中华医学会系列杂志

ISSN 1007-5232
CN 32-1463 / R

中华消化内镜杂志®

ZHONGHUA XIAOHUA NEIJING ZAZHI

2025年11月　第42卷　第11期

CHINESE JOURNAL OF DIGESTIVE ENDOSCOPY

Volume 42　Number 11
November 2025

ISSN 1007-5232
9 771007 523250 11





中华消化内镜杂志®

CHINESE JOURNAL OF DIGESTIVE ENDOSCOPY

月刊　1996年8月改刊　第42卷　第11期　2025年11月20日出版


微信:xhnjxw


新浪微博

主　管
中国科学技术协会

主　办
中华医学会
100710,北京市东四西大街42号

编　辑
中华消化内镜杂志编辑委员会
210003,南京市紫竹林3号
电话:(025)83472831,83478997
传真:(025)83472821
Email:xhnj@xhnj.com
http://www.zhxhnjzz.com
http://www.medjournals.cn

总编辑
张澍田

编辑部主任
唐涌进

出　版
《中华医学杂志》社有限责任公司
100710,北京市东四西大街42号
电话(传真):(010)51322059
Email:office@cmaph.org

广告发布登记号
广登32010000093号

印　刷
江苏省地质测绘大队

发　行
范围:公开
国内:南京报刊发行局
国外:中国国际图书贸易集团有限公司
(北京399信箱,100048)
代号 M4676

订　购
全国各地邮政局
邮发代号 28-105

邮　购
中华消化内镜杂志编辑部
210003,南京市紫竹林3号
电话:(025)83472831
Email:xhnj@xhnj.com

定　价
每期25.00元,全年300.00元

中国标准连续出版物号
ISSN 1007-5232
CN 32-1463/R

2025年版权归中华医学会所有

未经授权,不得转载、摘编本刊文章,不得使用本刊的版式设计

除非特别声明,本刊刊出的所有文章不代表中华医学会和本刊编委会的观点

本刊如有印装质量问题,请向本刊编辑部调换

目　次

专家论坛

单中心消化内镜诊疗镇静/麻醉20年经验总结与展望 …… 841
左赞　郭强

论　著

基于人工智能的内镜虚拟标尺测量食管静脉曲张直径的初步研究(含视频) …… 848
曹传坤　晋晶　张衡　蔡瑞　肖婷　梅雪灿　孔德润

基于卷积网络建立的小肠胶囊内镜人工智能辅助自动识别系统 …… 853
陈健　孙斌　王甘红　夏开建　徐晓丹

超声内镜引导胃肠吻合术直接法和双气囊封堵法治疗胃流出道梗阻疗效对比的回顾性队列研究(含视频) …… 864
吴兆荣　展薇　李文婷　田恬　殷芹　沈珊珊　王雷　李雯

内镜黏膜下剥离术在原发性胃肠道淋巴瘤诊断中的价值 …… 871
高华　杨丽萍　黄晓铨　江振健　何杰　魏征　王豆　马丽黎

自控式链圈联合组织夹牵引辅助技术在早期结直肠肿瘤内镜黏膜下剥离术中的应用 …… 875
余超　董弢　许远　荆琳　王耀辉　刘春祥　肖君

气囊辅助小肠镜诊治小肠Dieulafoy病变的效果和安全性初步分析(含视频) …… 881
赵东林　徐梦楠　江志锰　张静　于妍　肖年军　李白容　范崇熙　宁守斌　孙涛

新型内镜口鼻屏蔽罩在上消化道内镜检查中减少气溶胶传播的研究 …… 887
王俊雄　次仁伦珠　次央　德吉　支章　尼玛莎珍　次旦央宗　巴桑卓玛　伊比然恨　李鹏　张澍田

内镜逆行胰胆管造影术治疗儿童门冬酰胺酶相关胰腺炎的有效性及安全性 …… 892
林珑　杨凯华　邓朝晖

短篇论著

内镜诊治十二指肠静脉曲张破裂出血5例的临床分析 …… 898

隗永秋　王文海　孟凡冬　冀明　王拥军　李鹏　张澍田

病例报道

滑轮牵引联合肌切开辅助内镜黏膜下剥离术治疗早期十二指肠球部癌1例(含视频) …… 902

郑烨　王鹏　季锐

单气囊小肠镜检查后肠系膜血肿形成并腹腔出血1例 …… 904

杨明兰　管恩翠　朱慧云　田峰

会议纪要

第五十三届“友谊”消化内镜沙龙——EUS篇会议纪要 …… 907

张政　陈光勇　王拥军　冀明　李鹏　张澍田

综　　述

消化内镜碳足迹的研究进展 …… 909

陈俐娜　邵青梅　吴冰冰　金静芬

结直肠无蒂锯齿状病变伴异型增生的研究进展 …… 915

格旦　赵贝　孙文琦　李逸凡　王维维　吴欣荣　熊龙英　徐佳乐　窦晓坛　陈敏　王雷

超声内镜引导胃肠吻合术的临床应用进展 …… 921

陈航宇　张平平　王凯旋

读者·作者·编者

《中华消化内镜杂志》2025年可直接使用英文缩写的常用词汇 …… 863

《中华消化内镜杂志》对来稿中统计学处理的有关要求 …… 914

插页目次 …… 847

本刊稿约见第42卷第1期第82页、第7期第586页

本期责任编辑　顾文景

本刊编辑部工作人员联系方式

唐涌进，Email：tang@xhnj.com　　朱　悦，Email：zhuyue@xhnj.com

周　昊，Email：zhou@xhnj.com　　钱　程，Email：qian@xhnj.com

顾文景，Email：gwj@xhnj.com　　许文立，Email：xwl@xhnj.com

本刊投稿方式

登录《中华消化内镜杂志》官方网站http://www.zhxhnjzz.com进行在线投稿。


唐涌进


周　昊


顾文景


朱　悦


钱　程


许文立

（扫码添加编辑企业微信）

·论著·

基于卷积网络建立的小肠胶囊内镜人工智能辅助自动识别系统

陈健[1]　孙斌[2]　王甘红[3]　夏开建[4]　徐晓丹[1]

[1]常熟市第一人民医院消化内科，苏州 215500；[2]常熟市尚湖中心医院消化内科，苏州 215500；[3]常熟市中医院消化内科，苏州 215500；[4]常熟市医学人工智能与大数据重点实验室，苏州 215500

通信作者：徐晓丹，Email：xxddocter@gmail.com

【摘要】 目的　基于卷积神经网络开发一款能够自动识别多种病变的小肠胶囊内镜人工智能辅助系统，并测试其性能。**方法**　采用3个小肠胶囊内镜图像数据集进行深度学习模型的训练（*n*=26 638）、验证（*n*=6 652）和测试（*n*=1 013），每一数据集均包括了12类小肠病变（血管畸形、出血、糜烂、红斑、狭窄、淋巴管扩张、黏膜下肿瘤、息肉、淋巴滤泡、异物、静脉以及正常黏膜）。测试时，使用接收者操作特征曲线下面积（area under curve，AUC）、敏感度、特异度、精确度、准确率、F1分数等衡量模型的性能，并开展人机对比实验。**结果**　最优模型（EfficientNet-CE）诊断12类小肠病变的敏感度（加权平均）为86.28%，特异度（加权平均）为98.67%，AUC（加权平均）为0.987 4。人机大赛中，最优模型（EfficientNet-CE）在准确率（加权平均：86.28%）和处理速度（52.43帧/s）方面均表现出色，处理速度约为低年资内镜医师（阅片经验不足3年）的42.4倍、高年资内镜医师（具有超过5年阅片经验）的40.3倍。**结论**　基于卷积神经网络开发的深度学习模型能够快速、准确地识别12类小肠病变，通过高灵敏度的识别能力，能够有效辅助内镜医师进行小肠胶囊内镜检查的阅片工作。

【关键词】 胶囊内窥镜；　深度学习；　人工智能；　计算机辅助；　卷积神经网络

基金项目：苏州市第二十三批科技发展计划项目（SLT2023006）；常熟市医学人工智能与大数据重点实验室能力提升项目（CYZ202301）；苏州市临床重点病种诊疗技术专项项目（LCZX202334）；苏州市科技发展计划项目（SYW2025034）

Establishment of an artificial intelligence-assisted system for automatic lesion recognition in small intestinal capsule endoscopy based on convolutional networks

Chen Jian[1], Sun Bin[2], Wang Ganhong[3], Xia Kaijian[4], Xu Xiaodan[1]

[1]Department of Gastroenterology, Changshu No.1 People's Hospital, Suzhou 215500, China; [2]Department of Gastroenterology, Changshu City Shanghu Central Hospital, Suzhou 215500, China; [3]Department of Gastroenterology, Changshu Traditional Chinese Medicine Hospital, Suzhou 215500, China; [4]Department of Information Engineering, Changshu Key Laboratory of Medical Artificial Intelligence and Big Data, Changshu 215500, China

