

DOI:10.3969/j.issn.1005-202X.2026.02.015

医学人工智能

基于多序列MRI和Swin Transformer深度学习模型的肝细胞癌微血管侵犯预测

黄倩¹, 庄银平¹, 徐鹏², 巩萍¹

1. 徐州医科大学医学影像学院, 江苏 徐州 221004; 2. 徐州医科大学附属医院影像科, 江苏 徐州 221000

【摘要】目的: 开发并验证基于多序列MRI的Swin Transformer(ST)深度学习模型在肝细胞癌(HCC)微血管侵犯(MVI)预测中的应用价值, 为HCC患者的精准诊疗提供新的解决方案和客观、科学依据。**方法:** 回顾性纳入徐州医科大学附属医院174例经手术病理确诊的HCC患者, 收集患者的术前多序列MRI图像[动脉期(AP)、延迟期(DP)、脂肪抑制T₂加权成像(T₂WI-FS)、扩散加权成像(DWI)], 经过图像预处理和增强后, 构建基于ST的DL模型, 并将其与基于卷积神经网络(CNN)架构的DenseNet121、DenseNet169、ResNet34、ResNet50、VGG16、GoogleNet模型, 基于Transformer架构的Vision Transformer(ViT)模型以及影像组学模型进行对比, 计算模型的准确率、受试者工作特征曲线下面积(AUC)、敏感度、特异度, 通过混淆矩阵和决策曲线分析(DCA)评估性能。为提高模型的可解释, 采用梯度加权类激活映射(Grad-CAM)可视化模型关注区域。**结果:** ST模型在4个序列的MVI预测任务中总体性能最好, 特别是DP序列上预测结果最优, 其准确度、AUC(95%CI)、敏感度和特异度分别为0.944、0.993(0.985~0.998)、0.984和0.904, 显著优于其他深度学习模型和影像组学模型, DCA图进一步证实其卓越的临床应用价值。Grad-CAM结果显示ST模型能够有效关注肿瘤相关区域。**结论:** ST在不同序列MRI图像上的各项指标均表现出优异的预测性能, 可为临床诊断、治疗决策及预后评估提供有力的辅助工具。

【关键词】 肝细胞癌; 微血管侵犯; 多序列磁共振成像; Swin Transformer; 影像组学

【中图分类号】 R318; R735.7

【文献标志码】 A

【文章编号】 1005-202X(2026)02-0245-10

Predicting microvascular invasion in hepatocellular carcinoma with multi-sequence MRI and a Swin Transformer-based deep learning model

HUANG Qian¹, ZHUANG Yinping¹, XU Peng², GONG Ping¹

1. School of Medical Imaging, Xuzhou Medical University, Xuzhou 221004, China; 2. Department of Radiology, the Affiliated Hospital of Xuzhou Medical University, Xuzhou 221000, China

Abstract: Objective To develop and validate a Swin Transformer (ST)-based deep learning (DL) model using multi-sequence magnetic resonance imaging (MRI) for predicting microvascular invasion (MVI) in hepatocellular carcinoma (HCC), thereby providing a novel solution and objective scientific evidence to facilitate the precision diagnosis and treatment for HCC patients. **Methods** A retrospective analysis was performed on 174 patients with surgically and pathologically confirmed HCC who were admitted to the Affiliated Hospital of Xuzhou Medical University. Preoperative multi-sequence MRI images, including arterial phase, delayed phase (DP), fat-suppressed T₂-weighted imaging, and diffusion-weighted imaging, were collected for each patient. After image preprocessing and enhancement, an ST-based DL model was established and then compared with models based on convolutional neural network architectures, including DenseNet121, DenseNet169, ResNet34, ResNet50, VGG16, and GoogleNet, as well as a Transformer-based Vision Transformer model and a radiomics model. Model accuracy, area under the receiver operating characteristic curve (AUC), sensitivity, and specificity were calculated, and the model performance was evaluated using confusion matrices and decision curve analysis. To enhance model interpretability, gradient-weighted class activation mapping (Grad-CAM) was employed to visualize the regions of interest focused on by the models. **Results** The ST model demonstrated the optimal overall performance among the 4

【收稿日期】 2025-10-21

【基金项目】 国家自然科学基金(82001987); 江苏省研究生科研与实践创新计划(KYCX25_3266)

【作者简介】 黄倩, 硕士研究生, 研究方向: 肝癌人工智能分析, E-mail: 303103110587@stu.xzhu.edu.cn

【通信作者】 巩萍, 博士, 副教授, 硕士生导师, 研究方向: 医学图像处理与人工智能分析, E-mail: gongping@xzhu.edu.cn

sequences in MVI prediction, with particularly outstanding results on the DP sequence. Specifically, it achieved an accuracy of 0.944, an AUC of 0.993 (95% CI: 0.985-0.998), a sensitivity of 0.984, and a specificity of 0.904, significantly outperforming other DL models and the radiomics model. Decision curve analysis further confirmed its superior potential for clinical application. Grad-CAM visualization indicated that the ST model effectively concentrated on tumor-associated regions. **Conclusion** The ST model exhibits excellent predictive performance across multiple MRI sequences and can serve as a powerful auxiliary tool to support clinical diagnosis, treatment decision-making, and prognostic evaluation.

Keywords: hepatocellular carcinoma; microvascular invasion; multi-sequence magnetic resonance imaging; Swin Transformer; radiomics

前言

肝癌是严重危害人类生命健康的恶性肿瘤,其发病率和死亡率分别位居全球恶性肿瘤第4位和第2位^[1]。肝细胞癌(Hepatocellular Carcinoma, HCC)是肝癌最常见类型,手术切除和肝移植是HCC根治性治疗的首选方式^[2],但术后常面临易复发、转移,生存率低等问题^[3]。研究发现,微血管侵犯(Microvascular Invasion, MVI)是影响HCC术后复发、转移及预后不良的独立危险因素,也是决定是否需要术后辅助治疗和指导复发性HCC患者治疗方案选择的重要指标^[4-6]。病理学检查是临床诊断MVI的金标准,但存在侵入性,潜在采样偏差和耗时等缺陷,因此,采用无创方法术前准确评估MVI状态对HCC患者的个体化治疗具有重要意义。

相较于CT,磁共振成像(Magnetic Resonance Imaging, MRI)具有较好的软组织分辨率,能通过多序列成像提供肿瘤的全面信息,在HCC的诊断、治疗及预后评估中起着重要作用^[7]。近年来,影像组学作为医学影像领域快速发展的新兴技术,其在传统影像学检查的基础上,使用自动化高通量的特征提取算法,将感兴趣区域的影像数据转化为定量的图像特征,通过对特征进行挖掘分析,为MVI的术前无创预测提供了新方法^[8]。研究者使用MRI影像组学技术在HCC MVI预测方面开展了大量研究^[9-12]。但这些基于影像组学的研究方法首先需要从不同MRI序列图像中提取灰度、形状、纹理等人工定义特征,然后应用最小绝对收缩与选择算子(Least Absolute Shrinkage And Selection Operator, LASSO)、递归特征消除(Recursive Feature Elimination, RFE)等方法进行特征选择,最后使用机器学习分类器构建MVI预测模型。尽管取得一定进展,但预测结果受限于所提取特征类型、特征选择方法及所用分类器。

