

基于张量的多模态融合诊断微血管侵犯

汪少东, 李柳军, 李蕊, 苏中振, 陆遥

引用本文

汪少东, 李柳军, 李蕊, 苏中振, 陆遥. 基于张量的多模态融合诊断微血管侵犯[J]. 计算机科学, 2026, 53(4): 284-290.

WANG Shaodong, LI Liujun, LI Rui, SU Zhongzhen, LU Yao. [Tensor-based Multimodal Fusion Technique to Diagnose Microvascular Invasion](#) [J]. Computer Science, 2026, 53(4): 284-290.

相似文章推荐 (请使用火狐或 IE 浏览器查看文章)

Similar articles recommended (Please use Firefox or IE to view the article)

[跨模态融合的少样本勒索软件分类器:基于预训练模型的多模态编码](#)

Cross-modal Fusion Few-sample Ransomware Classifier:Multimodal Encoding Based on Pre-trained Models

计算机科学, 2026, 53(4): 435-444. <https://doi.org/10.11896/jsjcx.250500078>

[基于多任务学习的眼科视频特征融合与多维画像](#)

Multi-task Learning-based Ophthalmic Video Feature Fusion and Multi-dimensional Profiling

计算机科学, 2026, 53(3): 383-391. <https://doi.org/10.11896/jsjcx.260200058>

[大语言模型驱动的语言障碍评估方法综述](#)

Review of Speech Disorder Assessment Methods Driven by Large Language Models

计算机科学, 2026, 53(3): 307-320. <https://doi.org/10.11896/jsjcx.250300125>

[多模态水声图像目标视觉检测](#)

Multimodal Visual Detection for Underwater Sonar Target Images

计算机科学, 2026, 53(2): 227-235. <https://doi.org/10.11896/jsjcx.241200082>

[基于多模态体育教育数据的图空间融合动作识别方法](#)

Multimodal Physical Education Data Fusion via Graph Alignment for Action Recognition

计算机科学, 2026, 53(2): 89-98. <https://doi.org/10.11896/jsjcx.250800007>

基于张量的多模态融合诊断微血管侵犯

汪少东¹ 李柳军² 李蕊¹ 苏中振² 陆遥¹

¹ 中山大学计算机学院 广州 510006

² 中山大学附属第五医院超声科 广东 珠海 519000

(wangshd6@mail2.sysu.edu.cn)

摘要 微血管侵犯(MVI)作为肝细胞癌(HCC)术后复发和生存率低的关键预后因素,其术前精准定位对治疗决策至关重要。针对现有放射组学方法特征泛化弱、可解释性差且忽略瘤周 MVI 空间分布的问题,提出通过病理全切片(WSI)与三维超声(3D US)的空间融合实现 MVI 三维定位,并设计特征张量融合深度学习模型(融合多尺度特征、特征张量及正交化损失函数)提取瘤周 MVI 分布语义特征。在收集的数据集上开展了详细的对比分析和消融实验研究,使用受试者工作特征曲线下的面积(AUC)、准确度(Accuracy)和 F1 分数等指标证明了该模型的有效性。实验验证了该模型性能优异(AUC:0.910,ACC:0.930,F1 score:0.852),证实了其在术前 MVI 精确诊断中的临床潜力。

关键词: 微血管侵犯;多模态;张量融合;多尺度融合;肝细胞癌

中图分类号 TP183

Tensor-based Multimodal Fusion Technique to Diagnose Microvascular Invasion

WANG Shaodong¹, LI Liujun², LI Rui¹, SU Zhongzhen² and LU Yao¹

¹ School of Computer Science and Engineering, Sun Yat-sen University, Guangzhou 510006, China

² Department of Ultrasound, the Fifth Affiliated Hospital of Sun Yat-Sen University, Zhuhai, Guangdong 519000, China

Abstract Microvascular invasion(MVI) is a critical prognostic factor for postoperative recurrence and reduced survival in hepatocellular carcinoma(HCC), making its precise preoperative localization essential for treatment planning. To address limitations of existing radiomics approaches—including poor feature generalizability, weak interpretability, and neglect of spatial peritumoral MVI distribution crucial for surgical strategies—this study proposes: 1) multimodal fusion-based 3D localization of MVI via spatial alignment of whole-slide pathology images(WSI) with 3D ultrasound(3D US); 2) a feature tensor fusion deep learning model integrating multiscale features, tensor fusion, and orthogonal loss functions to extract semantic features of peritumoral MVI distribution. Model performance is evaluated on the curated dataset using metrics such as the area under the receiver operating characteristic curve(AUC), Accuracy and F1 score, demonstrating its effectiveness. Experimental validation demonstrates exceptional performance(AUC:0.910, Accuracy:0.930, F1-score:0.852), confirming the proposed model's clinical potential for precise preoperative MVI diagnosis.

