

• 临床研究 •

基于卷积神经网络的胃癌病理图像分类诊断与分级识别

吴邦钰^{1,2}, 王志霄³, 项景轩², 孙立³, 马玲^{1,2,4,5*}

¹南京医科大学第一附属医院肿瘤科, 江苏 南京 210029; ²南京医科大学第一临床医学院临床医学系, 江苏 南京 211166; ³东南大学能源与环境学院动力工程及自动化系, 江苏 南京 211189; ⁴南京医科大学部省共建肿瘤个性化医学协同创新中心, 江苏 南京 211166; ⁵睢宁县人民医院肿瘤科, 江苏 徐州 221200

[摘要] 目的: 基于深度学习技术, 建立胃癌病理切片的分类诊断模型与分级模型, 评估模型性能。方法: 基于公开网络资源收集胃癌和非癌组织分类诊断识别数据集与胃癌分级识别数据集。对数据进行数据增强, 划分为训练集、验证集和测试集。初始阶段, 构建17种卷积神经网络(convolutional neural network, CNN)模型, 统一设置初始训练参数, 对17个模型进行胃癌和非胃癌分类识别训练。训练结束后, 以模型在测试集上的识别准确率、训练耗时作为评价指标, 全面评估不同模型架构的效能。基于这些指标, 筛选出效能最优的架构, 进一步优化训练, 构建胃癌分类诊断识别模型。分类模型完成后, 基于分类模型的基础构建胃癌分级模型, 在胃癌分级模型的训练过程中, 训练17个分级网络, 根据性能指标筛选出适合作为基模型的网络。在基模型确定后, 分别采用Voting和Stacking方法进行集成学习, 并与单模型进行比较。探究集成学习对性能提升的影响, 构建胃癌分级识别模型。结果: 在胃癌分类诊断模型训练中, Xception经对比被选为最终分类诊断模型的网络。经过参数调整与训练后, 最终胃癌分类诊断模型在测试集上表现的准确率为98.13%、灵敏度为98.11%、特异度为98.11%、F1分数为98.12%、受试者工作特征(receiver operating characteristic, ROC)曲线下面积(area under the curve, AUC)为0.998; 胃癌分级模型训练中, 以随机森林为代表的堆叠法相对于以硬投票法为代表的投票法拥有较大的提升。以随机森林集成模型为最终分级模型, 其准确率为95.06%、灵敏度为94.77%、特异度为98.36%、F1分数为94.82%, 良性AUC为0.999, 管状低分化腺癌AUC为0.981, 管状中分化腺癌AUC为0.990, 管状高分化腺癌AUC为0.995。结论: 两个模型都拥有良好的识别性能, 证明了利用CNN实现对胃肿瘤病理图像进行高精度分类诊断及分级的可行性, 实现了迁移-集成联合框架对胃肿瘤图像的分级运用, 有望应用于医院智能化诊断辅助系统。

[关键词] 胃癌; 精准医疗; 卷积神经网络; 集成学习; 人工智能**[中图分类号]** R735.2**[文献标志码]** A**[文章编号]** 1007-4368(2026)04-520-13**doi:** 10.7655/NYDXBNSN250368

Classification and grading recognition of gastric cancer pathological images based on convolutional neural networks

WU Bangyu^{1,2}, WANG Zhixiao³, XIANG Jingxuan², SUN Li³, MA Ling^{1,2,4,5*}

¹Department of Oncology, the First Affiliated Hospital of Nanjing Medical University, Nanjing 210029; ²Department of Clinical Medicine, the First school of Clinical Medicine, Nanjing Medical University 211166; ³Department of Power Engineering and Automation, School of Energy and Environment, Southeast University, Nanjing 211189; ⁴Collaborative Innovation Center of Personalized Oncology Medicine, Ministry of Education and Jiangsu Province, Nanjing Medical University, Nanjing 211166; ⁵Department of Oncology, Sui Ning County People's Hospital, Xuzhou 221200, China

[Abstract] **Objective:** To establish classification and grading models for gastric cancer pathological sections based on deep learning

[基金项目] 国家自然科学基金(82102981); 江苏高校“青蓝工程”(KY102R202412); 中国博士后科学基金第75批面上资助(2024M751224); 南京市博士后项目[BSHNJ2023003(JC23)]; 北京希思科临床肿瘤学研究基金会(Y-Young2024-0152); 部省共建肿瘤个性化医学协调创新中心青年项目(2024CICCPMHR033); 胃癌免疫治疗专病队列研究(JZ214490202106052); 江苏省高等学校大学生创新创业训练计划(X2025103120042)

*通信作者(Corresponding author), E-mail: maling@njmu.edu.cn (ORCID: 0000-0002-2960-1956)

technology, and to evaluate the performance of these models. **Methods:** Classification and grading datasets for gastric cancer and non-cancerous tissues were collected from public online resources. Data augmentation was performed, and the dataset were divided into training, validation, and test sets. In the initial stage, 17 convolutional neural network (CNN) architectures were constructed, and the initial training parameters were uniformly set to train these 17 models for the classification of gastric cancer and non-cancerous tissues. After training, the recognition accuracy on the test set and the training time were used as evaluation indicators to comprehensively assess the efficacy of different model architectures. Based on these indicators, the optimal architecture was selected for further optimization and training to construct the gastric cancer classification model. After the completion of the classification model, the gastric cancer grading model was built based on the foundation of the classification model. During the training of the gastric cancer grading model, 17 grading networks were trained, and suitable base models were selected according to performance indicators. After the base model was determined, voting and stacking methods were applied for ensemble learning and compared with single models to explore the impact of ensemble learning on performance improvement and to construct the gastric cancer grading model. **Results:** In the training of the gastric cancer classification model, the Xception network was selected as the final classification model after comparison. After parameter adjustment and training, the final gastric cancer classification model achieved an accuracy of 98.13%, sensitivity of 98.11%, specificity of 98.11%, F1 score of 98.12%, and AUC of 0.998 on the test set. In the training of the gastric cancer grading model, the stacking method represented by random forest showed significant improvement compared to the voting method represented by hard voting. The ensemble model based on random forest was selected as the final grading model, with an accuracy of 95.06%, sensitivity of 94.77%, specificity of 98.36%, and F1 score of 94.82%. The area under the receiver operating characteristic (ROC - AUC) curve values were 0.999 for benign, 0.981 for poorly differentiated tubular adenocarcinoma, 0.990 for moderately differentiated tubular adenocarcinoma, and 0.995 for well - differentiated tubular adenocarcinoma. **Conclusion:** Both models demonstrated excellent recognition performance, proving the feasibility of using CNN to achieve high - precision classification and grading of gastric tumor pathological images. The transfer-learning and ensemble-learning framework was successfully applied to the grading of gastric tumor images and holds promise for integration into hospital intelligent diagnostic assistance systems.

[Key words] gastric cancer; precision medicine; convolutional neural network; ensemble learning; artificial intelligence

[J Nanjing Med Univ, 2026, 46(04): 520-532]

胃癌是全球的重大公共卫生挑战之一。根据国际癌症研究机构(International Agency for Research on Cancer)发布的GLOBOCAN 2022数据,全球胃癌年新增病例约109万例,死亡76万例,发病率与死亡率分别位居癌症谱第5位和第4位。东亚地区(中国、日本、韩国)的年龄标准化发病率(age-standardized rate of incidence, ASR)高达24.3/10万,显著高于全球平均水平(11.1/10万),这一差异与幽门螺杆菌高感染率(>50%)、高盐饮食及遗传多态性(如CDH1基因突变)密切相关^[1]。中国国家癌症中心2023年统计显示,胃癌新发病例47.9万例,死亡37.4万例,分别占恶性肿瘤发病和死亡的第3位与第2位,且约80%患者确诊时已进展至中晚期,5年生存率仅为28%^[2]。由于早期胃癌缺乏特异性症状,现行诊断体系高度依赖内镜活检与病理学评估,但传统病理诊断受限于医师经验差异与视觉疲劳,误诊率可达8.94%^[3]。因此,开发高精度、标准化的人工智能(artificial intelligence, AI)辅助诊断工具对提升胃癌早筛效率和减少误诊率具有重要临床价值。