为克服影像组学在特征提取和模型构建方面的局限性,近年来深度学习(Deep Learning, DL)逐渐被应用于MVI预测研究中^[13-15]。与影像组学方法相比,DL是一种端到端的学习过程,能够自动完成图像的

特征提取及模型训练和测试。现有基于DL的MVI预测研究大多基于卷积神经网络(Convolutional Neural Networks, CNN)架构,如ResNet50、DenseNet121等^[16-19],虽能取得较好的预测效能,但受限于卷积操作的局部感受野,难以充分捕捉图像中的全局上下文信息。而基于Transformer的架构通过自注意力机制突破了这一局限,其多头注意力层可建立像素级长程依赖关系,更好的定量分析肿瘤全域异质性^[20]。相较于Vision Transformer(ViT)^[21],Swin Transformer(ST)引入局部窗口注意力和跨窗口连接,同时采用移位窗口(Shifted Window, SW)策略,在显著降低计算复杂度的同时,兼具多尺度特征提取与全局建模能力,在医学影像处理中展现出卓越性能^[22],且该模型在MVI预测领域的相关研究较少。

基于此,本文构建一种基于ST网络的DL模型用于预测HCC患者的MVI状态,为验证模型性能,将其与基于CNN架构的DenseNet121、DenseNet169、ResNet34、ResNet50、VGG16、GoogleNet模型,基于Transformer架构的ViT模型以及影像组学模型在不同序列MRI图像上进行对比。针对DL作为“黑匣子”在临床可解释性不足的问题,进一步探索ST模型预测过程的可视化方法,以实现MVI的无创、智能、高效辅助预测,为HCC患者的精准诊疗提供新的解决方案。

1 资料与方法

1.1 研究对象

回顾性收集2019年1月~2023年7月在徐州医科大学附属医院接受治疗的HCC患者的MRI影像及临床资料。纳入标准:(1)术后经病理确诊为HCC的患者;(2)术前1个月内完成有效的MRI检查及具备完整的临床资料;(3)具备MVI的病理评估结果。排除标准:(1)术前接受过放疗、消融治疗或全身治疗等其他治疗的患者;(2)影像质量不达标;(3)临床资料不完整。最终,按照上述纳入和排除标准,共筛选出174例符合条件的HCC患者纳入本研究。其中,MVI

阴性患者 127 例, MVI 阳性患者 47 例。图 1 展示了本研究纳入患者流程图。

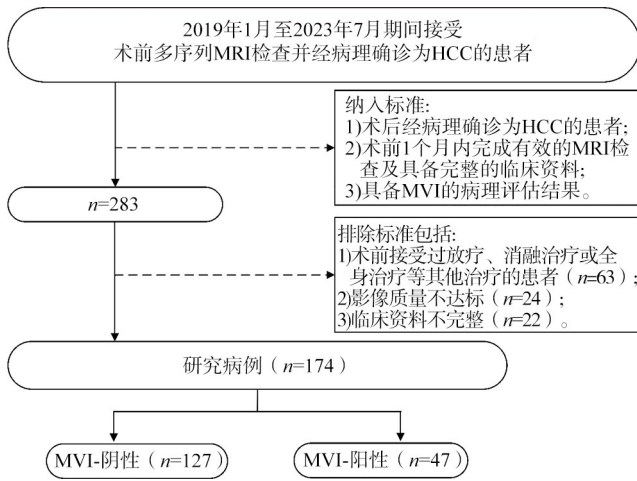


图 1 纳入患者流程图

Figure 1 Flowchart of patient inclusion

1.2 MRI 数据采集

采用 3.0T MR 设备 (GE Signa 或 Philips Ingenia Elition), 16 通道或 8 通道的体部相控阵线圈采集 MRI 图像。成像序列包括: 呼吸触发的 T₁ 加权序列 (T₁-Weighted, T₁WI)、T₂WI、同相位和反相位的 T₁WI 和脂肪抑制的 T₂WI 序列、b 值为 800 的扩散加权序列 (Diffusion-Weighted, DWI) 以及需要屏气的动态增强序列 (Dynamic Contrast-Enhanced Magnetic Resonance Imaging, DCE-MRI)。其中, DCE-MRI 以 1.0 mL/s 的速率经肘正中静脉注射 0.025 mmol/kg 钆赛酸二钠对比剂, 获得动脉期 (Arterial Phase, AP) (25~30 s)、门静脉期 (Portal Venous Phase, PVP) (55~65 s)、延迟期 (Default Phase, DP) (2~5 min) 以及肝胆期 (Hepatobiliary Phase, HBP) (20 min) 影像。由于部分病人的 PVP 和 HBP 影像存在缺失, 本文使用 DCE-MRI 的 AP 和 DP 期像、T₂WI-FS 和 DWI 4 种序列图像进行 MVI 预测研究。

(25~30 s)、门静脉期 (Portal Venous Phase, PVP) (55~65 s)、延迟期 (Default Phase, DP) (2~5 min) 以及肝胆期 (Hepatobiliary Phase, HBP) (20 min) 影像。由于部分病人的 PVP 和 HBP 影像存在缺失, 本文使用 DCE-MRI 的 AP 和 DP 期像、T₂WI-FS 和 DWI 4 种序列图像进行 MVI 预测研究。

1.3 数据预处理

为消除不同设备采集图像之间的差异, 对原始图像进行重新采样 [将每个体素采样为 (1×1×1) mm³], 并采用 Z-score 标准化进行归一化处理。对预处理后的 DCE-MRI (AP、DP、T₂WI-FS、DWI) 4 个序列的影像数据, 由两名具有 10 年经验的放射科医师在盲法条件下使用 ITK-SNAP 软件 (<http://www.itksnap.org>) 手动勾画肿瘤感兴趣区域 (Region Of Interest, ROI), 并达成一致意见。随后, 由 1 名具有超过 15 年 HCC 诊断经验的专家对所有已勾画的 ROI 进行复核确认。

由于所收集的病例中 MVI 阳性与阴性分布存在不平衡情况, 采用随机度数旋转、平移、翻转及添加随机高斯噪声的方式对每个序列的病灶图像进行数据增强。随后, 对 4 个序列的 MRI 图像分别在已勾画的病灶区域上进行裁剪, 选择病灶最大层及其上下第五层, 并向外扩展 5 个像素以包含病灶周边区域。最终, 每个序列共获得 1 251 张图像, 其中 MVI 阳性 625 张, 阴性 626 张。将病灶图像大小统一调整为 224×224 的尺寸。考虑到同一患者不同序列图像可能引发的数据泄露风险, 本研究以患者为单位对相应的图像数据按照 8:2 比例进行训练集 (n=1 000) 和测试集 (n=251) 划分。