Corresponding author: Xu Xiaodan, Email: xxddocter@gmail.com

【Abstract】 Objective　To develop and validate an artificial intelligence-assisted system based on convolutional neural networks (CNN) for automatic lesion recognition in small intestinal capsule endoscopy. **Methods**　Three small intestinal capsule endoscopy datasets were used for training (*n*=26 638), validating (*n*=6 652), and testing (*n*=1 013) the deep learning model, covering 12 lesion categories, including vascular malformations, hemorrhage, erosion, erythema, stenosis, lymphangiectasia, submucosal tumors,

DOI：10.3760/cma.j.cn321463-20240306-00103

收稿日期　2024-03-06　　**本文编辑**　顾文景

引用本文：陈健，孙斌，王甘红，等. 基于卷积网络建立的小肠胶囊内镜人工智能辅助自动识别系统[J]. 中华消化内镜杂志，2025，42(11)：853-863. DOI：10.3760/cma.j.cn321463-20240306-00103.

www.zhxhnjzz.com

polyps, lymphoid follicles, foreign bodies, veins, and normal mucosa. CNN performance was measured by area under receiver operating characteristic curve (AUC), sensitivity, specificity, precision, accuracy, and F1 score, with comparisons with endoscopists of different experience levels. **Results** The top-performing model (EfficientNet-CE) achieved 86.28% sensitivity, 98.67% specificity, and AUC of 0.987 4 across all categories. It demonstrated high accuracy (86.28%) and a processing speed of 52.43 frames per second, approximately 42.4 times faster than junior endoscopists (<3 years′ experience) and 40.3 times faster than senior endoscopists (>5 years′ experience). **Conclusion** The CNN-based model allows rapid, accurate identification of 12 small intestinal lesion types and effectively supports endoscopists in reviewing capsule endoscopy examinations due to its high sensitivity.

【**Key words**】 Capsule endoscopy; Deep learning; Artificial intelligence; Computer-assisted; Convolutional neural networks

Fund program: Suzhou Science and Technology Development Plan Project (SLT2023006); Changshu Key Laboratory of Medical Artificial Intelligence and Big Data Capability Enhancement Project (CYZ202301); Suzhou Special Project on Clinical Key Disease Diagnosis and Treatment Technologies (LCZX202334); Suzhou Science and Technology Key Project (SYW2025034)

随着医学技术的进步，胶囊内镜已经变革了消化道疾病的诊断方式。作为一种无创、无需麻醉的检查手段，胶囊内镜以其患者友好的特性和高依从性，已经成为小肠疾病诊断的金标准[1-2]。自2000年首次被引入临床以来，胶囊内镜已经证明其在多项适应证中的有效性，包括隐性出血、糜烂以及小肠息肉等[3]。然而，胶囊内镜的一个主要挑战在于其产生的大量图像数据，一次检查可以产生高达数万张图像，这不仅导致了医生在阅读图像时的时间压力，还因阅片疲劳而可能增加漏诊风险。

人工智能技术的兴起为胶囊内镜领域带来了新的希望。人工智能的应用已在消化内镜诊断中取得显著成果，特别是在处理传统内镜中的胃肠道病变方面[4-5]。在胶囊内镜这样一个图像密集型的领域，人工智能的利用尤其具有潜在优势。通过自动化图像分析，人工智能有助于优化病变的识别和辅助诊断过程，减轻医生的工作负担，提高诊断效率和准确性。

随着胶囊内镜在临床上的广泛应用，多个胶囊内镜品牌相继问世。在我国，常用的胶囊内镜品牌包括PillCam（Medtronic，美国）、EndoCapsule（Olympus，日本）和OMOM（金山科技，中国）。不同制造商的胶囊内镜系统在颜色和亮度上存在细微差异，这意味着基于单一品牌图像构建的人工智能模型在通用性和泛化能力上可能存在不足。此外，早期人工智能在胶囊内镜应用方面的研究主要聚焦于单一类型病变的识别，如出血点、血管畸形和息肉等，虽然这类模型构建相对容易，但能自动识别多种病变类型以确保消化道病变诊断全面性和准确性的研究仍较为缺乏。本研使用涵盖3种品牌胶囊内镜图像的数据集，旨在开发一个能识别12类小肠病变的人工智能模型，以提高小肠病变诊断的全面性和准确性。

资料与方法

一、数据集

本研究基于3个数据集，共计34 303张图像进行：数据集1（来自SEE-AI公开数据库），数据集2（来自Kvasir-Capsule公开数据库），数据集3（来自常熟市第一人民医院）。收集的图像包括12类小肠病变的胶囊内镜图像，由3种不同品牌的胶囊内镜设备拍摄，分别是PillCam SB3、EndoCapsule和OMOM。SEE-AI公开数据库[6]是由Medtronic（位于美国明尼苏达州明尼阿波利斯）的PillCam SB3小肠胶囊内镜获得的图像组成（https://www.kaggle.com/datasets/capsuleyolo/kyucapsule?resource=download）。Kvasir-Capsule公开数据库[7]是由来自挪威的Vestre Viken医院使用EndoCapsule系统收集的图像组成（https://osf.io/dv2ag/）。常熟市第一人民医院提供的是OMOM胶囊收集的小肠病变图像。上述3个数据集的图像被随机分入训练集（n=26 638）、验证集（n=6 652）和测试集（n=1 013）。数据集中具有代表性的注释图像详见图1，注释标签涵盖了多种病变及正常结构，包括血管畸形（疑似毛细血管病变的红斑区域）、出血（明显出血区域，排除变暗且远离出血源的部位）、糜烂（黏膜损伤区域，如侵蚀、溃疡和缺口）、红斑（可能与炎症相关的红斑和水肿区域）、狭窄（收缩和僵硬区域）、淋巴管扩张（包含大于一个点的淋巴管的区域）、黏膜下肿瘤（类似黏膜下肿瘤的区域）、息肉（有基底的隆起病变或疑似腺瘤区域）、淋巴滤泡（包含正常滤泡和疑似淋巴滤泡的区域）、异物（除食物外的异物）、静脉（带有静脉结构的区域）以及正常黏膜（未见上述病变的正常黏膜）。本项回顾性队列研究已获得常熟市第一人民

www.zhxhnjzz.com

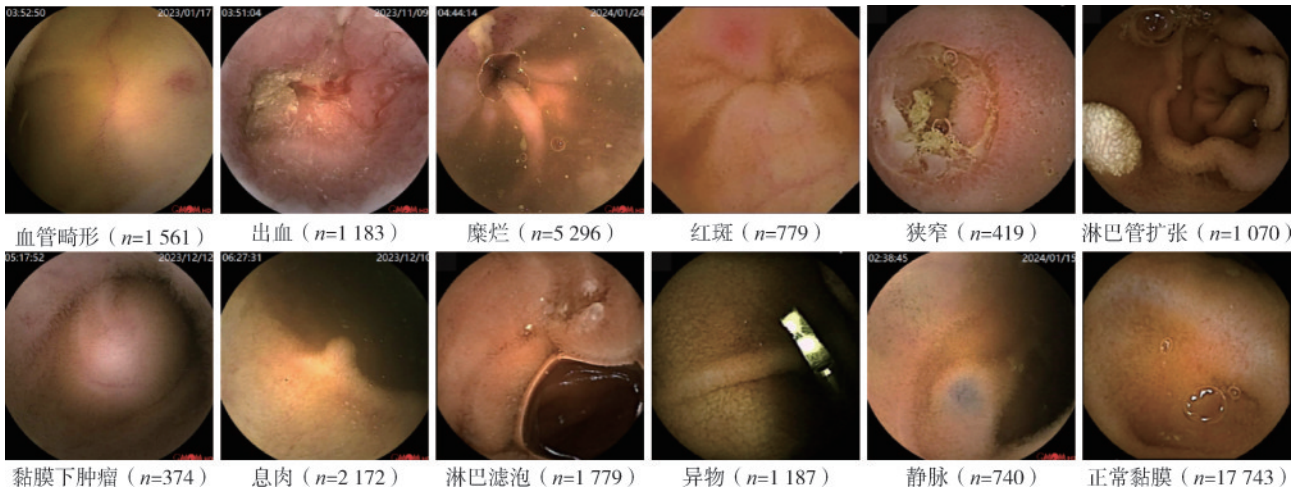


图1　代表性注释图像示例

医院伦理委员会审批通过（批准号：2024 L-003 号），由于研究的非干预性回顾性设计，知情同意被豁免。

二、深度学习网络

1. 图像预处理：为确保模型具备较强的泛化能力，研究中对图像数据进行了综合预处理和增强操作。特别地，我们采用了在线数据增强的方法[8]，增强操作是在训练过程中实时进行的，不会产生新的图像文件，确保每次训练时模型都能看到略有不同的图像版本。训练集：首先执行随机图像尺寸调整及裁剪至 224 像素×224 像素。为了增加数据集多样性，进一步进行随机水平翻转，随后图像从 PIL Image 或 numpy.ndarray 格式转换为 PyTorch Tensor，并归一化至[0，1]范围；红绿蓝（RGB）通道的标准化采用均值[0.485，0.456，0.406]及标准偏差[0.229，0.224，0.225]。验证集处理策略略有不同，首先将图像短边调整至 256 像素，然后中心裁剪出 224 像素×224 像素尺寸，后续转换和归一化步骤与训练集相同，亦采用相应的 RGB 通道标准化参数。预处理和增强步骤均通过 PyTorch 的 torchvision 库完成。