近年来,AI技术在医学图像分析中取得突破性进展^[4-5]。卷积神经网络(convolutional neural network, CNN)通过端到端特征学习,在乳腺癌^[6]、肺癌^[7]等病理图像分类任务中展现出超越传统方法的性能。轻量化架构如EfficientNet-V2^[8]与MobileNetV3^[9]通过复合缩放策略,在保持高精度的同时显著降低计算复杂度,为边缘设备部署奠定基础。然而,胃癌病理图像的异质性(如分化程度、染色差异)对模型泛化能力提出更高要求,亟需系统性探索多模型集成与迁移学习的协同优化策略^[10]。

本研究通过系统性评估17种CNN模型在胃癌病理图像分类诊断与分级任务中的性能,筛选最优架构并构建集成学习框架。在胃癌诊断中引入多模型横向对比,量化不同网络在特征提取能力、训练效率及泛化性能方面的差异。并提出一种基于Stacking策略的异构模型集成方法,通过多模型集成进行互补特征融合,提高模型准确率与泛化性。以期通过迁移-集成联合框架,促进AI模型在低资源医疗场景下的临床快速部署,为数字化病理诊断

系统的开发提供理论支持与技术方案。

1 材料和方法

1.1 材料

1.1.1 数据来源与处理

本研究基于公开网络收集胃癌病理图像数据,选取华威大学计算机科学系组织图像分析中心创建的,用于泛癌肿瘤检测的大型病理学图像数据库 HISTOPANTUM^[11],提取其中的胃癌数据集部分用于胃癌与正常组织分类识别诊断模型训练;选取由韩国首尔的 Kangbuk 三星医院提供的 KBSMC_gastric_cancer_grading_dataset 等^[12]数据集用于胃癌分级识别模型训练。HISTOPANTUM 数据集的构建基于对4种不同癌症类型(结直肠癌、子宫癌、卵巢癌和胃癌)的全面采样。研究团队从癌症基因组图谱(The Cancer Genome Atlas, TCGA)项目中获取了每种癌症类型的40张全切片图像(whole slide image, WSI),并确保样本涵盖多种肿瘤亚型、性别、种族和中心,以增强数据集的多样性。随后,经验丰富的病理学家对这些切片中的肿瘤和非肿瘤区域进行了细致标注,从而提取出肿瘤和非肿瘤的图像块,形成了包含281 142个图像块的 HISTOPANTUM 数据集。每个图像块的分辨率为512×512像素。本研究中取胃癌图像作为训练数据,包含胃正常组织图像15 323张,胃癌图像13 527张。划分后癌与非癌分类训练集样本量 $n=20\ 194$ 、验证集样本量 $n=4\ 327$ 、测试集样本量 $n=4\ 329$;KBSMC_gastric_cancer_grading_dataset 涵盖了2016—2020年收集的98例胃癌患者的WSI。这些图像经过40倍放大后,使用Aperio数字切片扫描仪进行扫描。随后,图像被分割成大小为512×512像素的补丁图像,以适应深度学习模型的输入要求。两位资深病理学家 Kim Kyungeun 和 Song Boram 对图像进行了标注,区分了良性(benign,

BN)和3种不同分化的腺癌,分为管状高分化腺癌(tubular well differentiated adenocarcinoma, TW)、管状中分化腺癌(tubular moderately differentiated adenocarcinoma, TM)、管状低分化腺癌(tubular poorly differentiated adenocarcinoma, TP),本研究中取其全部数据集作为训练数据,包含TP 17 995张, TW 13 068张, TM 16 456张,良性 18 777张。划分后癌分级训练集样本量 $n=46\ 405$ 、验证集样本量 $n=9\ 943$ 、测试集样本量 $n=9\ 948$ 。

在真实场景应用模拟时,从TCGA中挑选无标注的WSI进行切割与输入,以模拟真实场景的直接输入,检测模型能力。在挑选时,已确认所用WSI与HISTOPANTUM中数据无重复。

为提高训练速度,增强模型稳定性,并提升泛化能力,为模型训练提供更稳定的输入,训练前将图像处理为224×224像素,并进行像素归一化。对图像采用随机旋转、随机平移、随机缩放、随机剪切变换等数据增强,缓解过拟合,提升模型的泛化能力、鲁棒性、性能和适应性,提高成本效益(图1)。研究使用MATLAB中的Augmented Image Data Store方法进行数据增强,在训练过程中实时生成增强图像。增强后数据集样本量不变。

1.1.2 模型框架选择

为探究不同种类网络在病理图像识别中的效果,尽可能全面评估不同架构的性能,本研究选用SqueezeNet、GoogleNet、InceptionV3、DenseNet201、MobileNetV2、ResNet18、ResNet50、ResNet101、Xception、Inception - ResNetV2、ShuffleNet、NASNet - mobile、Darknet-19、Darknet-53、EfficientNetB0、VGG16、VGG19等17种CNN网络进行训练测试。通过多模型横向对比确定胃癌分类诊断的最优架构。

这些网络都拥有各自的特点,通过初步训练后,不仅可以筛选出针对本项目的最优网络,同时

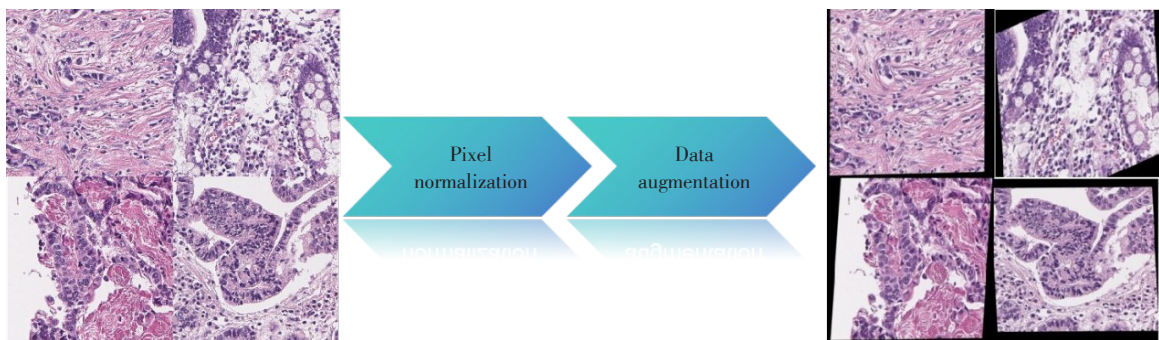


图1 图像数据增强示例

Figure 1 Example of image data augmentation

可以了解各网络在医学图像识别领域的大致效果,为后续研究提供便利。作为集成学习基模型时,也可提高基模型的多样性与独立性。本研究采用迁移学习进行模型训练,以节省计算资源,提升模型性能与泛化能力。

在集成学习方面,用基于硬投票法的投票法和基于随机森林的堆叠法分别完成模型整合。随机森林具有较高的准确性和泛化能力,通过集成多棵决策树有效降低过拟合风险,尤其适合高维数据和复杂场景。它对噪声、异常值和缺失值的鲁棒性强,能提供特征重要性评估,优化特征选择和模型解释。此外,随机森林也支持并行化训练,计算效率高,参数调优简单,使用灵活。相比单棵决策树,它更稳定;相比复杂模型(如XGBoost),它调参难度低,适合多种任务,是机器学习中的高效选择。而投票法是相对简单,稳定高效的集成学习方法。

1.1.3 环境与硬件选择

相对于其他深度学习工具,如PyTorch、TensorFlow等,MATLAB在矩阵运算方面经过底层优化,能够高效处理大规模矩阵数值运算,而矩阵运算是卷积神经网络的核心基础,因此,MATLAB在卷积特征提取时拥有独特优势。在网络构建方面,MATLAB拥有图形化交互式的拖拽式网络构建器,更适合对小型模型进行调整。但MATLAB在大型组合模型、分布式训练、多硬件支持、复杂模型模块处理方面逊色于PyTorch、TensorFlow,3者可综合应用环境来进行挑选。本研究主要针对单一卷积神经网络进行训练优化与集成,综合考虑,使用MATLAB进行代码编辑以及模型构建,版本为MATLAB2024b。硬件为个人笔记本电脑,GPU型号RTX4080laptop。在MATLAB中调用深度学习设计器,完成迁移学习模型构建及相关资源的本地部署。训练前,将网络全连接层与分类层替换,以适应病理图像分类任务,进行迁移学习。