1.4 模型构建

本文所构建的模型与性能评估流程如图 2 所示。

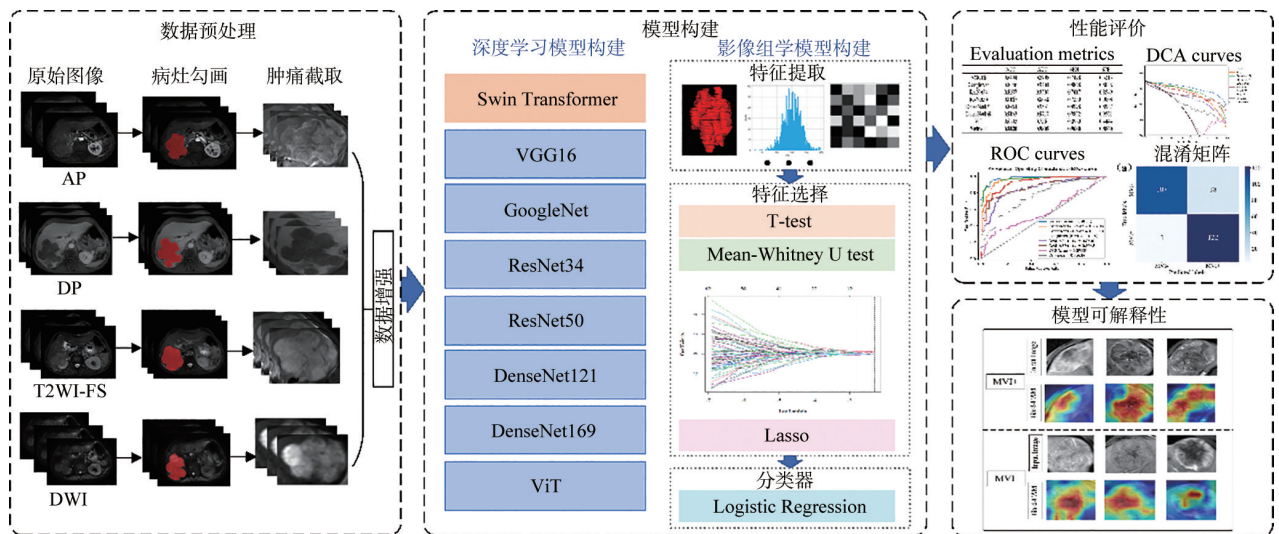


图 2 研究概述图

Figure 2 Study overview

1.4.1 ST模型构建 本研究构建的ST模型如图3所示,主要由特征提取模块和分类模块组成。特征提取模块首先利用补丁分割层将输入的病灶图像分割为大小为 4×4 的不重叠小块,每个小块记为“token”,其特征由补丁内所有像素的RGB值组合而成。然后特征进入由线性嵌入层和一系列ST块组成的“Stage 1”进行初始特征提取,通过线性嵌入层投影到C维特

征空间。接着通过“Stage2”进行局部到全局的特征聚合,通过补丁合并层对特征进行降采样。第一个补丁合并层将每组相邻的 2×2 补丁特征连接起来,形成具有4C维的特征向量,通过线性层将特征从4C降维至2C,从而减少token数量,重复“Stage 2”的过程两次为“Stage 3”和“Stage 4”。

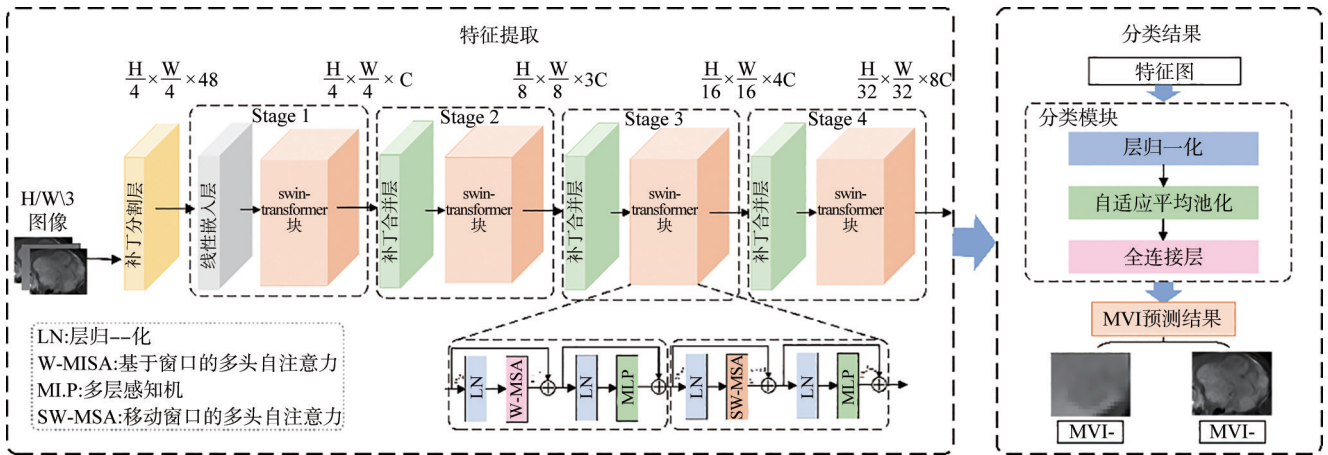


图3 基于ST模型MVI预测结构

Figure 3 Structure of MVI prediction based on ST model

特征提取模块的核心是两个连续的ST块,包含窗口自注意力模块(W-MSA)和滑动窗口自注意力模块(SW-MSA),通过窗口移动和填充实现不同窗口的信息交互,增强全局上下文捕捉能力,同时在每个窗口内进行注意力计算,大大减少计算复杂度。ST块的计算过程如式(1)表示。假设输入到ST块的特征为 z^{l-1} ,经过LN和W-MSA得到特征为 \hat{z}^l ,通过LN和MLP得到的特征为 z^l ,通过LN和SW-MSA得到 \hat{z}^{l+1} ,输出特征为 z^{l+1} 。

$$\begin{aligned} \hat{z}^l &= \text{W-MSA}(\text{LN}(z^{l-1})) + z^{l-1} \\ z^l &= \text{MLP}(\text{LN}(\hat{z}^l)) + \hat{z}^l \\ \hat{z}^{l+1} &= \text{SW-MSA}(\text{LN}(z^l)) + z^l \\ z^{l+1} &= \text{MLP}(\text{LN}(\hat{z}^{l+1})) + \hat{z}^{l+1} \end{aligned} \quad (1)$$

每个窗口的自注意力计算如式(2)所示,通过 Q 、 K 、 V 三者的交互捕获MRI图像中特征关系。

$$\text{Attention}(Q, K, V) = \text{softmax}\left(\frac{QK^T}{\sqrt{d}} + B\right)V \quad (2)$$

其中, Q 、 K 、 V 分别表示查询、键和值矩阵, B 为相对位置偏置,softmax归一化计算注意力权重,与 V 加权求和。

模型分类模块首先对来自Stage 4的特征图通过层归一化进行标准化,随后通过全局池化将特征图压缩为固定长度的全局特征向量,保留重要的全

局信息。最后经过全连接层将特征映射到分类空间,最后输出MRI图像的MVI预测结果。

1.4.2 模型对比及训练配置 为充分验证ST模型在MVI预测中的性能,本研究将其与基于CNN架构中的DenseNet121、DenseNet169^[23]、ResNet34、ResNet50^[24]、VGG16^[25]、GoogleNet^[26]模型,基于Transformer架构的ViT模型以及传统影像组学模型进行对比。所有DL模型均基于ImageNet数据集进行预训练,并采用迁移学习将其应用于MVI预测任务。本研究所使用的DL模型均来源于“torchvision.models”。在训练过程中,采用Adam优化器,训练总周期为200次,批量大小为16,初始学习率设为 $1e-3$,每20个周期将学习率按0.1倍进行衰减。所有DL模型的数据增强、训练及测试均在PyTorch框架下完成。

