (0.696-0.945)/0.792/0.722	0.572 (0.394-0.750)/0.556/0.556
	DenseNet121	0.822 (0.700-0.944)/0.771/0.718	0.531 (0.344-0.718)/0.778/0.666
	ResNet101	0.840 (0.718-0.962)/0.771/0.718	0.578 (0.396-0.760)/0.611/0.364
	EfficientNet-b5	0.852 (0.738-0.965)/0.771/0.732	0.611 (0.438-0.785)/0.667/0.625
T2-FLAIR+ADC	SVM	0.822 (0.700-0.944)/0.771/0.718	0.709 (0.546-0.872)/0.722/0.635
	LR	0.840 (0.718-0.962)/0.771/0.718	0.537 (0.359-0.716)/0.667/0.571
	GP	0.852 (0.738-0.965)/0.771/0.732	0.525 (0.336-0.715)/0.611/0.588
	NB	0.834 (0.710-0.958)/0.833/0.765	0.740 (0.585-0.895)/0.778/0.714
	DenseNet121	0.834 (0.710-0.958)/0.833/0.765	0.568 (0.392-0.745)/0.833/0.800
	ResNet101	0.793 (0.660-0.926)/0.750/0.700	0.525 (0.339-0.711)/0.722/0.615
	EfficientNet-b5	0.844 (0.722-0.966)/0.771/0.718	0.570 (0.388-0.753)/0.722/0.615
T1-CE	SVM	0.527 (0.338-0.717)/0.667/0.385	0.527 (0.338-0.717)/0.444/0.445
	LR	0.562 (0.377-0.748)/0.625/0.471	0.562 (0.377-0.748)/0.500/0.470
	GP	0.549 (0.367-0.730)/0.646/0.452	0.529 (0.342-0.717)/0.556/0.556
	NB	0.529 (0.342-0.717)/0.667/0.385	0.549 (0.367-0.730)/0.444/0.375
	DenseNet121	0.721 (0.565-0.876)/0.708/0.632	0.721 (0.565-0.876)/0.611/0.364
	ResNet101	0.705 (0.541-0.869)/0.729/0.552	0.705 (0.542-0.869)/0.444/0.375
	EfficientNet-b5	0.684 (0.516-0.851)/0.708/0.611	0.684 (0.517-0.852)/0.667/0.625
T2-FLAIR	SVM	0.703 (0.552-0.854)/0.708/0.611	0.518 (0.329-0.708)/0.722/0.666
	LR	0.703 (0.552-0.854)/0.688/0.615	0.562 (0.377-0.748)/0.778/0.714
	GP	0.717 (0.571-0.863)/0.667/0.619	0.525 (0.336-0.715)/0.611/0.588
	NB	0.678 (0.517-0.839)/0.729/0.587	0.553 (0.372-0.733)/0.722/0.666
	DenseNet121	0.711 (0.563-0.859)/0.708/0.611	0.611 (0.438-0.785)/0.722/0.615
	ResNet101	0.703 (0.552-0.854)/0.688/0.615	0.539 (0.365-0.713)/0.667/0.625
	EfficientNet-b5	0.711 (0.560-0.861)/0.729/0.606	0.539 (0.366-0.713)/0.667/0.666
ADC	SVM	0.779 (0.622-0.936)/0.750/0.690	0.713 (0.565-0.860)/0.778/0.667
	LR	0.807 (0.666-0.947)/0.833/0.661	0.723 (0.576-0.869)/0.833/0.769
	GP	0.836 (0.712-0.959)/0.750/0.714	0.691 (0.536-0.847)/0.667/0.625
	NB	0.819 (0.690-0.948)/0.833/0.734	0.725 (0.582-0.867)/0.556/0.556
	DenseNet121	0.807 (0.674-0.934)/0.771/0.703	0.723 (0.563-0.882)/0.833/0.800
	ResNet101	0.770 (0.614-0.925)/0.729/0.683	0.691 (0.528-0.855)/0.833/0.769
	EfficientNet-b5	0.834 (0.714-0.954)/0.792/0.706	0.719 (0.559-0.879)/0.722/0.706