1.2 方法

1.2.1 分类模型筛选与训练

在MATLAB中完成17个CNN网络的迁移与构建,取胃癌正异常病理图像分类数据集,按0.70、0.15、0.15的比例划分为训练集、验证集、测试集。训练集与验证集用于模型训练,测试集用于模型性能验证。初步模型训练评估时,统一设置训练参数,分别对17个模型进行胃癌分类识别训练。训练后,以识别准确率、训练耗时为评价指标,综合评估模型效能。挑选出相对最优架构,再进一步优化超

参数,提取全部数据完成训练,得到胃癌病理图像分类诊断模型(图2),评估模型性能。

1.2.2 分级模型构建

以单个CNN网络作为基模型,利用硬投票法和随机森林进行集成,随机森林为元模型。提取全部胃癌分级图像数据,按0.70、0.15、0.15的比例划分为训练集、验证集、测试集。其中训练集按0.8、0.2的比例划分为子训练集与子验证集,用于基模型训练。验证集用于元模型训练,测试集用于基模型与集成模型的性能评估。训练17个网络,并分别进行参数优化,挑选合适的网络作为基模型。基模型挑选完毕后,分别以Voting和Stacking方法进行集成学习(图2),最终将二者与单模型对比,探究集成学习带来的提升,得到胃癌分级模型,综合评估模型性能。

1.2.3 模型联合与真实应用场景模拟

在处理WSI时,导入OpenSlide外部库以处理SVS格式的整张切片,然后计算像素与切割patches数量的步长,按步长移动,遍历整个WSI,将整个WSI切割为多个512×512像素的patch并保存,保存时以坐标命名。对保存的patch进行初步处理,以高于75%白色组成为判别标准,剔除背景的白色patch。将处理后的patch输入胃癌与正常组织分类模型,标记癌与非癌,再将标记癌的patch输入分级模型,进行分级识别。最后提取patch的坐标与标签,以不同颜色代表分类结果,将结果拼接,进行识别结果的可视化(图3)。

1.3 统计学方法

使用MATLAB 2024b软件进行统计学分析。模型初步训练评估中,统计模型在验证集上的准确率和训练耗时,评估模型效能。分类模型及分级模型评估时,在MATLAB2024b中生成混淆矩阵,统计模型在测试集上的灵敏度、准确率、特异度和F1值,并以两个最终模型绘制受试者工作特征(receiver operating characteristic, ROC)曲线,计算曲线下面积(area under the curve, AUC),进行性能对比评估。对于每组数据,使用Bootstrap方法计算95%置信区间(confidence interval, CI)。通过从原始数据中随机抽取与原始数据相同大小的样本,计算每次抽样的目标参数,并重复1 000次,形成一个Bootstrap分布。然后根据其百分位数来确定数据的95%CI。

2 结果

2.1 诊断模型对比训练

经过初步测试,统计各模型的效能参数,见表

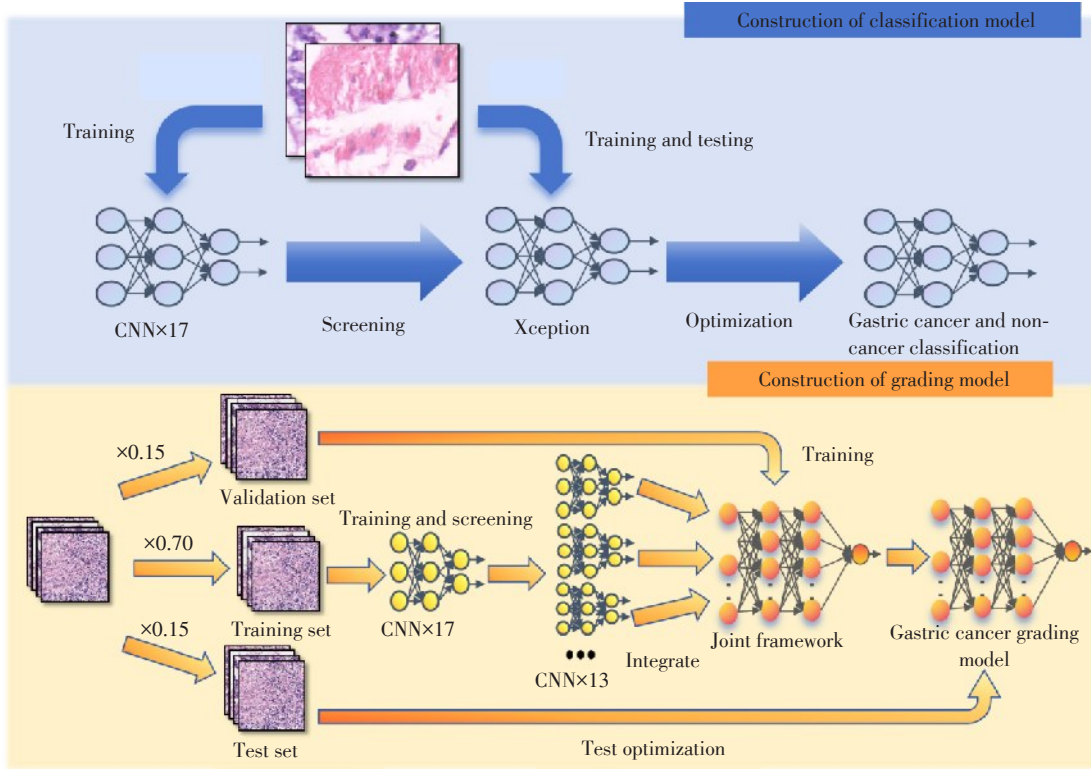


图2 模型构建流程图

Figure 2 Flowchart of model construction

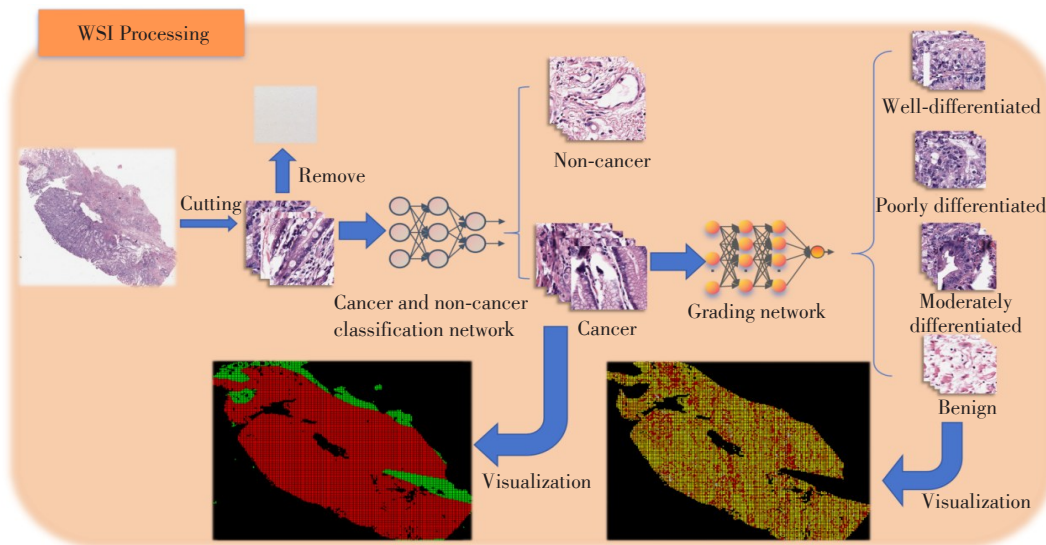


图3 真实应用场景处理流程图

Figure 3 Flowchart of the processing procedure in a real-world application scenario

1. 以准确度为纵轴, 训练耗时为横轴, 绘制效能对比图(图4), 综合评估各模型性能。SqueezeNet、Darknet19、VGG16、VGG19、GoogleNet等网络架构较老或结构过于简单, 在限定条件下均未收敛, 无对比意义, 因此未加入效能图。Darknet53相对Darknet19有所提升, 但精度和耗时并无明显优势; Inception-

ResNetV2相比同系列未及预期, 效能较差; DenseNet201精度高, 但训练耗时过长; ResNet101、NASNetMobile、EfficientNetB0在精度和耗能间取得了平衡; Xception、ShuffleNet、MobileNetV2、ResNet18/50、InceptionV3等网络都具有优秀的效能, 其中Xception在训练耗时35 min内的模型中达

到了最高的精度。经对比,将Xception选为最终分类诊断模型的网络进行参数调整与训练。

取完整数据集训练胃癌分类诊断模型,通过优化超参数提升模型性能,减少过拟合现象(图5)。最终分类诊断模型在测试集上的识别准确率达到98.13%(95%CI:97.74%~98.52%)、灵敏度为98.11%(95%CI:97.71%~98.52%)、特异度为98.11%(95%