在影像组学模型中,首先利用PyRadiomics工具包对AP、DP、 T_2 WI-FS、DWI 4个序列的病灶图像进行影像组学特征提取,提取的特征类型包括一阶统计特征、形状特征、纹理特征及图像变换特征等,4个序列分别提取1032个影像组学特征。首先,使用t检验和Mean-Whitney U检验对每个序列的影像组学特征进行初步筛选,然后利用LASSO算法对各个序列的影特征进行特征选择,最后利用逻辑回归(Logistic

Regression, LR)分类器进行模型训练与预测。

1.5 模型评价指标

为评价模型性能,采用准确度(Accuracy, ACC)、敏感度(Sensitivity, SEN)、特异度(Specificity, SPE)、受试者工作特征曲线下面积(Area Under the Receiver Operating Characteristic Curve, AUC)4个指标对不同模型测试集上的结果进行评价。此外,使用DeLong检验评估不同模型的AUC是否具有统计学差异, $P<0.05$ 为差异有统计学意义,使用决策曲线

分析(Decision Curve Analysis, DCA)验证各模型的临床应用效能。本研究的统计分析采用Python(版本3.9.13, <https://www.python.org>)软件实现。

2 结果

2.1 深度学习模型结果对比

表1给出AP、DP、T₂WI-FS、DWI 4个序列下ST模型与其他DL模型在测试集上的MVI预测结果。

表1 不同DL模型在4个序列下测试集上的MVI预测结果

Table 1 MVI prediction results of different DL models on the test set on 4 sequences

序列	模型	准确度	AUC(95%CI)	敏感度	特异度
AP	VGG16	0.649	0.571(0.495~0.644)	0.742	0.532
	GoogleNet	0.845	0.920(0.884~0.952)	0.883	0.805
	ResNet34	0.801	0.850(0.802~0.893)	0.742	0.855
	ResNet50	0.817	0.865(0.814~0.908)	0.725	0.901
	DenseNet121	0.896	0.953(0.926~0.974)	0.898	0.895
	DenseNet169	0.916	0.971(0.951~0.987)	0.883	0.956
	ViT	0.789	0.861(0.812~0.903)	0.832	0.746
	ST	0.912	0.980(0.966~0.991)	0.968	0.856
DP	VGG16	0.705	0.557(0.488~0.626)	0.597	0.811
	GoogleNet	0.884	0.944(0.917~0.967)	0.839	0.925
	ResNet34	0.797	0.851(0.804~0.895)	0.708	0.878
	ResNet50	0.841	0.894(0.852~0.932)	0.800	0.878
	DenseNet121	0.940	0.987(0.976~0.996)	0.920	0.965
	DenseNet169	0.940	0.984(0.968~0.996)	0.934	0.947
	ViT	0.837	0.886(0.842~0.926)	0.864	0.810
	ST	0.944	0.993(0.985~0.998)	0.984	0.904
T ₂ WI-FS	VGG16	0.653	0.610(0.537~0.680)	0.567	0.752
	GoogleNet	0.861	0.908(0.873~0.941)	0.935	0.789
	ResNet34	0.865	0.940(0.913~0.968)	0.908	0.824
	ResNet50	0.805	0.888(0.846~0.924)	0.758	0.847
	DenseNet121	0.924	0.987(0.977~0.994)	0.920	0.930
	DenseNet169	0.936	0.981(0.966~0.992)	0.912	0.965
	ViT	0.857	0.906(0.866~0.939)	0.888	0.825
	ST	0.944	0.985(0.966~0.998)	0.992	0.896
DWI	VGG16	0.797	0.544(0.473~0.615)	0.544	0.610
	GoogleNet	0.885	0.949(0.923~0.971)	0.938	0.826
	ResNet34	0.888	0.934(0.901~0.961)	0.908	0.870
	ResNet50	0.841	0.913(0.878~0.944)	0.833	0.847
	DenseNet121	0.904	0.969(0.950~0.984)	0.883	0.930
	DenseNet169	0.912	0.974(0.956~0.989)	0.883	0.947
	ViT	0.849	0.902(0.864~0.936)	0.888	0.825
	ST	0.908	0.977(0.959~0.992)	0.976	0.840

由表1可见,ST模型在各序列中均表现出最优的预测性能。在AUC(95%CI)方面,ST模型在AP、DP及DWI序列上分别达到0.980(0.966~0.991)、0.993(0.985~0.998)和0.977(0.959~0.992),均为各模型中的最高值。在准确度方面,ST模型在4个序列上的准确度值均超过0.90,且在敏感度指标上同样为最优。综合不同序列的表现,ST模型在DP和T₂WI-FS序列上的预测效果尤为突出,优于AP和DWI序列。上述结果验证ST模型在多序列MRI影像MVI预测中的优势与稳定性。

在CNN架构中,DenseNet系列总体表现优于其他模型,各模态下AUC值均超过0.90,准确度稳定在0.9左右,尤其在DP序列上,DenseNet121和DenseNet169的准确度均达0.940,AUC(95%CI)分别高达0.987(0.976~0.996)和0.984(0.968~0.996),敏感度与特异度均处于较高水平。ResNet系列在各序列中AUC值均超过0.85,除了DP上ResNet34的准确度

值低于0.8,其他序列上准确度均超过0.8,该系列总体性能在DWI序列上表现较好。GoogleNet在不同序列上的AUC均超过0.90,准确度亦均在0.80以上,且在DWI序列上的表现尤为突出。相比之下,VGG16在4个序列上的AUC值最低,预测性能显著逊于其他CNN模型。

在Transformer架构中,ViT在DP、DWI和T₂WI-FS 3个序列上的准确度均稳定在0.80以上。在AP和DP序列中,AUC值均超过0.85,在T₂WI-FS和DWI序列上的AUC均超过0.90。在不同序列中,ViT在T₂WI-FS序列上表现最佳,准确度达到0.857,AUC(95%CI)为0.906(0.866~0.939),敏感度和特异度分别为0.888和0.825。图4展示了ST模型与其他DL模型在AP、DP、T₂WI-FS和DWI序列上的ROC曲线对比,图5为不同深度学习模型的DCA图,图6给出了ST模型在各序列上的混淆矩阵。

