Hans Journal of Biomedicine 生物医学, 2025, 15(5), 1012-1023
Published Online September 2025 in Hans. https://www.hanspub.org/journal/hjbm
https://doi.org/10.12677/hjbm.2025.155108

Hans汉斯

结直肠癌病理类型的多模态融合分类模型

廖南清[1*]，张祁新[2]

[1]中山大学第一附属医院广西医院检验科，广西 南宁
[2]首都师范大学附属育新学校，北京

收稿日期：2025年8月22日；录用日期：2025年9月11日；发布日期：2025年9月19日

摘 要

结直肠癌发病率和死亡率持续上升，提高该疾病的诊断准确率成为重要议题。医学图像在诊断、治疗及健康管理等方面发挥着重要作用，深度学习技术在医学图像识别方面取得了显著进展，然而依靠单一的医学图像进行诊断则存在深层组织成像受限、人为判断差异导致等问题。深度学习结合多模态融合将成为未来的研究热点，而目前只有少数研究探索了将深度学习算法与组织病理学图像相结合的方向。本文构建了一种基于跨模态注意力机制的ResNet50双分支融合模型，将结直肠癌的组织病理切片图和MRI图进行多模态融合，以提升结直肠癌组织病理类型的识别准确率。该模型框架实现了多模态的中期和晚期融合，避免不同模态独有特征丢失，同时实现了不同模态间的高效信息交互。结果表明，进行多模态融合后，模型的整体分类性能、对关键类别的阳性样本识别能力、对9种不同类别的区分能力等都得到了提升(1.98%~14.70%)。t-SNE可视化也显示多模态融合模型在9种类别之间实现了更为有效地分离。该多模态融合模型有助于降低结直肠癌类别的误判和假阳性，有助于增强癌症区域的识别，模型及结果具有一定理论和应用价值。

关键词

结直肠癌，医学图像，多模态特征融合，注意力机制，深度学习

Multimodal Fusion Classification Model for Colorectal Cancer Pathological Types

Nanqing Liao[1*], Qixin Zhang[2]

[1]Department of Clinical Laboratory Medicine, Guangxi Hospital Division of The First Affiliated Hospital, Sun Yat-Sen University, Nanning Guangxi
[2]Capital Normal University Affiliated Yuxin School, Beijing

Received: Aug. 22nd, 2025; accepted: Sep. 11th, 2025; published: Sep. 19th, 2025

*通讯作者。

文章引用: 廖南清, 张祁新. 结直肠癌病理类型的多模态融合分类模型[J]. 生物医学, 2025, 15(5): 1012-1023.
DOI: 10.12677/hjbm.2025.155108

Abstract

The incidence and mortality rates of colorectal cancer continue to rise, making improving the diagnostic accuracy of this disease a critical issue. Medical imaging plays a vital role in diagnosis, treatment, and health management, and deep learning has achieved significant progress in medical image recognition tasks. However, relying on a single type of medical image for diagnosis is limited by constraints such as restricted deep tissue imaging and inter-observer variability. Deep learning combined with multimodal fusion will emerge as a future research hotspot, while currently only a few studies have explored combining deep learning algorithms with histopathological images. This study constructs a ResNet50 dual-branch fusion model based on a cross-modal attention mechanism, integrating colorectal cancer histopathological section images and MRI features to enhance the accuracy of histopathological type recognition. The model framework implements multimodal mid-term and late fusion, avoiding the loss of modality-specific features while enabling efficient information interaction between different modalities. Results show that after multimodal fusion, the model's overall classification performance, ability to identify positive samples of key categories, and capability to distinguish between 9 different categories were all improved (1.98%~14.70%). t-SNE visualization further demonstrated that the multimodal fusion model achieved more effective separation among the 9 categories. This multimodal fusion model helps reduce misdiagnosis and false positives of colorectal cancer pathological types, enhances the identification of cancerous regions, and holds theoretical and practical value.

Keywords

Colorectal Cancer, Medical Imaging, Multimodal Feature Fusion, Attention Mechanism, Deep Learning

Copyright © 2025 by author(s) and Hans Publishers Inc.
This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).
http://creativecommons.org/licenses/by/4.0/
Open Access

1. 引言

结直肠癌(colorectal cancer, CRC)是世界第三大发病率和第二大死亡率的恶性肿瘤[1]。根据世界卫生组织(World Health Organization, WHO)癌症统计数据，每年新诊断的直肠癌 CRC 病例超过 120 万，且相关死亡人数超过 60 万。在中国，CRC 的发病率和死亡率分别是第二位和第五位，每年新发现的 CRC 患者接近 100 万[2]。随着中国老龄化进程加快及生活方式的变化，其发病率和死亡率持续上升，因此，提高诊断的准确率成为重要议题。