CI:97.71%~98.52%)、F1值为98.12%(95%CI:97.72%~98.52%)。AUC为0.998(95%CI:0.998~0.999),具有非常优异的性能(图6)。

2.2 分级模型构建评估

采用子训练集与子验证集对17个模型进行训练,每个模型分别优化参数,使模型性能最大化提升。训练完成后,将所得网络保存在mat文件中,以

表1 初步模型训练结果
Table 1 Preliminary model training results

Network	Validation set accuracy(95%CI)(%)	Training time consumption
EfficientNetB0	96.07(95.47-96.63)	51 min 27 s
NASNetMobile	95.22(94.62-95.78)	66 min 23 s
Xception	95.22(94.56-95.84)	32 min 44 s
MobileNetV2	94.48(93.82-95.15)	18 min 56 s
DenseNet201	93.86(93.16-94.48)	150 min 30 s
ResNet18	92.28(91.48-93.02)	9 min 55 s
InceptionV3	92.03(91.29-92.87)	21 min 58 s
Darknet53	91.66(90.76-92.47)	34 min 43 s
ResNet50	91.04(90.22-91.86)	23 min 59 s
ResNet101	87.83(86.86-88.82)	36 min
ShuffleNet	83.09(82.02-84.18)	16 min 40 s
InceptionResNetV2	71.06(69.80-72.30)	62 min 16 s
VGG16	53.11(51.69-54.50)	20 min 6 s
SqueezeNet	53.11(51.63-54.56)	9 min 52 s
VGG19	53.11(51.61-54.60)	27 min 7 s
GoogleNet	53.11(51.59-54.70)	12 min 14 s
Darknet19	52.88(51.33-54.32)	13 min 13 s

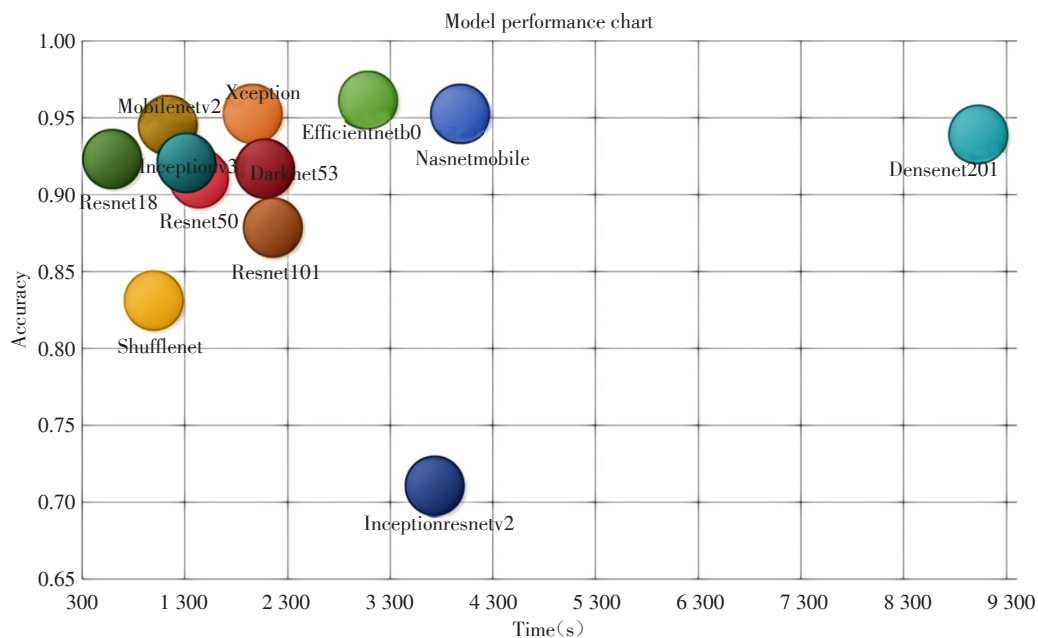


图4 模型效能对比图

Figure 4 Chart of model performance comparison

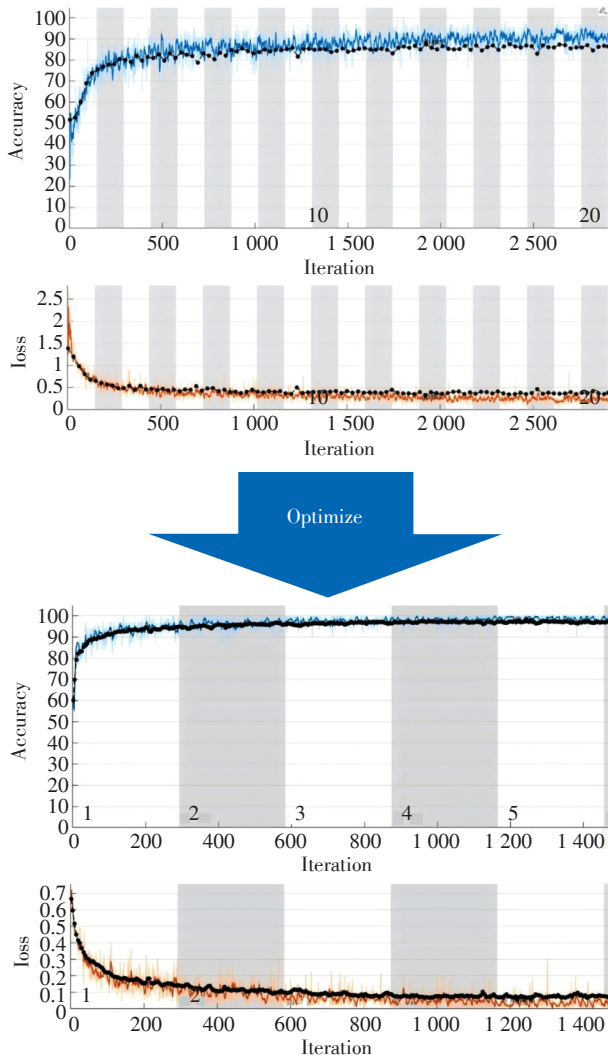


图5 Xception 优化前后准确率曲线与损失曲线

Figure 5 Accuracy and loss curves of Xception before and after optimization

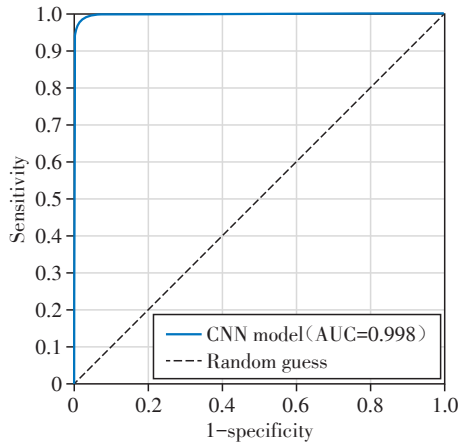


图6 分类模型 ROC 曲线

Figure 6 ROC curve of the classification model

便在已成熟的模型中筛选出所需的基模型,在集成学习时进行快速调用。测试各个模型,完成性能对

比(表2)。挑选准确度在90%以上的模型作为集成学习的基模型,以保证基模型输入质量,避免集成模型性能受限。

根据数据筛选,移除性能较差的 SqueezeNet、GoogleNet、VGG19、Darknet19 网络,选择 InceptionV3、DenseNet201、MobileNetV2、ResNet18、ResNet50、ResNet101、Xception、InceptionResNetV2、ShuffleNet、NASNetMobile、Darknet53、EfficientNetB0、VGG16 共 13 个网络作为基模型进行集成学习。进行集成学习时,通过投票法和堆叠法进行集成,堆叠法元模型选用随机森林,投票法选用硬投票进行集成。

硬投票法是集成学习中一种简单的分类决策方法,本研究的实现步骤如下:首先,调用筛选后的各个单一级模型作为基分类器,令每个基分类器对测试集数据进行独立预测,并输出类别标签。然后,对于每个待预测的样本,统计所有基分类器的预测结果,选择得票最多的类别作为最终的预测结果。如果出现票数相同的情况,会随机选择其中一个类别。

