



DOI: 10.11817/j.issn.1672-7347.2025.250082

基于多期相增强CT的肾细胞癌病理分级深度学习算法

陈皓中, 刘军, 邓凯, 梅习龙, 彭德红, 肖恩华

(中南大学湘雅二医院放射科, 长沙 410011)

[摘要] 目的: 肾细胞癌(renal cell carcinoma, RCC)是一种肾恶性肿瘤, 严重威胁患者健康, 术前病理分级的准确判断对确定RCC的治疗方法具有重大意义。目前, 深度学习技术已成为RCC病理分级的重要方法, 但是现有方法主要基于单期相计算机断层扫描(computed tomography, CT)影像进行分析与预测, 仍存在微小病灶的漏诊、评估片面性和局部聚焦等方面的不足。本研究提出一种联合多期相增强CT影像和临床变量数据的多模态深度学习算法, 以期为预测RCC病理分级提供依据。**方法:** 首先, 该算法以平扫期、动脉期、静脉期和延迟期4个期相增强CT影像与临床变量为输入模块; 然后, 使用嵌入编码模块提取临床变量中的异构信息, 采用三维ResNet50模型捕捉多期相增强CT影像数据的空间信息; 最后, 使用Fusion模块将临床变量的特征信息与各期相CT影像的特征信息进行深度融合, 并进一步利用交叉自注意力机制实现多期相特征融合, 从而更全面捕捉患者数据中的深层次语义信息, 充分利用多模态、多期相数据的信息互补优势。纳入1 229例RCC患者的资料以验证该算法的有效性。**结果:** 该算法在准确率(83.87%)、召回率(95.04%)和F1分数(82.23%)等指标上均优于传统的影像组学方法和其他深度学习算法。**结论:** 本研究提出的算法具有较好的稳定性和敏感性, 显著提升了RCC病理分级的预测能力, 为RCC的精准诊断提供了新的思路和方法。

[关键词] 病理分级; 深度学习; 跨模态融合; 肾细胞癌; 三维ResNet

Deep learning algorithm for pathological grading of renal cell carcinoma based on multi-phase enhanced CT

CHEN Haozhong, LIU Jun, DENG Kai, MEI Xilong, PENG Dehong, XIAO Enhua

(Department of Radiology, Second Xiangya Hospital, Central South University, Changsha 410011, China)

ABSTRACT

Objective: Renal cell carcinoma (RCC) is a malignant renal tumor that poses a significant threat to patient health. Accurate preoperative pathological grading plays a crucial role in

收稿日期(Date of reception): 2025-02-25

第一作者(First author): 陈皓中, Email: haozhong.chen@csu.edu.cn, ORCID: 0009-0001-3955-5690

通信作者(Corresponding author): 彭德红, 副主任医师, Email: pdh123456@csu.edu.cn, ORCID: 0009-0009-0395-2107; 肖恩华, 教授, Email: xiaoenhua64@csu.edu.cn, ORCID: 0000-0002-3281-6384

基金项目(Foundation item): 中南大学前沿交叉项目基金(2023QYJC020)。This work was supported by the Central South University Research Program of Advanced Interdisciplinary Studies Foundation, China (2023QYJC020).

开放获取(Open access): 本文遵循知识共享许可协议, 允许第三方用户按照署名-非商业性使用-禁止演绎4.0(CC BY-NC-ND 4.0)的方式, 在任何媒介以任何形式复制、传播本作品(<https://creativecommons.org/licenses/by-nc-nd/4.0/>)。

determining the appropriate treatment for this disease. Currently, deep learning technology has become an important method for pathological grading of RCC. However, existing methods primarily rely on single-phase computed tomography (CT) imaging for analysis and prediction, which has limitations such as missing small lesions, one-sided evaluation, and local focusing issues. Therefore, this study proposes a multi-modal deep learning algorithm that integrates multi-phase enhanced CT images with clinical variable data, aiming to provide a basis for predicting the pathological grading of RCC.

Methods: First, the algorithm took four-phase enhanced CT images from the plain scan, arterial phase, venous phase, and delayed phase, along with clinical variables, as inputs. Then, an embedding encoding module was used to extract heterogeneous information from the clinical variables, and a 3-dimensional (3D) ResNet50 model was employed to capture spatial information from the multi-phase enhanced CT image data. Finally, a Fusion module deeply integrated the feature information from clinical variables and each phase's CT image features, further utilizing a cross-self-attention mechanism to achieve multi-phase feature fusion. This approach comprehensively captures the deep semantic information from the patient data, fully leveraging the complementary advantages of multi-modal and multi-phase data. To validate the effectiveness of the proposed method, a total of 1 229 RCC patients were approved by ethics review were included to train the model.

Results: Experimental results demonstrated superior performance compared to traditional radiomics and state-of-the-art deep learning methods, achieving an accuracy of 83.87%, a recall rate of 95.04%, and an F1-score of 82.23%.

Conclusion: The proposed algorithm exhibits strong stability and sensitivity, significantly enhancing the predictive performance of RCC pathological grading. It offers a novel approach for accurate RCC diagnosis and personalized treatment planning.

KEY WORDS pathological grading; deep learning; cross-modal fusion; renal cell carcinoma; 3-dimensional ResNet

肾细胞癌(renal cell carcinoma, RCC)是肾恶性肿瘤中最为常见的一种疾病,其发病率在全球范围内正持续上升。不同时期的RCC的治疗效果差异明显。I期和II期RCC的预后明显好于III、IV期,I期RCC的5年生存率为80%~95%,II期RCC的5年生存率约80%,而III和IV期的5年生存率分别是40%~60%和10%~20%^[1]。RCC的病理分级是评估患者长期生存率及预测远处转移风险的关键信息,该分级的精确性对指导临床治疗方案的选择、优化治疗策略以及提升患者生存质量均具有极其重要的参考价值。

目前,计算机断层扫描(computed tomography, CT)与磁共振成像(magnetic resonance imaging, MRI)是RCC诊断与全面评估流程中的2种主流无创性检查手段。通过扩散加权成像和动态对比增强的MRI技术在RCC诊断中有高检出率与诊断精确度的优

势^[2-3],检出率在85%以上,但是其在钙化显示、流动相关伪影控制、扫描时长、禁忌证限制及维护成本等方面仍存在局限性。相对而言,CT对RCC钙化病变具有高敏感性,能更精准地显示肾中的钙化细节,这对于RCC的早期筛查、良恶性鉴别和术前规划具有参考价值^[4]。此外,CT因其扫描速度快和图像质量高而具有快速诊断的优势。但CT在针对等密度肿瘤与正常组织间的鉴别和在淋巴结状态评估方面仍然具有不足^[5],对肾癌的诊疗仍然高度依赖于病理医师的丰富经验与主观判断^[6]。

机器学习(machine learning, ML)与深度学习(deep learning, DL)方法的快速发展为RCC的智能辅助诊断带来了机遇与思路。ML方法主要以影像组学技术为主,通过从医学影像中提取大量定量特征,对肿瘤表型进行量化,进一步筛选和分类这些高通

量特征, 构建具有高预测价值的影像表型模型, 从而为肿瘤的诊断与预后提供关键信息。Ding等^[7]基于CT纹理与非纹理特征构建的回归模型, 成功预测了肾透明细胞癌(clear cell renal cell carcinoma, ccRCC)的Fuhrman分级。黎良山等^[8]基于多参数MRI影像组学模型在预测ccRCC的WHO/ISUP分级时, 准确率达到了87.5%。然而, 这些研究过度依赖人工特征(如形状、强度和纹理等特征), 无法全面捕捉图像中的信息, 从而影响模型性能。