Keywords Microvascular invasion, Multimodal, Tensor fusion, Multi-scale fusion, Hepatocellular carcinoma

1 引言

原发性肝癌位列全球恶性肿瘤发病率第六位,该疾病在全球癌症相关致死病因中位居第三。2020 年全球监测数据显示,该年度新确诊原发性肝癌患者达 90.6 万例,对应同期死亡病例约 83 万例^[1]。在原发性肝癌病例中,肝细胞癌

(HCC)占最多,为 75%~85%^[1],乙型肝炎病毒和丙型肝炎病毒感染是肝癌发展的主要危险因素。HCC 作为全球主要健康问题之一,是全球第二大致命肿瘤,同时也是癌症相关死亡的第四大原因^[1-2]。HCC 患者的 5 年生存率仅为 18%^[1],术后复发率最高可达 70%^[3-4]。超过 90% 的 HCC 相关死亡归因于转移和复发^[5]。有证据表明,大多数国家的 HCC 发

到稿日期:2025-06-26 返修日期:2025-09-08

基金项目:国家重点研发计划(2023YFE0204300);琶洲实验室(黄埔区)研发项目(2023K0606);国家自然科学基金(82441027,62371476);广州市科学技术局项目(2023B03J1237);湖南省卫生健康委卫生健康科研重大项目(W20241010);广东省计算科学重点实验室(中山大学)(2020B1212060032)

This work was supported by the National Key R&D Program of China (2023YFE0204300), the R&D Project of Pazhou Lab (HuangPu) (2023K0606), National Natural Science Foundation of China(82441027,62371476), Guangzhou Science and Technology Bureau(2023B03J1237), Health Research Major Projects of Hunan Health Commission(W20241010) and Guangdong Province Key Laboratory of Computational Science at the Sun Yat-sen University(2020B1212060032).

通信作者:陆遥(luyao23@mail.sysu.edu.cn)

病率都在增加,尤其是在东亚^[6]。流行病学研究揭示,中国 HCC 疾病负担尤为突出^[7]。在乙型及丙型病毒性肝炎高流行率、非酒精性脂肪性肝病、酒精性肝损伤等多重致病因素的协同作用下,该疾病患者群体规模已占据全球总病例数的 50%^[7]。当前临床对于 MVI 的诊断主要依赖于术后的组织学检查,这使其成为一种回顾性指标^[8]。因此,如何利用术前影像识别 MVI 对于 HCC 治疗决策至关重要。

临床研究已证实,微血管浸润(MVI)是肝细胞癌患者复发和低生存率的重要预后因素^[9-12],其是通过在显微镜下由内皮细胞排列的血管腔内的癌细胞巢来识别的^[13]。MVI 不仅反映了肿瘤的侵袭性,还被认为是 HCC 隐匿性转移的重要诱因^[14-15]。此外,肝内转移被认为是影响 HCC 患者预后的关键因素,而 MVI 阳性患者术后早期复发与肝内转移关系密切^[16]。临床研究表明,对于 MVI 阳性的 HCC 患者,相较于非解剖性肝切除,解剖性肝切除术可有效降低 MVI 阳性患者的术后复发率及死亡风险^[13,17];并且 MVI 阳性的 HCC 患者在接受肝移植后,肿瘤复发率往往较高^[18-19]。因此,对于 MVI 风险较高的 HCC 患者,应优先考虑解剖性肝切除术以改善预后^[20]。

先前的研究通常通过常规术前影像学^[1,3]、血清生物标志物^[4]和临床数据^[9]来预测 MVI 状态。然而,这些方法通常采用传统放射组学,导致其准确性有限。同时,放射组学识别的特征有限且固定,其普遍性和生物学解释性可能较弱。此外,还有一些研究强调了肿瘤周围特征在预测 MVI 中的重要性,并取得了不错的预测性能^[21]。然而,这些研究往往只关注 MVI 是否存在,忽略了对于术前治疗策略(如手术切除或热消融)至关重要的瘤周 MVI 的定位和分布。

为应对上述挑战,本文提出联合多模态融合成像技术,整合病理全切片图像(WSI)与三维超声(3D US)图像,建立 MVI 病理与 US 影像之间的空间对应关系,得到 MVI 在 3D US 图像中的准确位置信息。随后,设计了一个深度学习模型,提取融合后的 3D US 图像中与瘤周分布相关的语义特征,并结合临床形态学特征进行 MVI 诊断。此外,据调研,当前尚无临床研究使用 3D US 深度学习进行 MVI 预测,本研究首次采用术前影像进行 MVI 预测,为未来术前 MVI 的定位和分布奠定了方法论基础,优化了治疗计划,改善了患者预后,能更有效地管理肝细胞癌(HCC)。本文方法主要的创新点如下:

1)提出了一种基于自注意力机制的多尺度特征融合方法,通过捕捉不同尺度特征间的关联,提高模型对临床关注的瘤周分布信息相关的语义特征的学习能力。

2)为了促进不同模态特征的深度交互,采用张量外积方法融合多模态特征,不仅保留了各模态的互补信息,还使网络能够学习到更加稳健且多样化的特征表示。

3)引入了一种正交损失函数,通过促进双模态(图像特征与临床形态学特征)之间的正交性,同时有效约束特征之间的互补性,从而提升模型性能和泛化能力。

2 相关工作

微血管侵犯作为肝细胞癌术后早期复发的关键独立预测

因子,其临床评估对预后管理具有重要指导价值。当前临床实践中,获取术前 MVI 病理证据仍依赖于侵入性组织采样,但此类操作可能引发针道转移风险及取样误差,限制了其广泛适用性。在当前背景下,融合影像组学与分子标志物的非侵入性检测技术已成为实现 MVI 术前动态监测、优化 HCC 个体化治疗决策的重要研究方向。