集成学习中的堆叠法是一种通过组合多个基模型的预测结果来训练一个元模型的方法。本研究的实现步骤如下:首先,调用筛选后的各个单一级模型作为基模型,令基模型对验证基数据进行预测,并生成预测概率向量。将所得的各个模型的预测概率向量水平拼接,作为新的特征,组合成一个新的训练数据集。接下来,使用随机森林作为元模型,用这些新的特征进行训练,学习如何结合基模型的预测结果以最大化整体性能。最后,使用训练好的元模型对测试集数据进行预测,生成最终的预测结果。两种方式与平均单模型的性能对比见表3。

经过训练与测试,发现集成学习方法对于单模型来说有显著提升,硬投票法和随机森林的集成模型相对于平均基模型,准确度提升 2.42% 和 3.10%、平均灵敏度提升 2.51% 和 3.30%、平均特异度提升 0.80% 和 1.03%、平均 F1 分数提升 2.52% 和 3.26%。随机森林为代表的堆叠法相对于以硬投票法为代表的投票法有较大的提升。以随机森林的集成模型作为最终胃癌分级模型,其准确率为 95.06%、灵敏度为 94.77%、特异度为 98.36%、F1 分数为 94.82%。绘制多分类 ROC 曲线(图7),良性 AUC=0.999(95% CI: 0.999~0.999),管状低分化腺癌 AUC=0.981(95% CI: 0.979~0.983),管状中分化腺癌 AUC=0.990(95% CI: 0.988~0.991),管状高分化腺

表2 单模型分级性能
Table 2 Single-model grading performance

Network	Accuracy(95%CI)	Average sensitivity(95%CI)	Average specificity(95%CI)	Average F1-score(95%CI)
ResNet101	93.34(92.85-93.83)	92.91(92.41-93.44)	97.78(97.62-97.95)	93.02(92.52-93.53)
ResNet50	93.29(92.78-93.75)	92.84(92.31-93.34)	97.76(97.59-97.92)	92.97(92.43-93.45)
EfficientNetB0	93.11(92.60-93.57)	92.62(92.10-93.11)	97.71(97.54-97.87)	92.70(92.17-93.18)
DenseNet201	92.93(92.41-93.42)	92.55(92.00-93.10)	97.67(97.49-97.83)	92.54(91.98-93.06)
ResNet18	92.51(91.97-93.02)	92.04(91.49-92.58)	97.51(97.32-97.68)	92.16(91.61-92.69)
MobileNetV2	92.21(91.67-92.71)	91.71(91.16-92.24)	97.41(97.23-97.58)	91.80(91.25-92.32)
Darknet53	92.06(91.53-92.59)	91.47(90.90-92.03)	97.36(97.18-97.53)	91.67(91.11-92.23)
Xception	91.84(91.30-92.38)	91.47(90.90-92.04)	97.30(97.12-97.48)	91.50(90.93-92.06)
NASNetMobile	91.29(90.74-91.83)	90.57(89.96-91.15)	97.10(96.91-97.28)	90.80(90.22-91.37)
ShuffleNet	90.97(90.42-91.52)	90.34(89.74-90.93)	97.00(96.81-97.19)	90.49(89.89-91.07)
InceptionResNetV2	90.86(90.28-91.42)	90.49(89.91-91.07)	96.98(96.79-97.16)	90.43(89.84-91.01)
InceptionV3	90.61(90.04-91.16)	90.07(89.45-90.64)	96.88(96.69-97.06)	90.14(89.53-90.71)
VGG16	90.47(89.84-91.05)	90.01(89.42-90.61)	96.84(96.63-97.04)	90.07(89.47-90.66)
VGG19	89.97(89.40-90.56)	89.38(88.81-89.96)	96.66(96.46-96.85)	89.48(88.90-90.05)
GoogleNet	88.75(88.16-89.38)	88.15(87.53-88.82)	96.26(96.06-96.48)	88.26(87.66-88.91)
Darknet19	86.75(86.09-87.47)	85.65(84.95-86.38)	95.56(95.33-95.80)	85.98(85.26-86.71)
SqueezeNet	83.83(83.10-84.53)	82.97(82.22-83.72)	94.62(94.38-94.86)	83.18(82.41-83.91)

表3 集成模型性能指数
Table 3 Average performance index of ensemble models

Network	Accuracy(95%CI)	Average sensitivity(95%CI)	Average specificity(95%CI)	Average F1-score(95%CI)
Hard voting method	94.38(93.93-94.84)	93.98(90.23-97.72)	98.13(97.24-99.25)	94.08(91.14-97.34)
Random forest	95.06(94.64-95.47)	94.77(94.32-95.21)	98.36(98.23-98.50)	94.82(94.38-95.26)

癌 AUC=0.995 (95% CI: 0.994~0.996)。平均 AUC=0.991 (95%CI: 0.990~0.992), 显示优秀的识别性能。

2.3 真实场景模拟评估

挑选TCGA中的病理切片模拟真实场景处理, 进行端到端的诊断识别流程演示。输入需要识别的整张切片, 模型开始切割, 去除背景, 保留组织部分。将组织部分细分切割后, 第一个模型开始识别, 区分癌与非癌, 用绿色表示非癌区块, 红色表示癌区块, 按识别结果重新拼接为大图, 将识别结果可视化。癌与非癌区分完成后, 依照第一个模型的分标记提取癌区块部分, 输入第2个模型, 区分高分化、中分化、低分化, 用红色表示低分化, 橙色表示中分化, 黄色表示高分化, 同样按识别结果重新拼接为大图, 将识别结果可视化。由此完成整张切片的癌与非癌分类诊断与癌区域的分级识别, 结果见图8。

在实地场景应用时, 也需考虑模型计算成本, 前文中提到使用了 RTX4080 Laptop 训练模型, 主要是使用 RTX4080 Laptop GPU 进行训练过程的加速, 训练完成后, 保存已经训练好的模型。当模型运行

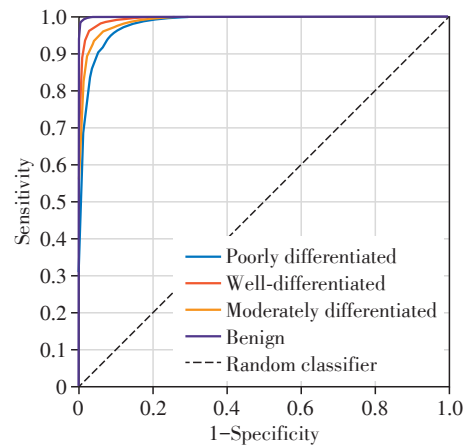


图7 分级模型的ROC曲线

Figure 7 ROC curve of grading model

时, 可直接调用训练好的模型识别, 计算成本与模型本身可学习的参数总数有关。若要利用GPU进行加速, 也可通过调整模型输入图片的batchsize来调整显存占用。为评估计算成本, 本研究统计每个模型的可学习参数总数对比图(图9)。由数据可知, 大部分模型参数占用均在200 MB以下, 部署计算成