2. 模型训练配置：为实现图像分类任务，采用了基于卷积神经网络（convolutional neural network，CNN）的预训练模型进行迁移学习。选用的模型包括 DenseNetn-121、EfficientNet、ResNet101、ResNet50 和 VGG19。这些 CNN 模型均由卷积层、平均池化层以及带 ReLU 激活函数的全连接层组成。为了适应 12 分类的数据集，对每个预训练模型增加了 2 个带 ReLU 激活函数的致密层和 1 个用于分类的带 Softmax 激活函数的输出层。输出层的节点数设置为 12，匹配分类任务的需求。模型训练采用交叉熵损失函数和 Adam 优化器，设定 35 个训练周期。为防止过拟合，引入早停策略，即若验证集性能在连续 6 个周期内未见提升，则终止训练。此外，实行学习率调度，每 5 个周期减半 1 次学习率。所有操作在 PyTorch 框架下完成。相关的神经网络架构见图 2。

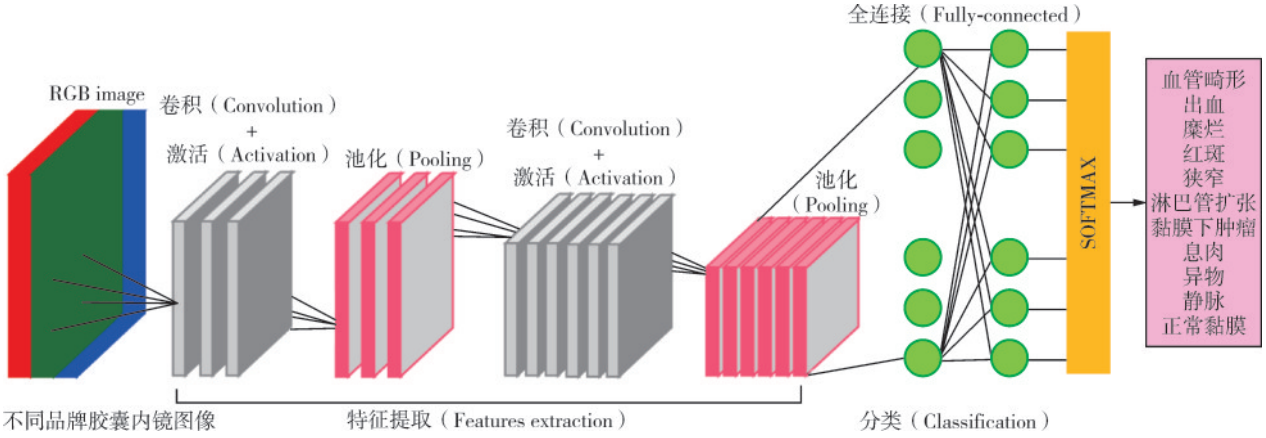


图2　相关神经网络架构示意图

www.zhxhnjzz.com

3. 模型可解释性分析：高计算成本、数据获取难题以及深度学习的不透明性限制了计算机视觉在医疗领域内的广泛应用。针对这些挑战，可解释性人工智能技术应运而生，旨在提高模型的透明度。本研究采用 Grad-CAM 和 SHAP 解释性技术分析 CNN 模型[9-10]，解决深度学习中的“黑盒”问题，使模型的决策过程更加明晰和可解释。Grad-CAM 通过生成热图揭示了模型决策过程中关键的图像区域；而 SHAP 则为图像分类任务中的每个像素分配了重要性分数，清晰地标识了其在模型决策中的作用。这些技术共同提升了对小肠病变自动分类过程的理解深度。

在本研究中，深度学习技术被应用于小肠胶囊内镜检查图像的自动化分类，覆盖了 12 种病变类型，包括血管畸形、出血、糜烂、红斑、狭窄、淋巴管扩张、黏膜下肿瘤、息肉、淋巴滤泡、异物、静脉以及正常黏膜。为深入了解模型在语义层面的分类能力，从图像分类模型中抽取了中间层输出作为语义特征。通过在目标层注册前向钩子，实现了这些特征的捕获。继而，应用 t-SNE 技术将高维特征降维到二维空间[11]，并利用 plotly 库对这些特征进行可视化分析。

4. 多设备终端模型部署：为实现胶囊内镜图像的自动化诊断，构建并在本地计算机及浏览器上部署了深度学习模型，该模型能实时、精准地监控胶囊内镜视频，满足最低帧率要求，确保全程检查覆盖。通过迁移学习，对基于 PyTorch 的模型进行了性能优化。进一步，模型转换为 ONNX 格式，以支持跨平台部署，确保在 Linux、Windows、MacOS 等操作系统上通过 ONNX Runtime 高效执行，并对不同硬件（CPU、GPU）进行了优化。ONNX 作为开放标准，保障了模型的互操作性并增加了部署灵活性，显著提高胶囊内镜视频中病变实时检测的准确性。研究流程见图 3。

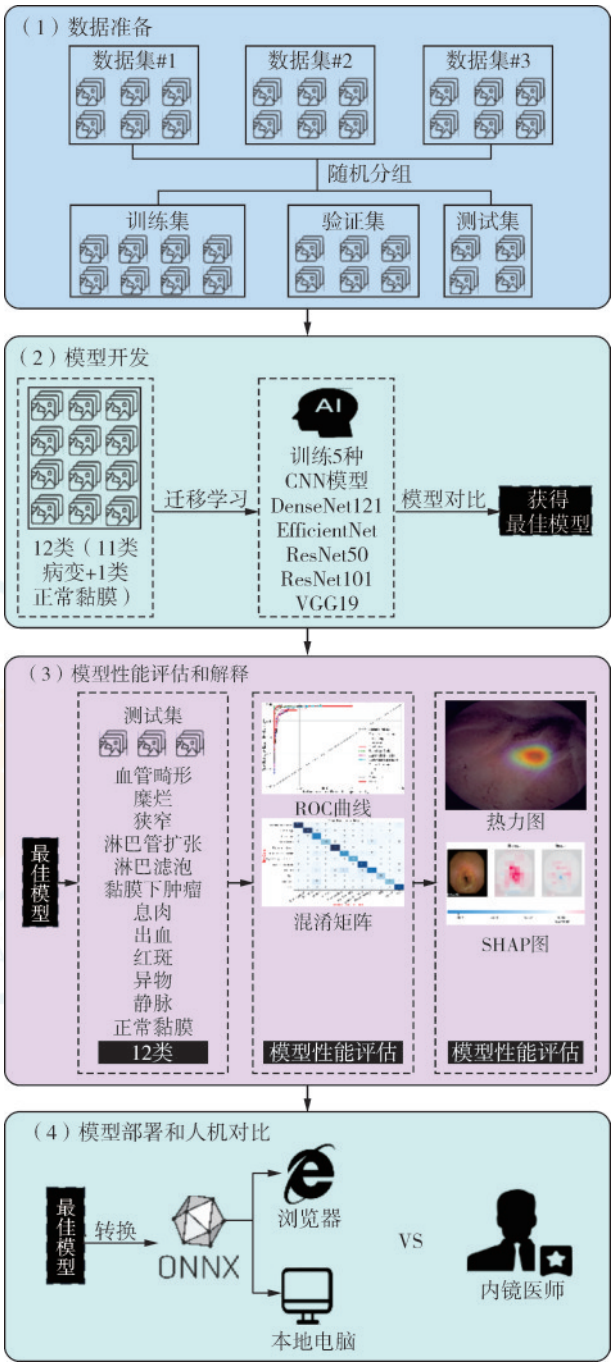


图 3　研究流程图

三、人机对比

在人机对比实验中，常熟市第一人民医院消化内镜中心邀请了 2 名高年资内镜医师（具有超过 5 年阅片经验）和 2 名低年资内镜医师（阅片经验不足 3 年）独立对测试集图像（n=1 013）进行了评估。之后，将这些内镜医师的评估结果与模型的阅片结果进行了对比分析。在分析中，对比了 5 种 CNN 模型和不同资历内镜医师诊断的准确性和速度。