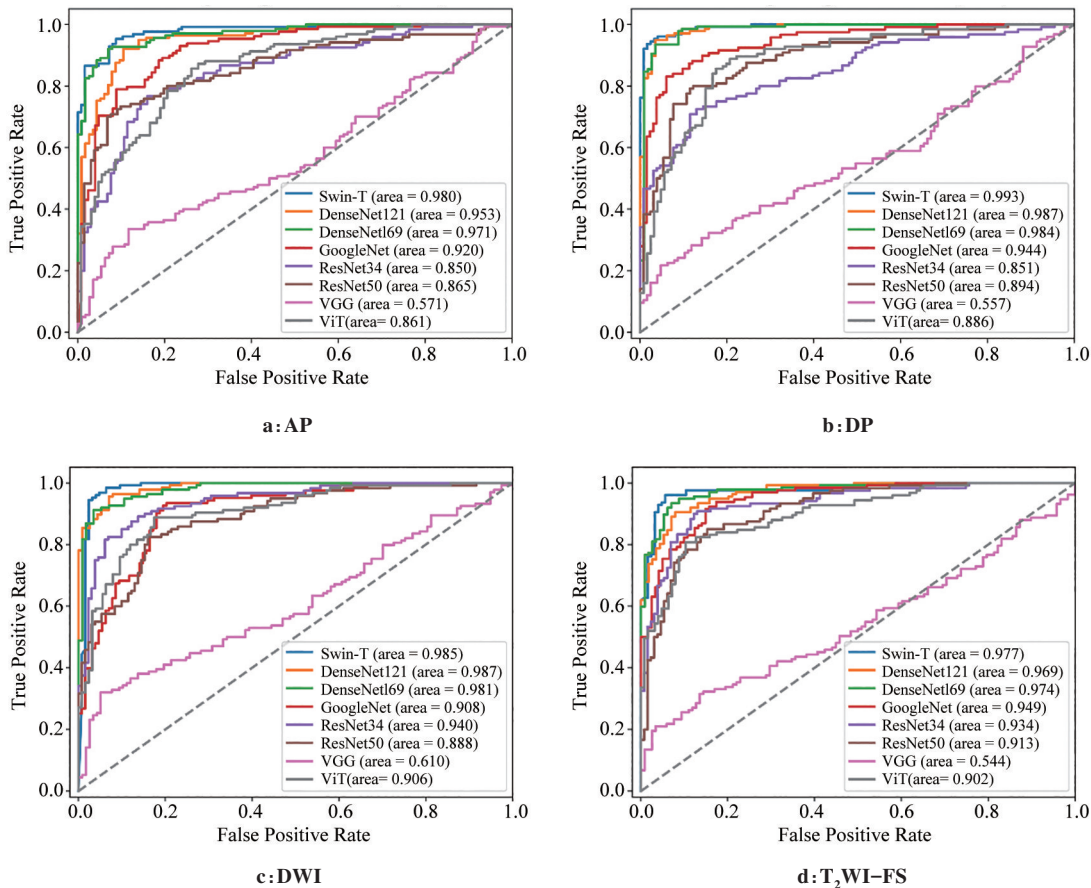


图4 ST模型与其他DL模型在不同序列上的ROC曲线

Figure 4 ROC curves of ST model and other DL models on different sequences

2.2 ST模型与影像组学模型对比

图7给出4个序列下ST模型与影像组学模型在测试集上的实验结果对比。

由图7可知,影像组学模型在DP序列上的预测性能最优,其准确度为0.821,AUC为0.896,敏感度为0.841,特异度为0.800。值得注意的是,ST模型在所

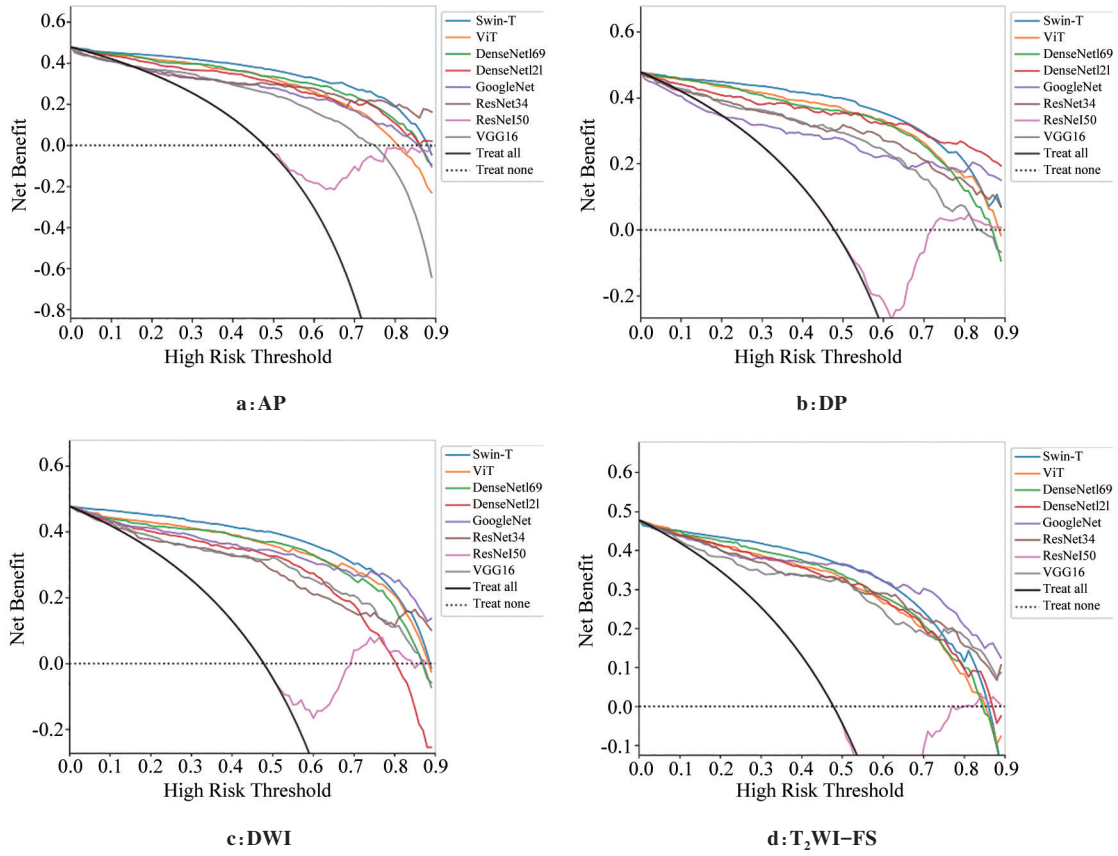


图5 ST模型与其他DL模型DCA曲线对比图

Figure 5 Comparison of DCA curves of ST model with other DL models

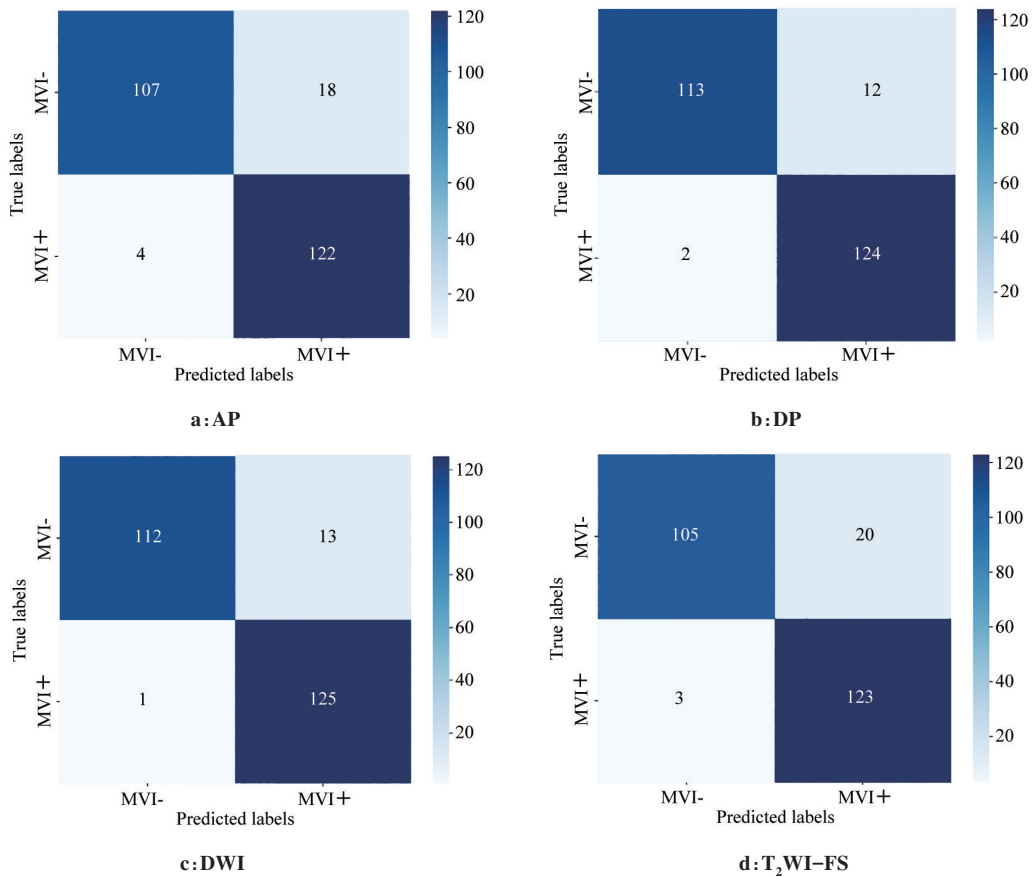


图6 ST模型在不同序列上的混淆矩阵图

Figure 6 Confusion matrices of the ST model on different sequences

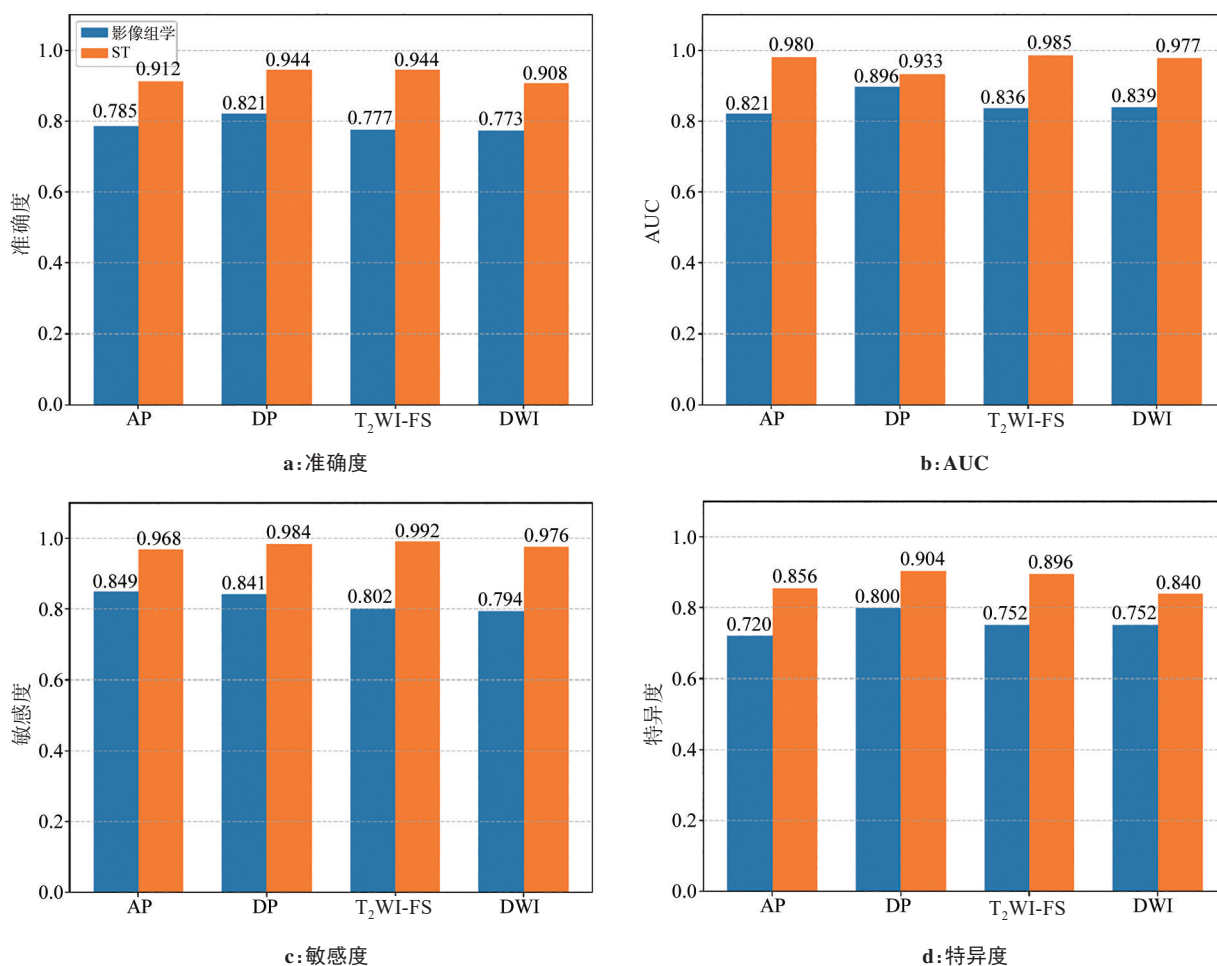


图 7 ST 与影像组学模型在 4 个序列上的结果对比

Figure 7 Comparison of results between ST and radiomics models on 4 sequences

