

基于深度学习的多亚型腹膜后软组织肉瘤诊断*

姜永军¹, 李红玲², 阮萍³, 陈路⁴, 李艳春⁵, 胡庆⁵, 谢功勋⁵, 孟云鹤¹

1. 中山大学人工智能学院, 广东 珠海 519082
2. 佛山市中医院, 广东 佛山 528000
3. 广西中医药大学附属瑞康医院, 广西 南宁 530000
4. 宜昌市中心人民医院, 湖北 宜昌 443000
5. 湖南省人民医院, 湖南 长沙 410000

摘要: 在仅依赖组织病理图像且缺乏额外辅助检测的情况下, 腹膜后软组织肉瘤小体积活检样本易导致观察者之间的判断差异, 影响疾病亚型的整体诊断准确性。为了解决这一问题, 从多中心收集了157张全切片图像(WSI), 涵盖去分化脂肪肉瘤、平滑肌肉瘤、恶性周围神经鞘瘤、未分化多形性肉瘤和高分化脂肪肉瘤五种疾病类别。基于上述WSIs, 提出了基于单尺度图像与多尺度图像的两模型集成方法, 并利用ResNet18、EfficientNet B7和EfficientNet V2等深度学习模型进行训练。结果表明: 两种模型集成方法均取得了较高的分类准确率, 最佳模型在块级分析中达到82.27%的总体准确率, 在全切片分析中达到80.95%。因此, 所提方法能够有效辅助病理学家在临床实践中诊断腹膜后软组织肉瘤。

关键词: 腹膜后软组织肉瘤; 深度学习; 全切片图像; 亚型诊断

中图分类号: TP391 **文献标志码:** A **文章编号:** 2097-0137(2025)03-0156-09

Deep learning-based diagnosis of multi-subtype retroperitoneal soft tissue sarcomas

JIANG Yongjun¹, LI Hongling², RUAN Ping³, CHEN Lu⁴,
LI Yanchun⁵, HU Qing⁵, XIE Gongxun⁵, MENG Yunhe¹

1. School of Artificial Intelligence, Sun Yat-sen University, Zhuhai 519082, China
2. Foshan Traditional Chinese Medicine Hospital, Foshan 528000, China
3. Ruikang Hospital Affiliated to Guangxi University of Chinese Medicine, Nanning 530000, China
4. Yichang Central People's Hospital, Yichang 443000, China
5. Hunan Provincial People's Hospital, Changsha 410000, China

Abstract: In the absence of additional auxiliary tests and relying solely on histopathological images, small-volume biopsy samples of retroperitoneal soft tissue tumors often lead to interobserver variability, impacting the overall diagnostic accuracy of disease subtypes. To address this issue, 157 whole-slide images (WSIs) were collected from multiple centers, encompassing five disease categories: dedifferentiated liposarcoma (DDL), leiomyosarcoma (LMS), malignant peripheral nerve

* 收稿日期: 2024-10-24 录用日期: 2025-01-03 网络首发日期: 2025-01-24

基金项目: 国家自然科学基金(61673390)

作者简介: 姜永军(1998年生), 男; 研究方向: 深度学习; E-mail: jiangyj8@mail2.sysu.edu.cn;
李红玲(1981年生), 女; 研究方向: 病理学; E-mail: 522918824@qq.com
(姜永军、李红玲并列第一作者)

通信作者: 孟云鹤(1978年生), 男; 研究方向: 人工智能; E-mail: mengyh7@mail.sysu.edu.cn;
谢功勋(1978年生), 男; 研究方向: 病理学; E-mail: 278180296@qq.com
(孟云鹤、谢功勋并列通信作者)

全文阅读



ZR20240309

sheath tumor (MPNST), undifferentiated pleomorphic sarcoma (UPS), and well-differentiated liposarcoma (WDLP). Based on these WSIs, two model ensemble methods were proposed: one based on single-scale images and the other on multi-scale images. Deep learning models, such as ResNet18, EfficientNet B7, and EfficientNet V2, were trained on the collected data. Results showed that both ensemble methods achieved high classification accuracy, with the best model achieving an overall accuracy of 82.27% in patch-level analysis and 80.95% in WSI-level analysis. Therefore, the proposed methods can effectively assist pathologists in the clinical diagnosis of retroperitoneal soft tissue tumors.

Key words: retroperitoneal soft tissue sarcomas; deep learning; whole slide images; subtyping

腹膜后软组织肉瘤(RPS)是一种罕见的肿瘤,约70%的病例为恶性,涵盖70多种组织学类型,主要包括脂肪肉瘤、未分化多形性肉瘤、恶性周围神经鞘瘤和平滑肌肉瘤等(Mack et al., 2022; Drohan et al., 2023)。由于其非特异性的临床表现和隐匿的发病特点, RPS往往在疾病晚期才能确诊。RPS的组织学类型和亚型多样,这种多样性不仅影响发病机制(Trojani et al., 1984),也导致了临床表现和治疗方案的显著差异。使用大体切除标本来区分含脂肪细胞成分的高分化脂肪肉瘤、去分化脂肪肉瘤与其他三种肉瘤可能不会构成显著的挑战。然而,当高分化脂肪肉瘤主要表现为硬化增生时,去分化脂肪肉瘤的主要特征在于其去分化成分,而平滑肌肉瘤、恶性外周神经鞘瘤和未分化多形性肉瘤则分别主要表现为梭形细胞或多形性形态学特征。这些实体肿瘤的局部区域可能与其他肿瘤相似,从而增加了准确诊断的难度。尽管MDM2和CDK4的FISH扩增检测以及NF1和TP53的缺失检测等,能够更精确地诊断(Zhao et al., 2023),但并非所有机构都具备这些分子检测的条件。且后腹膜肉瘤位置较深且通常体积较大,临床中越来越多地使用微创活检技术,而这些技术仅能获取有限的组织样本以进行恶性肿瘤的组织学分类和分级(Almond et al., 2019)。很多时候,这些小活检标本对进一步的判断研究来说并不足够,迫使诊断主要依赖形态学特征。在缺乏辅助检测且活检标本较小的情况下,不同肿瘤类型之间重叠的形态特征可能导致观察者间的显著判断差异,从而降低整体诊断的准确性。