基于临床精准诊疗的迫切需求,MVI 的术前无创评估技术研发已成为肝癌研究领域的重点攻关方向。其中,一些团队选择使用常规术前成像^[1,3]、血清学标志物^[4-5]或临床资料^[9-10]来预测 MVI 状态,然而这些方法的准确性有限。

近年来,放射组学作为术前预测 MVI 的前沿技术,展现出广阔的应用前景。该方法依托医学影像技术,如计算机断层扫描(CT)、磁共振成像(MRI)和超声(US),构建多模态影像与组织病理学的深度关联,从而通过特征解码技术实现病理表型的逆向推演。影像组学技术体系包含三大核心环节:1)采集多模态影像大数据构建标准化数据库;2)应用信号解析算法进行高维度定量特征谱的构建;3)基于多维度特征矩阵的深度分析,实现微观病理生物学特征的智能提取。其中,肿瘤最大径线及瘤周界面形态等影像学参数在 MVI 术前评估中具有重要参考价值。基于常规影像组学技术提取上述特征指标,已逐步建立 MVI 无创预测的初步模型体系。Dong 等^[11]采用基于原始射频信号的超声放射组学。Xu 等^[12]采用增强 CT 放射组学。Yang 等^[8]采用多序列、多参数 MRI 放射组学,共同显示了 MVI 的术前预测性能。此外,一些研究强调了瘤周特征在 MVI 预测中的重要性^[21-27]。然而,传统放射组学技术在 MVI 评估中的准确性仍有提升空间。此外,传统放射组学特征提取依赖于人为设定的模式匹配方式,这种方法在全面性和精度上存在一定局限,同时,特征之间的相关性可能会被进一步放大,影响最终结果的可靠性。因此,该方法难以构建端到端的自动化模型来实现术前 MVI 评估。

近年来,研究者对深度学习的深入探索,为这一领域带来了新的突破。Liu 等^[25]基于体素内不相干运动(IVIM)成像技术,采用 LeNet 架构的卷积神经网络模型进行肝细胞癌 MVI 预测分析。该研究揭示了肿瘤边缘纹理异质性参数在 MVI 状态判别中的关键作用,为影像组学技术路径的创新提供了方法学参考。此外,Wang 等^[26]开发的新型深度学习架构创新性地整合病灶分割与 MVI 预测功能,该端到端模型通过编码-解码结构同步执行感兴趣区域(ROI)自动提取和 MVI 状态智能判别,显著提升了影像辅助诊断系统的临床转化价值。然而,这些研究往往只关注 MVI 是否存在,忽略了对于术前治疗策略(如手术切除或热消融)至关重要的瘤周 MVI 的定位和分布。

3 材料和方法

3.1 数据采集与预处理

为探究术前影像与术后病理微血管侵犯定位的空间关联性,本研究采用 VX2 原位肝肿瘤模型,通过获取离体三维超声图像与病理全切片数字图像,融合二者实现 3D-US 图像中 MVI 区域的精准定位。本研究的数据采集及多模态图像融合定位处理流程如图 1 所示,下面分别介绍建立兔子肿瘤模

型、体外 US 成像协议、WSI 协议及 US-WSI 图像融合。

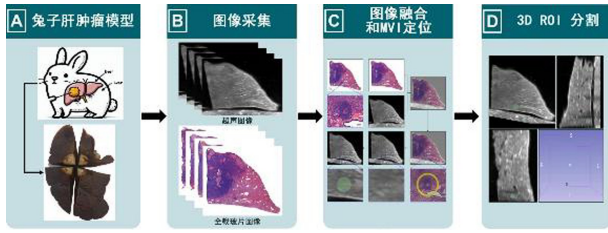


图1 数据采集流程

Fig. 1 Process of data acquisition

3.1.1 动物和 VX2 肝肿瘤模型

本研究的动物实验严格遵循《动物实验报告统一标准指南》(ARRIVE Guidelines 2.0)^[27], 实验方案经珠海百试通生物科技有限公司实验动物管理与使用委员会 (Institutional Animal Care and Use Committee, IACUC) 批准 (许可证号: SYXK2020-0229, 协议编号: L202104001)。实验选用 9 只健康新西兰白兔 (3~4 月龄, 体重 2~3 千克), 在超声引导下将单枚 VX2 肿瘤组织块 (约 1mm³) 植入肝脏左叶实质内构建肝肿瘤模型。术后通过连续监测肿瘤生长情况, 待肿瘤最大径达到 2~3cm 时确认建模成功。

3.1.2 体外 3D US 成像协议

高分辨率超声成像对精准定位并提取微血管侵犯的影像特征具有关键作用。完成肿瘤模型构建后, 将含瘤肝叶组织完整切除, 并置于 4% 多聚甲醛溶液中固定 48 小时以上。随后以肿瘤为中心将肝叶切割为 4 块标本, 每块标本边缘组织修剪量控制在最小范围。为减少超声伪影干扰, 将标本置于以海绵为基底的生理盐水中浸没。采用机械臂辅助的 Acuson Sequoia 超声系统 (德国西门子) 及 18L6 探头 (频率 18MHz), 以 0.1mm 层厚获取 3D-US 图像。关键超声参数 (包括扫描深度、声窗宽度、聚焦位置及增益) 经优化设置以确保高分辨率成像。所获图像在 3DSlicer 软件中以 53×53μm 间距显示, 并以 DICOM 格式存储。