有序列上的表现均显著优于影像组学模型。进一步分析发现,两种模型在 DP 序列上的准确度均达到峰值(影像组学:0.821;ST:0.944),综合比较 4 个指标,ST 模型展现出更稳定的跨序列性能,表明其在 MVI 预测任务中具有更高的临床应用价值。

2.3 模型可解释性

为深入探索 ST 模型的可解释性,直观展示病灶区域特征对 MVI 预测任务的影响,使用梯度加权类激活映射(Grad-CAM)生成网络的热力图^[27],结果如图 8 所示。热力图中红色区域表示模型在做决策时重点关注的部分,该区域对模型的决策贡献度较大,蓝色与绿色的区域表示模型对于该区域的关注度较小。

可视化结果显示,模型的高响应区域主要集中在肿瘤主体,并逐渐扩展至肿瘤边缘,提示其能够捕捉与 MVI 相关的微血管浸润征象。这进一步表明,模型已学习到基于肿瘤内部及边缘影像特征的空间分布模式,从而有效辅助 MVI 状态的预测。

3 讨论与展望

MVI 被认为是导致 HCC 术后早期复发及肝内转

移的关键病理因素^[28]。因此,术前对 MVI 状态进行早期、精准的预测,对于指导个体化治疗策略和评估预后具有重要意义^[29]。近年来,随着人工智能技术的迅速发展,DL 因其卓越的特征提取与端到端学习的能力,为 MVI 的无创预测提供了新的技术途径。已有多项研究证实 DL 在 MVI 预测中的潜力。如 Wei 等^[30]基于增强 MRI 和 CT 图像采用 ResNet 模型预测 MVI 状态,显示基于增强 MRI 的 DL 模型性能优于 CT (AUC: 0.812 vs 0.736)。Xu 等^[31]使用 ResNet 与 ViT 对术前 CT 图像进行分析,结果显示 ViT 取得了与 ResNet 相当的预测性能 (AUC 分别为 0.882 和 0.875),提示 Transformer 架构在该任务中的可行性。