深度学习通过仿生神经网络架构从数据中提取特征(Lecun et al., 2015),能够有效学习数据中的复杂非线性模式(Bengio et al., 2013)。近年来,随着全片数字扫描技术的快速发展和广泛应用,病理学与人工智能的交汇逐渐成为热点(van der Laak et al., 2021)。全切片图像包含大量信息,为深度学习在计算病理学中的应用奠定了基础(Cui

et al., 2021)。例如,深度学习已成功应用于肌层浸润性膀胱癌(Woerl et al., 2020)、横纹肌肉瘤(Milewski et al., 2023)、粘液性软组织病变(Yeung et al., 2022)和骨肉瘤(Mishra et al., 2018; Arunachalam et al., 2019)等方面。虽然已有软组织肉瘤的相关研究(Foersch et al., 2021),但其研究主要集中于软组织肉瘤的整体分类和诊断。相比之下,本文研究聚焦于腹膜后软组织肉瘤,旨在深入探讨其亚型诊断与分类。因此,针对腹膜后软组织肉瘤的特异性,本文进一步细化了疾病亚型分类,提出了两种不同的模型集成方法并成功应用于该领域。

1 数据准备与研究方案

1.1 研究伦理

本研究严格遵循《赫尔辛基宣言》的伦理原则,所有研究程序均通过佛山市中医院伦理委员会审查批准(伦理批件号: KY-2023-332-1)。数据的收集、处理和分析均严格遵循匿名化和去识别化程序,以最大限度保护参与者的个人隐私。本研究仅基于学术研究目的,相关数据不会用于其他任何非研究用途或公开传播。

1.2 数据准备

原始全切片图像(WSI, whole slide image)数据集包含157张切片,均由专业临床医生精心标注。其中,第1批数据包含115张ZYP格式图像,使用ZYP VIEWER软件转换为TIFF格式;第二批数据包含42张KFB格式图像,来源独立于第一批数据,并通过自编写的Python脚本转换为TIFF格式。如图1所示,随机从第一批切片中选择部分作为测试集1,第二批切片全部指定为测试集2。测试集1和测试集2用于切片级分析,剩余的切片则经裁剪为图像块后,随机分配至训练集、验证集或测试集(块级)中的任意一个,用于模型训练和块级分析。测试集和验证集的所有图像均采用20倍放大倍率进行转换,训练集的图像则包含10、20和40倍放大倍率。

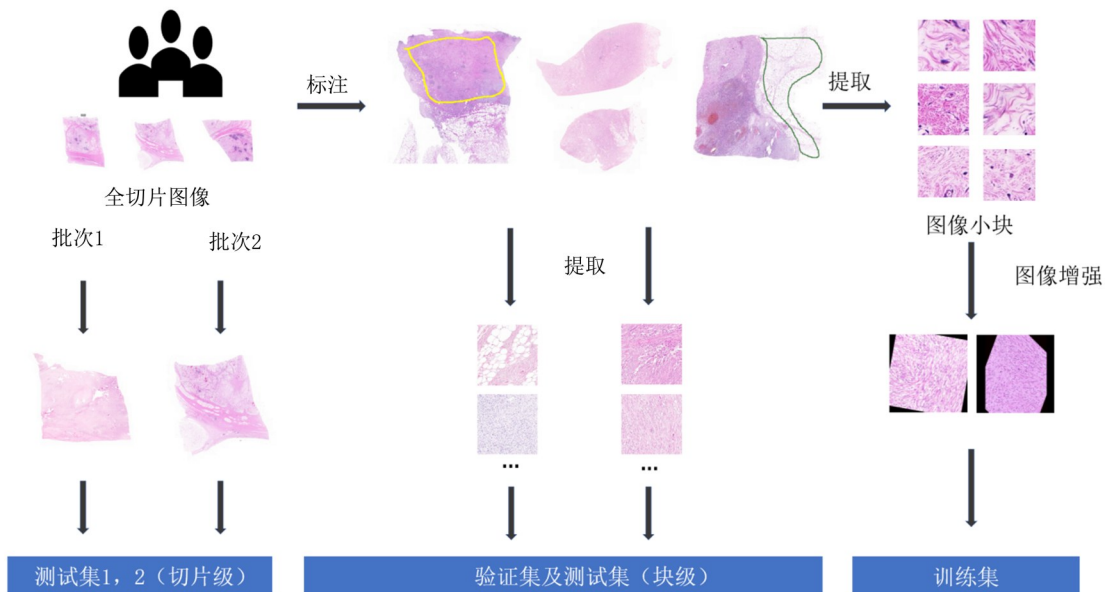


图1 数据准备流程图

Fig. 1 Diagram of the data preparation process

在数据集划分过程中, 为确保训练集、验证集和测试集三者之间的数据独立性, 除MPNST数据集外, 其他所有数据集都严格遵守每张切片生成的图像块不会同时出现在不同数据集中的原则。对于MPNST数据集, 由于切片数量有限(仅3张), 为保证训练样本的多样性, 采用的方法是将切割后的图像块先混合再随机分配至训练集、验证集和测试集中, 并确保图像没有重复使用。虽然这样MPNST在块级分析中的准确率可信度会低于其他4种类型, 但可以通过观察其他疾病类型是否被误分类为MPNST来得到间接信息。