组织病理切片图像是通过显微镜观察病理组织的切片后获取的图像，它揭示了病变组织细胞形态、结构等微观形态结构，是 CRC 诊断的“金标准”。但该类疾病的病理诊断复杂且存在观察者间差异。医学图像能从宏观到微观辅助医生识别人体内部结构和病变细节，在诊断、治疗及健康管理等方面发挥着重要作用，是现代医疗体系的核心组成部分。其中核磁共振成像(Magnetic Resonance Imaging, MRI)是现在常用的对 CRC 进行诊断的最重要的影像学检查手段之一。MRI 是用强磁场和射频波生成人体内部结构的高清图像，能从宏观层面显示肿瘤的位置、大小、与周围组织的解剖关系等影像特征，在疾病诊断中凭借其无创无辐射、高软组织分辨率等优势，成为了现代医学不可或缺的工具。然而，传统上这些医学图像都是由资深医生进行人工判读，不可避免存在效率低、主观差异、信息整合有限等问题，这是临床诊断面临的重大挑战。

近年来，深度学习技术在医学病理和图像识别方面取得了显著进展。大量深度学习研究证明了病理组学图像在肿瘤等疾病领域的诊断及预后预测等方面的潜力。Wang 等人[3]开发了一种使用迁移学习和Inception-V3 卷积神经网络(CNN)的方法对正常和 CRC 癌变的瓦片进行分类。A. Ben Hamida 等人[4]基于CRC 组织病理学图像 WSI，通过引入 Att-UNet 的新型增强模型进行图像分割，利用注意力门实现结肠癌组织病理学图像分割的弱监督学习，检测到阳性肿瘤组织。Mohsin 等人[5]基于弱监督深度学习，从 H&E 染色的 CRC 全玻片图像中预测 BRAF、TP53、KRAS 等关键分子突变状态，取得了良好的预测结果。这些工作都极大地提升了诊断的效率和准确率。传统的图像诊断都依赖于资深医生，而深度学习具有超越人眼的特征提取能力以及小样本学习能力，有助于实现从经验医学到数据智能医学的跨越。

目前的大多数深度学习模型都是在单一范式中运行的，然而，使用单一的医学图像进行诊断存在深层组织成像受限、人为判断差异导致的主观误差等问题。基于多模态医疗数据，深度学习方法提取并结合不同模态之间的互补、冗余和协作信息，能从多维视角的差异化特征观测同一目标。这些多维视角极大拓展了人类对疾病的认知，比单模态方法更少的错误来解决这些问题，显著提高了疾病诊断的精准度与可靠性[6]。Chen 等人[7]用深度学习模型对组织学图像和基因组特征进行多模态融合，使肿瘤分级和分子分型的准确性高于仅基于组织学和基因组数据训练的单峰深度网络。He [8]等人用深度学习算法整合放射组学(如肺癌 CT 影像特征)与分子组学数据，构建了预后预测模型，在治疗反应评估和生存期预测中表现优异。He 等人[9]探索了一种多模态深度学习框架(采用多模态紧凑双线性池化技术)，整合组织病理学图像、临床数据和分子信息，用于预测结直肠癌患者的五年生存率。深度学习结合多模态融合将成为未来的研究热点，而目前只有少数研究探索了将深度学习算法与组织病理学图像相结合的方向[9]。

本文将构建一种基于跨模态注意力机制的 ResNet50 [10]双分支融合模型，将 CRC 的组织病理切片图和 MRI 图进行多模态融合。二者结合可从宏观和微观特征出发，更全面反映疾病全貌，减少因单一图像数据信息不足导致的误诊或漏诊。

2. 材料与方法

2.1. 数据获取与预处理

下载 MedMNIST v2 数据集(https://medmnist.com)的 PathMNIST 和 OrganMNIST_axial 子集，基本信息如表 1 所示。PathMNIST 包含 9 类 CRC 的组织病理分类(表 2)，OrganMNIST 包含 11 类 MRI 腹部器官定位与识别(表 3)。PathMNIST 与 OrganMNIST 数据集在病理与解剖维度上具层级对应关系，二者联合可提升模型对癌变区域的结构化识别能力。

Table 1. Dataset information
表 1. 数据集

数据集	模态	图像类型	图像数量	类别数	应用场景描述
PathMNIST	病理切片	RGB	107,180	9	CRC 组织病理分类
OrganMNIST_axial	MRI	灰度图	58,850	11	腹部器官定位与识别

Table 2. Histopathological classification of CRC in the PathMNIST dataset
表 2. PathMNIST 数据集的 CRC 组织病理分类

类别	训练集	验证集	测试集	总计
adipose (脂肪)	9366	1041	1338	11,745
background (背景)	9509	1057	847	11,413

续表

debris (碎片)	10,360	1152	339	11,851
lymphocytes (淋巴细胞)	10,401	1156	634	12,191
mucus (黏液)	8006	890	1035	9931
smooth muscle (平滑肌)	12,182	1354	592	14,128
normal colon mucosa (正常结肠黏膜)	7886	877	741	9504
cancer-associated stroma (癌相关基质)	9401	1045	421	10,867
colorectal adenocarcinoma epithelium (结直肠腺癌上皮)	12,885	1432	1233	15,550
总计	89,996	10,004	7180	107,180

Table 3. Abdominal organ MRI localization in the OrganMNIST dataset
表 3. OrganMNIST 数据集的腹部器官 MRI 定位