3.1.3 序列病理全载玻片切片

为建立超声影像与组织学的精确空间对应关系, 标本经精密切割为等厚平行组织块, 其切割平面与超声扫描平面严格对齐。组织块依次编号后, 进行梯度脱水与石蜡包埋处理。采用间隔 300μm 的连续切片方案, 制备 5μm 厚度切片并进行苏木精/伊红 (Hematoxylin/Eosin, H&E) 染色。染色切片使用 Panoramic P250 Flash 全自动切片扫描仪 (匈牙利 3D Histech 公司) 以 20 倍光学放大倍率进行数字化扫描, 并通过 CaseViewer 2.4 软件 (匈牙利 3D Histech 公司) 进行全景图像浏览与分析。

3.1.4 三维超声与全切片数字图像融合

图像融合操作严格遵循既定标准流程执行。首先, 通过 CaseViewer 软件以 1.0 倍放大倍率获取病理全切片数字图像 (WSIs) 的全景视图。基于组织学轮廓特征, 使用 Photoshop 22.1.1 (美国 Adobe 公司) 对 WSIs 进行空间对齐并保存为 JPG 格式。继而将各标本的 WSIs 与 3D-US 图像导入 3DSlicer 4.11.0 开源平台¹⁾, 重建轴状位、矢状位及冠状位三

维容积数据。在 3DSlicer 中, 采用标志点配准法将固定体积 (WSIs) 与移动体积 (3D-US) 进行空间匹配: 先通过刚体配准 (Affine Registration) 校正位置偏移与旋转偏差, 再应用非刚体配准 (Nonrigid Registration) 补偿组织脱水与切片导致的形变误差。由经验丰富的放射科医师手工选取至少 10 对特征标志点 (包括小血管分叉、特定轮廓及解剖结构特征) 作为稳定参考点。针对每个标本, 额外记录 3 对标志点的空间坐标, 计算其平均欧氏距离作为配准误差评估指标。当配准误差小于 1mm 时判定融合成功, 确保方法的高精度特性。

3.1.5 形态学特征获取

所有形态学参数 (见表 1) 均通过双盲复核确保数据可靠性, 测量结果导入定制化数据库进行后续统计分析。

表 1 形态学特征

Table 1 Morphological characteristics

形态学特征	备注
ROI 所在区域主瘤边缘	光滑=0, 不平整=1
ROI 所在肝脏背景	有肝硬化=1, 无肝硬化=0
ROI 方位	近心端=1, 肿瘤区=2, 远心端=3
ROI 所在区域主瘤边界	清晰=0, 模糊=1
ROI 超出肝边缘	1=有, 0=无
ROI 与主瘤边缘距离	
主瘤最大径	
MVI 最大径	

3.2 数据预处理

3.2.1 图像标准化

由于原图的体素间隔 (Voxel Spacing) 不统一, 为了研究的严谨性, 本文使用 Python 中的 SimpleITK 工具将所有患者的影像通过近邻插值^[28] (NearestNeighbor) 的方法从原来较小的体素间隔统一到一个较大的体素间隔 (0.05263 * 0.05263 * 0.07), 该操作不会产生伪体素信息。另外为了方便模型输入和训练且不丢失信息, 将所有样本的 ROI 裁剪到相同的规格 (48 * 48 * 48)。在将图像输入网络前, 首先对图像进行 Z-score 标准化处理。

3.2.2 类平衡差异化数据增强

由于数据存在样本量不足和比例不均衡的问题, 因此设计了一种类平衡差异化数据增强机制。具体地, 首先设计一个数据增强方法池, 如图 2 所示, 其中包括翻转、平移和中心旋转等 7 种方法。针对样本量不足和比例不均衡的问题, 训练模型时会给定分别对各类别样本增广的倍数, 进而生成一个样本索引队列, 队列中包含原样本和增广样本的标记。在模型读取队列中每一个标记时, 如果识别到标记为原样本, 就直接加载原图像数据; 如果识别到标记为增广样本, 就会计算出一个由样本编号和增强编号组成的数字作为随机种子, 以该种子得到 1-7 中的一个随机整数, 接下来便会去数据增强方法池中选择对应的方法对原图像数据进行增强, 得到变换后的图像数据输入模型。该方法不仅能够显著增加样本量, 还能实现动态平衡类别, 避免了模型学习到“针对不同类别用不同的增强手段”的特点。另外, 由于引入了确定的随机种子, 确保了实验可复现性。数据集的具体分布和划分情况如表 2 所列, 括号中的数据表示数据增强后的样本量。