1.2.1 图像提取 利用Histolab(Marcolini et al., 2022), 对TIFF格式的图像进行裁剪, 裁剪后的图像为512×512的PNG格式。如图2所示, 裁剪过程分为基于网格框架和基于评分机制两种。基于网络框架的如图2(a)-(b)所示, 根据预设网格布局对原始WSI进行裁剪。裁剪后, 根据图像块的组织含量进行筛选, 删除组织成分低于设定阈值的图像块, 每个全切片图像可生成约50至1000张图像块。基于评分机制的如图2(c)-(d)所示, 对裁剪后的图像块细胞核密度进行评分, 并按评分保留前 N 张图像, 丢弃其余部分, 该方法每个WSI可生成约50~300张图像块, 特别适用于高分化脂肪肉瘤的处理。由于该肿瘤中脂肪空泡较多, 组织密度较低, 这种选择性过程能够优化图像获取过程, 确保在组织学特征较为复杂的情况下获得高质量的样本。通过上述两种方法提取图像块后, 进行

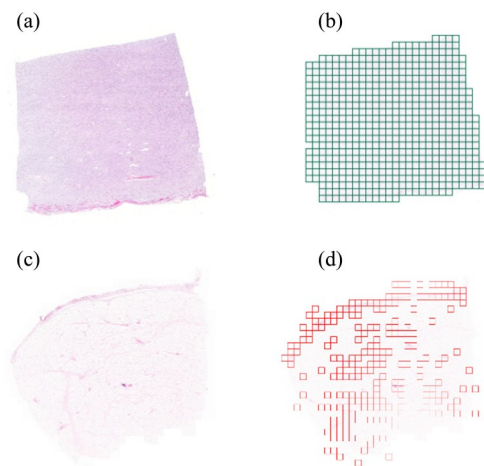


图2 图像块提取方法示意图

Fig. 2 Diagram of the image patch extraction method

人工筛选, 进一步剔除与疾病诊断无关的图像块。所有图像都经过亮度和对比度调整, 达到对样本质量进行平衡、提高图像细节可见性的目的。

1.2.2 数据增强 为了增加训练样本多样性, 减轻过拟合问题并提高模型的鲁棒性, 采用2种图像增强策略。第1种策略包括对图像进行翻转、旋转、裁剪和遮挡等操作。针对5类图像分别选择[6, 12, 30]的差异化增强倍数, 从而确保训练集中各类样本的相对均匀分布。第2种策略采用mix-up技术(Zhang et al., 2018), 在训练过程中随机选择2个不同的样本及其对应的标签, 然后按照一定比例(通常从Beta分布中采样)对其特征和标签进行线性组合, 从而生成混合样本。具体来说,

给定两个样本 x_i 、 x_j (标签为 y_i 、 y_j) 以及从 Beta 分布中采样的随机数 λ , 混合样本 x_{mix} 及其标签 y_{mix} 的数学表达式为 $x_{\text{mix}} = \lambda x_i + (1 - \lambda)x_j$, $y_{\text{mix}} = \lambda y_i + (1 - \lambda)y_j$ 。这一策略通过引入标签和特征的连续性, 帮助模型在训练时学会处理样本之间的更广泛变化, 从而提高对新数据的适应性。

1.3 基于深度学习的诊断方案

基于深度学习的诊断方案如图3所示。先对两

批次的病理切片图像进行预处理, 将其分割成多个图像涵盖多种腹膜后软组织肉瘤类型的图像块。接着, 将这些图像小块输入不同的深度学习模型进行训练, 并通过模型集成方法来融合多个模型的优势, 提高分类性能。最后, 使用训练好的模型对测试数据进行预测, 并通过混淆矩阵、饼图和分类地图等评估工具评估模型的性能。