相较于上述模型,ST 通过引入 SW 机制构建层次化特征表示,能够有效兼顾局部与全局特征的捕捉,显著增强模型对多尺度上下文信息的整合能力。Sun 等^[32]在肺癌分类研究中同样证实 ST 的优越性,其准确率较 ViT 提高 2.53%。在本研究中,ST 在不同 MRI 序列上的 AUC、准确度、敏感度和特异度均居首位,Grad-CAM 热图直观显示了其可准确关注肿瘤区域,DCA 曲线进一步验证了其卓越的临床应用价值。

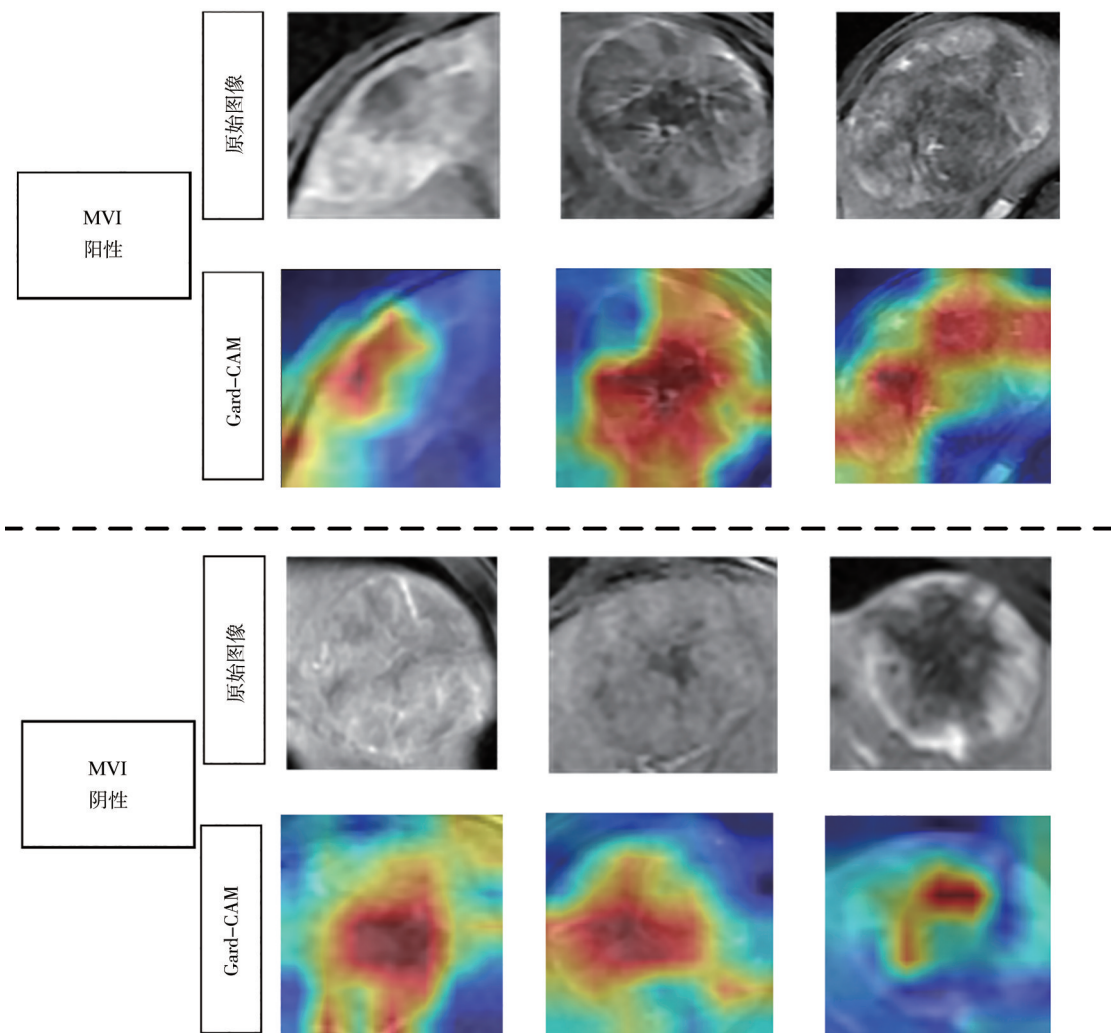


图8 ST模型生成的Gard-CAM样本示例

Figure 8 Example of Gard-CAM samples generated by the ST model

与DL模型相比,影像组学方法主要依赖于人工预先定义的浅层纹理、形状及强度等特征,其特征提取与选择过程较大程度上受制于先验知识和主观经验,限制模型对复杂高阶特征的捕捉能力^[33]。本研究结果进一步验证了这一差异,显示DL模型在MVI预测中的总体表现显著优于基于LR的传统影像组学模型。这与Liu等^[34]的研究结果一致,他们基于增强MRI构建的ResNet18模型在预测性能上明显优于基于支持向量机(Support Vector Machine, SVM)的影像组学模型(AUC: 0.82 vs 0.78),从侧面印证了DL在自动学习深层次特征、提高预测准确性方面的优势。

综上所述,本研究通过构建基于ST的DL模型,系统性地评估了不同DL模型及影像组学模型在MVI预测任务中的性能,充分探讨了ST模型在该领域的应用潜力。研究结果表明,ST模型在多序列MRI图像上的预测表现显著优于其他模型,显示出较高的准确度、敏感度及特异度,凸显了其在术前无创评估MVI中的临

床应用价值。但是研究仍然存在数据规模较小,未深入探讨多序列融合对MVI预测的影响等局限性。未来,计划在更大样本、多中心的数据基础上,进一步结合DL与影像组学技术,探索多序列特征的高效融合框架,为临床决策提供更为精准和可靠的辅助工具,从而推动MVI的精准诊疗进程。

【参考文献】

- [1] Sung H, Ferlay J, Siegel RL, et al. Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries[J]. CA Cancer J Clin, 2021, 71(3): 209-249.
- [2] 张继云, 张学琴, 张涛, 等. 基于钆塞酸二钠增强MRI联合微血管侵犯预测HBV相关HCC患者术后早期复发的价值[J]. 临床放射学杂志, 2023, 42(1): 66-71.
Zhang JY, Zhang XQ, Zhang T, et al. The value of predicting early tumor recurrence in HBV-Related hepatocellular carcinoma based on Gd-EOB-DTPA-enhanced MRI combined with microvascular invasion [J]. Journal of Clinical Radiology, 2023, 42(1): 66-71.
- [3] Mazzaferro V, Llovet JM, Miceli R, et al. Predicting survival after liver transplantation in patients with hepatocellular carcinoma beyond the Milan criteria: a retrospective, exploratory analysis[J]. Lancet Oncol, 2009, 10(1): 35-43.