以卷积神经网络(convolutional neural network, CNN)为核心的DL方法具有强大的图像信息捕捉能力, 在医学影像分析领域中展现出巨大潜力^[9]。一项研究^[10]针对ccRCC分级预测构建了融合自监督预训练方法的DL框架, 该框架引入混合损失策略和样本重加权技术来识别高分级ccRCC患者, 结果显示该框架的曲线下面积(area under the curve, AUC)高达0.864, 验证了DL方法在ccRCC分级评估中的有效性。Wang等^[11]的研究也验证了CNN模型在肾肿瘤识别与分类中的有效性, 体现了DL方法在良恶性肿瘤识别方面的巨大潜力。此外, Aziz等^[12]基于DL方法和CT图像数据实现了肾癌病理分级及分子亚型的精准预测, 为患者的个体化治疗提供了科学依据。Chanchal等^[13]提出一种改进的DL框架, 采用半自动与全自动方法分割肾肿瘤, 并使用CNN模型进行分类, 结果显示该框架在良恶性肿瘤识别方面具有较高的准确性。这些方法有效推动了RCC临床决策模式从经验判断向数据驱动的转变。

随着医疗设备的更新, 多模态影像数据的获取和利用成本随之降低, 使用多模态或多源数据已成为当前预测研究的焦点, 这为RCC的临床治疗与管理提供了技术支撑^[14]。一方面, 基于不同设备获取的多模态数据(如CT与MRI)可以整合解剖结构(CT钙化敏感性)与功能代谢(MRI扩散加权成像)特征, 覆盖肿瘤形态、血供、组织微环境等多维度信息, 弥补单一设备在钙化显示、伪影控制或软组织分辨率上的局限性。如Wang等^[15]融合CT与MRI图像数据、Luo等^[16]集成PET与CT图像数据, 均在医学图像和疾病预测任务中取得较好的结果。另一方面, 针对同一设备获取的不同期相的CT影像数据(如平扫期、动脉期、静脉期、延迟期)可以捕捉肿瘤在造影剂动态分布过程中的互补特征, 从而揭示肿瘤血管生成、渗透性及代谢活性的动态变化规律, 避免单期相影像因静态成像导致的微小病灶漏诊或病理分级评估片面性问题。一项研究^[17]以3期CT影像数据为基础, 采用ResNet50网络从中提取语义特征, 并使用稀疏

贝叶斯极限学习机作为分类器, 结果显示融合3期图像特征的模型优于单期模型。刘华云^[18]使用ResNet50的预训练模型提取影像组学特征和DL特征, 结合临床因子构建了动态融合模型, 并取得较好的结果。上述研究结果表明, 利用多设备或多模态数据可以有效提升疾病诊断的准确性, 为术前风险分层提供了有力依据。

尽管DL技术在多模态影像融合领域取得显著进展, 但其在RCC诊疗中的实际应用仍存在单期相CT影像的信息局限性。单期相CT影像是单次造影剂注射后在特定时间窗内完成CT扫描, 仅获取平扫期、动脉期、静脉期或延迟期中单一期相的影像数据。在早期诊断环节, 平扫期CT影像中微小RCC病灶与正常组织密度相近、特征隐匿, 因此极易导致漏诊。在病情评估阶段, 不同患者的RCC即使在单期相CT影像中表现出相似的形态学特征, 但实际病理分级可能存在显著差异。上述局限性的原因在于, 单期相CT影像仅捕捉肿瘤某一时刻的静态信息, 无法动态呈现肿瘤生长过程中的血供变化及微环境特征, 导致基于单期相数据构建的模型易出现过拟合, 泛化能力受限, 难以实现病情的精准诊断与治疗方案的科学制订。此外, 单期相CT影像作为解剖学特征的单维度表征, 其空间信息局限于肾局部结构, 无法整合患者家族遗传易感性、代谢综合征等生物学信息, 进一步制约了对肿瘤发展趋势的全面评估。因此, 本研究提出一种联合多期相增强CT影像和临床变量数据的多模态DL算法, 以期用于RCC的病理分级预测。

1 资料与方法

1.1 伦理声明

本研究已获得中南大学湘雅二医院(以下简称“我院”)伦理委员会批准(审批号: LYF2023086)。

1.2 资料

回顾性分析我院2015年至2024年期间RCC患者资料。纳入标准: 1)经完整手术病理确诊为RCC; 2)病理TNM分期为I期(T1N0M0)或II期(T2N0M0); 3)肿瘤病理分级依据FWHO/ISUP分级体系; 4)术前接受多期相增强CT的检查, 其中包含平扫期、动脉期、静脉期、延迟期等期相CT检查; 5)CT图像质量符合分析标准。本研究使用的CT设备厂家型号分别为: 西门子SOMATOM Force CT、西门子SOMATOM Flash CT和GE Revolution CT, 其具体参

数见表 1。

本研究共纳入 1 229 例经手术病理确诊的 RCC 病例，临床和病理资料详尽，涵盖年龄、性别、TNM 分期、病理分级等关键信息。肿瘤病理分级依据 FWHO/ISUP 标准，细分为低级别(1~2 级)与高级别(3~4 级)^[19]。数据集随机分为训练集(861 例)、验证集(124 例)和测试集(244 例)，比例为 7:1:2。经统计分

析，3 组在年龄、性别、肿瘤大小、偏侧性、位置、病理分级、分期及数据来源上的差异均无统计学意义(均 $P>0.05$)。图 1 展示了某患者的多期相 CT 图像及其掩膜(MASK)结果，同一患者的肿瘤形状在不同期相存在差异，这表明不同期相的 CT 图像提供的信息在一定程度上可以互补。

表 1 CT 相关参数

Table 1 CT related parameters

Parameters names	Single energy device	Dual energy device
Detector collimation/mm	64×0.625	14.0×1.2 or 32.0×0.6
Tube voltage/kV	80/140	80/140 or 100/140
FOV/cm	50	26/50 or 33/50
Tube current-time product/mAs	640	499/126 or 300/232
Gantry revolution time/s	0.6	0.5
Acquisition mode	Single source, helical	Dual source, helical
Pitch	1.375	0.55 or 0.60
Reconstructed section thickness/mm	5.00	5.00
Reconstructed section interval/mm	5.00	5.00
Matrix size	512×512	512×512
Reconstruction algorithm	FBP/ASIR	FBP/SAFIRE

FOV: Field of view; FBP: Filtered back projection; ASIR: Adaptive statistical iterative reconstruction; SAFIRE: Sinogram affirmed iterative reconstruction.

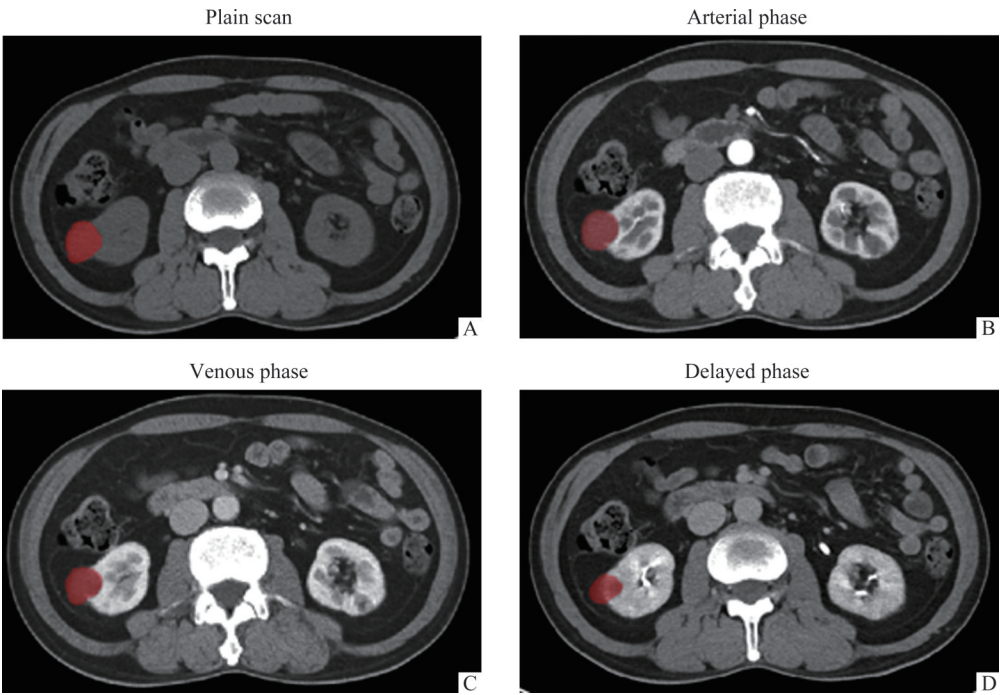


图 1 某患者的多期相 CT 图像及其 MASK 结果

Figure 1 Multi-phase CT images with MASK of a patient

1.3 方法

本研究提出一种基于多期相增强CT影像与临床特征的RCC病理分级深度学习框架,如图2所示。该框架由特征增强、特征提取和特征融合3部分组成。首先,通过旋转、缩放、翻转等多种变换操作,丰富数据特征的多样性。然后,利用具有残差结构