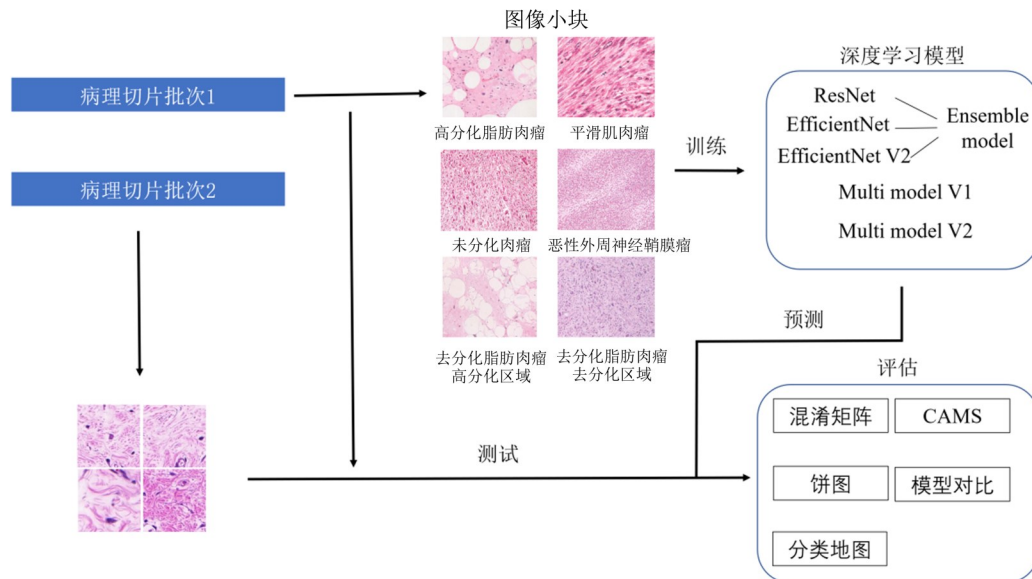


图3 基于深度学习的诊断方案

Fig. 3 Deep learning-based diagnostic framework

近年, 图像分类领域涌现了众多表现优异的深度学习模型, 如 ResNet (He et al., 2016)、EfficientNet (Tan et al., 2019) 和 EfficientNet V2 (Tan et al., 2021)。本文选取这3种模型进行训练, 均使用基于 ImageNet 数据集的预训练权重进行初始化。在初始化后, 首先仅解冻模型的顶层, 根据训练效果逐步解冻更多层, 以确定最佳的冻结策略。通过验证集多次迭代调整学习率、mix-up 混合因子等超参数, 并采用早停策略保存最优模型。经过多轮实验, 保留了多个模型的权重文件, 并基于这些权重文件提出了两种模型集成方案。

(1) Ensemble model。在 Ensemble 策略中, 采用 ResNet18、EfficientNet B7 和 EfficientNet V2 的权重进行各种组合。结果表明, 最优方案为基于不同超参数训练的3个 EfficientNet B7 模型。通过对相同放大倍数下训练出来的模型进行决策层融合, 构建了 Ensemble model。

(2) Multi model。Multi model 策略通过集成

不同放大倍数下训练的模型实现。在图4所示的结构中, 每条路径代表一个 EfficientNet B7 网络。Multi model V1 基于 10 和 20 倍放大倍率的图像进行训练并融合; Multi model V2 则基于 10、20 和 40 倍的图像进行训练并融合。通过对不同放大倍数下训练的模型输出在决策层进行融合, 最终构建了 Multi model。该策略通过整合病理图像在不同尺度下的信息, 有效提升了分类性能。

2 试验结果及评估

本文从两个层面对模型进行评估: 块级分析直接对提取出的图像块进行评估, 每个图像块的标签默认继承其来源的全切片图像的标签; 切片级分析评估整个切片, 通过网格化方式提取图像块, 随后空白、模糊或组织代表性不足的图像将被剔除。然后, 将各块的结果汇总, 出现频率最高的预测标签作为该 WSIs 的总体标签, 从而将局部预测整合为全局诊断(对于包含 WDLP 成分的

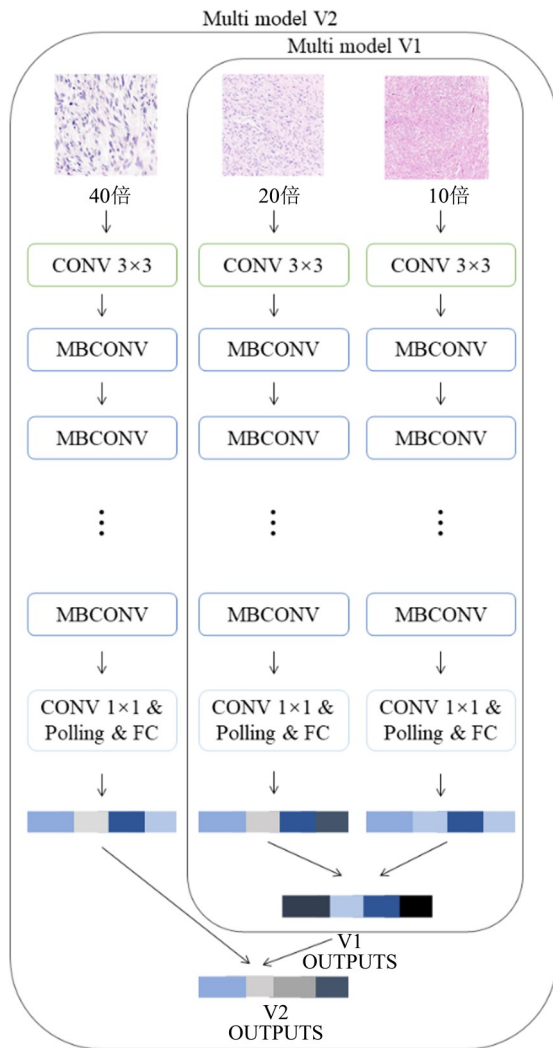


图4 Multi model 结构示意图

Fig. 4 Diagram of the Multi model structure

DDLp切片, 判断标准是最频繁出现的两个预测标签是否为DDLp和WDLp)。

2.1 块级分析

如表1所示, 不同类别的分类表现有所差异: LMS和WDLp的分类准确率较高, 与其他类别的

混淆较少, 表明这些类别具有较强的区分能力。MPNST在EfficientNet V2、EfficientNet B7和Ensemble model中表现良好, 但在其他模型中的表现较差。如图5所示, DDLp和UPS是最难区分的类别, 两者的界限模糊, 容易在各个模型中发生相互误分类。DDLp和UPS的误分类是导致总体准确率下降的主要原因之一, 这种误分类现象进一步凸显了在精确区分这两个类别时所面临的挑战, 及其对整体性能指标的影响。

2.2 切片级分析

在全切片图像层面对肿瘤进行分类, 分类地图和切片分类结果的比例如图6-7所示。由于MPNST数据有限, 分类仅针对其余4种肿瘤类型进行。如表2所示, EfficientNet B7、Ensemble model、Multi model V1和Multi model V2的总体准确率分别为70.18%、71.93%、78.95%和80.70%。且LMS和WDLp的分类表现非常出色, 所有测试切片均得到了正确的预测。然而, DDLp和UPS之间的区分在切片层面依然存在问题, 许多DDLp和UPS切片被错误地预测为对方的标签。在临床实践中, 当DDLp缺乏脂肪成分时, 的确很难将其与UPS区分开来。为了提高分类的准确性, 对分类结果进行了优化。在一些WSIs层面的DDLp切片中, 存在高分化脂肪瘤样区域。基于这一特征, 将同时含有DDLp、UPS和WDLp证据的切片(如图6(a)和(c))中的UPS预测成分视为DDLp, 并重新进行了预测。修正后的结果如表2所示, 4个模型的总体准确率分别提升到了77.19%、77.19%、91.23%和87.72%。