T1-CE: Contrast-enhanced T1-weighted imaging; T2-FLAIR: T2-weighted fluid-attenuated inversion recovery; ADC: Apparent diffusion coefficient; AUC: Area under the curve; ACC: Accuracy; F1: F1-score; SVM: Support vector machine; LR: Logistic regression; GP: Gaussian process; NB: Naive Bayes; DenseNet121: Densely connected convolutional networks 121; ResNet101: Residual network 101.

3 讨 论

本研究基于多参数MRI，通过49种“序列-分类器”组合筛选出最优影像组学（GP）与深度学习（DenseNet121）模型，为术前无创性区分DLBCL的GCB与non-GCB亚型提供了新策略。

结果显示，在49种“序列-模型”组合中，T1-CE+T2-FLAIR+ADC三序列联合的GP影像组学模型（内部验证AUC为0.900，ACC为0.896，F1-score为0.840；外部验证AUC为0.834，ACC为0.889，F1-score为0.833）与DenseNet121深度学习模型（内部验证AUC为0.846，ACC为0.854，F1-score为0.774；外部验证AUC为0.807，ACC为0.889，F1-score为0.875）均取得最佳表现，优于经验放射科医师（最高AUC为0.678）。与既往仅聚焦PCNSL与胶质瘤/转移瘤鉴别或Ki-67增殖指数、双重表达状态预测的研究不同^[11, 14-15]。本研究结果显示，多参数MRI可在术前无创区分DLBCL的GCB与non-GCB亚型。进一步分析发现：① 序列层面，ADC单序列虽优于T1-CE或T2-FLAIR单序列，但其AUC仍低于三序列联合，提示三序列互补信息至关重要；② 算法层面，GP模型在交叉验证中波动较小，显示其对高维小样本数据的稳健性，与其核函数灵活、可量化不确定性的特点一致^[16]；而NB虽在外部验证AUC较高，但ACC与F1略低于GP，提示其对数据分布偏移更敏感。值得注意的是，多发性硬化研究^[17]显示GP表现欠佳，再次说明模型效能受任务与数据集影响，需具体问题具体分析。

与手工勾画ROI的影像组学相比，深度学习直接以整幅图像为输入，理论上可捕获更广域的信息。结果部分显示，在三序列（T1-CE+T2-FLAIR+ADC）条件下，DenseNet121取得内部验证最高AUC为0.846，ACC为0.854，F1-score为0.774，且外部验证保持AUC为0.807，ACC为0.889，F1-score为0.875，均优于ResNet101与EfficientNet-b5，表明其密集连接结构在多模态数据中具备更强的特征整合能力。单序列比较中，以ADC为输入的DenseNet121亦表现最佳（内部验证AUC为0.807，外部验证AUC为0.723），与既往“ADC比值提示MYC/BCL2基因重排”^[18]一致，提示ADC对细胞密度及微观结构变化最敏感。T1-CE与T2-FLAIR则主要反映解剖边界及水肿范围；三序列结合后，DenseNet121在ACC与F1-score上的提升最为显著，说明其密集连接

网络不仅能减少梯度消失，还可高效复用跨序列特征，从而更全面地刻画DLBCL的病理学异质性^[19]。既往研究亦验证了DenseNet121在脑肿瘤分类中的优势。Mao等^[20]在2025年报道，融合扩张卷积与SE模块的Dilated SE-DenseNet在公开脑肿瘤MRI数据集上的AUC达0.94，显著优于传统CNN，进一步支持DenseNet121在神经系统肿瘤影像分析中的卓越性能。综上，“多参数MRI+GP/DenseNet121”可能作为DLBCL分子亚型无创分型的可靠方案，为后续确立个体化治疗策略提供客观依据。

