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10月 第42卷 第10期

CHINESE JOURNAL OF DIGESTIVE ENDOSCOPY

Volume 42 Number 10
October 2025

ISSN 1007-5232
9 771007 523250 10>





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10期　2025年10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

共识与指南

胃静脉曲张超声内镜引导弹簧圈栓塞术专家共识(2025，杭州)…… 757

中华医学会消化内镜学分会超声内镜学组

中华医学会消化内镜学分会肝病协作组

胰岛素瘤内镜诊疗专家共识(2025年) …… 768

中华医学会消化内镜学分会

论　　著

超声内镜诊断结直肠黏膜下肿瘤准确性的影响因素分析 …… 780

崔晓丙　袁葵　凌琳　徐春玲　郭培　杨根华　包崇举　胡威

龚伟

基于深度学习的门静脉高压性胃病识别与分级系统的开发与验证 …… 789

古皓文　杨杰　肖勇　万新月　胡伟　谢先木　黄定鹏　姚承明

施信良　刘仕倩　黄丽　张弛　郑碧清　陈明锴

胆总管结石经十二指肠乳头自行排石的预测因素分析 …… 796

袁广众　汪翰英　毛丽娟　孙仁虎　吴大鹏　张其德　凌亭生

金海林

聚乙二醇电解质散与硫酸盐溶液在肠道准备及肠镜检查质量中的效果对比 …… 803

付越　田原　年碧霄　贺琰　郭新月　戎龙

异物钳辅助内镜黏膜切除术治疗长径≤10 mm直肠神经内分泌瘤的疗效分析(含视频) …… 809

张洪实　林燕娟　王木清　黄耀奎

离体纽扣电池致食管腐蚀伤动物模型构建及植物油的干预影响 … 817

李昭飞　赵德安　李明阳　王凌超　马唯唯

短篇论著

不含染色剂的黏膜下注射液在胃内镜黏膜下剥离术中的应用效果及安全性 …… 823
陆婉　丁永刚　张婷　毛丽娟　陈晶　周玉宏　肖君　李文杰　王耀辉　张其德
可分离式系线磁控胶囊内镜在高原地区食管疾病中的临床应用研究 …… 828
马雪琴　蔡甲慧　杨爱荣　马颖才

病例报道

内镜诊治直肠浅表锯齿状腺瘤合并传统锯齿状腺瘤1例(含视频) …… 831
侯雯雯　姚静静　邱月虹　刘洋　邵青　王晓　付金栋
转化治疗联合内镜黏膜下剥离术治疗颈段进展期食管癌1例 …… 833
宋宝回　钱杨　王燕　贺东黎　王健　钟芸诗　蔡明琰

综　　述

胆囊结石合并胆总管结石微创治疗应用进展 …… 836
陈璇桦　张筱凤

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 …… 827
插页目次 …… 802

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　钱程

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·

基于深度学习的门静脉高压性胃病识别与分级系统的开发与验证

古皓文[1]　杨杰[2]　肖勇[1]　万新月[1]　胡伟[3]　谢先木[4]　黄定鹏[5]　姚承明[6]　施信良[7]　刘仕倩[1]　黄丽[1]　张弛[1]　郑碧清[2]　陈明锴[1]

[1]武汉大学人民医院消化内科 消化疾病湖北省重点实验室，武汉 430060；[2]武汉楚精灵医疗科技有限公司，武汉 430000；[3]武汉市中西医结合医院消化内科，武汉 430022；[4]荆州市第二人民医院消化内科，荆州 434000；[5]中国人民解放军南部战区总医院消化内科，广州 510405；[6]湖南中医药高等专科学校附属第一医院消化内科，株洲 421000；[7]竹山县人民医院消化内科，十堰 442200