为了验证模型的泛化能力, 在测试集2上进行了相同的测试。测试集2独立于训练数据, 且来自不同中心, 结果如表3所示, 尽管在新的测试集上准确率有所下降, 但表现最好的模型Multi model V1仍然达到了80.95%的总体准确率。

表1 块级分析准确率
Table1 Patch-level analysis accuracy

肿瘤类别	ResNet18	EfficientNet V2	EfficientNet B7	Ensemble model	Multi model V1	Multi model V2
DDLp	47.15	84.62	85.25	83.26	75.20	67.24
LMS	80.31	83.07	97.00	93.16	83.79	79.59
MPNST	62.75	99.02	95.10	100.00	73.53	65.69
UPS	66.51	31.93	53.89	63.40	77.13	78.41
WDLp	90.75	99.71	100.00	100.00	99.86	100
总体准确率	68.21	72.34	81.78	83.22	82.04	79.01
总体准确率(无MPNST)	68.36	71.61	81.41	82.76	82.27	79.37

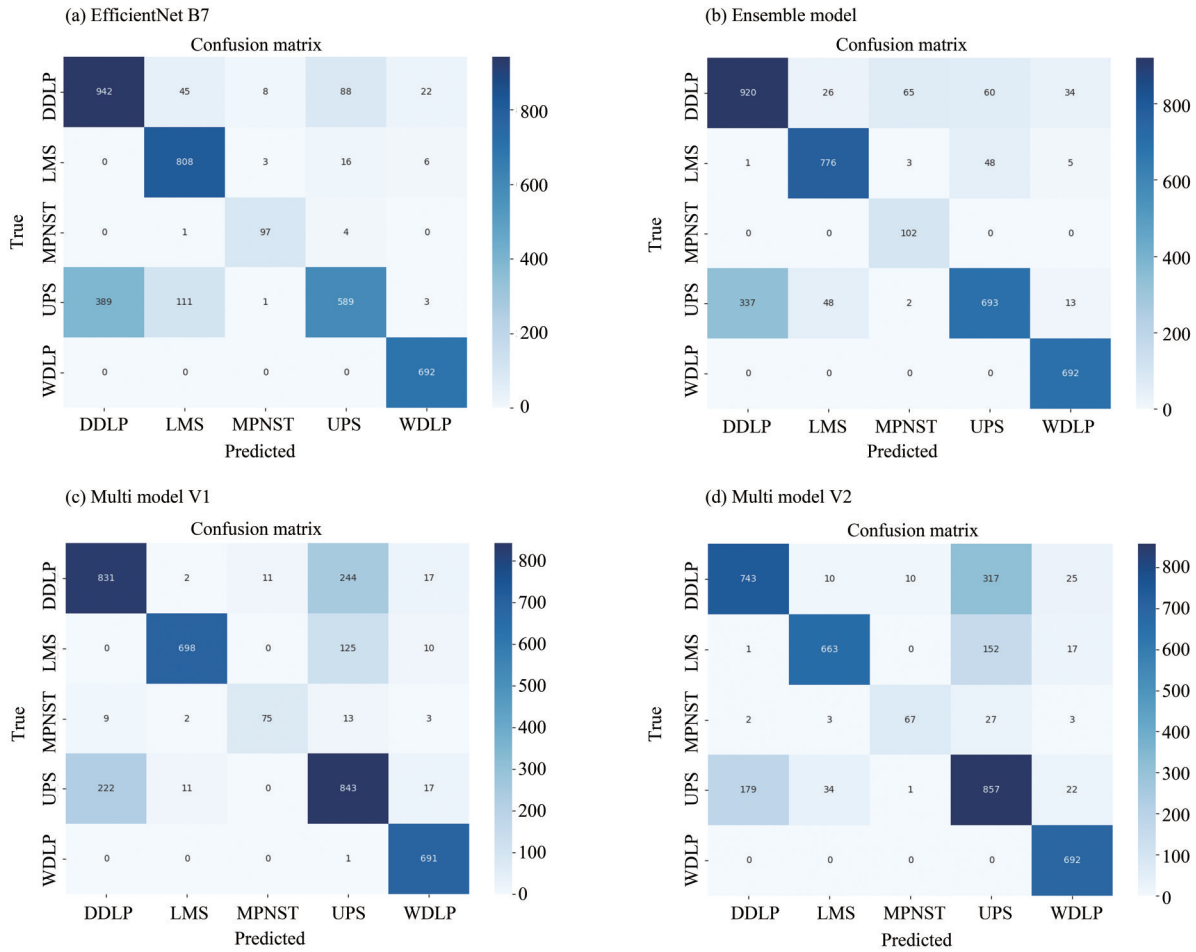


图 5 混淆矩阵
Fig. 5 Confusion matrix

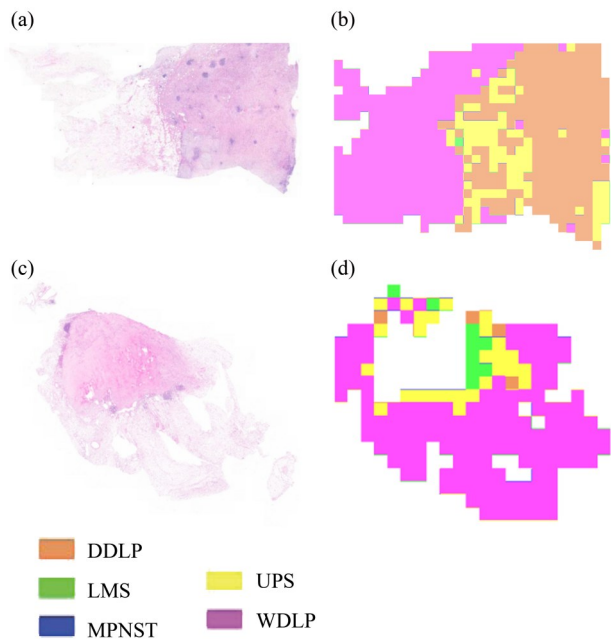


图 6 切片分类地图
Fig. 6 Classification map of WSIs

3 讨论与分析

本研究提出两种不同的模型集成方法, 分别在图像块和全切片图像层面进行应用。在块级分析中, 本文提出的方法并没有展现出相比常规单体模型的优势。但是在更考验模型泛化能力的切片级分析中, 本文提出的集成模型相比单体模型有性能的提高, 其中表现最优的是 Multi model V1。尽管 V2 相比 V1 增加了 40 倍放大倍数的信息融合, 但在某些情况下准确率不如 V1, 这可能是由于 40 倍大小图像块所包含的信息反而造成了亚型的混淆, 因为在更大的放大倍数下, WDLP 的部分图像块与 DDLP 的图像块相似。无论是块级分析还是切片级分析, 图像块和切片被误分类为 MPNST 的情况较少, 但都遇到区分去分化脂肪肉瘤(DDLp)和未分化多形性肉瘤(UPS)的困难。这是由于两者在显微镜下均显示出高度的细胞多形性和异形性, 特别是在 DDLp 样本缺乏明确脂肪细胞分化时, 诊断更加困难。