的三维(3-dimensional, 3D)ResNet模型捕捉各期相影像中的多视角深层语义信息,并采用异构信息提取模块对患者的异构临床变量信息进行编码。最后,将编码后的临床特征分别融入各期相影像数据的语义信息中,并进一步利用注意力机制动态融合多期相信息。

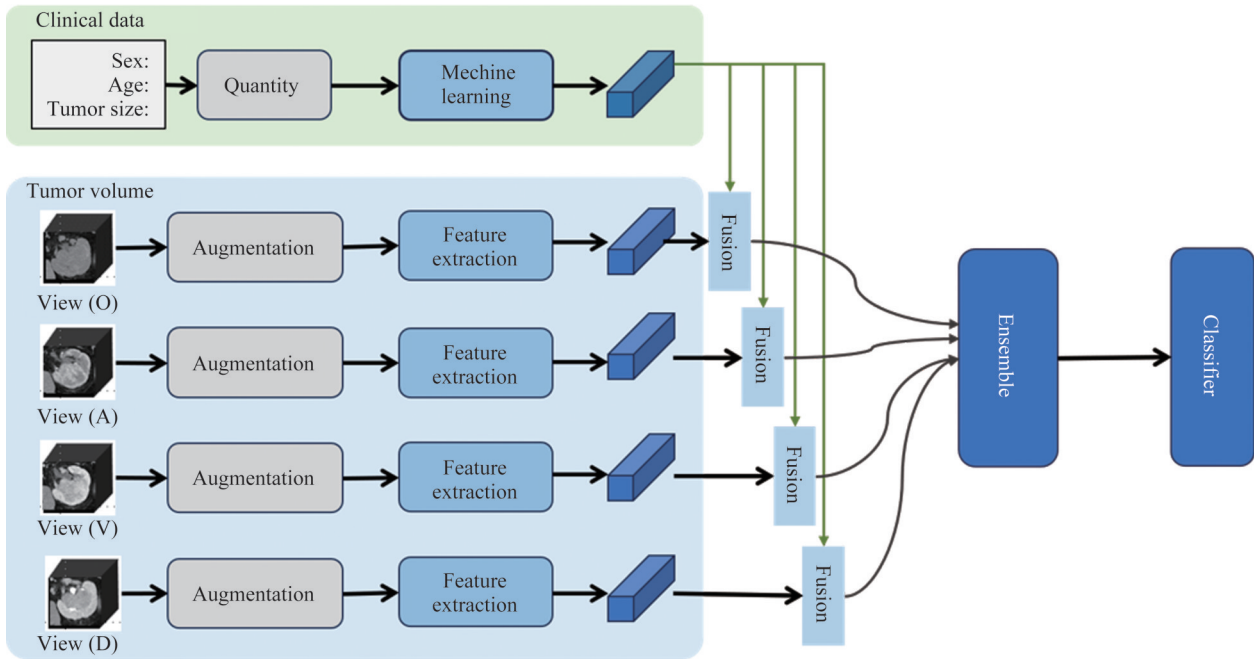


图2 本研究模型的整体架构

Figure 2 Overall architecture of this research model

This framework takes clinical data and multi-phase imaging omics data as input. First, data augmentation and feature extraction are performed separately, then the feature vectors from both parts are fused, and finally, classification prediction is achieved.

1.3.1 数据预处理

由1位具有5年CT诊断经验的放射科专家对导入3D Slicer软件(版本4.6)的RCC患者的CT影像进行精确的手动分割。为验证分割结果的准确性,所有病例完成勾画后,由1位拥有超过10年CT诊断经验的高级医师随机抽取10%的患者数据进行复核。一旦发现不合格,即会重新进行分割。值得注意的是,感兴趣区勾画者在整个分割过程中未接触患者的临床及病理资料,从而确保研究结果的客观性与准确性。

为增强模型的稳定性与泛化能力,采用多种方式对影像数据进行增强处理,具体如下:1)几何变换,包括旋转、翻转、缩放与裁剪。旋转角度设为(-15°~15°),模拟病灶在不同体位的角度变化,辅助模型学习病灶在角度偏差下的特征。水平与垂直翻转以0.5概率随机执行,增加数据多样性。缩放比例在0.8~1.2随机设定,提升模型对病灶尺寸变化的适

应性。按5%~10%的比例在图像边缘随机裁剪,引导模型关注病灶的内部特征。2)颜色变换,对图像亮度、对比度与饱和度进行调整。亮度调整采用0.8~1.2的随机因子,模拟医学影像采集时因设备、光线等因素导致的亮度差异。对比度因子在0.7~1.3随机取值,突出或弱化病灶与周围组织的差异,提升模型对病灶边界与特征的识别能力。饱和度因子控制在0.5~1.5,模拟不同成像设备的色彩还原差异,增强模型对颜色变化的鲁棒性。3)添加噪声与随机擦除。向图像中添加高斯噪声,模拟实际成像干扰。随机擦除部分图像,引导模型学习整体图像信息,提升泛化能力。

1.3.2 多模态特征提取

1.3.2.1 基于嵌入编码模块的临床变量特征提取

患者临床变量包括患者的统计学信息、症状描述、检验指标、疾病诊断等多个维度的异构信息内容,这些信息有助于医护人员决策。为此,本研究

设计了一种轻量化的嵌入-多层感知器(embedding-multiple layer perceptron, Embedding-MLP)作为嵌入编码模块组件,用于处理临床变量的异构数据 $X =$

$[X_C; X_D]$, 其中, $X_C = \{x_{c_1}, \dots, x_{c_n}\}$ 为类别型变量(共 n 维), $X_D = \{x_{d_1}, x_{d_2}, \dots, x_{d_m}\}$ 为数值型变量(共 m 维)。其结构如图3所示。