通信作者：陈明锴，Email：kaimingchen@163.com

【摘要】　目的　开发一款基于深度学习的门静脉高压性胃病（portal hypertensive gastropathy，PHG）实时辅助识别与分级系统并评估其辅助初级内镜医师判断内镜下 PHG 特征的效果。**方法**　回顾性收集 2015 年 1 月至 2023 年 10 月武汉大学人民医院、武汉市中西医结合医院、荆州市第二人民医院的消化内镜中心数据库中 832 例肝硬化患者的 2 848 张胃镜图片。本系统参考 Baveno Ⅱ 评分系统的 3 个内镜特征：胃窦毛细血管扩张症（gastric antral vascular ectasia，GAVE）、马赛克图案（mosaic-like pattern，MLP）及红色征（red marks，RM），分别开发了 3 个模型，具体分类指代如下：（1）GAVE 模型：0 无、1 有；（2）MLP 模型：0 无、1 轻度、2 重度；（3）RM 模型：0 无、1 孤立、2 融合。以 3 名内镜专家对内镜下 PHG 特征的分类结果作为金标准。使用 yolov8-m 模型训练，以 8∶1∶1 分配训练集、验证集和测试集。测试集用于测试模型性能和模型辅助初级内镜医师诊断 PHG 内镜特征的效果，计算准确率、召回率、精确率、特异度和 *Kappa* 系数。**结果**　GAVE 模型的准确率、召回率、特异度为 96.0%（48/50）、87.5%（7/8）、97.6%（41/42），其准确率与金标准相比，差异无统计学意义（χ^2=316.226，*P*=1.000）。GAVE1 与 GAVE0 的精确率分别为 87.5%（7/8）和 97.6%（41/42）。MLP 模型的准确率为 84.1%（132/157），与金标准相比，差异无统计学意义（χ^2=3.286，*P*=0.193）；MLP2 的精确率和召回率分别为 88.2%（15/17）和 75.0%（15/20），MLP1 的精确率和召回率为 77.9%（60/77）和 88.2%（60/68），MLP0 的精确率和召回率为 90.5%（57/63）和 82.6%（57/69）。RM 模型的准确率为 87.9%（123/140），与金标准相比，差异无统计学意义（χ^2=2.891，*P*=0.409）；RM2 的精确率和召回率为 94.7%（18/19）和 78.3%（18/23），RM1 的精确率和召回率为 72.2%（26/36）和 81.3%（26/32），RM0 的精确率和召回率为 92.9%（79/85）和 92.9%（79/85）。3 名初级医师的平均准确率，在 GAVE 模型、MLP 模型、RM 模型辅助前后，分别从 95.3% 升至 99.3%、从 83.9% 升至 91.9%、从 81.9% 升至 83.1%。3 名初级医师整体与金标准的一致性分析提示，GAVE 模型辅助前后的一致性均极强（整体 *Kappa* 均为 1.000）；MLP 模型由辅助前的中等（整体 *Kappa*=0.601）升至辅助后的极强（整体 *Kappa*=0.964）；RM 模型辅助前（整体 *Kappa*=0.792）后（整体 *Kappa*=0.798）均较强。**结论**　本系统能对胃镜下 PHG 的内镜特征准确识别并分级，可提升初级内镜医师诊断 PHG 的效能。

【关键词】　人工智能；　门静脉高压性胃病；　深度学习；　诊断；　计算机辅助

基金项目：国家重点研发计划（2023YFC2507400）

DOI：10.3760/cma.j.cn321463-20250227-00210

收稿日期　2025-02-27　　**本文编辑**　钱程

引用本文：古皓文，杨杰，肖勇，等. 基于深度学习的门静脉高压性胃病识别与分级系统的开发与验证[J]. 中华消化内镜杂志，2025，42(10)：789-795. DOI：10.3760/cma.j.cn321463-20250227-00210.

www.zhxhnjzz.com

Development and validation of a recognition and classification system for portal hypertensive gastropathy based on deep learning

Gu Haowen[1], Yang Jie[2], Xiao Yong[1], Wan Xinyue[1], Hu Wei[3], Xie Xianmu[4], Huang Dingpeng[5], Yao Chengming[6], Shi Xinliang[7], Liu Shiqian[1], Huang Li[1], Zhang Chi[1], Zheng Biqing[2], Chen Mingkai[1]

[1]Department of Gastroenterology, Renmin Hospital of Wuhan University, Hubei Key Laboratory of Digestive Diseases, Wuhan 430060, China; [2]Wuhan EndoAngel Medical Technology Co., Ltd., Wuhan 430000, China; [3]Department of Gastroenterology, Wuhan Hospital of Traditional Chinese and Western Medicine, Wuhan 430022, China; [4]Department of Gastroenterology, The Second Hospital of Jingzhou, Jingzhou 434000, China; [5]Department of Gastroenterology, General Hospital of Southern Theater Command of the PLA, Guangzhou 510405, China; [6]Department of Gastroenterology, The First Affiliated Hospital of Hunan Traditional Chinese Medical College, Zhuzhou 421000, China; [7]Department of Gastroenterology, People's Hospital of Zhushan County, Shiyan 442200, China