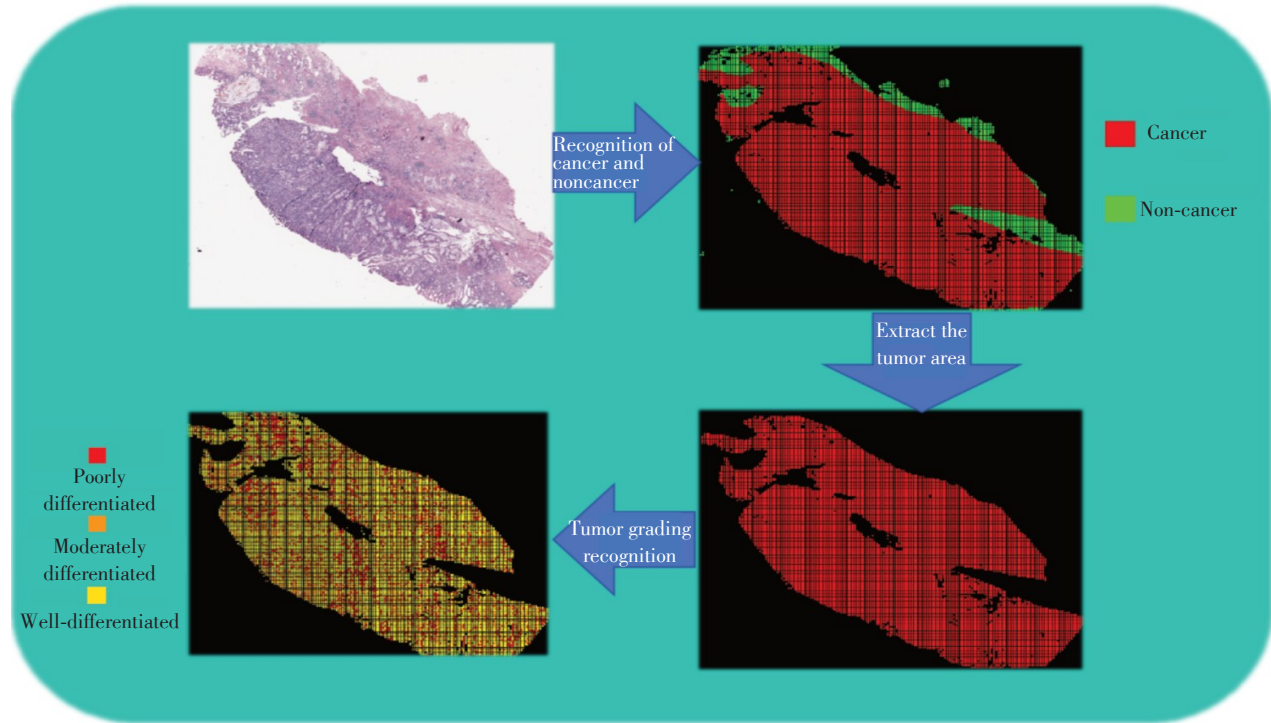


图8 可视化结果示例

Figure 8 Example of visualization results

本较低,即使是拥有较高参数的VGG系列,现代计算资源也可以轻松部署。而集成模型部署时可分别调用子模型,综合来看,部署可行性同样较高。

3 讨论

胃癌是全球范围内常见的恶性肿瘤之一,早期胃癌的治愈率较高,但多数患者在确诊时已处于晚期,治疗难度大且预后差。因此,早期筛查和诊断对降低胃癌的病死率至关重要。然而,传统的胃癌筛查方法存在诸多挑战。例如,内镜检查依赖于内镜医师的经验和技能,且检查过程较为痛苦,患者

依从性较低。此外,病理诊断也存在一定的主观性,不同病理学家对早期病变的判断可能存在差异。这些问题都限制了胃癌筛查的普及和准确性。在这样的背景下,开发高精度、标准化的AI辅助诊断工具能够减少人为误差,提高诊断准确性,同时提高筛查速度、促进筛查诊疗规范化,对提升胃癌早筛效率具有重要的临床价值。近期研究也进一步强调标准化AI工具在提升早期胃癌检测效率中的关键作用。Wang等^[13]通过深度学习系统显著提高了内镜图像中早期胃癌病变的诊断准确率(AUC=0.96),并将内镜医师间的观察差异性降低了

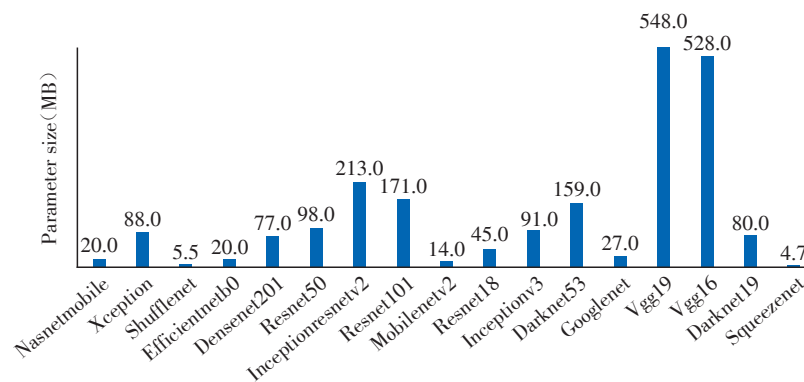


图9 模型参数图

Figure 9 Model parameter diagram

38%。类似地, Park 等^[14]的一项多中心随机试验表明,与传统方法相比, AI 辅助框架在医疗资源有限的环境中(尤其是专科医生不足的地区)将筛查效率提升了 25%。这些进展与临床需求高度契合,突显了发展标准化 AI 驱动工作流程在大规模胃癌筛查中的必要性。Xia 等^[15]利用 ResNet50 提取特征进行多任务分类,加入多实例学习框架与注意力机制,在 3 个外部测试队列中 AUC 均>0.90,实现了对慢性胃炎组织学改变的精准、高效分类。Lan 等^[16]提出一种名为 LHFF-MIL 的新型多示例学习框架,用 Transformer 架构捕获不同分辨率 WSI 块间的相关性,筛选关键区域,引入切片 Wasserstein 距离(sliced-Wasserstein distance)解决高维计算难题,通过层次注意力机制动态融合多分辨率特征,避免无关区域干扰,提升 WSI 分类的准确性。Zubair 等^[17]构建了一种基于多通道注意力机制和迁移学习的可解释性框架,通过加权投票机制融合三通道的分类决策值,用于胃镜病理图像的胃癌分类任务。Li 等^[18]提出了一种针对胃癌组织病理学图像诊断的深度学习框架 DeepNCCNet,在 6 种预训练 CNN 模型中进行对比,最终基于 MobileNetV2 加权平均整合癌与非癌区域特征,提升诊断性能。Zhao 等^[19]结合预训练的 ResNet-34 提取深度学习特征与手工特征,引入多头图注意力网络与层级结构,构建名为 SurvGraph 的新型混合图注意力网络模型,用于基于胃癌 WSI 的生存预测。以往研究多依赖单个或少数模型进行拓展深化(如 CNN、Transformer、图网络),通过引入注意力机制,调整数据利用策略,融合不同层级以优化模型性能。未考虑利用不同 CNN 网络的特征提取差异性,将多个模型集成融合,进行优势互补,从而进行整合优化。本研究系统性构建大规模 CNN 进行横向对比,探究不同模型间特点,并利用模型深度学习特征提取差异互补的特点,通过集成学习进行整合优化,提升模型对胃癌的分级性能。

本研究使用 40 倍扫描倍率切割图像并提取深度学习特征,相比于 10 倍、20 倍扫描倍率,其所提取的数据量较多,可学习更细微的特征数据,对分级等复杂任务,有望达成更好的效果^[20],获得更高的识别灵敏度^[21]。在现实世界病理切片观察中,40 倍倍率通常用于观察细胞级别的细节,如细胞核的形态特征、核浆比、染色质结构等,深度学习模型比起传统手动提取特征的机器有着更好的性能,更优的识别效率,但深度学习所提取的特征往往可解释性较低,难以解释模型学习特征与细胞细节的因果与

关联性,这也是推动人工智能在医疗领域应用的一大阻碍^[22],下一步研究将尝试拓展模型可解释性,增强模型辅助诊断的严谨性与可靠性。

本研究系统性构建并评估了 17 种(包含经典网络、超轻量级网络和多方位优化的新型网络)CNN 架构,最大程度地对比发掘出网络优势。揭示了不同模型优化特点与胃癌病理图像特征匹配之间的动态平衡关系。其中, Xception 因综合耗时与准确率最优,被选为最终分类诊断模型网络架构。经过训练优化,以准确率 98.13%、灵敏度 98.11%、特异度 98.11%、F1 分数 98.12%、AUC 0.998 3 的成绩展现出优秀的识别性能。Xception 模型在分类诊断任务中的优异表现可归因于其核心组件——深度可分离卷积(depthwise separable convolution)。这种结构将标准卷积分解为逐通道的空间卷积(depthwise convolution)和点卷积(pointwise convolution),显著减少了模型参数数量和计算复杂度(约降低 1 个数量级)。这种设计迫使网络更高效地学习空间特征与通道特征之间的解耦关系。在病理图像分析中,表现为模型能更专注地捕获多尺度的组织学关键特征(如腺体结构的扭曲度、细胞核的异型性、核分裂像的密度),同时有效抑制由 HE 染色批次差异、组织折叠或切片伪影等引入的无关噪声信息,从而在保持高精度的同时避免了过拟合风险^[23]。