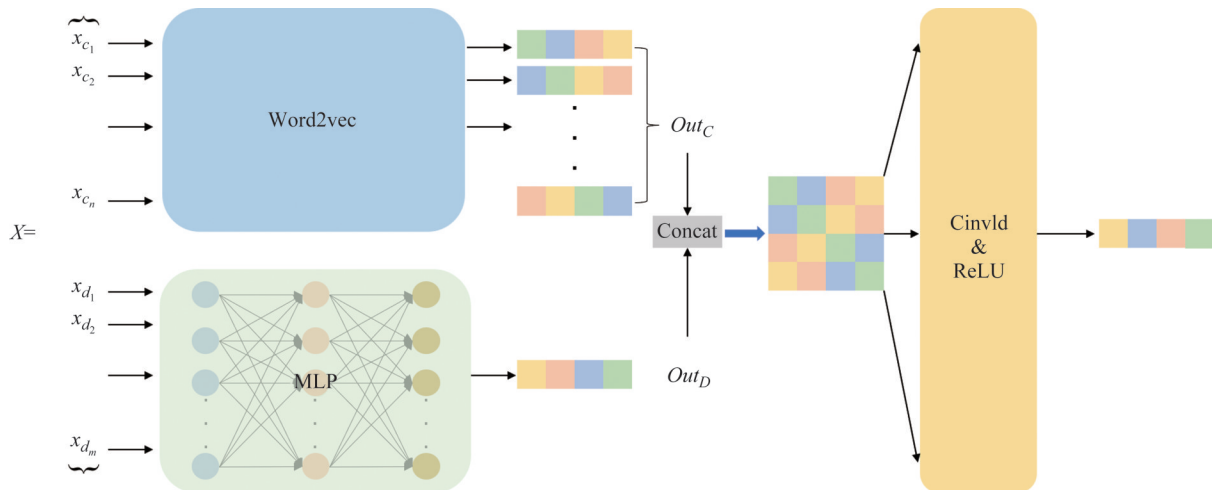


图3 Embedding-MLP模型组件的结构示意图

Figure 3 Schematic diagram of the Embedding-MLP

This module uses Word2vec to extract features from textual variables and employs MLP to process numeric variables' features. Then, the features are concatenated and fused, and finally, they are processed by a convolutional network. MLP: Multiple layer perceptron.

使用自然语言处理技术的Word2Vec词向量模型将文本信息转换为低维稠密向量,以捕捉文本特征的深层次语义信息,并在训练过程中对其进行微调,计算公式见式1。

$$Out_C = Word2vec(X_C) \quad (1)$$

针对患者的数值型变量,采用多层感知器(multiple layer perceptron, MLP)对其进行特征提取。MLP基于其非线性特征提取能力,能够深入挖掘数据背后隐藏的复杂模式,计算公式见式2。

$$Out_D = MLP(X_D) \quad (2)$$

最后,将文本特征向量与MLP提取的数值特征向量进行拼接,进一步使用卷积层融合2类特征向量,见式3。该融合过程使得模型能够充分整合患者的异构信息,从而为RCC病理分级模型提供更全面的语义信息。

$$Output_{Clinical} = ReLU(Conv1d(Concat(Out_C, Out_D))) \quad (3)$$

1.3.2.2 基于3D ResNet模型的各期相CT影像数据进行特征提取

为充分提取各期相CT影像数据中的3D特征,本研究采用3D ResNet50^[20]模型作为图像特征提取

器,分别对4个期相(平扫期、动脉期、静脉期和延迟期)的CT影像数据进行特征提取,从而捕捉影像中的多视角深层语义信息,最终得到各期相的特征向量 $Output_{CT}$ 。该模型的结构图如图4A所示。该模型采用堆叠的方式构建了含有多个残差块的特征提取器,其提取步骤如下:1)粗粒度特征提取阶段(Stage1)。该阶段从各期相增强CT影像的整体结构与全局信息中抽取出具有关键代表性的特征。首先,使用卷积核对输入影像进行卷积操作,提取初步特征;其次,对卷积后的结果进行归一化处理,加速模型收敛,减少内部协变量偏移;然后,使用ReLU激活函数捕捉特征的非线性关系;最后,使用最大池化(max pool)进行下采样,从而减少冗余信息对模型的干扰。2)细粒度特征提取阶段(Stage2~Stage5)。该阶段主要由卷积块(convolutional block, Conv Block)与恒等块(identity block, ID Block)交替堆叠实现,深度挖掘CT影像数据中的语义信息。Conv Block在结构上包含一系列卷积层,其通过不同尺寸的卷积核在影像数据上进行滑动卷积操作,能够有效提取影像不同层次与尺度的局部特征,促进特征图在空间维度与通道维度的信息融合与变换,其结构见图4B。ID Block则在保留输入特征图基本信息

的同时,通过残差连接的方式,将输入直接叠加到经过卷积操作后的输出上,缓解了深度神经网络在训练过程中的梯度消失与梯度爆炸问题,增强了模

型对深层次特征的学习与表达能力,其结构见图4C。二者相互配合,提升了细粒度特征的提取质量。

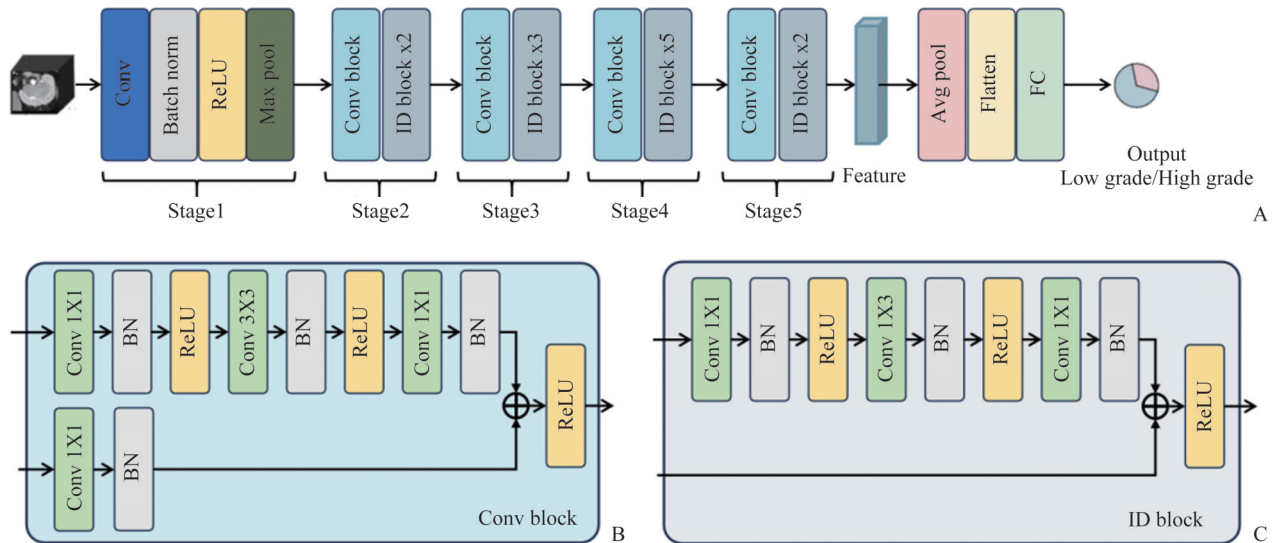


图4 3D ResNet网络图

Figure 4 3D ResNet network diagram

A: Backbone structure of the 3DResNet 50; B: Convolution module structure within A; C: Identity block structure.

Conv Block用于改变网络维度,其信息处理过程为:首先,通过 1×1 卷积调整通道数,再经批量归一化和ReLU激活函数进行非线性变换;其次,使用 3×3 卷积进一步提取特征,同样地进行批量归一化和ReLU非线性变换操作;然后,再次使用 1×1 卷积调整通道数,并进行批量归一化处理;最后,对该Block的输入信息进行 1×1 卷积处理,再与上一步的输出信息跳跃连接,通过ReLU激活输出捕捉到的深度非线性特征。ID Block的结构与Conv Block类似,但不改变输入输出维度,主要是加深网络,学习更抽象、高阶的特征。

1.3.3 跨模态信息融合

考虑到提取的特征向量来源于多模态、多期相的数据,为更有效地耦合这些特征信息,本研究设计了Fusion和Ensemble模块对上述步骤捕捉的特征进行分级融合。

1.3.3.1 针对临床特征与各期相的多模态信息融合

针对从临床变量和各期相CT影像的多模态数据中提取的特征向量 $Output_{Clinical}$ 和 $Output_{CT}$,采用Fusion模块实现特征对齐与融合,其结构如图5所示。首先,将临床特征与CT影像特征作为输入,各自进入对应的Conv Block。Conv Block内部包含卷积