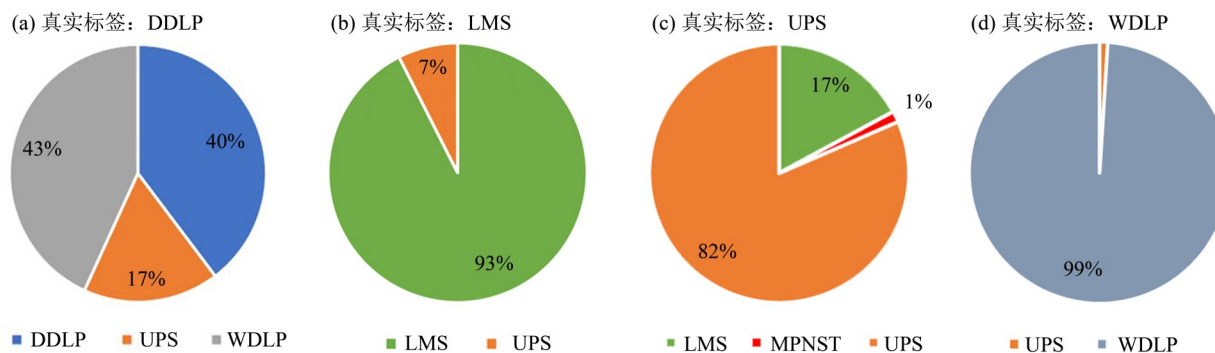


图7 切片分类结果比例

Fig. 7 Diagram of slide classification outcome ratios

表2 基于测试集1的全切片图像准确率对比(原始与优化方法)

Table 2 Comparison of accuracy on whole slide images based on test set 1 (original vs refined methods)

模型	DDLDP	LMS	UPS	WDLP	Overall
	(True/False/ACC)	(True/False/ACC)	(True/False/ACC)	(True/False/ACC)	(True/False/ACC)
EfficientNet B7(原始)	13/14/48.15%	6/0/100%	5/3/62.50%	16/0/100%	40/17/70.18%
Ensemble model(原始)	15/12/55.56%	6/0/100%	6/2/75.00%	16/0/100%	41/16/71.93%
Multi model V1(原始)	15/12/55.56%	6/0/100%	8/0/100%	16/0/100%	45/12/78.95%
Multi model V2(原始)	16/11/59.26%	6/0/100%	8/0/100%	16/0/100%	46/11/80.70%
EfficientNet B7(优化)	17/10/62.96%	6/0/100%	5/3/62.50%	16/0/100%	44/13/77.19%
Ensemble model(优化)	16/11/59.26%	6/0/100%	6/2/75.00%	16/0/100%	44/13/77.19%
Multi model V1(优化)	22/05/81.48%	6/0/100%	8/0/100%	16/0/100%	52/05/91.23%
Multi model V2(优化)	20/07/74.07%	6/0/100%	8/0/100%	16/0/100%	50/07/87.72%

表3 基于测试集2的全切片图像准确率对比(原始与优化方法)

Table 3 Comparison of accuracy on whole slide images based on test set 2 (original vs refined methods)