在初步训练效能对比中, SqueezeNet 等超轻量模型对于病理图像识别的潜力较弱,参数量较大的模型如 DenseNet、InceptionResNetV2 等未达到预期, EfficientNet、ResNet、NASNetMobile、ShuffleNet、Xception 等资源需求程度不大的中或轻量模型已表现出足够的识别精度。本研究选取此类模型是兼顾资源与性能的较优方案,可为之后相关肿瘤病理图像分类研究提供一定参考。然而,部分轻量化模型(如 MobileNetV2)在训练速度与精度间的权衡表明,边缘设备部署需进一步优化网络结构与量化策略^[24]。同时值得注意的是,在分类模型初步训练与分级模型基模型训练中,同一模型的性能指标会有所波动,这表明在不同训练集上,不同模型的性能表现可能会有所不同,且经过参数调整,可以缩小不同模型之间的性能差距。

迁移学习的成功应用与 Raghu 等^[25]的研究结论一致,即自然图像预训练特征在医学域中具有强泛化潜力。然而,领域自适应(domain adaptation)的不足可能导致模型对染色差异(如 HE 染色批次效应)敏感^[26]。未来研究可结合对比学习(contrastive

learning)增强跨域特征对齐,以提升模型对异质数据集的鲁棒性^[27]。

在分类模型构建完成的基础上,本研究进一步探究CNN在肿瘤病理图像分级识别的应用。训练结果显示,相对于分类识别,分级识别的任务复杂度明显提高,模型的性能均出现了不同程度的下降。为了提升分级识别性能,本研究尝试通过集成学习达成相比单一模型更优的性能。结果显示,集成学习可在单一模型的基础上继续突破,拥有更高的图像识别精度上限与潜力。其中,堆叠法的提升效果优于投票法。最终分级识别模型准确率达95.06%、灵敏度达94.77%、特异度达98.36%、F1分数达94.82%,BN AUC达0.999 4,TP AUC达0.981 1, TM AUC达0.989 6, TW AUC达0.995 1。相比单一模型,其准确率、灵敏度、特异度都拥有一定提升,达成更高的F1分数。Stacking集成框架在分级任务中展现出的显著性能提升(较平均单模型准确率提高3.10%,较硬投票法提高0.68%),其本质原因在于它有效利用了基模型之间的多样性和特征互补性。如表2所示,本研究筛选的13个基模型(如ResNet101、EfficientNetB0、DenseNet201、ResNet50、MobileNetV2、Darknet53、Xception等)虽然整体性能较好,但它们在架构设计、特征提取偏好和错误模式上存在显著差异(如ResNet系列可能擅长捕捉深层语义特征,而MobileNetV2/EfficientNetB0在纹理和局部模式识别上可能更优)。这种多样性至关重要,它确保了基模型对同一张病理图像可能产生不同视角或侧重面的预测(概率向量)。Stacking方法的核心优势在于其元模型(本研究中为随机森林)能够学习如何最优组合这些异构基模型的预测输出(即概率向量拼接形成的新特征空间)。随机森林作为强大的非线性分类器,能够从基模型的“集体智慧”中挖掘更深层次的、单一模型难以捕捉的模式关联,特别是对于胃腺癌分级这种需要识别细微组织学差异(如腺管形成程度、细胞极性、核异型性程度)的连续谱系任务。元模型通过学习基模型预测之间的复杂非线性关系,能够校正单个模型的系统性偏差或特定类别的误判,从而实现更鲁棒和准确的最终预测^[28]。相比之下,硬投票法仅进行简单的多数表决,无法充分利用基模型预测中蕴含的丰富概率信息和互补性。

为模拟真实应用场景,本研究构建了一个输入全切片图像-分割-癌与正常组织分类识别-癌分级诊断-可视化结果的端到端处理流程,将识别结果进行可视化,进一步加强研究的可用性,并且根据模型参数评

估了计算资源。综合来看,模型应用较高的可部署性与一定的实用性,为真实世界的应用提供了理论基础与试验框架,未来研究可据此进一步深化验证。

本研究的高分类诊断准确率(98.13%)与高分级准确率(95.06%)为胃癌个性化治疗提供了关键技术支持。例如,精准识别中分化腺癌(TM类)可辅助临床医生制定靶向治疗方案,降低过度化疗风险^[29]。此外,AI辅助诊断系统可缓解病理学家的工作负荷,尤其适用于医疗资源匮乏地区。Tomita等^[30]研究表明,AI与病理学家的协同诊断可将误诊率降低至2.1%。然而,临床落地仍需解决模型可解释性问题,如通过Grad-CAM可视化关键病理区域,增强医生对AI决策的信任度^[31]。

在DeepSeek带来的浪潮下^[32],多家医院开始接入语言大模型,构建医疗智能应用平台。创建“大模型中枢+专科智能体”协同体系,调用自有AI小模型形成诊疗链(chatflow)的方式拥有巨大的应用潜力。本研究的迁移-集成联合策略对于AI小模型的快速部署实现具有一定辅助意义。

本研究的主要局限性包括:①由于数据集限制,未能进一步识别病理图片的其他内容,如亚型分类等,应用范围受限;②研究模型仅进行了胃腺癌的分级,识别种类较少,更多的胃癌亚型分级有待研究;③数据来源集中于东亚人群,可能限制模型对欧美胃癌图像识别的泛化能力^[33];④未纳入多模态数据(如基因组或免疫组化染色),限制了图像识别预测的全面性^[34]。

未来研究可从以下方向拓展:①构建跨种族多中心数据集,结合联邦学习(federated learning)在保护数据隐私的同时增强泛化性^[35];②开发多任务学习框架,同步实现分类、分级与生存分析^[36];③结合自监督或弱监督学习,探究深度学习特征提取与病理图像现实世界特征的关联,并拓展其可解释性^[37];④进一步加入细胞核的形态特征、核浆比、染色质结构等手动提取特征,对比、融合以提升模型性能。未来工作将聚焦于多模态数据融合与轻量级实践部署,增强模型的可解释性与准确度,以实现胃癌诊断从“辅助工具”向“决策核心”的转化。本研究的方法论也可扩展至其他癌症类型,推动AI在精准医学中的纵深应用。

综上,本研究针对胃肿瘤病理图像的分类诊断与分级识别构建模型,通过多模型横向对比与异构集成策略,系统优化了胃癌病理图像模型的分类与分级性能,并创新性地融合迁移学习与集成学习,模拟现实世界的端到端识别辅助流程,验证了该联

合框架的临床适用性。实验表明, Xception 模型在分类诊断任务中表现最优(准确率 98.13%), 而 Stacking 集成框架显著提升了胃腺癌的分级识别精度(较单一模型平均提高 3.1%)。两个模型均拥有优秀的识别性能, 为数字化病理系统的开发提供了理论辅助依据与技术蓝图。

利益冲突声明:

全体作者声明没有利益冲突。

Conflict of Interests:

The authors declared no conflict of interests.

作者贡献声明:

吴邦钰负责数据收集、模型构建、试验方案与数据分析、论文撰写与修改; 王志霄负责代码修改、模型优化、研究方法优化指导、论文修改; 项景轩负责论文撰写与修改、文献查找、资料收集、方案设计、数据分析; 孙立负责提供研究方案、进行技术支持与指导、论文修订; 马玲负责提供研究方案与项目基金支持, 进行技术支持与指导、论文修订。

Author's Contributions:

WU Bangyu was responsible for data collection, model construction, experiments and data analysis, as well as writing and revising the manuscript; WANG Zhixiao was in charge of code modification, model optimization, and guidance on the optimization of research methods; XIANG Jingxuan was responsible for writing and revising the manuscript, literature search, data collection, scheme design, and data analysis; SUN Li was responsible for providing the research plan, offering technical support and guidance, and revising the manuscript; MA Ling was responsible for providing the research plan and funding support, offering technical support and guidance, and revising the manuscript.