层、归一化层及激活函数等组件,进一步将临床特征与CT影像特征映射至统一的特征维度,确保2种模态特征在后续融合中具备结构适配性。然后,将处理后的临床特征与CT影像特征通过拼接操作整合,在特征维度上完成多模态信息的并行合并,最大限度保留临床数据的统计特性与CT影像数据的空间特征。最后,拼接后的特征输入卷积层(convolution),借助卷积核的滑动运算进一步交叉融合2类特征,弱化模态差异,最终输出临床-CT影像融合特征CCTFF,为后续分析提供兼具临床属性与图像语义的综合特征表示,计算公式见式4。

$$CCTFF = \text{Conv}(\text{Concat}(\text{Conv}(\text{Output}_{Clinical}), \text{Conv}(\text{Output}_{CT}))) \quad (4)$$

1.3.3.2 针对多期相增强CT影像的信息融合

为实现多期相的临床-CT影像融合特征有机耦合,此处采用Ensemble模块进行特征向量的自适应动态加权融合,其结构如图5所示。具体而言,该模块以4个期相的临床-CT影像融合特征向量 $X = \{CCTFF_i, i \in [1, 4]\}$ 为输入模块,将 M 个串联堆叠的交叉自注意力机制(cross self-attention mechanism)模块用于多期相输入的自适应动态加权融合。

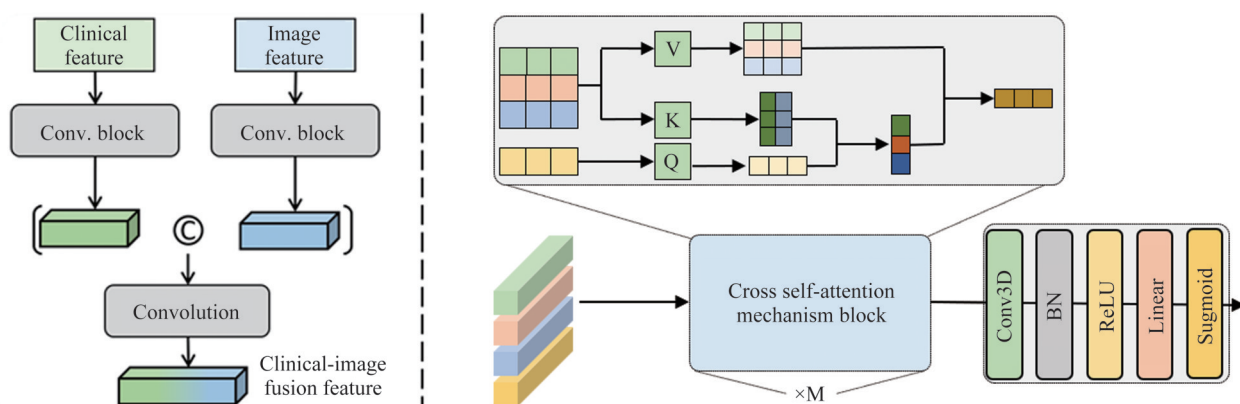


图5 Fusion模块和Ensemble模块的结构图

Figure 5 Structural diagram of the Fusion module and the Ensemble module

单个交叉自注意力机制层的计算过程如下:

1)对输入特征向量 $CCTFF_i$ 进行线性变换, 得到 Query(查询)、Key(键)和 Value(值)向量, 分别用 Q_i 、 K_i 和 V_i 表示, 见式5。其中, W^Q 、 W^K 和 W^V 分别为线性变换矩阵。

$$\begin{aligned} Q_i &= CCTFF_i \cdot W^Q \\ K_i &= CCTFF_i \cdot W^K \\ V_i &= CCTFF_i \cdot W^V \end{aligned} \quad (5)$$

2)计算第 i 个期相的查询向量 Q_i 与第 j 个期相的键向量 K_j 的相似度, 从而得到第 i 个期相相对于第 j 个期相的注意力分数(或权重), 计算公式见式6。其中, 使用 $\sqrt{d}$ 对注意力分数进行缩放, 防止其在大规模矩阵运算中数值过大或过小, d 为注意力机制中设定的维度。

$$Score(Q_i, K_j, V_i) = \text{softmax}\left(\frac{Q_i \cdot K_j^T}{\sqrt{d}}\right) \quad (6)$$

3)将所有期相的值向量 $V_j (j \in [1, 4])$ 进行加权求和, 得到第 i 个期相的自适应融合向量, 见式7。

$$Output_i = \sum_{j=1}^4 Score(Q_i, K_j, V_i) \cdot V_j \quad (7)$$

以此类推, 经过 M 次自适应加权融合后, 得到加权融合特征矩阵 $Output = [Output_1; Output_2; Output_3; Output_4]^T$, 分别对应平扫期、动脉期、静脉期和延迟期的特征向量, 这些向量已充分融合了其他期相的深层语义信息。对融合后的特征矩阵 $Output$ 采用多层卷积神经网络进行特征提取, 该网络由卷积层、批归一化层、ReLU层、Linear层和Sigmoid层依次堆叠组成。经过多层特征提取后, 最终输出RCC病理分级的概率分布, 计算公式见式8。

$$P(Output) = \text{Sigmoid}(\text{Linear}(\text{ReLU}(\text{BN}(\text{Conv1d}(Output)))))) \quad (8)$$

1.4 评估指标

本研究采用准确度、精确度、召回率和F1分数评估模型性能^[14, 21-25], 并与其他方法进行对比。计算公式见式9~12。

1)准确率(accuracy): 作为衡量模型预测正确性的直接指标, 反映了模型在全部测试样本中正确分类样本所占的比例。

2)精确度(precision): 用于衡量模型预测为阳性的样本中, 实际为阳性样本的比例。精确率越高, 意味着模型预测结果的可靠性越强。

3)召回率(recall): 该指标用于衡量模型识别实际阳性样本的能力。召回率越高, 表明模型识别出所有阳性样本的能力越强。

4)F1分数(F1 score): 是一个综合指标, 由Precision和Recall 2个指标构成, 旨在平衡模型的准确性和完整性。

$$Accuracy = \frac{TP + TN}{TP + FP + TN + FN} \quad (9)$$

$$Precision = \frac{TP}{TP + FP} \quad (10)$$

$$Recall = \frac{TP}{TP + FN} \quad (11)$$

$$F1 = 2 \times \frac{Precision \times Recall}{Precision + Recall} \quad (12)$$

其中, 真阳性(true positive, TP)表示模型正确预测为阳性的样本数, 真阴性(true negative, TN)表示模型正确预测为阴性的样本数, 假阳性(false positive, FP)表示模型错误预测为阳性的样本数, 假阴性(false negative, FN)表示模型错误预测为阴性的样本数。

除了上述指标, 还使用受试者操作特征(receiver operator characteristic, ROC)曲线的AUC和精确度-召回率AUC(precision-recall AUC, PR-AUC)来进一步评估模型性能。ROC的AUC值反映了模型在不同阈值下分类的表现, PR-AUC值则重点评估模型在处理不平衡数据集时的能力, 这2个指标越接近1, 表示模型分类效果越好。

1.5 实验设置

本研究最佳参数如表2所示。本研究基于Pycharm软件进行编码, 采用PyTorch(版本1.19.0)深度学习框架搭建模型, 并使用2个NVIDIA 3090 GPU加速计算。在Embedding-MLP中, 采用腾讯开源的Word2vec模型将各类别变量分别转换为1组200维的特征向量, 以表征其高维空间中的特征信息。针对数值型变量, 利用具有3层隐藏层的MLP对其进行编码, 神经元个数分别为[256, 512, 200]。在多期相增强CT影像的信息融合时, 利用8个交叉自注意力机制层堆叠而成的自注意力模块进行特征融合。在模型训练阶段, batch size设定为16, 最大训练轮数为500, 采用随机梯度下降法优化模型权重和偏置, 初始学习率为0.001。

此外, 为防止模型震荡不收敛, 本研究使用了学习率衰减机制逐渐减小学习率, 每训练100轮衰减1次, 衰减率为0.9。同时, 为了预防模型过拟合, 本研究使用了早停机制, patience为25。