在本研究中，采用了配备 RTX A4000 显卡（16 GB 显存）、5×E5-2680 v4 CPU 和 350 GB 硬盘空间的计算机。通过 TensorFlow（2.7.0）、Keras（2.7.0）和 OpenCV（4.5.4.60）等 Python 库，构建、训练深度学习模型并处理图像数据。数据的整理、分析及可视化通过 Pandas（1.3.4）、NumPy（1.21.4）、Matplotlib（3.5.0）和 Plotly（5.4.0）完成。模型优化采用 PyTorch（1.10.0+cu113），模型的保存和加载依赖于 H5py（3.6.0）。

本研究采用多元化的评估指标以全面评估胶囊内镜图像分类模型的性能。评估指标包括接收者操作特征曲线下面积（area under curve，AUC）、敏感度（Sensitivity）、特异度（Specificity）、精确度

www.zhxhnjzz.com

(Precision)、准确率(Accuracy)、F1 分数(F1 Score)、宏观平均(Macro Average)和加权平均(Weighted Average)。敏感度或真阳性率(true positive rate, TPR)=TP/(TP+FN),特异度或真阴性率(true negative rate, TNR)=TN/(TN+FP),精确度或阳性预测值(positive predictive value, PPV)=TP/(TP+FP),准确率=(TP+TN)/(TP+TN+FP+FN),F1 分数=2(精确度×敏感度)/(精确度+敏感度);宏观平均=$\frac{1}{k}\sum_{i=1}^{k}P_i$,加权平均=$\sum_{i=1}^{k}w_i \cdot P_i$。其中,TP(True Positives)表示正确预测为阳性的样本数量,TN(True Negatives)表示正确预测为阴性的样本数量,FP(False Positives)表示错误预测为阳性的样本数量,FN(False Negatives)表示错误预测为阴性的样本数量。P_i是第i个类别的性能指标值,w_i是第i个类别的权重,通常定义为$w_i=\frac{n_i}{N}$,n_i是第i个类别的样本量,而 N 是总样本量。

结　果

一、神经网络构建情况

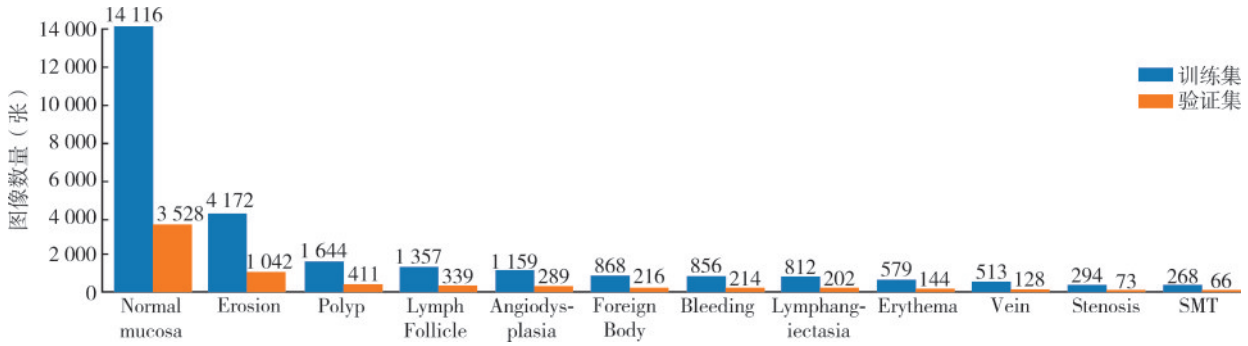
模型开发使用了 33 290 张胶囊内镜图像(包括训练集和验证集),覆盖了 12 类小肠病变,这些胶囊内镜图像的小肠病变种类及数量详见图 4。在模型的训练初期,损失函数的快速下降揭示了模型对数据特征的迅速捕捉。随着训练周期的增长,损失函数的下降速率减缓并最终稳定,表明了模型在学习过程中达到饱和状态。性能指标在初始提升后维持平稳,未出现性能衰退或显著波动,这暗示模型避免了过度拟合数据的风险。相关的训练动态见图 5。

二、不同模型的胶囊内镜图像诊断性能比较

表 1 展示了在含有 6 652 张胶囊内镜图像的验证集上,5 种经过迁移学习训练的深度学习模型在 12 类分类任务上的对比情况。EfficientNet 模型因在验证集上获得最高的准确率以及在其他重要性能指标上的优异表现(在精确度指标上排名第二)而被选为最优模型,我们将其命名为 EfficientNet-CE。

表 1　不同模型在验证集(n=6 652)上的性能比较

模型名称	准确率	精确度	敏感度	F1 分数
DenseNet121	84.92%	73.36%	73.21%	72.65%
EfficientNet	91.55%	84.14%	83.14%	83.57%
ResNet50	87.13%	76.72%	72.61%	74.05%
ResNet101	87.37%	76.87%	73.20%	74.26%
VGG19	91.24%	84.43%	82.04%	83.10%



注:Normal mucosa(正常黏膜),Erosion(糜烂),Polyp(息肉),Lymph Follicle(淋巴滤泡),Angiodysplasia(血管畸形),Foreign Body(异物),Bleeding(出血),Lymphangiectasia(淋巴管扩张),Erythema(红斑),Vein(静脉),Stenosis(狭窄),SMT(黏膜下肿瘤)

图 4　模型开发使用的 12 类小肠病变的胶囊内镜图像分布情况

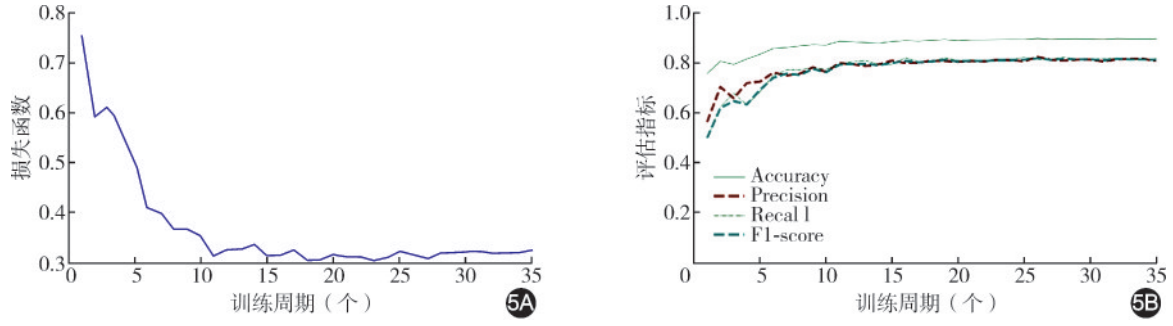


图 5　训练过程中损失函数和性能指标变化情况　5A:训练周期增加时的损失函数变化;5B:训练周期增加时的评估指标变化

www.zhxhnjzz.com

三、最佳模型在测试集上的性能评估

表2展示了在含有1 013张胶囊内镜图像的测试集上，最优模型EfficientNet-CE的性能评估结果。该表格为12个类别的图像在模型中的识别性能提供了全面的指标分析，涵盖了精确度、敏感度（也称召回率）、特异度、F1分数、准确率、平均精度（AP）、接收者操作特征曲线下面积（AUC）、马修斯相关系数（MCC）以及科恩卡帕系数（Cohen′s Kappa）。此外，表中还提供了宏观平均和加权平均两种汇总统计指标，为模型整体性能提供了量化的视角。

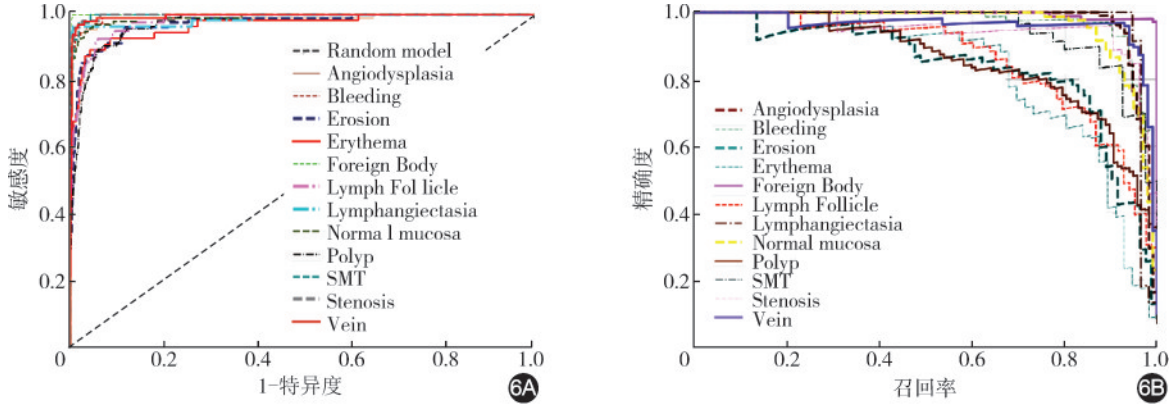
图6展示了EfficientNet-CE模型在测试集上对不同类别的预测性能的2个关键评估曲线：接收者操作特性（ROC）曲线和精确度-召回率（PR）曲线。在ROC曲线图中，如“淋巴管扩张”和“黏膜下肿瘤”的曲线紧邻图表的左上角，表明模型在这些类别上有出色的表现；同样的，在PR曲线图中，趋向于右上角的曲线表示在这些类别上模型同样表现良好。相反，在ROC曲线中远离左上角的曲线，例如“息肉”和“红斑”，以及在PR曲线中远离右上角的曲线，都表明模型在这些类别上的表现相对较差。