类别	训练集	验证集	测试集	总计
bladder (膀胱)	1956	321	1036	3313
femur-left (左股骨)	1390	233	784	2407
femur-right (右股骨)	1357	225	793	2375
heart (心脏)	1474	392	785	2651
kidney-left (左肾)	3963	568	2064	6595
kidney-right (右肾)	3817	637	1965	6419
liver (肝脏)	6164	1033	3285	10,482
lung-left (左肺)	3919	1033	1747	6699
lung-right (右肺)	3929	1009	1813	6751
pancreas (胰腺)	3031	529	1622	5182
Spleen (脾脏)	3561	511	1884	5956
总计	34,561	6491	17,778	58,830

采用了 Z-score 标准化方法进行了图像的预处理，以确保图像数据在模型训练过程中的一致性和可比性。将 MRI 的灰度图像转为伪 RGB。对所有图像采用了调整尺寸、随机翻转、旋转(±10˚)的数据增强。

2.2. 构建基于跨模态注意力机制的 ResNet50 双分支融合模型

本文构建的多模态医学图像融合网络框架如图 1 所示，包含基于 ResNet50 [9]的双分支用于从两个模态分别提取特征，其中包含 4 个卷积层，并修改初始卷积层(3 × 3 卷积，步长 1)以保留更多空间信息，取消最大池化避免特征过早下采样。在 layer3 和 layer4 引入双向跨模态注意力模块，通过 1 × 1 卷积生成查询、键、值向量，计算跨模态注意力权重，将注意力权重应用于特征融合，增强相关特征表示，实现特征交互。双分支的输出特征图通过全局平均池化和上采样统一尺寸。最后进行特征拼接并通过全连接层学习特征，并用 Softmax 函数完成 CRC 组织病理的 9 分类预测。

用 ResNet50 的双分支网络从两个模态分别提取特征，表 4 详细展示了该特征提取模块的主要结构的参数。该模型对各模态数据首先进行独立特征提取，之后再进行中期融合及晚期融合。具体来讲，首先在第 1 和第 2 卷积层分别独立提取各模态的特征，以避免早期特征融合导致丢失不同模态独有的特征。

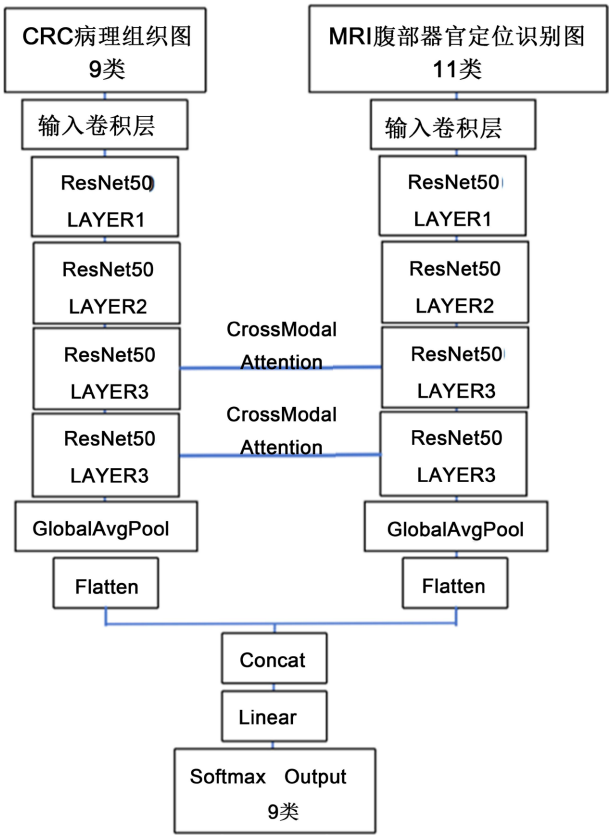


Figure 1. The workflow diagram of ResNet50 dual-branch multi-modal fusion model based on cross-modal attention mechanism
图 1. 基于跨模态注意力机制的 ResNet50 双分支多模态融合模型结构图

随后，在第 3 和第 4 卷积层使用双向轻量级注意力模块，该模块计算两种模态特征图之间的交互权重，对特征进行自适应融合(中期融合)。中期融合能保留不同模态数据间复杂的相互关系信息。注意力模块内部的可学习参数(如缩放系数 γ)通过反向传播自动优化，用以动态调整各模态特征在融合中的贡献度。该注意力模块嵌入残差块内部，在不破坏梯度流的前提下增强了跨模态交互。然后，对两个分支最终输出的特征进行拼接(晚期融合)，避免了对单模态分类结果进行后期加权平均导致的模态间交互程度较低等问题，并使得不同模态数据的特异性得以充分体现。最后，通过线性全连接层学习跨模态特征交互，并用 Softmax 函数实现 CRC 的 9 个类别(对应表 2)的组织病理分类预测。

Table 4. Convolutional layer parameters
表 4. 卷积层参数

名称	输入维度	输出维度	网络参数	**特征图尺寸**	残差块
输入卷积层	3 × 3	64 × 64	卷积核 3 × 3 步长 1	64*64	
卷积层 1	64 × 64	256 × 256	1 × 1/3 × 3/1 × 1 步长 1	64*64	3
卷积层 2	256 × 256	512 × 512	1 × 1/3 × 3/1 × 1 步长 2/1	32 × 32	4