本研究将两位放射科医师的诊断结果与最优模型进行了直接比较：经验较少者AUC为0.589，ACC为0.583，F1-score为0.524；经验较多者AUC为0.678，ACC为0.667，F1-score为0.619。均显著低于三序列GP影像组学模型（AUC为0.900，ACC为0.896，F1-score为0.840）和DenseNet121深度学习模型（AUC为0.846，ACC为0.854，F1-score为0.774）。这一差距提示，传统影像判读易受主观经验、视觉疲劳及认知偏差影响，而基于多参数MRI的影像组学与深度学习技术能够量化微观结构差异，提供客观、可重复的无创分型。因此，将上述AI工具嵌入临床工作流程，可在术前精准识别DLBCL分子亚型，减少不必要的侵入性活检，并为作出个体化治疗决策提供可靠依据。

本研究存在以下局限：样本量偏小且GCB比例低，外部验证仅13例，易导致过拟合及外推受限；回顾性单中心设计带来扫描参数、病例选择偏倚；ROI手工分割存在主观误差；不同设备/序列的扫描参数差异可能降低模型稳健性。需前瞻性多中心大样本临床试验进一步验证。

综上所述，影像组学和深度学习方法可以利用多参数MRI特征有效地区分原发颅内DLBCL的亚型，使优化治疗方案和改善患者预后成为可能。其中，GP和DenseNet121模型在处理复杂图像数据以实现DLBCL亚型的精准分类方面表现优异，特别是基于多序列特征组时。

第一作者：

曾延玮（ORCID: 0000-0002-0014-2661），博士研究生在读。

通信作者：

耿道颖（ORCID: 0000-0002-1707-1521），主任医师，博士研究生导师，E-mail: gdy_2019@163.com。

作者贡献：

曾延玮：实验设计，数据分析，文章撰写；徐智

坚: 数据分析, 文章撰写; 曹鑫: 数据分析; 吕锬: 数据分析; 李惠明: 外部验证队列数据采集; 高敏: 外部验证队列数据采集; 居胜红, 外部验证队列数据采集; 刘军: 外部验证队列数据采集; 耿道颖: 实验设计, 文章审核定稿。

[参 考 文 献]