1.6 实验内容

开展3项实验验证本研究框架的性能。1)对比实验验证模型的有效性。与多种传统影像组学方法以及其他先进的DL方法进行比较。在传统影像组学方法中, 需要提取形态学特征与纹理特征作为模型的输入模块; 在DL模型方法, 则直接使用CT影像数据作为模型的输入模块, 利用DL模型捕捉深层次的语义信息。2)消融实验用于解释利用多期相特征互补优势的必要性。此处对比分析使用多期相和单期相CT影像数据的方法在验证集上的性能。其中, 临床变量模型是指仅以临床变量为输入、以多层感知机为分类器的模型, 记为Clinical。基于平扫期、动脉期、静脉期和延迟期的CT图像训练的模型分别对应View1、View2、View3与View4 4个模型。整合模型则为本研究提出的联合多期相CT影像数据和临床变量的DL方法。3)泛化实验验证本方法的泛化能力。由于医疗数据具有高度的隐私性, 以及多期相CT数据获取的复杂性, 导致本机构以外的多期相CT数据

获取难度极高。鉴于此, 本研究采集我院2025年1~3月满足1.2纳入标准的新入院患者的数据, 对本研究方法进行泛化性验证。

表2 模型重要参数设置
Table 2 Setting important parameters for the model

参数	细节
Word2vec输出维度	200
MLP的隐藏层参数	层数为3 维度为[256, 512, 200]
Fusion模块中Conv. Block的参数	卷积核个数为256 卷积核大小为2×2 卷积核步长为2
Ensemble模块中交叉自注意力块的参数	层数为8 维度d为128
patience	25
批大小Batch size	16
3D ResNet参数	参考3D ResNet50
Embedding-MLP中的卷积核参数	卷积核个数为256 卷积核大小为4×4 卷积核步长为2
Fusion模块中Convolution层的参数	卷积核个数为256 卷积核大小为2×2 卷积核步长为2
优化器参数	优化器为SGD 初始学习率为0.001 动量为0.9 衰减率为0.9
损失函数	交叉熵损失
最大训练轮数	500

2 结 果

2.1 模型对比分析

本研究方法对比模型的评估结果如表3所示, 本研究提出的基于多期相增强CT和临床变量的DL方法具有最优的性能。除精确度外, 其余3个指标均明显优于所有对比模型, 尤其是召回率超出其他对比方法13.54%以上。

图6展示了各模型的推理速度和模型内存消耗情况。从推理速度角度分析, 传统的影像组学方法具有明显的优势, 其速度是其他DL方法的2倍以上。

表3 对比模型的评估结果

Table 3 Evaluation results of the comparison model

方法类别	模型	准确率/%	精确度/%	召回率%	F1 分数/%
传统影像组学方法	SVM	73.56	66.93	74.88	70.68
	Logistical	74.96	68.63	79.55	73.69
	RF	75.13	69.34	80.11	74.34
	MLP	74.99	70.12	77.34	73.55
深度学习方法	RCNN	78.58	69.38	80.47	74.51
	ResNet34	79.10	72.31	81.50	76.63
	sECANet	81.65	74.32	80.87	77.46
	RepVGG-A0	80.48	75.12	81.02	77.96
	DC-Net	82.13	75.68	81.07	78.28
	本研究方法	83.87	72.46	95.04	82.23

SVM：支持向量机；RF：随机森林；MLP：多层感知机。

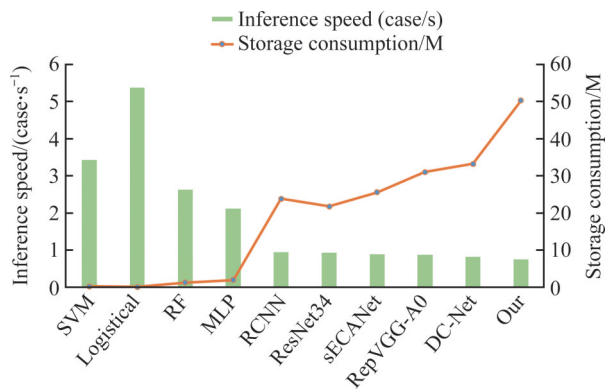


图6 各模型的推理速度及模型存储消耗量的对比
Figure 6 Inference speed and model storage consumption for each model

2.2 消融实验分析

性能分析结果显示：本研究方法明显优于仅使用单期相CT影像数据或临床变量的方法(表4)。临床变量模型准确率为82.26%，优于4个单一视角模型(View1、View2、View3和View4)的准确率，但低于整合模型的准确率(83.87%)。在精确度方面，各模型差异较大，使用平扫期CT影像数据时模型的精确度最大，达到了88.57%，本研究方法的精确率相对较低。在召回率方面，本研究方法取得最优值(95.04%)，超过其他方法10.73%以上。在关键的F1分数上，本研究方法的取得最大值，为82.23%。ROC曲线和PR曲线结果显示：本研究方法的ROC AUC(0.97)和PR AUC(0.97)均为最优，且明显高于其

他模型(图7)。这表明整合模型在区分RCC病理分级方面具有更高的准确性和鲁棒性。

表4 临床变量、View1、View2、View3、View4以及整合模型在验证集中的性能

Table 4 Performance of clinical variables, View1, View2, View3, View4, and the integrated model on the validation set

Method	Accuracy/%	Precision/%	Recall/%	F1 score/%
Clinical	82.26	75.44	84.31	79.63
View1	80.65	88.57	60.78	72.09
View2	79.03	85.71	58.82	69.77
View3	75.00	66.13	80.39	72.57
View4	81.96	81.25	50.98	62.65
Ours	83.87	72.46	95.04	82.23

2.3 模型泛化能力分析

泛化实验的有效患者数量为20例，其中低级别患者12例，高级别患者8例，年龄为(55±22)岁。本研究方法的分类结果混淆矩阵如图8所示。该批患者预测结果的准确率、精确率、召回率与F1分数分别为85%、77.78%、87.5%和82.35%。经比较，该评估结果与2.1的结果基本一致，这验证了本研究方法的泛化能力与有效性。

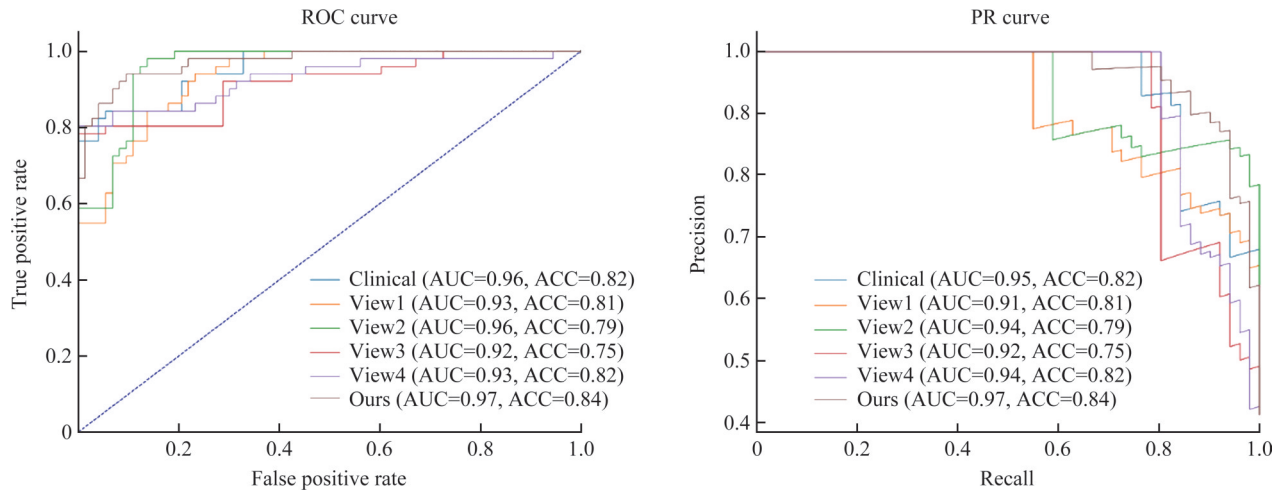


图7 单视角及多视角模型在验证集中的ROC曲线和PR曲线

Figure 7 ROC curve and PR curve of the single-phase mutual influence and multi-phase models on the validation set

ROC: Receiver operator characteristic; PR: Precision-recall; AUC: Area under the curve; ACC: Accuracy.