- [4] Nakashima Y, Nakashima O, Tanaka M, et al. Portal vein invasion and intrahepatic micrometastasis in small hepatocellular carcinoma by gross type[J]. *Hepatol Res*, 2003, 26(2): 142-147.
- [5] D'Amico F, Schwartz M, Vitale A, et al. Predicting recurrence after liver transplantation in patients with hepatocellular carcinoma exceeding the up-to-seven criteria[J]. *Liver Transpl*, 2009, 15(10): 1278-1287.
- [6] Shah SA, Cleary SP, Wei AC, et al. Recurrence after liver resection for hepatocellular carcinoma: risk factors, treatment, and outcomes[J]. *Surgery*, 2007, 141(3): 330-339.
- [7] 卓越, 梅昕, 李学栋. CT与MRI检查在肝细胞癌临床诊断中的应用价值[J]. *影像研究与医学应用*, 2025, 9(1): 166-168.
Zhuo Y, Mei X, Li XD. Clinical value of CT and MRI in the diagnosis of hepatocellular carcinoma[J]. *Journal of Imaging Research and Medical Applications*, 2025, 9(1): 166-168.
- [8] Aerts HJWL, Velazquez ER, Leijenaar RTH, et al. Decoding tumour phenotype by noninvasive imaging using a quantitative radiomics approach[J]. *Nat Commun*, 2014, 5: 4006.
- [9] Hong SB, Choi SH, Kim SY, et al. MRI features for predicting microvascular invasion of hepatocellular carcinoma: a systematic review and meta-analysis[J]. *Liver Cancer*, 2021, 10(2): 94-106.
- [10] Lü K, Cao X, Du P, et al. Radiomics for the detection of microvascular invasion in hepatocellular carcinoma[J]. *World J Gastroenterol*, 2022, 28(20): 2176-2183.
- [11] Chong HH, Yang L, Sheng RF, et al. Multi-scale and multi-parametric radiomics of gadoxetate disodium-enhanced MRI predicts microvascular invasion and outcome in patients with solitary hepatocellular carcinoma ≤ 5 cm[J]. *Eur Radiol*, 2021, 31(7): 4824-4838.
- [12] Meng XP, Wang YC, Zhou JY, et al. Comparison of MRI and CT for the prediction of microvascular invasion in solitary hepatocellular carcinoma based on a non-radiomics and radiomics method: which imaging modality is better?[J]. *J Magn Reson Imaging*, 2021, 54(2): 526-536.
- [13] Dai HJ, Lu MH, Huang BS, et al. Considerable effects of imaging sequences, feature extraction, feature selection, and classifiers on radiomics-based prediction of microvascular invasion in hepatocellular carcinoma using magnetic resonance imaging[J]. *Quant Imaging Med Surg*, 2021, 11(5): 1836-1853.
- [14] Feng ST, Jia YM, Liao B, et al. Preoperative prediction of microvascular invasion in hepatocellular cancer: a radiomics model using Gd-EOB-DTPA-enhanced MRI[J]. *Eur Radiol*, 2019, 29(9): 4648-4659.
- [15] Liu QP, Xu X, Zhu FP, et al. Prediction of prognostic risk factors in hepatocellular carcinoma with transarterial chemoembolization using multi-modal multi-task deep learning[J]. *EClinicalMedicine*, 2020, 23: 100379.
- [16] 李鑫, 陈金明, 刘蕊, 等. 基于深度学习CT影像术前预测肝癌患者微血管侵犯[J]. *中国医学影像学杂志*, 2025, 33(4): 390-395.
Li X, Chen JM, Liu R, et al. Prediction of the occurrence of microvascular invasion in hepatocellular carcinoma *via* preoperative prediction model using deep learning on CT image[J]. *Chinese Journal of Medical Imaging*, 2025, 33(4): 390-395.
- [17] Qin XC, Zhu JH, Tu ZZ, et al. Contrast-Enhanced ultrasound with deep learning with attention mechanisms for predicting microvascular invasion in single hepatocellular carcinoma[J]. *Acad Radiol*, 2023, 30 (Suppl 1): S73-S80.
- [18] Zhang YX, Lv XF, Qiu JL, et al. Deep learning with 3D convolutional neural network for noninvasive prediction of microvascular invasion in hepatocellular carcinoma[J]. *J Magn Reson Imaging*, 2021, 54(1): 134-143.
- [19] Wei JW, Jiang HY, Zeng MS, et al. Prediction of microvascular invasion in hepatocellular carcinoma *via* deep learning: a multi-center and prospective validation study[J]. *Cancers (Basel)*, 2021, 13(10): 2368.
- [20] Dosovitskiy A, Beyer L, Kolesnikov A, et al. An image is worth 16x16 words: transformers for image recognition at scale [EB/OL]. (2021-06-03). <https://arxiv.org/abs/2010.11929>.
- [21] Xu YL, Chen YY, Wu JM, et al. The utility of Vision Transformer in preoperatively predicting microvascular invasion status of hepatocellular carcinoma[J]. *HPB (Oxford)*, 2023, 25(5): 533-542.
- [22] Liu Z, Lin YT, Cao Y, et al. Swin transformer: hierarchical vision transformer using shifted windows [C]//2021 IEEE/CVF International Conference on Computer Vision (ICCV). Piscataway, NJ, USA: IEEE, 2021: 9992-10002.
- [23] Huang G, Liu Z, Van Der Maaten L, et al. Densely connected convolutional networks [C]//2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Piscataway, NJ, USA: IEEE, 2017: 2261-2269.
- [24] He KM, Zhang XY, Ren SQ, et al. Deep residual learning for image recognition [C]//2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Piscataway, NJ, USA: IEEE, 2016: 770-778.
- [25] Simonyan K, Zisserman A. Very deep convolutional networks for large-scale image recognition[EB/OL]. (2015-04-10). <https://arxiv.org/abs/1409.1556>.
- [26] Szegedy C, Liu W, Jia YQ, et al. Going deeper with convolutions [C]//2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Piscataway, NJ, USA: IEEE, 2015: 1-9.
- [27] Selvaraju RR, Cogswell M, Das A, et al. Grad-CAM: visual explanations from deep networks via gradient-based localization [C]//2017 IEEE International Conference on Computer Vision (ICCV). Piscataway, NJ, USA: IEEE, 2017: 618-626.
- [28] Van Calster B, Wynants L, Verbeek JFM, et al. Reporting and interpreting decision curve analysis: a guide for investigators[J]. *Eur Urol*, 2018, 74(6): 796-804.
- [29] Erstad DJ, Tanabe KK. Prognostic and therapeutic implications of microvascular invasion in hepatocellular carcinoma[J]. *Ann Surg Oncol*, 2019, 26(5): 1474-1493.
- [30] Wei JW, Jiang HY, Zeng MS, et al. Prediction of microvascular invasion in hepatocellular carcinoma *via* deep learning: a multi-center and prospective validation study[J]. *Cancers (Basel)*, 2021, 13(10): 2368.
- [31] Xu YL, Chen YY, Wu JM, et al. The utility of Vision Transformer in preoperatively predicting microvascular invasion status of hepatocellular carcinoma[J]. *HPB (Oxford)*, 2023, 25(5): 533-542.
- [32] Sun RN, Pang YX, Li WF. Efficient lung cancer image classification and segmentation algorithm based on an improved swin transformer [J]. *Electronics*, 2023, 12(4): 1024.
- [33] 张歌, 钟淑媛, 胡根文, 等. 基于平扫MRI的影像组学及深度学习模型预测肝细胞癌微血管侵犯: 一项双中心研究[J]. *实用放射学杂志*, 2025, 41(3): 424-428.
Zhang G, Zhong SY, Hu GW, et al. Radiomics and deep learning models based on unenhanced MRI to predict microvascular invasion in hepatocellular carcinoma: a two-center study [J]. *Journal of Practical Radiology*, 2025, 41(3): 424-428.
- [34] Liu SC, Lai J, Huang JY, et al. Predicting microvascular invasion in hepatocellular carcinoma: a deep learning model validated across hospitals[J]. *Cancer Imaging*, 2021, 21(1): 56.

(编辑: 薛泽玲)