- [1] SCHAFF L R, GROMMES C. Primary central nervous system lymphoma [J] . *Blood*, 2022, 140(9): 971–979.
- [2] 中华医学会血液学分会淋巴细胞疾病学组, 中国临床肿瘤学会 (CSCO) 淋巴瘤专家委员会. 原发性中枢神经系统淋巴瘤诊断及治疗专家共识 (2024年版) [J] . *白血病·淋巴瘤*, 2024, 33(3): 129–137.
- The Lymphocytic Disease Group of the Hematology Branch of the Chinese Medical Association and the Lymphoma Expert Committee of the Chinese Society of Clinical Oncology (CSCO) Expert. Consensus on diagnosis and treatment of primary central nervous system lymphoma (2024 edition) [J] . *J Leuk Lymph*, 2024, 33 (3): 129–137.
- [3] HANS C P, WEISENBURGER D D, GREINER T C, et al. Confirmation of the molecular classification of diffuse large B-cell lymphoma by immunohistochemistry using a tissue microarray [J] . *Blood*, 2004, 103(1): 275–282.
- [4] WILSON W H, WRIGHT G W, HUANG D W, et al. Effect of ibrutinib with R-CHOP chemotherapy in genetic subtypes of DLBCL [J] . *Cancer Cell*, 2021, 39(12): 1643–1653.e3.
- [5] MARCUS C, MARAGKOS G A, ALTERMAN R L, et al. GCB-type is a favorable prognostic factor in primary CNS diffuse large B-cell lymphomas [J] . *J Clin Neurosci*, 2021, 83: 49–55.
- [6] NIPARUCK P, BOONSAKAN P, SUTTHIPPINGKIAT T, et al. Treatment outcome and prognostic factors in PCNSL [J] . *Diagn Pathol*, 2019, 14(1): 56.
- [7] 曾延玮, 曹鑫, 吕锬, 等. 多参数磁共振成像鉴别中枢神经系统弥漫大B细胞淋巴瘤的亚型 [J] . *中国医学计算机成像杂志*, 2025, 31(2): 156–161.
- ZENG Y W, CAO X, LÜ K, et al. Non-invasive discrimination of cerebral diffuse large B-cell lymphoma subtypes using multi-parametric MRI [J] . *Chin Comput Med Imag*, 2025, 31(2): 156–161.
- [8] FERRARI R, TRINCI M, CASINELLI A, et al. Radiomics in radiology: what the radiologist needs to know about technical aspects and clinical impact [J] . *Radiol Med*, 2024, 129(12): 1751–1765.
- [9] JOO B, AHN S S, AN C, et al. Fully automated radiomics-based machine learning models for multiclass classification of single brain tumors: glioblastoma, lymphoma, and metastasis [J] . *J Neuroradiol*, 2023, 50(4): 388–395.
- [10] LIU J P, TU J Q, XU L H, et al. MRI-based radiomics signatures for preoperative prediction of Ki-67 index in primary central nervous system lymphoma [J] . *Eur J Radiol*, 2024, 178: 111603.
- [11] LIU G L, ZHANG X Y, ZHANG N, et al. Detecting double expression status in primary central nervous system lymphoma using multiparametric MRI based machine learning [J] . *J Magn Reson Imaging*, 2024, 59(1): 231–239.
- [12] NASER P V, MAURER M C, FISCHER M, et al. Deep learning aided preoperative diagnosis of primary central nervous system lymphoma [J] . *iScience*, 2024, 27(2): 109023.
- [13] LIU X, DU P, DAI Z G, et al. SRTRP-Net: a multi-task learning network for segmentation and prediction of stereotactic radiosurgery treatment response in brain metastases [J] . *Comput Biol Med*, 2024, 175: 108503.
- [14] GARABA A, ASLAM N, PONZIO F, et al. Radiomics for differentiation of gliomas from primary central nervous system lymphomas: a systematic review and meta-analysis [J] . *Front Oncol*, 2024, 14: 1291861.
- [15] ZHAO L M, HU R, XIE F F, et al. Radiomic-based MRI for classification of solitary brain metastases subtypes from primary lymphoma of the central nervous system [J] . *J Magn Reson Imaging*, 2023, 57(1): 227–235.
- [16] CHALLIS E, HURLEY P, SERRA L, et al. Gaussian process classification of Alzheimer's disease and mild cognitive impairment from resting-state fMRI [J] . *Neuroimage*, 2015, 112: 232–243.
- [17] PONTILLO G, TOMMASIN S, CUOCOLO R, et al. A combined radiomics and machine learning approach to overcome the clinicoradiologic paradox in multiple sclerosis [J] . *AJNR Am J Neuroradiol*, 2021, 42(11): 1927–1933.
- [18] OZTURK K, SOYLU E, CAYCI Z. Differentiation between primary CNS lymphoma and atypical glioblastoma according to major genomic alterations using diffusion and susceptibility-weighted MR imaging [J] . *Eur J Radiol*, 2021, 141: 109784.
- [19] ZHANG Y, NING C Y, YANG W J. An automatic cervical cell classification model based on improved DenseNet121 [J] . *Sci Rep*, 2025, 15(1): 3240.
- [20] MAO Y N, KIM J, PODINA L, et al. Dilated SE-DenseNet for brain tumor MRI classification [J] . *Sci Rep*, 2025, 15(1): 3596.

(收稿日期: 2025-05-06 修回日期: 2025-06-13)

(责任编辑: 王琳辉)