续表

卷积层 3	512 × 512	1024 × 1024	1 × 1/3 × 3/1 × 1 步长 2/1	16 × 16	6
卷积层 4	1024 × 1024	2048 × 2048	1 × 1/3 × 3/1 × 1 步长 2/1	8 × 8	3
平均池化层	2048 × 2048	2048 × 2048	卷积核 8 × 8	1 × 1	

2.3. 模型训练

训练分为单模态预训练与多模态融合训练两个阶段。单模态训练阶段旨在让模型分别学习各模态的有效特征表示，设定训练周期为 50 (epoch)。多模态融合阶段专注于学习模态间的交互信息，设定训练周期为 40。批量大小(batch size)固定为 32。以 Adam 优化器为基础优化框架。考虑到 CRC 组织病理分类数据存在类别不平衡问题，损失函数采用交叉熵损失函数(Categorical Cross Entropy, CCE)结合标签分布感知边界(Label-distribution-aware Margin, LDAM)损失，驱动模型参数更新。采用梯度累积策略(累积步数 GRAD_ACCUM_STEPS = 4)，即每 4 个批次更新一次参数，以在有限硬件资源下模拟更大批量的训练效果。学习率调整采用 ReduceLROnPlateau 策略，当验证集准确率连续若干轮没有提升时，自动降低学习率，以避免过拟合并促进模型收敛。训练过程中持续监控验证集表现，仅保留验证准确率最高的模型作为最终结果，确保模型泛化能力。

2.4. 评价指标

模型的分类性能通过以下几个主要指标进行评估，分别是：准确率(Accuracy)、召回率(Recall)、精确率(Precision)、F1 分(F1 score)。针对多分类任务，进一步绘制 One-vs-Rest (OvR)和 One-vs-One (OvO)两种策略下的多类别 ROC 曲线，量化模型对不同类别的区分能力。

此外，通过提取模型中期跨模态注意力模块输出的特征(如 ResNet 第 3 卷积层的融合特征)，利用 t-SNE 算法进行降维可视化，分析不同模态特征在低维空间的分布聚类情况，验证跨模态注意力模块是否有效实现了多模态信息的融合与对齐。

3. 实验结果

3.1. 多模态融合模型与单模态模型的整体分类性能评价

首先对比了多模态融合模型与单模态模型的分类预测性能。多模态模型的分类性能比单模态取得明显提升。表 5 汇总了单模态与多模态融合模型的主要分类指标对比。

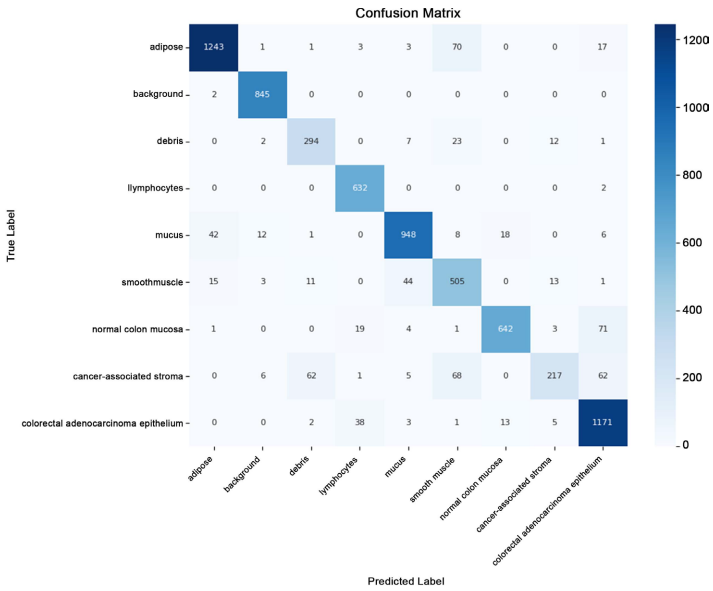
Table 5. Comparison of classification performance between single-modal and multi-modal models
表 5. 单模态与多模态模型的分类性能对比

指标	单模态模型	多模态融合模型	相对提升
Accuracy	0.9029	0.9249	+2.20%
Precision	0.8794	0.9087	+2.93%
Recall	0.871	0.9012	+3.02%
F1-Score	0.8705	0.9021	+3.16%

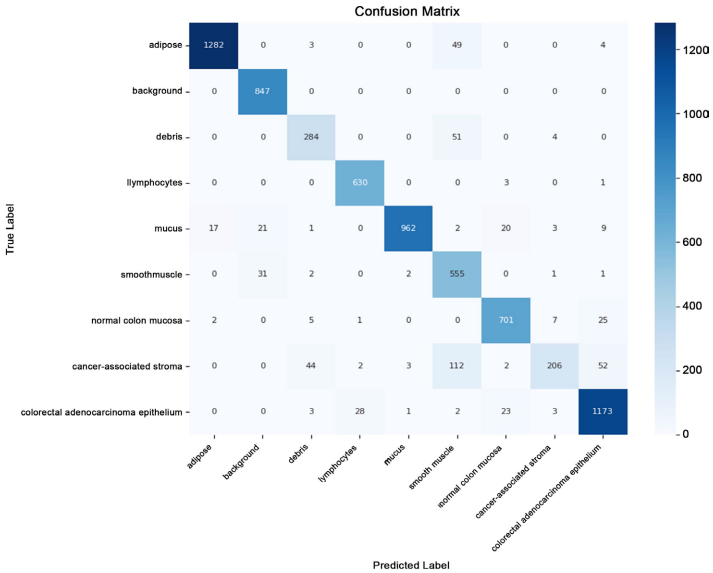
由表 5 可见，多模态融合模型的所有指标均比单模态模型的性能得到了提升，具体来说，准确率提