[参考文献]

- [1] BRAY F, LAVERSANNE M, SUNG H, et al. Global cancer statistics 2022: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries [J]. *CA Cancer J Clin*, 2024, 74(3): 229-263
- [2] HAN B F, ZHENG R S, ZENG H M, et al. Cancer incidence and mortality in China, 2022 [J]. *J Natl Cancer Cent*, 2024, 4(1): 47-53
- [3] LEI C D, SUN W Q, WANG K, et al. Artificial intelligence-assisted diagnosis of early gastric cancer: present practice and future prospects [J]. *Ann Med*, 2025, 57(1): 2461679
- [4] 徐思雨, 张永杰, 田水, 等. 深度学习重建算法在胰腺 HASTE-T2WI 序列中的临床应用价值[J]. *南京医科大学学报(自然科学版)*, 2025, 45(6): 810-815, 825
- [5] XU S Y, ZHANG Y J, TIAN S, et al. Clinical application value of deep learning reconstruction algorithm in pancreatic HASTE-T2WI sequence [J]. *Journal of Nanjing Medical University (Natural Sciences)*, 2025, 45(6): 810-815, 825
- [6] 唐蛰雨, 李绍钦, 贾中芝. 基于 CT 图像的深度学习在主动脉夹层中的应用进展[J]. *南京医科大学学报(自然科学版)*, 2024, 44(8): 1174-1178
- [7] TANG Z Y, LI S Q, JIA Z Z. Research progress on the application of deep learning based on CT images in aortic dissection [J]. *Journal of Nanjing Medical University (Natural Sciences)*, 2024, 44(8): 1174-1175
- [8] LEE C F, LIN J, HUANG Y L, et al. Deep learning-based breast MRI for predicting axillary lymph node metastasis: a systematic review and meta-analysis [J]. *Cancer Imaging*, 2025, 25(1): 44
- [9] WU Y J, MA J, HUANG X S, et al. DeepMMSA: a novel multimodal deep learning method for non-small cell lung cancer survival analysis [C]//2021 IEEE International Conference on Systems, Man, and Cybernetics (SMC). Melbourne, Australia. IEEE, 2022: 1468-1472
- [10] TAN M X, LE Q V. EfficientNetV2: smaller models and faster training [C]//International Conference on Machine Learning., 2021
- [11] HOWARD A, SANDLER M, CHEN B, et al. Searching for MobileNetV3 [C]//2019 IEEE/CVF International Conference on Computer Vision (ICCV). Seoul, Korea. IEEE, 2019: 1314-1324
- [12] WANG W C, AHN E, FENG D G, et al. A review of predictive and contrastive self-supervised learning for medical images [J]. *Mach Intell Res*, 2023, 20(4): 483-513
- [13] ZAMANITAJEDDIN N, JAHANIFAR M, XU K, et al. Benchmarking domain generalization algorithms in computational pathology [EB/OL]. [2025-09-14]. <https://arxiv.org/abs/2409.17063>
- [14] LE VUONG T T, KIM K, SONG B, et al. Joint categorical and ordinal learning for cancer grading in pathology images [J]. *Med Image Anal*, 2021, 73: 102206
- [15] MA L Y, SU X F, MA L Y, et al. Deep learning for classification and localization of early gastric cancer in endoscopic images [J]. *Biomed Signal Process Control*, 2023, 79: 104200
- [16] JIN P, JI X Y, KANG W Z, et al. Artificial intelligence in gastric cancer: a systematic review [J]. *J Cancer Res Clin Oncol*, 2020, 146(9): 2339-2350
- [17] XIA K, HU Y H, CAI S T, et al. GastritisMIL: an interpretable deep learning model for the comprehensive histological assessment of chronic gastritis [J]. *Patterns*, 2025, 6(8): 101286
- [18] LAN Q Z, WU Y Z, DING W P, et al. Effective latent hierarchical feature fusion in multiple instance learning for whole slide image classification [J]. *Appl Soft Comput*, 2025, 177: 113191

- [17] ZUBAIR M, OWAIS M, HASSAN T, et al. An interpretable framework for gastric cancer classification using multi-channel attention mechanisms and transfer learning approach on histopathology images [J]. *Sci Rep*, 2025, 15(1): 13087
- [18] LI L L, GENG Y, CHEN T, et al. Deep learning model targeting cancer surrounding tissues for accurate cancer diagnosis based on histopathological images [J]. *J Transl Med*, 2025, 23(1): 110
- [19] ZHAO Y S, LI L S, YU X, et al. SurvGraph: a hybrid-graph attention network for survival prediction using whole slide pathological images in gastric cancer [J]. *Neural Netw*, 2025, 189: 107607
- [20] TOĞAÇAR M, ÖZKURT K B, ERGEN B, et al. Breast-Net: a novel convolutional neural network model through histopathological images for the diagnosis of breast cancer [J]. *Phys A Stat Mech Appl*, 2020, 545: 123592
- [21] 金怀平, 薛飞跃, 李振辉, 等. 基于病理图像集成深度学习的胃癌预后预测方法 [J]. *电子与信息学报*, 2023, 45(7): 2623–2633
- JIN H P, XUE F Y, LI Z H, et al. Gastric cancer prognosis prediction method based on integrated deep learning of pathological images [J]. *Journal of Electronics & Information Technology*, 2023, 45(7): 2623–2633
- [22] 陈园琼, 邹北骥, 张美华, 等. 医学影像处理的深度学习可解释性研究进展 [J]. *浙江大学学报(理学版)*, 2021, 48(1): 18–29, 40
- CHEN Y Q, ZOU B J, ZHANG M H, et al. Research progress on deep learning interpretability in medical image processing [J]. *Journal of Zhejiang University (Science Edition)*, 2021, 48(1): 18–29, 40
- [23] CHOLLET F. Xception: deep learning with depthwise separable convolutions [C]//2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Honolulu, HI, USA. IEEE, 2017: 1800–1807
- [24] SANDLER M, HOWARD A, ZHU M L, et al. MobileNetV2: inverted residuals and linear bottlenecks [C]//2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Salt Lake City, UT, USA. IEEE, 2018: 4510–4520
- [25] RAGHU M, ZHANG C Y, KLEINBERG J M, et al. Transfusion: understanding transfer learning for medical imaging [C]//Neural Information Processing Systems., 2019
- [26] TELLEZ D, LITJENS G, VAN DER LAAK J, et al. Neural image compression for gigapixel histopathology image analysis [J]. *IEEE Trans Pattern Anal Mach Intell*, 2021, 43(2): 567–578
- [27] CHEN T, KORNBLITH S, NOROUZI M, et al. A simple framework for contrastive learning of visual representations [C]. *Proceedings of the 37th International Conference on Machine Learning*, 2020, 119: 1597–1607
- [28] NAIMI A I, BALZER L B. Stacked generalization: an introduction to super learning [J]. *Eur J Epidemiol*, 2018, 33(5): 459–464
- [29] AJANI J A, D'AMICO T A, BENTREM D J, et al. Gastric cancer, version 2.2022, NCCN clinical practice guidelines in oncology [J]. *J Natl Compr Canc Netw*, 2022, 20(2): 167–192
- [30] TOMITA N, ABDOLLAHI B, WEI J, et al. Attention-based deep neural networks for detection of cancerous and precancerous esophagus tissue on histopathological slides [J]. *JAMA Netw Open*, 2019, 2(11): e1914645
- [31] SELVARAJU R R, COGSWELL M, DAS A, et al. Grad-CAM: visual explanations from deep networks via gradient-based localization [J]. *Int J Comput Vis*, 2020, 128(2): 336–359
- [32] SANDMANN S, HEGSELMANN S, FUJARSKI M, et al. Benchmark evaluation of DeepSeek large language models in clinical decision-making [J]. *Nat Med*, 2025, 31(8): 2546–2549
- [33] CRISTESCU R, LEE J, NEBOZHYN M, et al. Molecular analysis of gastric cancer identifies subtypes associated with distinct clinical outcomes [J]. *Nat Med*, 2015, 21(5): 449–456
- [34] GAO F, DING J X, GAI B W, et al. Interpretable multi-modal fusion model for bridged histology and genomics survival prediction in pan-cancer [J]. *Adv Sci*, 2025, 12(17): 2407060
- [35] ZOLKOV E, GINZBERG N, COHEN E. A mm-wave blocker-tolerant receiver achieving <4 dB NF and -3.5 dBm B1dB in 65-nm CMOS [C]. *2023 IEEE Radio Frequency Integrated Circuits Symposium (RFIC)*, 2023: 297–300
- [36] LEE K, CAVALCANTI T C, KIM S, et al. Multi-task and few-shot learning-based fully automatic deep learning platform for mobile diagnosis of skin diseases [J]. *IEEE J Biomed Health Inform*, 2023, 27(1): 176–187
- [37] 张建新, 高程阳, 孙 鉴, 等. 混合监督学习的乳腺癌全切片病理图像分类 [J]. *中国图象图形学报*, 2024, 29(9): 2825–2836
- ZHANG J X, GAO C Y, SUN J, et al. Breast cancer whole slide pathological image classification based on hybrid supervised learning [J]. *Journal of Image and Graphics*, 2024, 29(9): 2825–2836

(收稿: 2025-03-26; 修回: 2025-09-17; 录用: 2025-09-19)

(本文编辑: 戴玉娟)