¹⁾ <http://www.slicer.org>

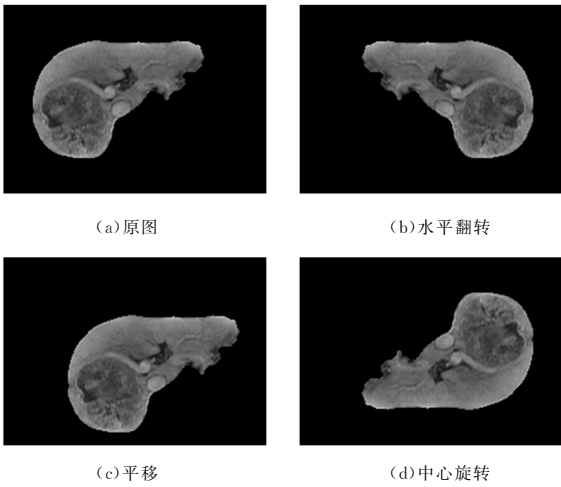


图2 数据增强示例

Fig. 2 Example of data augmentation

表2 数据集的分布和划分

Table 2 Distribution and division of datasets

	Positive	Negative	Total
Training	96(576)	292(584)	388
Validation	33	97	130
Test	33	97	130
Total	162	486	648

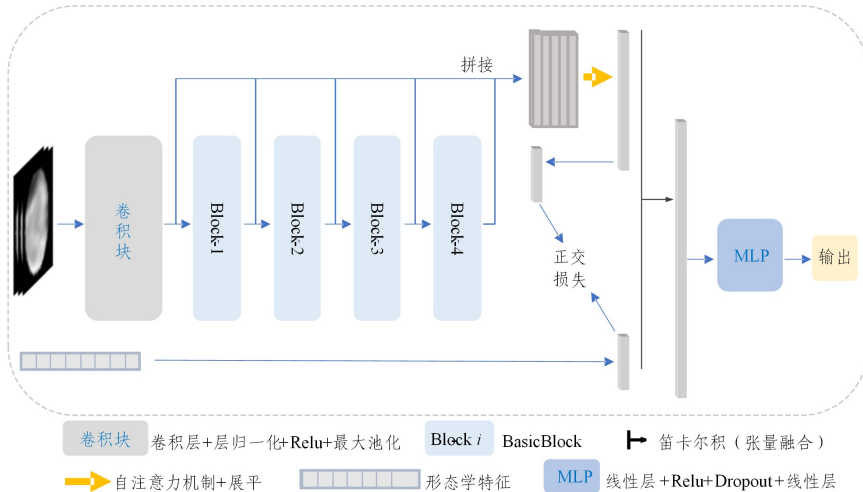


图3 模型整体框架

Fig. 3 Overall framework of the proposed model

3.2.4 基于自注意力的多尺度特征融合

为了更好地捕捉与瘤周信息相关的不同尺度特征,本文采用多尺度特征融合与自注意力机制^[29]协同优化的设计思路,如图4所示。具体而言,针对骨干网络提取的4个尺度特征图,首先跨空间维度执行全局平均池化(Global Average Pooling, GAP)操作,分别生成4个一维特征向量。随后沿空间维度进行张量拼接(Tensor Concatenation),构建维度为 $5 \times N$ 的特征矩阵,形成图像的多尺度序列化表征。在此基础上,基于Transformer架构中的自注意力机制^[23]对序列化特征进行动态权重分配与上下文建模,经过全连接层映射后,最终生成维度为 N 的紧凑型图像表征向量。值得注意的是,该设计通过构建层次化的特征交互体系,不仅实现了跨尺度特征的有机融合,更通过注意力权重动态捕捉不同区域的重

3.2.3 模型方法设计

本研究将三维超声与WSI融合实现MVI的具体定位,提出了一种特征张量融合深度学习模型TensorFusion-Net(见图3),首次将MVI的空间分布特征与形态学特征进行跨模态融合,三维超声图像提供空间信息及细胞级别的细节,而形态学特征能捕捉到特定形状和纹理等特征。

该模型采用了多种特征融合策略,如多尺度融合和特征级张量融合,并设计了多模态特征融合正交损失和基于类别中心的判别性对比学习损失函数。微血管侵犯诊断需同时分析肝血管的连续走行模式(三维超声)与局部异常结构(WSI)。传统CNN的固定感受野难以捕捉血管网络的分形特征,而自注意力机制通过动态权重分配,实现了跨尺度的空间依赖建模。三维超声-WSI融合影像提供空间拓扑信息,形态学特征补充定量诊断参数。传统早期融合会导致模态特异性信息丢失,而张量融合通过高阶交互建模,保留了各模态的独立诊断价值。该方法模拟了放射科医师的联合诊断思维:先独立评估各模态特征,再综合判断侵犯程度。微血管侵犯的不同亚型(如门静脉/肝动脉侵犯)在影像表现和形态特征上具有独立性。传统方法易导致特征冗余,而正交化约束强制不同模态特征向量解耦,使超声特征主要响应脉管空间异常,形态特征聚焦边缘浸润参数。