升了 2.2%，且提升具有极显著统计学意义($p < 0.001, z = 4.70$)，与之关联的精确率提升了 2.93%，召回率提升了 3.02%，F1 分数提高了 3.16%。这都表明，多模态融合模型对于 9 种 CRC 类型整体的正确分类能力显著升高。

模型对 9 种类别分类的混淆矩阵见图 2。由混淆矩阵可见，单模态模型错分了 683 个样本，而多模态模型将错分样本数量降低到 540 个。与单模态模型相比，多模态模型对大部分类别的样本的整体的正确识别能力都有所提升，尤其是对平滑肌和脂肪的识别能力提高显著，说明融合 MRI 的信息有助于定位平滑肌和脂肪的空间分布，弥补了单病理特征的不足。但对碎片和癌相关基质的整体样本识别略有下降，可能因这两类样本量较少(339 例和 421 例)。



(a) 单模态模型的混淆矩阵



(b) 多模态模型的混淆矩阵

Figure 2. Confusion matrix
图 2. 混淆矩阵

3.2. 多模态融合模型与单模态模型对关键类别的阳性样本识别能力评价

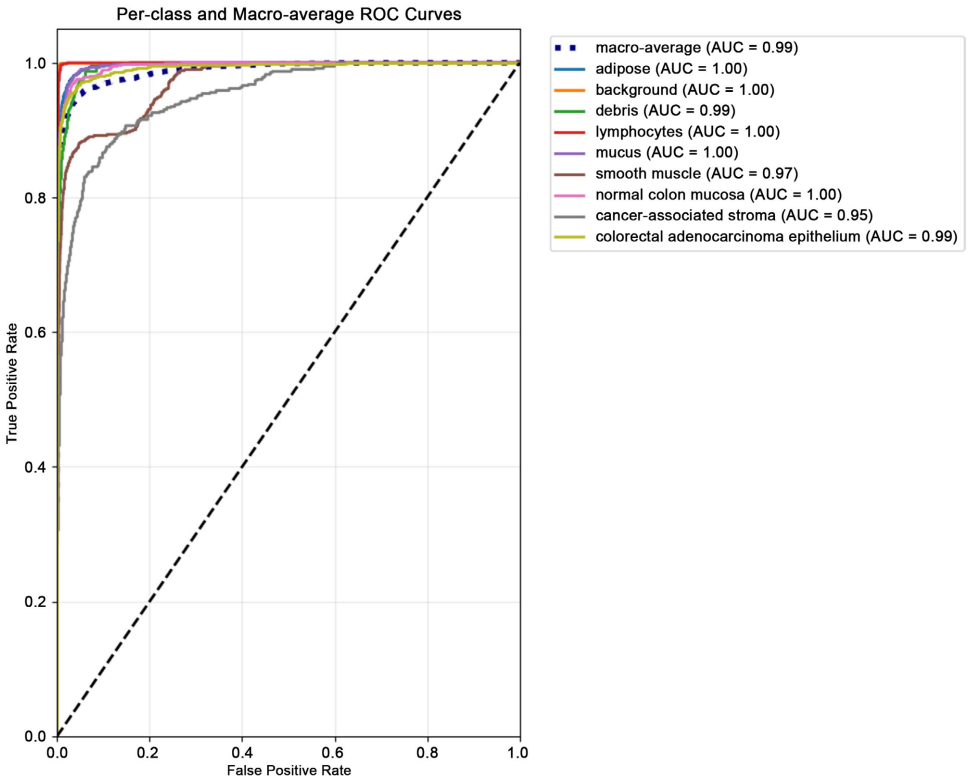
多模态模型在 CRC 关键类别(如癌相关基质、结直肠腺癌上皮、碎片、淋巴细胞)上的识别效果提升更显著(表 6)。F1-Score 是平衡精确率和召回率的综合指标，能综合反映模型对阳性样本的准确且全面的识别能力，在医疗诊断中能平衡误判与漏判。由表 6 可见，多模态模型对 CRC 相关基质的识别能力提升尤为显著，F1-Score 从 0.6097 提升至 0.7567，提升幅度达 14.70%，且具有统计学显著性($p < 0.01$, $\chi^2 = 8.42$)，这能显著提升对肿瘤微环境的检测能力以预测癌症转移风险。对结直肠腺癌上皮的识别能力提升有助于识别早期微浸润灶。对碎片的准确识别能排除干扰减少误判风险。对淋巴细胞的识别能提升对炎症反应程度的识别以及评估微卫星不稳定性。