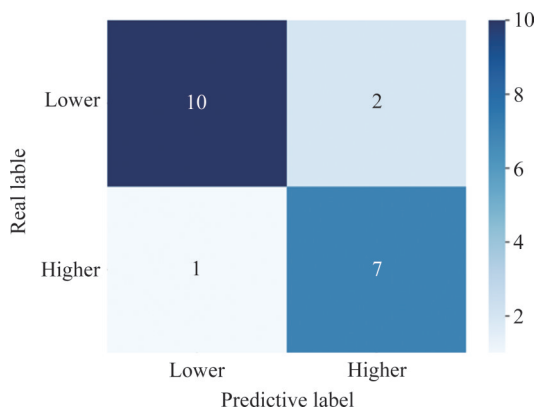


图8 20例患者的预测结果混淆矩阵

Figure 8 Confusion matrix of the predicted results for the 20 patients

3 讨论

RCC作为常见肾恶性肿瘤，准确的术前病理分级对治疗方案选择、患者预后意义重大。针对现有研究在单期相CT影像中存在的易漏诊微小病灶、评估片面、局部聚焦等局限性，本研究提出了一种基于多期相增强CT影像与临床特征的深度学习算法。具体而言，该方法以多期相(平扫期、动脉期、静脉期和延迟期)CT影像数据和临床数据为输入，分别利用嵌入模块提取临床数据的异构信息、采用3D ResNet50提取多期相CT影像数据，深度捕捉CT影像数据的语义信息，再分层融合多模态、多期相信息，最终实现RCC病理分级。本研究的主要贡献如下：1)首次整合了CT 4期相影像与临床变量数据，

通过构建跨模态特征体系解决单期相CT影像信息片面性与临床信息缺失的问题。2)通过“多模态特征提取-跨模态深度融合-自注意力加权”的分层机制，实现多模态、多期相数据的语义互补与全局关联建模。3)基于真实数据进行方法验证，实验结果验证了多期相影像与临床变量融合的必要性、模型架构的有效性及预测结果的稳健性，为模型的临床转化提供了可靠证据。

本研究探索了该方法的有效性。对比实验结果显示DL方法普遍优于传统的影像组学方法。这归因于传统的影像组学方法主要依赖预先提取的人工特征，而这类特征因无法全面涵盖RCC病理分级所需的所有复杂信息而存在局限性，造成信息损失，从而影响模型性能。相比之下，DL方法直接从大量影像数据中自动学习特征，能够挖掘出数据中隐藏的、更复杂的模式和关系，对数据的表征能力更强。通过对比DL方法的结果显示：1)以二维CNN为基础的RCNN模型在所有指标上均低于其他使用3D CNN的DL模型，这是因为与RCNN相比，3D CNN模型能够更好地捕捉CT影像的3D空间信息；2)随着使用CT影像的期相数量增加，绝大部分指标也呈现出上升趋势，这归因于它们逐渐联合使用更多期相的CT影像数据，可以有效地利用各期相CT影像的特征信息互补优势；3)使用迁移学习技术的RepVGG-A0和DC-Net取得了较好的性能，原因在于它们基于大规模图像数据提升了捕捉通用特征表示的能力，将其迁移到RCC病理分级任务中可使模型快速适应新任务，可有效提取影像中的相关特征。此外，尽管本模型结构复杂度提升，导致推理速度相对较慢以及

内存消耗量较大,但是其性能得到显著提升,这有效地推动了人工智能技术在临床中的应用。

本研究进一步探索了联合使用多期相CT和临床数据的必要性,以及该方法在新数据中的泛化能力。一方面,消融实验结果表明联合多模态、多期相数据有利于模型捕捉和利用更全面的局部信息,这使得模型的综合性能均优于其他模型。另一方面,泛化性实验结果证明了该方法的泛化能力,仅有少量新样本被模型误分类,经分析,误分类的主要原因在于异常的肿瘤细胞呈现出不规则的形状,从而对CT影像产生了噪声,导致模型捕捉局部信息时出现偏差。

总之,本研究结合临床信息与多期相增强CT影像构建深度学习模型,有效地提升了RCC病理分级预测能力,为其精准诊断提供了新方法,有助于优化临床治疗决策、改善患者预后。

在未来研究中,本研究团队将着力从以下几个关键方向推进。一方面,引入更先进的深度学习架构,如使用Transformer的自注意力机制捕捉图层间的依赖关系,进一步提升模型对CT影像特征的挖掘能力,从而提升模型性能。另一方面,探索更多临床变量对该任务的价值,如基因组数据、实验室检查结果等,以丰富患者信息维度,增强模型预测能力。此外,本研究团队将关注模型可解释性的分析技术,增强预测结果的因果分析,从而促进研究成果更好地应用于临床实践。

作者贡献声明: 陈皓中 研究设计, 论文撰写与修改, 数据分析; 刘军、邓凯、梅习龙 数据采集; 彭德红 研究设计与指导; 肖恩华 论文指导。所有作者阅读并同意最终的文本。