EfficientNet-CE模型的分类效果通过混淆矩阵进行了分析，以验证其在各个类别上的分类准确性和鲁棒性，结果见图7A。结果表明，深度学习模型在大多数情况下表现优异。然而，存在一些分类错误，如图7B和7C中所示的错误诊断案例。这些分类错误可能由图像类别间的特征重叠、意外的光反射、过远的拍摄距离以及图像模糊等因素引起。

表2　EfficientNet-CE模型在测试集上的各项性能评估指标

图像类别	精确度	敏感度	特异度	F1分数	准确率	AP	AUC	MCC	Cohen′s Kappa
血管畸形（n=113）	0.962 6	0.911 5	0.995 6	0.936 4	0.911 5	0.971 6	0.989 1	0.929 1	0.928 6
出血（n=113）	0.952 8	0.893 8	0.994 4	0.922 4	0.893 8	0.971 2	0.994 4	0.913 6	0.913 0
糜烂（n=82）	0.642 9	0.878 0	0.957 0	0.742 3	0.878 0	0.837 1	0.973 1	0.726 3	0.715 7
红斑（n=56）	0.883 7	0.678 6	0.994 8	0.767 7	0.678 6	0.828 4	0.967 9	0.763 3	0.756 0
异物（n=103）	0.919 6	1.000 0	0.990 1	0.958 1	1.000 0	0.995 7	0.999 6	0.954 2	0.953 2
淋巴滤泡（n=83）	0.764 7	0.783 1	0.978 5	0.773 8	0.783 1	0.859 9	0.980 8	0.753 4	0.753 4
淋巴管扩张（n=56）	0.871 0	0.964 3	0.991 6	0.915 3	0.964 3	0.968 4	0.988 4	0.911 4	0.910 0
正常黏膜（n=99）	0.817 4	0.949 5	0.977 0	0.878 5	0.949 5	0.961 5	0.991 4	0.867 3	0.864 2
息肉（n=117）	0.840 4	0.675 2	0.983 3	0.748 8	0.675 2	0.854 7	0.974 6	0.725 4	0.720 0
黏膜下肿瘤（n=40）	0.911 8	0.775 0	0.996 9	0.837 8	0.775 0	0.950 0	0.997 2	0.834 7	0.831 7
狭窄（n=52）	0.953 5	0.788 5	0.997 9	0.863 2	0.788 5	0.947 3	0.996 9	0.860 7	0.856 5
静脉（n=99）	0.930 0	0.939 4	0.992 3	0.934 7	0.939 4	0.961 2	0.995 2	0.927 6	0.927 6
宏观平均[a]	0.870 9	0.853 1	0.987 5	0.856 6	0.853 1	0.925 6	0.987 4	0.847 2	0.844 2
加权平均[b]	0.871 7	0.862 8	0.986 7	0.862 4	0.862 8	0.928 3	0.987 4	0.852 0	0.849 2

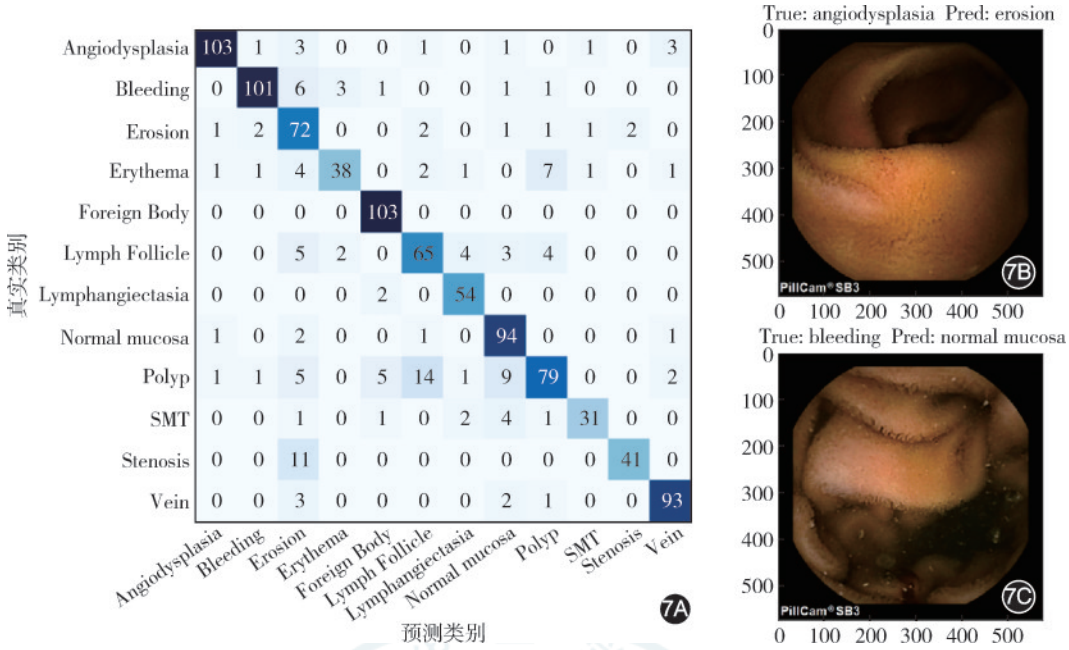
注：AP（平均精度），AUC（接收者操作特征曲线下面积），MCC（马修斯相关系数），Cohen’s Kappa（科恩卡帕系数）；[a]宏观平均指标平等地考虑了所有类别，确保即使是样本量较小的类别也能对模型性能评估有相等的贡献；[b]加权平均指标则顾及到了每个类别的样本量，对于样本量大的类别赋予了更高的权重



注：Angiodysplasia（血管畸形），Bleeding（出血），Erosion（糜烂），Erythema（红斑），Foreign Body（异物），Lymph Follicle（淋巴滤泡），Lymphangiectasia（淋巴管扩张），Normal mucosa（正常黏膜），Polyp（息肉），SMT（黏膜下肿瘤），Stenosis（狭窄），Vein（静脉）

图6　EfficientNet-CE模型在外部测试集（n=1 013）上的预测性能　6A：接收者操作特性（ROC）曲线；6B：精确度-召回率（PR）曲线

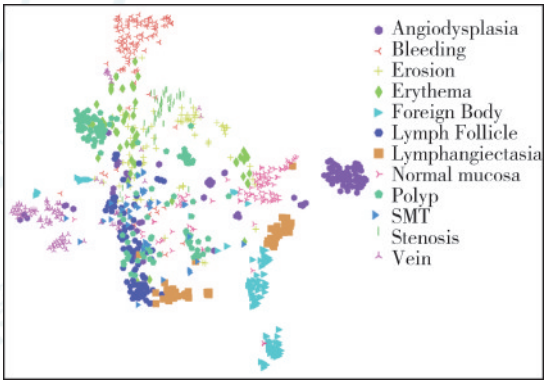
www.zhxhnjzz.com



注：Angiodysplasia（血管畸形），Bleeding（出血），Erosion（糜烂），Erythema（红斑），Foreign Body（异物），Lymph Follicle（淋巴滤泡），Lymphangiectasia（淋巴管扩张），Normal mucosa（正常黏膜），Polyp（息肉），SMT（黏膜下肿瘤），Stenosis（狭窄），Vein（静脉）

图 7　EfficientNet-CE 模型在测试集（n=1 013）上的表现　7A：混淆矩阵，展示模型的分类准确性；7B：误判的图像示例，模型将真实标签为血管畸形的图像误判为糜烂；7C：误判的图像示例，真实标签为出血，误判为正常黏膜