Table 6. Key classes' F1-Score
表 6. 关键类别的 F1-Score

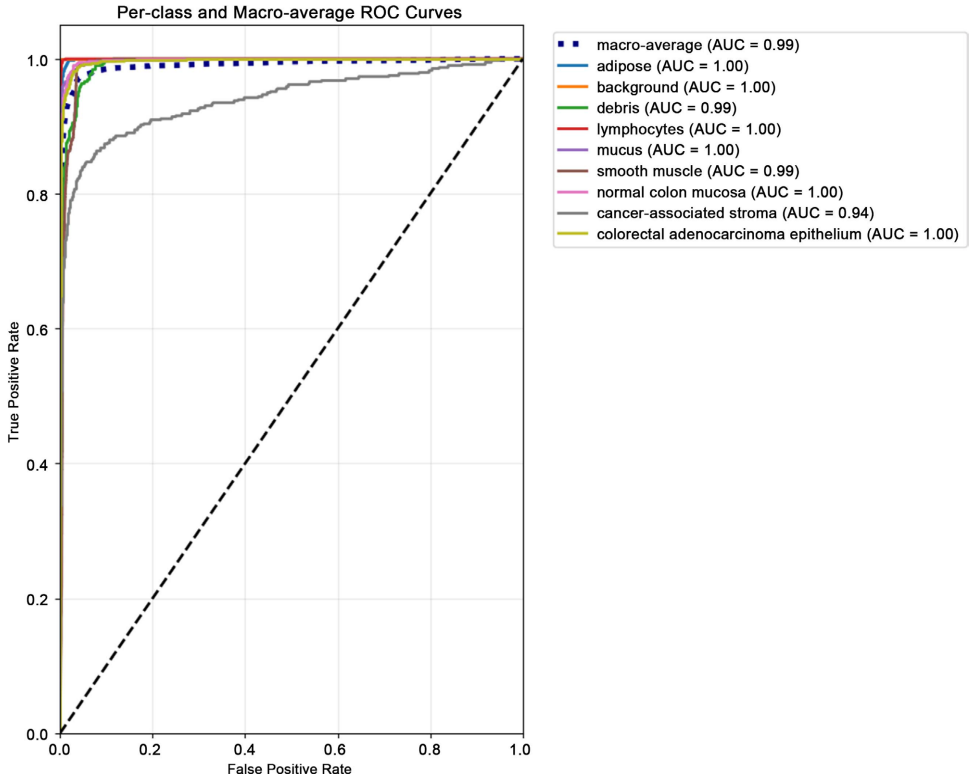
类别	单模态 F1	多模态 F1	提升幅度
cancer-associated stroma 癌相关基质	0.6097	0.7567	+14.70%
colorectal adenocarcinoma epithelium 结直肠腺癌上皮	0.9188	0.9649	+4.61%
Debris 碎片	0.8	0.8219	+2.19%
Lymphocytes 淋巴细胞	0.9593	0.9797	+2.04%

3.3. 多模态融合模型与单模态模型对 9 个类别的区分能力评价

绘制了各 CRC 类型在测试集上的 ROC 曲线，如图 3 所示。



(a) 单模态模型的 ROC 曲线



(b) 多模态融合模型的 ROC 曲线

Figure 3. Test set ROC curve
图 3. 测试集 ROC 曲线

由图 3 可见，用病理组织切片单模态进行病理分类，其疾病类型的区分能力为 90.49%，而将 MRI 腹部器官定位识别图与病理组织切片图进行特征融合后，模型的分类能力达到 92.48%，多模态融合模型的分类性能比单模态模型得到了提升(+1.98%)，且具有统计学显著性($p < 0.001$, $z = 4.28$)，宏平均指标也全部有提升，说明将 MRI 腹部器官定位识别图与病理组织切片图进行特征融合，确实可增强模型对 CRC 病理组织类别的区分能力。