利益冲突声明: 作者声称无任何利益冲突。

参考文献

- [1] Wu Y, Kwon YS, Labib M, et al. Magnetic resonance imaging as a biomarker for renal cell carcinoma[J]. *Dis Markers*, 2015, 2015: 648495. <https://doi.org/10.1155/2015/648495>.
- [2] 余薇, 杨阳, 邬颖华. 肾癌的多模态MRI研究进展[J]. *磁共振成像*, 2021, 12(9): 113-115. <https://doi.org/10.12015/issn.1674-8034.2021.09.029>.
YU Wei, YANG Yang, WU Yinghua. Research progress of multimodal MRI in renal cell carcinoma[J]. *Chinese Journal of Magnetic Resonance Imaging*, 2021, 12(9): 113-115. <https://doi.org/10.12015/issn.1674-8034.2021.09.029>.
- [3] Chang K, Bai HX, Zhou H, et al. Residual convolutional neural network for the determination of *IDH* status in low- and high-grade gliomas from MR imaging[J]. *Clin Cancer Res*, 2018, 24(5): 1073-1081. <https://doi.org/10.1158/1078-0432.CCR-17-2236>.
- [4] Ishigami K, Pakalniskis MG, Leite LV, et al. Characterization of renal cell carcinoma, oncocytoma, and lipid-poor angiomyolipoma by unenhanced, nephrographic, and delayed phase contrast-enhanced computed tomography[J]. *Clin Imaging*, 2015, 39(1): 76-84. <https://doi.org/10.1016/j.clinimag.2014.09.008>.
- [5] Rossi SH, Prezzi D, Kelly-Morland C, et al. Imaging for the diagnosis and response assessment of renal tumours[J]. *World J Urol*, 2018, 36(12): 1927-1942. <https://doi.org/10.1007/s00345-018-2342-3>. [PubMed]
- [6] 焦攀攀, 杨瑞, 郑庆源, 等. 深度学习技术在肾癌诊断中的应用[J]. *武汉大学学报(医学版)*, 2025, 46(3): 321-326. <https://doi.org/10.14188/j.1671-8852.2023.0287>.
JIAO Panpan, YANG Rui, ZHENG Qingyuan, et al. Application of deep learning technology in diagnosis of renal cell carcinoma[J]. *Medical Journal of Wuhan University*, 2025, 46(3): 321-326. <https://doi.org/10.14188/j.1671-8852.2023.0287>.
- [7] Ding J, Xing Z, Jiang Z, et al. CT-based radiomic model predicts high grade of clear cell renal cell carcinoma[J]. *Eur J Radiol*, 2018, 103: 51-56. <https://doi.org/10.1016/j.ejrad.2018.04.013>.
- [8] 黎良山, 洪敏萍, 吕海娟, 等. 基于MRI影像组学模型术前预测肾透明细胞癌WHO/ISUP分级的研究[J]. *现代实用医学*, 2023, 35(6): 803-806.
LI Liangshan, HONG Minping, LÜ Haijuan, et al. Study on preoperative prediction of WHO/ISUP grade of renal clear cell carcinoma based on MRI imaging model[J]. *Modern Practical Medicine*, 2023, 35(6): 803-806.
- [9] 刘蕾蕾, 柳林. 影像组学和深度学习在胰腺癌诊断中的研究进展[J]. *中国实验诊断学*, 2024, 28(6): 754-757. <https://doi.org/10.3969/j.issn.1007-4287.2024.06.026>.
LIU Xulei, LIU Lin. Research progress of imaging and deep learning in the diagnosis of pancreatic cancer[J]. *Chinese Journal of Laboratory Diagnosis*, 2024, 28(6): 754-757. <https://doi.org/10.3969/j.issn.1007-4287.2024.06.026>.
- [10] Xu L, Yang C, Zhang F, et al. Deep learning using CT images to grade clear cell renal cell carcinoma: development and validation of a prediction model[J]. *Cancers (Basel)*, 2022, 14(11): 2574. <https://doi.org/10.3390/cancers14112574>.
- [11] Wang Z, Zhang X, Wang X, et al. Deep learning techniques for imaging diagnosis of renal cell carcinoma: current and emerging trends[J]. *Front Oncol*, 2023, 13: 1152622. <https://doi.org/10.3389/fonc.2023.1152622>.
- [12] Aziz MA, Javadian F, Mathew SS, et al. Deep learning approach for renal cell carcinoma detection, subtyping, and grading[C]//2024 IEEE International Conference on Image Processing (ICIP). October 27-30, 2024, Abu Dhabi, United Arab Emirates. IEEE, 2024: 3861-3867. <https://doi.org/10.1109/ICIP51287.2024.10647236>.

- [13] Chanchal AK, Lal S, Kumar R, et al. A novel dataset and efficient deep learning framework for automated grading of renal cell carcinoma from kidney histopathology images[J]. Sci Rep, 2023, 13(1): 5728. <https://doi.org/10.1038/s41598-023-31275-7>.
- [14] Gu Y, Wu Q, Zou J, et al. Multi-modal clear cell renal cell carcinoma grading with the segment anything model[J]. Multimed Syst, 2024, 31(1): 19. <https://doi.org/10.1007/s00530-024-01602-7>.
- [15] Wang F, Chen Q, Chen Y, et al. A novel multimodal deep learning model for preoperative prediction of microvascular invasion and outcome in hepatocellular carcinoma[J]. Eur J Surg Oncol, 2023, 49(1): 156-164. <https://doi.org/10.1016/j.ejso.2022.08.036>.
- [16] Luo Y, Yang Q, Hu J, et al. Preliminary study on detection and diagnosis of focal liver lesions based on a deep learning model using multimodal PET/CT images[J]. Eur J Radiol Open, 2024, 14: 100624. <https://doi.org/10.1016/j.ejro.2024.100624>.
- [17] 姚嫦吟. 基于CT图像的深度学习模型精准预测肾透明细胞癌手术预后的研究[D]. 湛江: 广东医科大学, 2024. YAO Changyin. Predicting the prognosis of clear cell renal cell carcinoma after surgery using CT-based deep learning model [D]. Zhanjiang: Guangdong Medical University, 2024.
- [18] 刘华云. 基于人工智能的CT影像组学模型对肾透明细胞癌Leibovich风险组的预测研究[D]. 重庆: 重庆医科大学, 2024. <https://doi.org/10.27674/d.cnki.gcyku.2024.001171>. LIU Huayun. Artificial intelligence-based CT radiomics models for prediction of the LEIBOVICH RISK groups in renal clear cell carcinoma[D]. Chongqing: Chongqing Medical University, 2024. <https://doi.org/10.27674/d.cnki.gcyku.2024.001171>.
- [19] Moch H, Cubilla AL, Humphrey PA, et al. The 2016 WHO classification of tumours of the urinary system and male genital organs-part A: renal, penile, and testicular tumours[J]. Eur Urol, 2016, 70(1): 93-105. <https://doi.org/10.1016/j.eururo.2016.02.029>.
- [20] He K, Zhang X, Ren S, et al. Deep residual learning for image recognition[C]//2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Las Vegas, NV, USA. IEEE, 2016: 770-778. <https://doi.org/10.1109/CVPR.2016.90>.
- [21] 赵经纬. 基于CT的影像组学和深度学习在肾癌术前评估中的研究[D]. 北京: 北京协和医学院, 2024. <https://doi.org/10.27648/d.cnki.gzxhu.2024.000328>. ZHAO Jingwei. CT-based radiomics and deep learning in preoperative evaluation of renal cell carcinoma[D]. Beijing: Peking Union Medical College, 2024. <https://doi.org/10.27648/d.cnki.gzxhu.2024.000328>.
- [22] 张文娟, 张利文, 邓娟, 等. 进展期胃癌生存预测: 基于增强CT深度学习模型的构建[J]. 放射学实践, 2024, 39(4): 488-495. <https://doi.org/10.13609/j.cnki.1000-0313.2024.04.010>. ZHANG Wenjuan, ZHANG Liwen, DENG Juan, et al. The value of deep learning in the construction of survival prediction models for advanced gastric cancer based on enhanced CT[J]. Radiologic Practice, 2024, 39(4): 488-495. <https://doi.org/10.13609/j.cnki.1000-0313.2024.04.010>.
- [23] 李冬明. 基于CT的深度学习模型预测肾透明细胞癌病理分级的研究[D]. 保定: 河北大学, 2022. <https://doi.org/10.27103/d.cnki.ghebu.2022.001646>. LI Dongming. Prediction of pathological grading of clear cell renal cell carcinoma using deep learning model based on CT images[D]. Baoding: Hebei University, 2022. <https://doi.org/10.27103/d.cnki.ghebu.2022.001646>.
- [24] 杨昆, 常世龙, 王尉丞, 等. 基于sECANet通道注意力机制的肾透明细胞癌病理图像ISUP分级预测[J]. 电子与信息学报, 2022, 44(1): 138-148. YANG Kun, CHANG Shilong, WANG Yucheng, et al. Predict the ISUP grade of clear cell renal cell carcinoma using pathological images based on sECANet channel attention[J]. Journal of Electronics & Information Technology, 2022, 44(1): 138-148.
- [25] Bai Y, An ZC, Li F, et al. Deep learning using contrast-enhanced ultrasound images to predict the nuclear grade of clear cell renal cell carcinoma[J]. World J Urol, 2024, 42(1): 184. <https://doi.org/10.1007/s00345-024-04889-3>.

(责任编辑 田朴)

本文引用: 陈皓中, 刘军, 邓凯, 梅习龙, 彭德红, 肖恩华. 基于多期相增强CT的肾细胞癌病理分级深度学习算法[J]. 中南大学学报(医学版), 2025, 50(4): 651-663. DOI:10.11817/j.issn.1672-7347.2025.250082

Cite this article as: CHEN Haozhong, LIU Jun, DENG Kai, MEI Xilong, PENG Dehong, XIAO Enhua. Deep learning algorithm for pathological grading of renal cell carcinoma based on multi-phase enhanced CT[J]. Journal of Central South University. Medical Science, 2025, 50(4): 651-663. DOI:10.11817/j.issn.1672-7347.2025.250082