模型	DDLDP	LMS	UPS	WDLP	Overall
	(True/False/ACC)	(True/False/ACC)	(True/False/ACC)	(True/False/ACC)	(True/False/ACC)
EfficientNet B7(原始)	4/8/33.33%	7/0/100%	4/4/50.00%	14/1/93.33%	29/13/69.05%
Ensemble model(原始)	5/7/55.56%	7/0/100%	3/5/37.50%	15/0/100%	30/12/71.43%
Multi model V1(原始)	7/5/58.33%	7/0/100%	4/4/50.00%	15/0/100%	33/09/78.57%
Multi model V2(原始)	7/5/58.33%	7/0/100%	3/5/37.50%	15/0/100%	32/10/76.19%
EfficientNet B7(优化)	4/8/33.33%	7/0/100%	4/4/50.00%	14/1/93.33%	29/13/69.05%
Ensemble model(优化)	6/6/50.00%	7/0/100%	3/5/37.50%	15/0/100%	30/12/73.81%
Multi model V1(优化)	8/4/58.33%	7/0/100%	4/4/50.00%	15/0/100%	34/08/80.95%
Multi model V2(优化)	7/5/58.33%	7/0/100%	3/5/37.50%	15/0/100%	32/10/76.19%

在测试集1中,基于DDLDP全切片图像中是否含有一定比例WDLP成分存在的优化方法,显著提升了模型在DDLDP分类上的表现。与未使用该方法相比,DDLDP的分类准确率得到了显著提高(增加5%~10%)。尽管LMS和DDLDP、UPS在通常情况下有明确的形态学界限,但这些肿瘤在某些特征上仍然存在相似性,可能导致预测时的混淆。高分化脂肪肉瘤(WDLP)在各个模型中的分类准确

率始终较高,这可能是由于其较为显著的形态学特异性,使其在组织病理学图像中的区分更加容易。

测试集2呈现了与测试集1类似的情况,LMS和WDLP的分类表现依然非常出色,而DDLDP和UPS之间容易混淆。与测试集1的区别是,有更多DDLDP或UPS的切片被错误预测为LMS,这也是使用优化方法后,准确率提升不明显的原因。测

试集1的最高总体准确率达到了91.23%, 表现非常优异; 但在测试集2中, 最高的总体准确率则下降至80.95%。这一差异主要归因于测试集2的数据独立于测试集1, 其数据来源、切片格式、所使用的软件版本以及扫描设备均有所不同, 从而对模型的泛化能力造成了一定影响。尽管如此, 模型在测试集2中的表现依然令人满意, 尤其是LMS和WDLP的分类结果依旧稳健。而DDLp和UPS的分类结果在测试集2中的表现不如测试集1, 这进一步表明模型在处理数据不同的样本时, 尤其需要提升其泛化能力, 以及在形态学上相似的肿瘤间的区分能力。

此外, 本文利用深度学习模型最后一个卷积

层的特征图生成了类别激活图(CAMs), 如图8所示。其中, a-b为DDLp及其热力图, c-d为LMS及其热力图, e-f为MPNST及其热力图, g-i为UPS及其热力图, k-l为WDLP及其热力图。这些图突出了模型对每个类别中不同模式和感兴趣区域的解读。可以发现, 在区分DDLp、LMS、MPNST和UPS时, 模型的关注点主要集中在细胞核, 这表明细胞核的特定形态或结构是区分这些肿瘤类型的重要依据。相较之下, 在WDLP的分类过程中, 模型更倾向于关注细胞的轮廓特征, 表明整体形状或细胞间的关系在识别该肿瘤类型时起到了关键作用。