要性差异,其理论依据符合人类视觉系统的多尺度感知特性。

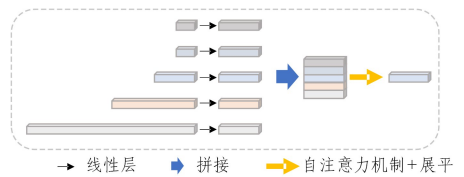


图4 基于自注意力的多尺度特征融合模块

Fig. 4 Multi-scale feature fusion module based on self-attention

3.2.5 特征级张量融合

本文提出基于张量分解的双模态特征交互融合方法,如图5所示,定义 $x_1 \in \mathbb{R}^{D_1}$ 与 $x_2 \in \mathbb{R}^{D_2}$ 分别为双模态特征向量,通过构建双线性交互张量实现模态互补。具体而言,首先对原始特征进行偏置增强处理,即:

$$\tilde{\mathbf{x}}_i = [\mathbf{x}_i; 1] \in \mathbb{R}^{D_i+1} \quad (1)$$

该操作在数学上等价于在特征空间引入偏置项,其必要性在于保持原始特征的线性完备性。继而通过外积运算生成二阶交互张量:

$$\mathbf{T} = \tilde{\mathbf{x}}_1 \otimes \tilde{\mathbf{x}}_2 \in \mathbb{R}^{(D_1+1) \times (D_2+1)} \quad (2)$$

该张量在理论层面构建了双模态特征的高阶交互空间,其秩分解特性可有效捕捉模态间的非线性关联。最终通过向量化操作 $\text{vec}(\mathbf{T}) \in \mathbb{R}^{(D_1+1)(D_2+1)}$ 获得融合特征表示。与传统的特征拼接相比,该双线性融合策略在跨模态检索任务中有效地增强了不同模态特征之间的互补性,能够捕捉它们的非线性交互信息。

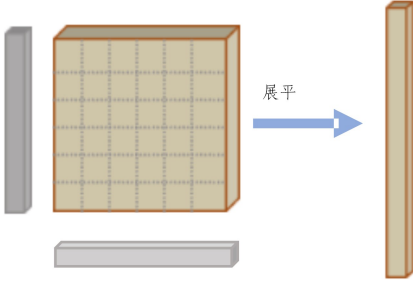


图5 张量融合

Fig. 5 Tensor fusion

3.2.6 特征正交化损失函数

本文提出基于正交约束的多模态特征优化框架。受张量融合中特征解耦理论的启发,构建了特征正交化正则项,如式(3)所示:

$$\mathcal{L}_{\text{orth}} = \frac{1}{n} \sum_{i=1}^n |\mathbf{x}_{i1}^T \mathbf{A} \mathbf{x}_{i2}|^2 \quad (3)$$

其中,投影矩阵 $\mathbf{A} \in \mathbb{R}^{m \times n}$ 为可学习参数, $\|\mathbf{A}\|_F = 1$; $\mathbf{x}_{i1} \in \mathbb{R}^m$, $\mathbf{x}_{i2} \in \mathbb{R}^n$ 分别表示双模态特征向量; $|\cdot|$ 表示绝对值。该约束项构建了正交优化空间,其几何解释为:当 $\mathcal{L}_{\text{orth}} \rightarrow 0$ 时,特征

向量在投影空间内满足 $\mathbf{A} \mathbf{x}_{i1}^T \perp \mathbf{x}_{i2}$,从而保证模态间信息的独立性与互补性。另外,用于分类任务的优化目标是交叉熵损失:

$$\mathcal{L}_{\text{CE}} = -\frac{1}{n} \sum_{i=1}^n [y_i \log(p_i) + (1-y_i) \log(1-p_i)] \quad (4)$$

其中, y_n 为真实标签, p_n 为预测概率。最终的目标函数为二者的线性组合:

$$\mathcal{L}_{\text{total}} = \mathcal{L}_{\text{orth}} + \mathcal{L}_{\text{CE}} \quad (5)$$

4 实验

4.1 实验和评估设置

本文中,按 6:2:2 的比例将 648 张图像划分为训练集(388 张)、验证集(130 张)和测试集(130 张)。本文方法基于 PyTorch 开发,并在配备 4 块 Nvidia GeForce GTX 1080 Ti 显卡和一个 Intel Xeon E5-2678 v3 处理器的系统上运行,所有网络参数均使用 Adam 优化器进行了更新。训练时,批量大小设为 128,学习率设定为 10^{-6} ,并采用了指数衰减策略。

本文使用 3 个指标来评估模型的性能,分别是 Accuracy, F1-score 和 AUC。

$$\text{Accuracy} = \frac{TP + TN}{TP + FN + TN + FP} \quad (6)$$

$$\text{F1-score} = \frac{2 \cdot TP}{2 \cdot TP + FN + FP} \quad (7)$$

其中, T 表示真, F 表示假, P 表示阳性, N 表示阴性。

4.2 实验结果分析

表 3 列出了本文方法与多种主流方法在微血管侵犯诊断任务上的分类性能对比。可以看出,本文方法在 Acc 和 AUC 指标上显著优于传统 VGG, Inception 等基线模型。本文方法 AUC 达到 0.910,较次优方法提升了 4.6 个百分点,表明模型在处理边界模糊、类间不均衡等实际问题中具备更强的判别能力。

表3 对比实验结果

Table 3 Results of comparative experiment

模型	Acc	AUC	F1-score
3D VGG-11	0.922 (0.876~0.969)	0.862(0.763~0.947)	0.821(0.706~0.921)
3DInception-V3	0.889(0.845~0.946)	0.844(0.732~0.939)	0.772(0.640~0.877)
3DResNet-18	0.891(0.837~0.946)	0.864 (0.766~0.952)	0.750(0.612~0.870)
3D DenseNet-121	0.889(0.845~0.946)	0.827(0.713~0.928)	0.745(0.605~0.863)
3D EfficientNet-B0	0.907(0.853~0.953)	0.836(0.732~0.928)	0.769(0.634~0.877)
TensorFusion-Net	0.930(0.884~0.969)	0.910(0.860~0.953)	0.852(0.745~0.939)