此外，多模态模型的 OvR (+0.21%)和 OvO (+0.37%)也均有所提升，表明模型在多分类场景下的排序能力和鲁棒性增强。

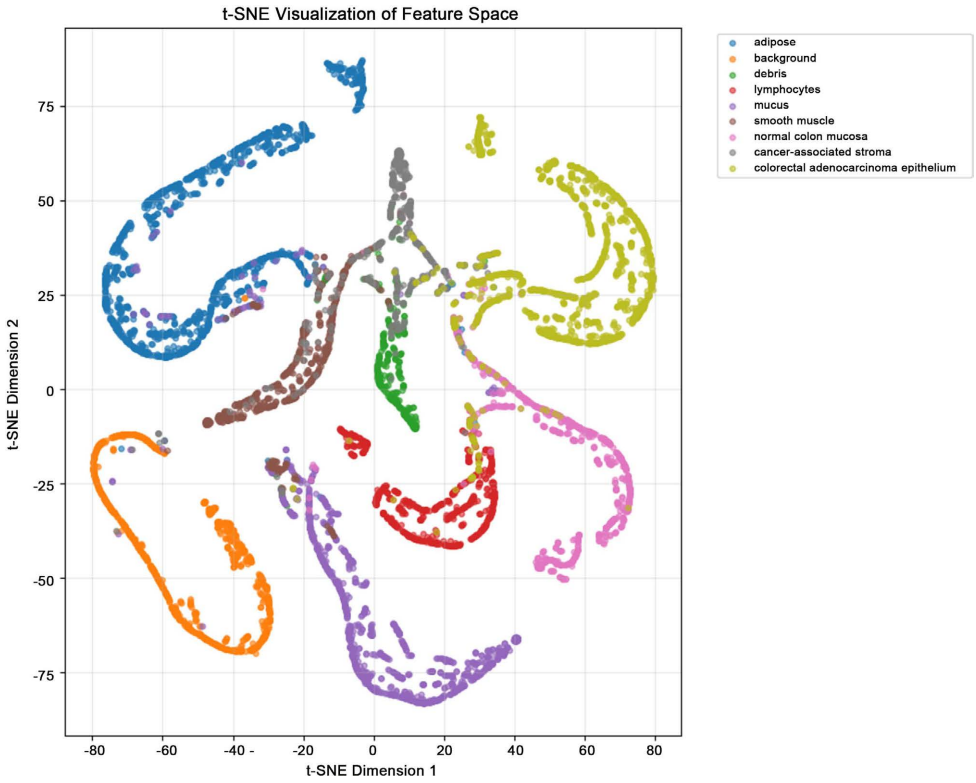
3.4. t-SNE 可视化

使用 t-SNE 可视化特征空间分布，验证融合后的特征是否有效分离不同类别样本。图 4 展示了单模态模型(a)和多模态模型(b)在特征空间上的 t-SNE 可视化对比。

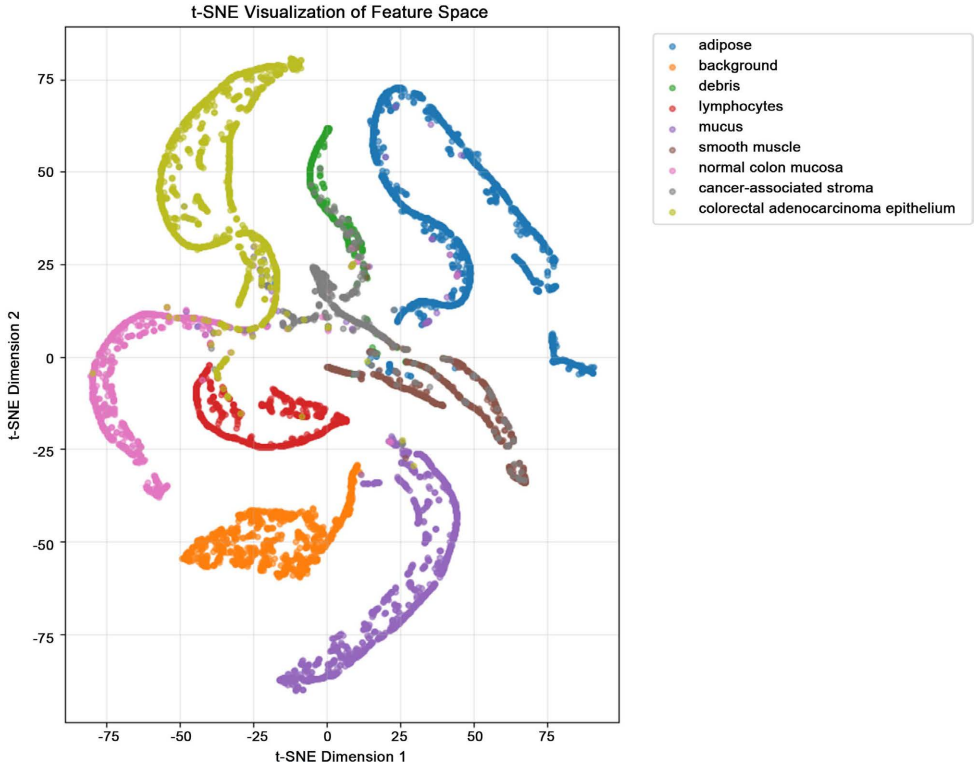
由该图可明显看到，单模态模型所得特征空间分布图 4(a)显示，部分类别样本点分散，甚至出现不同类别样本点交叉重叠，聚类簇边界模糊，难以直观识别样本所属类别。而多模态模型所得特征空间分布图 4(b)显示，9 类基本形成相对独立清晰的聚类簇，不同类别有较为明显的间隔，边界相对清晰，无明显远离所属聚类的孤立点。以上可视化结果表明，多模态融合模型在 9 种类别之间实现了更为有效地分离。

4. 总结与讨论

本文构建了一种多模态特征融合的结直肠癌组织病理分类模型，采用了跨模态注意力机制的



(a) 单模态模型的 t-SNE 图



(b) 多模态融合模型的 t-SNE 图

Figure 4. t-SNE dimensionality reduction visualization plot
图 4. t-SNE 降维可视化图

ResNet50 双分支融合框架。该模型框架实现了多模态的中期和晚期融合。首先通过双分支的 ResNet50 对两个模态进行独立特征提取，避免不同模态独有特征丢失，然后用 CrossModal Attention 模块实现了中期不同模态间的高效信息交互，同时保持训练资源开销可控。最后还对两个分支输出的特征进行了晚期拼接融合，使得不同模态数据的特异性得到充分体现。