为分析模型在胶囊内镜图像分类中误判的原因，采用 t-分布随机邻域嵌入（t-SNE）技术，将高维数据有效映射至二维平面，使不同类别之间的分离程度能够直观展现，详见图 8。此种可视化方法有助于辨识哪些类别的图像容易区分，哪些类别可能需要复杂的特征提取技术或更细致的分析来提高分类的准确性。例如，图中“血管畸形”和“糜烂”之间存在少量重叠，这解释了模型误分类的原因。进一步地，我们采用 plotly 库构建了二维空间的交互式语义特征图。该方法允许用户直接点击图中的任意点，即可查看其对应的测试集图像及其在语义特征空间中的位置（链接：https://pan.baidu.com/s/1eOQ2DDvdoqda0BMVu0vaYA?pwd=8k8l；提取码：8k8l）。



注：Angiodysplasia（血管畸形），Bleeding（出血），Erosion（糜烂），Erythema（红斑），Foreign Body（异物），Lymph Follicle（淋巴滤泡），Lymphangiectasia（淋巴管扩张），Normal mucosa（正常黏膜），Polyp（息肉），SMT（黏膜下肿瘤），Stenosis（狭窄），Vein（静脉）

图 8　测试集胶囊内镜图像（n=1 013）的二维语义特征图　每种颜色和形状代表不同的胶囊内镜图像类别，具有相似特征的图像类别倾向于在图形空间中聚集在一起；另一方面，如果某些点群彼此之间的边界不清晰，这可能表明模型在这些区域的分类性能可能存在重叠，导致分类困难

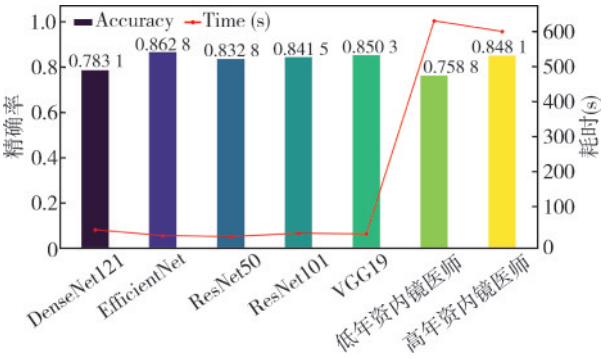
四、模型与内镜医师诊断性能比较

在包含 1 013 张胶囊内镜图像的测试集上对 5 种人工智能模型与不同资历的内镜医师进行了诊断性能对比，主要评估了诊断准确率和诊断耗时（单位：s），见图 9。EfficientNet-CE 模型在所有模型中表现最佳，诊断准确率达到 86.28%，超过了不同资历的内镜医师。在诊断速度方面，人工智能模型模型仅需 14.9 s 即可完成 1 013 张胶囊内镜图像的诊断（相当于 52.43 帧/s），远超低年资和高年资内镜医师的诊断速度，分别为低年资内镜医师的约 42.4 倍和高年资内镜医师的约 40.3 倍。

五、模型可解释性分析

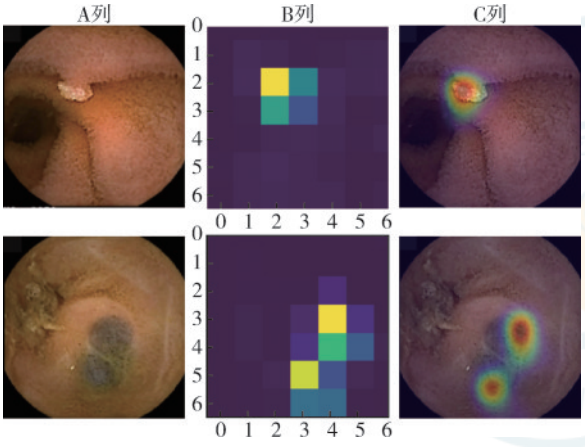
图 10 展示了应用 Grad-CAM 技术于人工智能模型决策过程的可视化。A 列呈现了原始内镜图像；B 列展示了基于 EfficientNet-CE 模型的特征提取，生成的像素激活热图，这些热图突出了模型决策过程中的关键区域；C 列则是激活热图与原始图像的叠加视图，其中的暖色调区域，如红色和黄色，指示了 EfficientNet-CE 模型用以识别的关键病变区域。

www.zhxhnjzz.com



注：条图为诊断准确率比较，线图为诊断用时比较；左侧纵轴代表诊断准确率，右侧纵轴代表诊断用时；低年资内镜医师阅片经验不足3年，高年资内镜医师具有超过5年阅片经验

图9　5种人工智能模型与不同资历内镜医师在测试集（n=1 013）上的诊断准确率及耗时比较



注：A列为原始的内镜图像；B列为使用Grad-CAM的像素激活热图；C列为原始图像与激活热图的组合

图10　人工智能模型决策过程的Grad-CAM可视化

图11通过应用SHAP分析法展示了模型预测逻辑的内部机制。在2个子图中，EfficientNet-CE模型的预测分别对应于出血和狭窄2种真实分类。图中的像素颜色深浅指示了对模型预测的贡献程度：红色代表对预测结果有正向的贡献，而蓝色代表有负向的贡献。在子图A中，与糜烂、红斑和血管畸形相比，红色区域更加明显，使得模型能够准确将其分类为出血；同样，子图B中的特征使得模型正确诊断为狭窄。

六、模型对视频实时预测

为了方便在不同的推理引擎上部署，本研究采用的PyTorch深度学习模型通过迁移学习后，已转换为开放式神经网络交换格式（ONNX）。这一标准化格式的使用，便于人工智能研究者跨平台共享和部署模型。EfficientNet-CE模型的这一转换，使其可以在多种环境中部署，包括本地电脑和Web前端，实现在多种终端设备上对胶囊内镜图像进行实时病变诊断。

图12展示了基于OpenCV实现的ONNX模型在胶囊内镜视频帧上的实时预测能力。A和B部分展现了模型对单帧图像中“血管畸形”和“淋巴管扩张”2种病变的预测结果，左图为原始图像，左下角红色字体标注了模型预测的前两个最可能的类别及其置信度，右图为不同类别置信度的柱状图表示。C部分以二维码形式提供了3个不同病变类别的胶囊内镜视频实时预测链接，视频左下角实时显示模型的预测类别和置信度，用户可扫描二维码以观看视频中的预测效果。

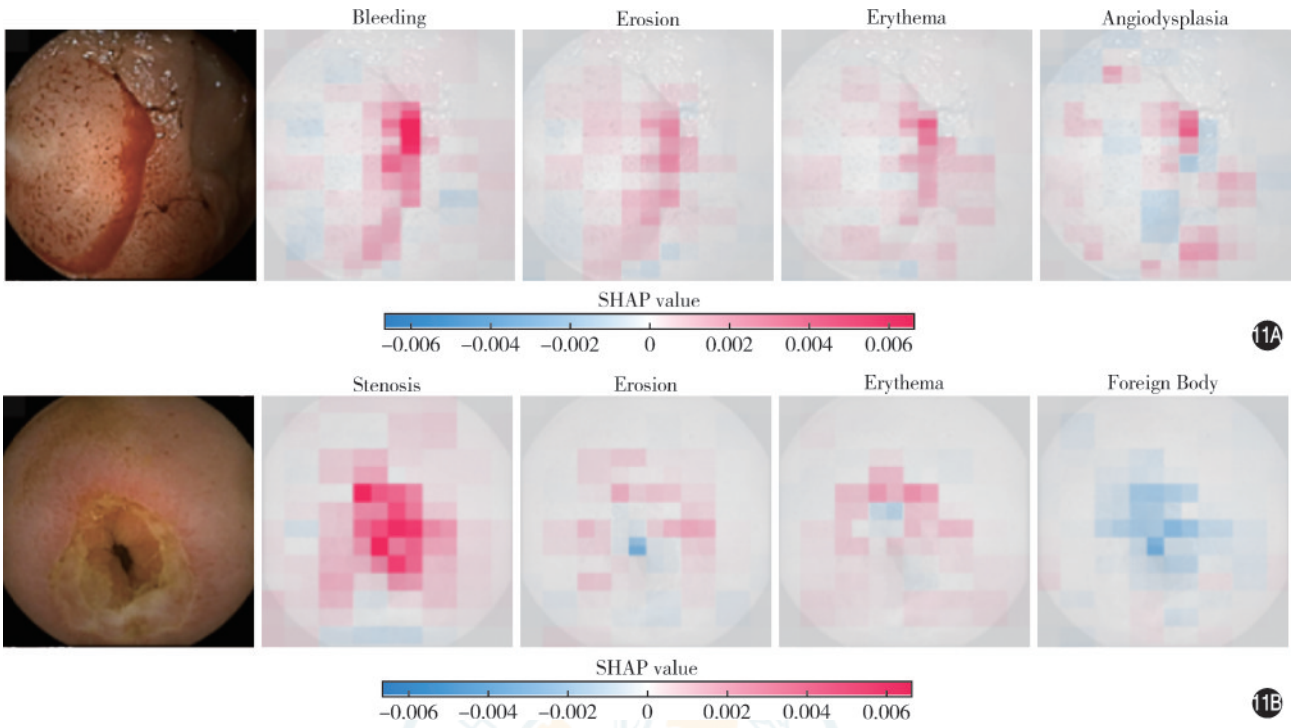
讨　论

本研究建立了包含糜烂、出血、淋巴管扩张等12种病变的胶囊内镜图像数据集，涉及3种不同品牌的胶囊内镜设备。基于该数据集，采用迁移学习方法开发了5种CNN架构的人工智能模型，旨在自动化诊断多类别的小肠病变。EfficientNet-CE模型在验证和测试过程中显示了最佳的性能，并在人机对比实验中展现了其泛化能力和可靠性。该模型转换为ONNX格式后，实现了本地设备上对胶囊内镜视频的实时预测功能。本研究覆盖了从数据集的构建、模型的开发与测试、可解释性分析，以及终端部署的完整流程。