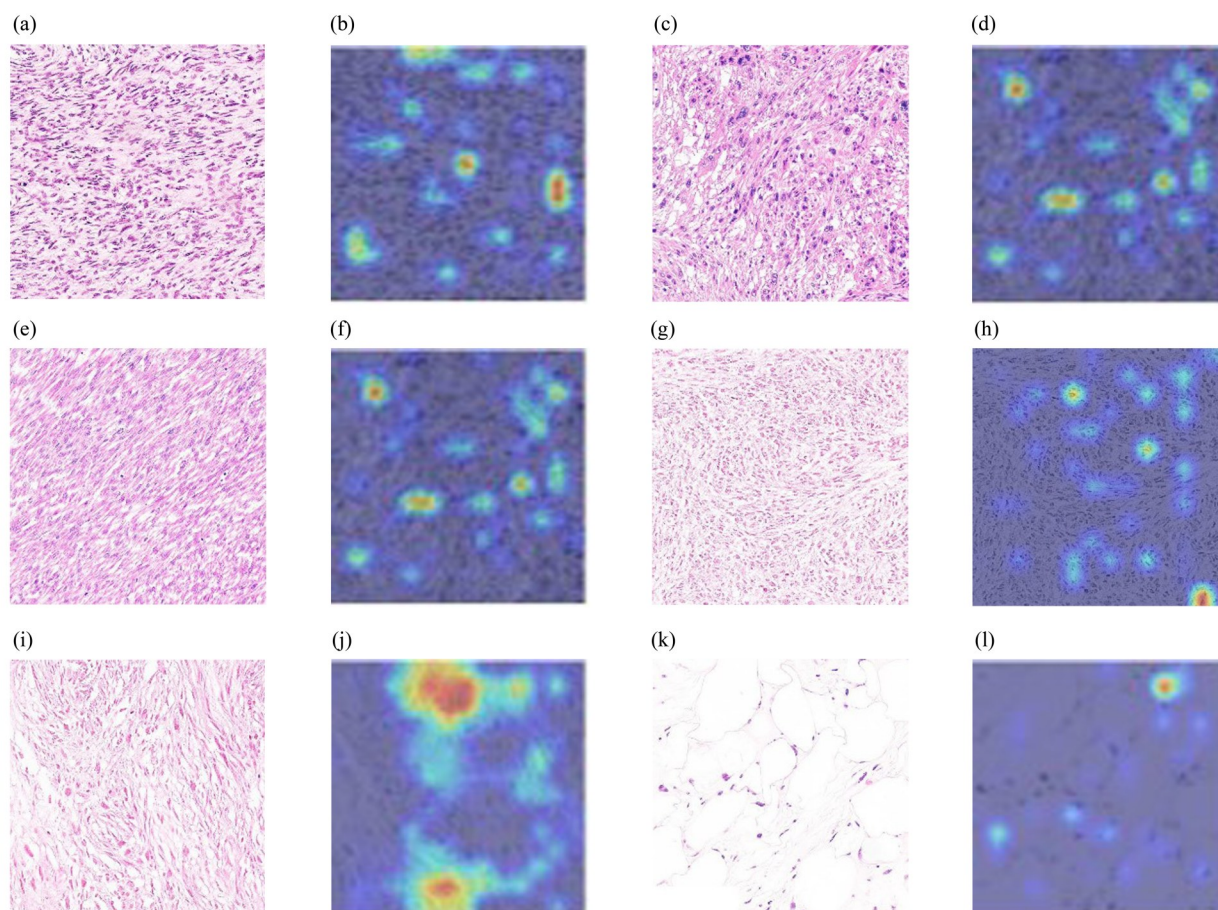


图8 类别激活图

Fig. 8 Class activation mapping

4 结论

本文基于多种深度学习模型开展了多亚型的腹膜后软组织肉瘤诊断研究。所提出的Ensemble model和Multi model在腹膜后软组织肉瘤的分类任务中表现优异, 尤其是Multi model V1在块级分类

上达到82.27%, 在全切片图像上的分类准确率达到80.95%。尽管数据格式和扫描设备等有所差异, 模型在不同测试集上表现稳定, 证明了其泛化能力。该模型在提高小型活检标本诊断过程中的准确性方面具有重要意义, 有望简化腹膜后肉瘤的诊断流程。

参考文献:

- ALMOND L, TIROTTA F, TATTERSALL H, et al, 2019. Diagnostic accuracy of percutaneous biopsy in retroperitoneal sarcoma[J]. *Br J Surg*, 106(4):395-403.
- ARUNACHALAM H B, MISHRA R, DAESCU O, et al, 2019. Viable and necrotic tumor assessment from whole slide images of osteosarcoma using machine-learning and deep-learning models[J]. *PLoS One*, 14(4): 0210706.
- BENGIO Y, COURVILLE A, VINCENT P, 2013. Representation learning: A review and new perspectives [J]. *IEEE Trans Pattern Anal Mach Intell*, 35(8): 1798-1828.
- CUI M, ZHANG D Y, 2021. Artificial intelligence and computational pathology [J]. *Lab Invest*, 101(4): 412-422.
- DROHAN A, GRONCHI A, 2023. Next questions in the management of retroperitoneal sarcoma [J]. *Curr Opin Oncol*, 35(4):301-308.
- FOERSCH S, ECKSTEIN M, WAGNER D C, et al, 2021. Deep learning for diagnosis and survival prediction in soft tissue sarcoma[J]. *Ann Oncol*, 32(9):1178-1187.
- HE K, ZHANG X, REN S, et al, 2016. Deep residual learning for image recognition [C]//IEEE Conference on Computer Vision and Pattern Recognition. Las Vegas, NV, USA:770-778.
- LECUN Y, BENGIO Y, HINTON G, 2015. Deep learning[J]. *Nature*, 521(7553):436-444.
- MACK T, PURGINA B, 2022. Updates in pathology for retroperitoneal soft tissue sarcoma[J]. *Curr Oncol*, 29(9): 6400-6418.
- MARCOLINI A, BUSSOLA N, ARBITRIO E, et al, 2022. Histolab: A python library for reproducible digital pathology preprocessing with automated testing [J]. *SoftwareX*, 20:101237.
- MILEWSKI D, JUNG H, BROWN G T, et al, 2023. Predicting molecular subtype and survival of rhabdomyosarcoma patients using deep learning of H&E images: A report from the children's oncology group[J]. *Clin Cancer Res*, 29(2):364-378.
- MISHRA R, DAESCU O, LEAVEY P, et al, 2018. Convolutional neural network for histopathological analysis of osteosarcoma [J]. *J Comput Biol*, 25(3): 313-325.
- TAN M, LE Q, 2019. EfficientNet: Rethinking model scaling for convolutional neural networks [C]//36th International Conference on Machine Learning. Long Beach, CA, USA:6105-6114.
- TAN M, LE Q, 2021. EfficientNet V2: Smaller models and faster training [C]//38th International Conference on Machine Learning. Graz, AT, Austria:10096-10106.
- TROJANI M, CONTESSO G, COINDRE J, et al, 1984. Soft-tissue sarcomas of adults; study of pathological prognostic variables and definition of a histopathological grading system[J]. *Int J Cancer*, 33(1):37-42.
- van der LAAK J, LITJENS G, CIOMPI F, 2021. Deep learning in histopathology: The path to the clinic [J]. *Nat Med*, 27(5):775-784.
- WOERL A C, ECKSTEIN M, GEIGER J, et al, 2020. Deep learning predicts molecular subtype of muscle-invasive bladder cancer from conventional histopathological slides [J]. *Eur Urol*, 78(2):256-264.
- YEUNG M C, CHENG I S, 2022. Artificial intelligence significantly improves the diagnostic accuracy of deep myxoid soft tissue lesions in histology [J]. *Sci Rep*, 12(1):6965.
- ZHANG H, CISSE M, DAUPHIN Y N, et al, 2018. Mixup: Beyond empirical risk minimization [EB/OL]. arxiv.1710.09412.
- ZHAO Z, CHEN X, XU J, et al, 2023. Whole exome sequencing of well-differentiated liposarcoma and dedifferentiated liposarcoma in older woman: A case report[J]. *Front Med*, 10:1237246.

(责任编辑 王海蓉)