用该模型对 CRC 的组织病理学图像和 MRI 图像进行了多模态融合，以提升 CRC 的多种病理类型的分类性能。结果表明，进行多模态融合后，模型的整体分类性能(准确率提升 2.2%)、对关键类别的阳性样本识别能力(癌相关基质识别能力提升 14.70%，结直肠腺癌上皮识别能力提升 4.61%)、对 9 种不同类别的区分能力(ROC 提升 1.98%)等都得到了提升。t-SNE 可视化也显示多模态融合模型在 9 种类别之间实现了更为有效地分离。该模型将有助于降低 CRC 类别的误判和假阳性，有助于增强对癌症区域的识别。该模型及其结果具有一定理论和应用价值。

然而，本研究尚存在以下局限：仅基于 MedMNIST v2 标准化数据集，图像尺寸限制可能无法完全反映高分辨率医学影像下的多模态耦合关系，且 PathMNIST 和 OrganMNIST_axial 子集数据不是对应样本，因此本研究仅验证了模态互补的有效性，临床应用时还需进一步建模；缺乏真实临床环境的外部验证集；多模态注意力机制对特征维度设置较敏感，需在更多或不同任务中调优参数；工作分析主要聚焦在模型的全面分类能力及与 CRC 密切相关的类别的分类能力方面，对脂肪、黏液、黏膜、平滑肌、背景等分析不足。

未来工作将聚焦引入多源高分辨率图像与非结构化数据，扩展模态维度；基于真实病例数据构建多中心、多病种的通用验证集，验证模型泛化能力；并联合医学专家优化模型反馈机制，实现可视化与诊疗建议的临床联动。

参考文献

[1] Bray, F., Laversanne, M., Sung, H., Ferlay, J., Siegel, R.L., Soerjomataram, I., *et al.* (2024) Global Cancer Statistics 2022: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries. *CA*: *A Cancer Journal for Clinicians*, **74**, 229-263. https://doi.org/10.3322/caac.21834

[2] Li, N., Lu, B., Luo, C., Cai, J., Lu, M., Zhang, Y., *et al.* (2021) Incidence, Mortality, Survival, Risk Factor and Screening of Colorectal Cancer: A Comparison among China, Europe, and Northern America. *Cancer Letters*, **522**, 255-268. https://doi.org/10.1016/j.canlet.2021.09.034

[3] Wang, K.S., Yu, G., Xu, C., Meng, X.H., Zhou, J., Zheng, C., *et al.* (2021) Accurate Diagnosis of Colorectal Cancer Based on Histopathology Images Using Artificial Intelligence. *BMC Medicine*, **19**, Article No. 76. https://doi.org/10.1186/s12916-021-01942-5

[4] Ben Hamida, A., Devanne, M., Weber, J., Truntzer, C., Derangère, V., Ghiringhelli, F., *et al.* (2022) Weakly Supervised Learning Using Attention Gates for Colon Cancer Histopathological Image Segmentation. *Artificial Intelligence in Medicine*, **133**, Article 102407. https://doi.org/10.1016/j.artmed.2022.102407

[5] Bilal, M., Raza, S.E.A., Azam, A., Graham, S., Ilyas, M., Cree, I.A., *et al.* (2021) Development and Validation of a Weakly Supervised Deep Learning Framework to Predict the Status of Molecular Pathways and Key Mutations in Colorectal Cancer from Routine Histology Images: A Retrospective Study. *The Lancet Digital Health*, **3**, e763-e772. https://doi.org/10.1016/s2589-7500(21)00180-1

[6] Boehm, K.M., Khosravi, P., Vanguri, R., Gao, J. and Shah, S.P. (2021) Harnessing Multimodal Data Integration to Advance Precision Oncology. *Nature Reviews Cancer*, **22**, 114-126. https://doi.org/10.1038/s41568-021-00408-3

[7] Chen, R.J., Lu, M.Y., Wang, J., Williamson, D.F.K., Rodig, S.J., Lindeman, N.I., *et al.* (2022) Pathomic Fusion: An Integrated Framework for Fusing Histopathology and Genomic Features for Cancer Diagnosis and Prognosis. *IEEE Transactions on Medical Imaging*, **41**, 757-770. https://doi.org/10.1109/tmi.2020.3021387

[8] He, X., Liu, X., Zuo, F., Shi, H. and Jing, J. (2023) Artificial Intelligence-Based Multi-Omics Analysis Fuels Cancer Precision Medicine. *Seminars in Cancer Biology*, **88**, 187-200. https://doi.org/10.1016/j.semcancer.2022.12.009

[9] He, B., Wang, L., Zhou, W., Liu, H., Wang, Y., Lv, K., *et al.* (2025) A Fusion Model to Predict the Survival of Colorectal Cancer Based on Histopathological Image and Gene Mutation. *Scientific Reports*, **15**, Article No. 9677. https://doi.org/10.1038/s41598-025-91420-2

[10] He, K., Zhang, X., Ren, S. and Sun, J. (2016) Deep Residual Learning for Image Recognition. 2016 *IEEE Conference on Computer Vision and Pattern Recognition* (*CVPR*), Las Vegas, 27-30 June 2016, 770-778. https://doi.org/10.1109/cvpr.2016.90