先前开发的胶囊内镜人工智能系统研究多基于单一品牌的胶囊内镜图像数据集来训练模型。具体而言，Yokote等[6]和de Maissin等[12]的研究利用了由美国PillCam SB3设备拍摄的图像，Smedsrud等[7]的研究使用了由日本EndoCapsule设备获取的图像，Xie等[13]的研究则基于中国OMOM品牌的图像数据。这种做法可能导致模型对于特定品牌的图像采集方式、图像质量和光学特性存在偏差，因为不同品牌的胶囊内镜在图像分辨率、对比度和色彩饱和度等方面可能存在差异。然而，Urban等[14]的研究表明，在混合使用不同图像增强技术的数据集上训练的模型能够优于只采用单一技术的数据集，因为混合数据集能够提供更丰富的训练数据和知识的协同效应。借鉴此发现，EfficientNet-CE模型通过结合3种不同品牌胶囊内镜设备拍摄的图像进行训练，有效地利用了这一效应，从而训练出的模型不仅展现了更高的品牌通用性，还拥有优秀的性能表现。

早期胶囊内镜领域的人工智能研究主要聚焦

www.zhxhnjzz.com



注：Stenosis（狭窄），Erosion（糜烂），Erythema（红斑），Foreign Body（异物）

图 11　使用 SHAP 技术的可解释性分析　11A：EfficientNet-CE 模型使用 SHAP 值正确预测“出血”标签的胶囊内镜图像；11B：模型使用 SHAP 值正确预测“狭窄”标签的胶囊内镜图像

于识别单一病变类型，如出血[15-16]、隆起性病变[17]。这些模型由于仅关注一种病变，结构较为简单，易于构建。然而，开发可识别多类别病变的模型，不仅能够显著增强模型的通用性，而且还提高了对小肠病变诊断的全面性。本研究开发的 EfficientNet-CE 模型能够识别包括血管畸形、出血、糜烂和息肉在内的 12 类常见小肠病变，在所有类别的综合评估中表现出色，其加权平均 AUC 和准确率分别达到了 98.74% 和 86.28%，能够辅助阅片医师高效筛查和诊断小肠病变，提升医疗服务效率。

在过去 10 年里，多种基于计算机的人工智能辅助系统已经被开发出来，专门用于分析胶囊内镜图像。一项系统综述及荟萃分析验证了可疑出血指示软件在胶囊内镜应用中的有效性，揭示了其在识别出血或潜在出血病变方面的总体敏感度为 55.3%，特异度为 57.8%[18]。此外，一项采用 Intromedic 胶囊系统的“快速查看”阅读模式的多中心前瞻性研究表明，在维持 82.2% 敏感度的同时，该模式显著提高了阅读效率，将平均阅读时间从 39.7 min 减少至 19.7 min[19]。相比之下，本研究开发的 EfficientNet-CE 模型在测试集上展现了 86.28% 的加权平均敏感度和 98.67% 的特异度，特别是在出血类别上达到了 89.38% 的敏感度和 99.44% 的特异度，能够在仅需 14.9 s 的时间内完成 1 013 张胶囊内镜图像的诊断，其速度约是低年资内镜医师的 42.4 倍。

在 EfficientNet-CE 模型识别的 12 个病变类别中，“息肉”和“红斑”类别的检测具有挑战性，表现为这两个类别的敏感度和精确度指标相比其他类别较低。对于“息肉”，区分肠道蠕动所引起的皱襞隆起和真正的息肉是一大难题；而对于“红斑”，光照条件和胶囊的运动可能导致红斑区域的图像信息模糊，从而增加了预测的难度。我们使用 t-SNE 技术将高纬数据降维可视化，更加直观地洞察模型误分类的原因，并据此针对性地进行数据采集、改进数据标注方法和调整模型训练策略，以期优化和提升模型的性能。

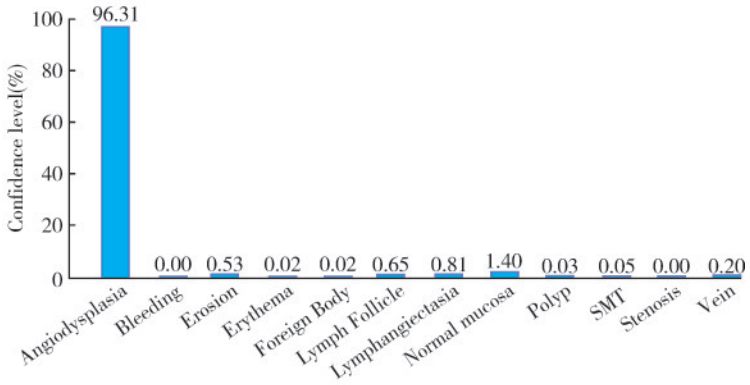
综上，本研究使用包含 3 个不同品牌胶囊内镜图像的数据集，通过训练、验证、测试到部署的全流程，构建了一个能够自动诊断 12 类小肠病变的胶囊内镜人工智能辅助系统。通过与不同资历内镜医师在诊断准确率和速度方面进行比较，展现出了该模型很大的临床应用潜力。

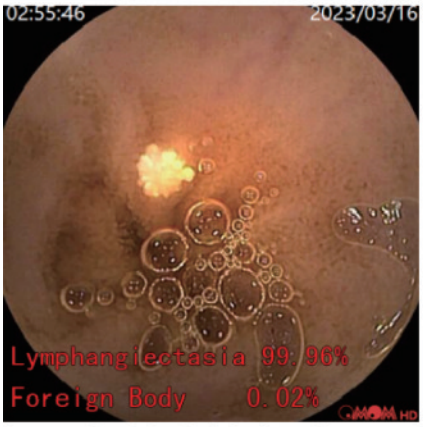
利益冲突　所有作者声明不存在利益冲突

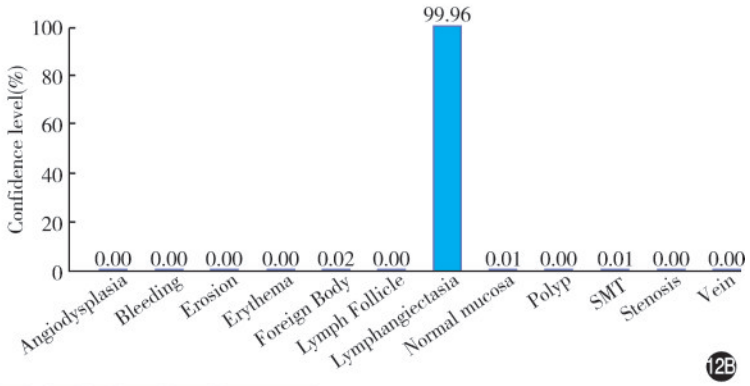
作者贡献声明　陈健：实施研究、分析数据、撰写论文；孙斌：采集数据；王甘红：采集数据、整理数据；夏开建：代码报错解决、模型可视化解释；徐晓丹：研究指导、论文修改、经费支持

www.zhxhnjzz.com









视频1　　视频2　　视频3　　12C

注：Angiodysplasia（血管畸形），Bleeding（出血），Erosion（糜烂），Erythema（红斑），Foreign Body（异物），Lymph Follicle（淋巴滤泡），Lymphangiectasia（淋巴管扩张），Normal mucosa（正常黏膜），Polyp（息肉），SMT（黏膜下肿瘤），Stenosis（狭窄），Vein（静脉）

图 12　EfficientNet-CE 模型部署后对单帧图像和视频的预测结果　12A、12B：模型对单帧胶囊内镜图像的预测，左侧图为原始图像，其中左下角的红色字体呈现了模型对前两名分类的预测及对应置信度，右侧图为各分类的置信度柱状表示；12C：以二维码形式提供了 3 种不同病变类别的胶囊内镜视频实时预测链接

参　考　文　献

[1] ASGE Technology Committee; Wang A, Banerjee S, et al. Wireless capsule endoscopy[J]. Gastrointest Endosc, 2013, 78(6):805-815. DOI: 10.1016/j.gie.2013.06.026.