本研究通过递进式实验评估各模块对模型性能的影响。如表 4 所列,消融实验结果既验证了模型设计的科学性与合理性,也明确显示每个改进模块均带来显著的性能提升。这种逐步叠加的实验设计确保每个调整都具有可量化价值,同时揭示了深度学习模型优化过程中细节设计与策略选择的重要性,为模型改进提供了可验证的方法论参考。相较于基础模型,基于注意力的多尺度特征融合方法有效整合了不同尺度的图像信息,显著提升了图像表征的全面性与表达能力。与传统的特征拼接相比,张量融合策略在跨模态检索任务中有效增强了不同模态特征之间的互补性,能够捕捉它们的非线性交互信息。

如表 4 所列,与常规的数据增强策略相比,本研究提出的差异化数据增强机制对模型性能的提升(AUC 提升 0.02)是

有显著作用的,其核心在于给任一样本采取的增强方法引入了随机性。最后,引入正交损失函数后,3 个关键指标都显示了显著的改进(Acc 提升 0.015, AUC 提升 0.026, F1-score 提升 0.045)。这表明正交性约束使网络能够学习更多独立和独特的特征表示,从而进一步提高了模型的表达能力和泛化能力。

表中每个指标括号的区间代表通过 bootstrap 自助抽样方法计算出的置信区间。具体来说,从测试集中随机抽取了 1000 个样本,每次生成一个与测试集大小相同的样本集。在这些样本上评估了指标,从中位数的 95% 范围内确定置信区间。表 4 的结果不仅充分展示了本文方法的有效性,而且显示了显著的性能提升。

表4 消融实验结果

Table 4 Results of ablation experiment

基于自注意力的 多尺度特征融合	特征级 张量融合	特征正交化 损失函数	差异化数据 增强机制	Acc	AUC	F1-score
×	×	×	√	0.891(0.837~0.946)	0.864(0.766~0.952)	0.750(0.612~0.870)
√	×	×	√	0.884(0.829~0.938)	0.876(0.786~0.961)	0.741(0.600~0.840)
√	√	×	√	0.915(0.868~0.961)	0.884(0.820~0.946)	0.807(0.687~0.909)
√	√	×	×	0.907(0.853~0.953)	0.849(0.749~0.932)	0.769(0.634~0.877)
√	√	√	√	0.930(0.884~0.969)	0.910(0.860~0.953)	0.852(0.745~0.939)

结束语 本文首先结合 WSI 与 3D US 图像之间的空间对应关系,构建多模态融合成像下的 3D US 图像,并提出了一种基于多尺度特征张量融合的深度学习方法,用于准确诊断肝细胞癌(HCC)中的微血管侵犯(MVI)。所设计的模块整合了多尺度特征,以更全面的方式增强了多模态信息的交互。通过融合不同尺度的特征,模型能够捕捉到细微细节和广泛模式,从而实现数据的更稳健和全面表示。研究结果表明了多尺度特征融合的有效型,显著提升了诊断性能。与单模态方法相比,本文的多模态融合方法在准确性上表现出显著提升。此外,实验结果表明,本研究中采用的张量外积融合方法是一种有潜力的方法,值得进一步研究,为基于术前预测 MVI 提供了新的研究思路。

但本研究仍存在不足之处,主要体现在实验数据维度层面,当前研究基于 VX2 兔肝肿瘤模型构建,受限于医学伦理审查要求及临床数据获取规范,尚未纳入人类肝癌病例进行跨物种验证,其生物学特性差异可能导致模型泛化能力受限。未来研究将构建人类肝癌数据集,推动模型向临床转化应用。

参考文献

- [1] SUNG H, FERLAY J, SIEGEL R L, et al. Global cancer statistics 2020; GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries[J]. CA: A Cancer Journal for Clinicians, 2021, 71(3): 209-249.
- [2] VILLANUEVA A. Hepatocellular carcinoma. reply [J]. The New England journal of medicine, 2019, 381(1): e2.
- [3] VITALE A, PECK-RADOSAVLJEVIC M, GIANNINI E G, et al. Personalized treatment of patients with very early hepatocellular carcinoma[J]. Journal of Hepatology, 2017, 66(2): 412-423.
- [4] ZHOU J, SUN H, WANG Z, et al. Guidelines for the diagnosis and treatment of hepatocellular carcinoma (2019 edition) [J]. Liver Cancer, 2020, 9(6): 682-720.
- [5] NJEI B, ROTMAN Y, DITAH I, et al. Emerging trends in hepatocellular carcinoma incidence and mortality [J]. Hepatology, 2015, 61(1): 191-199.
- [6] MARENCO A, ROSSO C, BUGIANESI E. Liver cancer: connections with obesity, fatty liver, and cirrhosis [J]. Annual Review of Medicine, 2016, 67(1): 103-117.
- [7] LAI C L, RATZIU V, YUEN M F, et al. Viral hepatitis B [J]. The Lancet, 2003, 362(9401): 2089-2094.
- [8] YANG L, GU D, WEI J, et al. A radiomics nomogram for preoperative prediction of microvascular invasion in hepatocellular carcinoma [J]. Liver Cancer, 2019, 8(5): 373-386.
- [9] CHAN A W H, ZHONG J, BERHANE S, et al. Development of pre and post-operative models to predict early recurrence of hepatocellular carcinoma after surgical resection [J]. Journal of Hepatology, 2018, 69(6): 1284-1293.
- [10] CHEN Z H, ZHANG X P, FENG J K, et al. Actual long-term survival in hepatocellular carcinoma patients with microvascular invasion: a multicenter study from China [J]. Hepatology International, 2021, 15: 642-650.
- [11] DONG Y, WANG Q M, LI Q, et al. Preoperative prediction of microvascular invasion of hepatocellular carcinoma: radiomics algorithm based on ultrasound original radio frequency signals [J]. Frontiers in Oncology, 2019, 9: 1203.
- [12] XU X, ZHANG H L, LIU Q P, et al. Radiomic analysis of contrast-enhanced CT predicts microvascular invasion and outcome in hepatocellular carcinoma [J]. Journal of Hepatology, 2019, 70(6): 1133-1144.
- [13] NITTA H, ALLARD M A, SEBAGH M, et al. Prognostic value and prediction of extratumoral microvascular invasion for hepatocellular carcinoma [J]. Annals of Surgical Oncology, 2019, 26: 2568-2576.
- [14] LI S, ZENG Q, LIANG R, et al. Using systemic inflammatory markers to predict microvascular invasion before surgery in patients with hepatocellular carcinoma [J]. Frontiers in Surgery, 2022, 9: 833779.
- [15] MAO S, YU X, YANG Y, et al. Preoperative nomogram for microvascular invasion prediction based on clinical database in hepatocellular carcinoma [J]. Scientific Reports, 2021, 11(1): 13999.
- [16] SHI L, QU Z, YANG Y, et al. Analysis of Related Risk Factors of Microvascular Invasion in Hepatocellular Carcinoma [J]. Computational and Mathematical Methods in Medicine, 2022, 2022(1): 8195512.
- [17] ZHOU Q, ZHOU C, YIN Y, et al. Development and validation of a nomogram combining hematological and imaging features for preoperative prediction of microvascular invasion in hepatocellular carcinoma patients [J]. Annals of Translational Medicine, 2021, 9(5): 402.
- [18] COMEN E A, BOWMAN R L, KLEPPE M. Underlying causes and therapeutic targeting of the inflammatory tumor microenvironment [J]. Frontiers in Cell and Developmental Biology, 2018, 6: 56.
- [19] WANG H, LU Y, LIU R, et al. A non-invasive nomogram for preoperative prediction of microvascular invasion risk in hepatocellular carcinoma [J]. Frontiers in Oncology, 2021, 11: 745085.

- [20] ZHAN J, WANG J, ZHANG Z, et al. Noninvasive diagnosis of significant liver inflammation in patients with chronic hepatitis B in the indeterminate phase[J]. *Virulence*, 2023, 14(1): 2268497.
- [21] CHONG H H, YANG L, SHENG R F, 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]. *European Radiology*, 2021, 31: 4824-4838.
- [22] ABI-JAOUDEH N, KRUECKER J, KADOORY S, et al. Multi-modality image fusion-guided procedures: technique, accuracy, and applications[J]. *Cardiovascular and Interventional Radiology*, 2012, 35: 986-998.
- [23] CARCINOMA VILLANUEVA A. Hepatocellular[J]. *The New England Journal of Medicine*, 2019, 380: 1450-1462.
- [24] ZHANG R, XU L, WEN X, et al. A nomogram based on bi-regional radiomics features from multimodal magnetic resonance imaging for preoperative prediction of microvascular invasion in hepatocellular carcinoma[J]. *Quantitative Imaging in Medicine and Surgery*, 2019, 9(9): 1503.
- [25] LIU B, ZENG Q, HUANG J, et al. IVIM using convolutional neural networks predicts microvascular invasion in HCC[J]. *European Radiology*, 2022, 32(10): 7185-7195.
- [26] WANG L, WU M, LI R, et al. MVI-Mind: a novel deep-learning strategy using computed tomography(CT)-based radiomics for end-to-end high efficiency prediction of microvascular invasion in hepatocellular carcinoma[J]. *Cancers*, 2022, 14(12): 2956.
- [27] PERCIE DU SERT N, HURST V, AHLUWALIA A, et al. The ARRIVE guidelines 2.0: Updated guidelines for reporting animal research[J]. *Journal of Cerebral Blood Flow & Metabolism*, 2020, 40(9): 1769-1777.
- [28] BEUTEL J. *Handbook of medical imaging* [M]. Spie Press, 2000.
- [29] VASWANI A, SHAZEER N, PARMAR N, et al. Attention is all you need[C]// *Proceedings of the 31st International Conference on Neural Information Processing Systems*. 2017: 6000-6010.



WANG Shadong, born in 2000, post-graduate. His main research interests include deep learning and medical imaging.



LU Yao, born in 1979, Ph.D, professor. His main research interests include medical artificial intelligence, computer-aided diagnosis, medical image analysis, machine learning, radiomics, and medical big data analysis.

(责任编辑:柯颖)