[2] Hosoe N, Takabayashi K, Ogata H, et al. Capsule endoscopy for small-intestinal disorders: current status[J]. Dig Endosc, 2019, 31(5):498-507. DOI: 10.1111/den.13346.

[3] Takada K, Yabuuchi Y, Kakushima N. Evaluation of current status and near future perspectives of capsule endoscopy: summary of Japan Digestive Disease Week 2019[J]. Dig Endosc, 2020, 32(4):529-531. DOI: 10.1111/den.13659.

[4] Sinonquel P, Eelbode T, Bossuyt P, et al. Artificial intelligence and its impact on quality improvement in upper and lower gastrointestinal endoscopy[J]. Dig Endosc, 2021, 33(2):242-253. DOI: 10.1111/den.13888.

[5] Yu H, Singh R, Shin SH, et al. Artificial intelligence in upper GI endoscopy - current status, challenges and future promise [J]. J Gastroenterol Hepatol, 2021, 36(1):20-24. DOI: 10.1111/jgh.15354.

www.zhxhnjzz.com

[6] Yokote A, Umeno J, Kawasaki K, et al. Small bowel capsule endoscopy examination and open access database with artificial intelligence: the SEE-artificial intelligence project[J]. DEN Open, 2024, 4(1):e258. DOI: 10.1002/deo2.258.
[7] Smedsrud PH, Thambawita V, Hicks SA, et al. Kvasir-capsule, a video capsule endoscopy dataset[J]. Sci Data, 2021, 8(1):142. DOI: 10.1038/s41597-021-00920-z.
[8] Athalye C, Arnaout R. Domain-guided data augmentation for deep learning on medical imaging[J]. PLoS One, 2023, 18(3): e0282532. DOI: 10.1371/journal.pone.0282532.
[9] Zhang Y, Hong D, McClement D, et al. Grad-CAM helps interpret the deep learning models trained to classify multiple sclerosis types using clinical brain magnetic resonance imaging[J]. J Neurosci Methods, 2021, 353: 109098. DOI: 10.1016/j.jneumeth.2021.109098.
[10] Kikutsuji T, Mori Y, Okazaki K, et al. Explaining reaction coordinates of alanine dipeptide isomerization obtained from deep neural networks using Explainable Artificial Intelligence (XAI)[J]. J Chem Phys, 2022, 156(15):154108. DOI: 10.1063/5.0087310.
[11] Linderman GC, Steinerberger S. Clustering with t-SNE, provably[J]. SIAM J Math Data Sci, 2019, 1(2):313-332. DOI: 10.1137/18m1216134.
[12] de Maissin A, Vallée R, Flamant M, et al. Multi-expert annotation of Crohn′s disease images of the small bowel for automatic detection using a convolutional recurrent attention neural network[J]. Endosc Int Open, 2021, 9(7):E1136-1144. DOI: 10.1055/a-1468-3964.
[13] Xie X, Xiao Y, Zhao X, et al. Development and validation of an artificial intelligence model for small bowel capsule endoscopy video review[J]. JAMA Netw Open, 2022, 5(7): e2221992. DOI: 10.1001/jamanetworkopen.2022.21992.
[14] Urban G, Tripathi P, Alkayali T, et al. Deep learning localizes and identifies polyps in real time with 96% accuracy in screening colonoscopy[J]. Gastroenterology, 2018, 155(4): 1069-1078.e8. DOI: 10.1053/j.gastro.2018.06.037.
[15] Musha A, Hasnat R, Mamun AA, et al. Computer-aided bleeding detection algorithms for capsule endoscopy: a systematic review[J]. Sensors (Basel), 2023, 23(16). DOI: 10.3390/s23167170. [published online ahead of print].
[16] 施慧英, 樊梦科, 王玮珺, 等. 基于深度卷积神经网络的小肠胶囊内镜智能辅助系统识别小肠出血[J]. 中华消化杂志, 2020, 40(11): 763-767. DOI: 10.3760/cma. j. cn311367-20200618-00392.
[17] Kim HJ, Gong EJ, Bang CS, et al. Computer-aided diagnosis of gastrointestinal protruded lesions using wireless capsule endoscopy: a systematic review and diagnostic test accuracy meta-analysis[J]. J Pers Med, 2022, 12(4): 644. DOI: 10.3390/jpm12040644.
[18] Yung DE, Sykes C, Koulaouzidis A. The validity of suspected blood indicator software in capsule endoscopy: a systematic review and meta-analysis[J]. Expert Rev Gastroenterol Hepatol, 2017, 11(1): 43-51. DOI: 10.1080/17474124.2017.1257384.
[19] Saurin J, Jacob P, Heyries L, et al. Multicenter prospective evaluation of the express view reading mode for small-bowel capsule endoscopy studies[J]. Endosc Int Open, 2018, 6(5): E616-621. DOI: 10.1055/a-0587-4788.

·读者·作者·编者·

《中华消化内镜杂志》2025年可直接使用英文缩写的常用词汇

ERCP(内镜逆行胰胆管造影术)
EST(经内镜乳头括约肌切开术)
EUS(超声内镜检查术)
EUS-FNA(超声内镜引导细针穿刺抽吸术)
EMR(内镜黏膜切除术)
ESD(内镜黏膜下剥离术)
ENBD(经内镜鼻胆管引流术)
ERBD(经内镜胆道内支架放置术)
APC(氩离子凝固术)
EVL(内镜下静脉曲张套扎术)
EIS(内镜下硬化剂注射术)
POEM(经口内镜食管下括约肌切开术)
NOTES(经自然腔道内镜手术)
MRCP(磁共振胰胆管成像)
GERD(胃食管反流病)
RE(反流性食管炎)
IBD(炎症性肠病)
UC(溃疡性结肠炎)
NSAIDs(非甾体抗炎药)
PPI(质子泵抑制剂)
HBV(乙型肝炎病毒)
HBsAg(乙型肝炎病毒表面抗原)
Hb(血红蛋白)
PaO_2(动脉血氧分压)
$PaCO_2$(动脉血二氧化碳分压)
ALT(丙氨酸转氨酶)
AST(天冬氨酸转氨酶)
AKP(碱性磷酸酶)
IL(白细胞介素)
TNF(肿瘤坏死因子)
VEGF(血管内皮生长因子)
ELISA(酶联免疫吸附测定)
RT-PCR(逆转录-聚合酶链反应)

(本刊编辑部)

www.zhxhnjzz.com

CMS 康哲药业 CHINA MEDICAL SYSTEM | 莱芙兰® LUMEBLUE®

本广告仅供医学药学专业人士阅读
粤药广审(文)第290610-02733号
批准文号:国药准字HJ20240037

广告

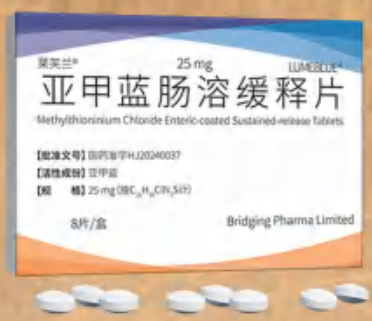
隐匿病变 及早发现



结肠镜提高病变检出率的便捷方案

适应症:
本品作为一种诊断剂，适用于在接受结肠镜检查筛查或监测的成人患者中增强结直肠病变的可视化

禁忌、不良反应，注意事项详见说明书



【生产企业】Cosmo S.p.A.
【中国唯一官方商业化合作伙伴】深圳市康哲药业有限公司

MB MMX DIFFICULT TO DETECT | Methylthioninium Chloride Enteric-coated Sustained-release Tablets

地址:深圳市南山区南头街道大新路198号创新大厦B座6-8楼 电话:0755-82416868 传真:0755-82416622 邮编:518052

Vedkang 唯德康

电子上消化道内窥镜
电子下消化道内窥镜
电子内窥镜图像处理器



Vedkang 唯德康

联系电话：0519-85071234
0519-69877759

产品注册证及名称：
粤械注准20252060765-电子上消化道内窥镜（UG-500、UG-500T、UG-700、UG-700T）
粤械注准20252060766-电子下消化道内窥镜（500系列、700系列）
粤械注准20252060897-电子内窥镜图像处理器（U7系列、U5系列）
粤广审准许字【2025】第019083号

禁忌内容或者注意事项详见说明书
以上仅指本公司产品

广告