Corresponding author: Chen Mingkai, Email: kaimingchen@163.com

【**Abstract**】 **Objective** To develop a deep learning-based system for real-time recognition and classification of portal hypertensive gastropathy (PHG) and evaluate its ability to assist junior endoscopists. **Methods** A total of 2 848 gastroscopy images from 832 patients with liver cirrhosis were selected from Digestive Endoscopy Center databases of Renmin Hospital of Wuhan University, Wuhan Hospital of Traditional Chinese and Western Medicine, and the Second Hospital of Jingzhou from January 2015 to October 2023. This system referred to 3 endoscopic features of Baveno Ⅱ scoring system. Three models were developed respectively for gastric antral vascular ectasia (GAVE), mosaic-like pattern (MLP), and red marks (RM). The specific classification references were as follows: (1) GAVE model: 0 no, 1 yes; (2) MLP model: 0 no, 1 mild, 2 severe; (3) RM model: 0 no, 1 isolated, 2 fused. The classification results for endoscopic characteristics of PHG of 3 endoscopy experts were taken as the gold standard. The yolov8-m model was used for training. The training dataset, validation dataset, and test dataset were allocated at a ratio of 8∶1∶1. The test dataset was used to evaluate the performance of models and their auxiliary effects on endoscopists. The accuracy, recall, precision, specificity and *Kappa* coefficient were calculated. **Results** The accuracy, recall, specificity of GAVE model were 96.0% (48/50), 87.5% (7/8) and 97.6% (41/42). There was no significant difference between its accuracy and the gold standard (χ^2=316.226, P=1.000). The precision of GAVE1 and GAVE0 were 87.5% (7/8) and 97.6% (41/42) respectively. The accuracy of MLP model was 84.1% (132/157), and there was no significant difference compared with the gold standard (χ^2=3.286, P=0.193). The precision and recall of MLP2 were 88.2% (15/17) and 75.0% (15/20). The precision and recall of MLP1 were 77.9% (60/77) and 88.2% (60/68). The precision and recall of MLP0 were 90.5% (57/63) and 82.6% (57/69). The accuracy of RM model was 87.9% (123/140), and there was no significant difference compared with the gold standard (χ^2=2.891, P=0.409). The precision and recall of RM2 were 94.7% (18/19) and 78.3% (18/23). The precision and recall of RM1 were 72.2% (26/36) and 81.3% (26/32). The precision and recall of RM0 were 92.9% (79/85) and 92.9% (79/85). The mean accuracy of the three junior endoscopists, with and without the assistance of the GAVE model, MLP model, and RM model, respectively increased from 95.3% to 99.3%, from 83.9% to 91.9%, and from 81.9% to 83.1%. The overall consistency analysis of the 3 junior endoscopists with the gold standard indicated that the consistency of the GAVE model before and after assistance was extremely strong (both an overall *Kappa* of 1.000); the consistency before assistance of the MLP model was moderate (with an overall *Kappa* of 0.601), which increased to extremely strong after assistance (with an overall *Kappa* of 0.964); and the consistency of the RM model before and after assistance was also relatively strong (with an overall *Kappa* of 0.792 before and 0.798 after). **Conclusion** The deep learning system accurately identifies and classifies PHG features and significantly enhances diagnostic performance of junior endoscopists.

【**Key words**】 Artificial intelligence; Portal hypertensive gastropathy; Deep learning; Diagnosis; Computer assisted

Fund program: National Key Research and Development Program of China (2023YFC2507400)

门静脉高压性胃病(portal hypertensive gastropathy,PHG)是肝硬化的常见并发症,患病率为20%~80%[1]。PHG与肝硬化的发展密切相关,肝静脉压力梯度(hepatic venous pressure gradient,HVPG)≥12 mmHg(1 mmHg=0.133 kPa)、肝功能Child-Pugh B/C级,以及较大直径的静脉曲张都是PHG的独立危险因素[2]。因此,PHG的早期诊断及分级,对肝硬化的全病程管理十分重要。

目前PHG的诊断主要依靠胃镜,典型特征是蛇皮样的马赛克图案[3]。常用的分级方法有:Baveno Ⅱ评分系统[4]、McCormack法[5]、NIEC法[6]及Tanoue分类[7]法;其中Baveno Ⅱ评分系统是唯一一

www.zhxhnjzz.com

项半定量标准，通过引入胃窦毛细血管扩张症（gastric antral vascular ectasia，GAVE），提高了对PHG慢性出血的预测能力[8]。具体标准如下：GAVE（存在2分）、马赛克图案（mosaic-like pattern，MLP）（轻度1分、重度2分）、红色征（red marks，RM）（孤立1分、融合2分），分数之和≤3分判定为轻度PHG，≥4分判定为重度PHG。

然而，PHG的临床报告普遍存在描述不准确、分级方法不统一等不足。Cucchiarelli等[9]从意大利罗马的拉齐奥大区抽取了120份肝硬化患者的胃镜报告，发现66%描述了有无PHG，其中79%描述了PHG的分布，42%描述了PHG的严重程度，最终评价为“良好”的报告仅有19%。除了分级的主观性以外，较大直径的静脉曲张和消化道出血等客观因素都会妨碍PHG的识别，尤其是重度PHG[10]。因此，PHG的准确识别与分级，是当前临床研究必须面对的首要问题。

人工智能技术已应用于肝硬化食管胃静脉曲张的识别分级、出血风险预测等方面[11-12]，但罕见PHG的相关研究。鉴于Baveno Ⅱ评分系统已被证明具有较好的观察者间一致性，并且与短期出血风险相关[8]。本研究基于该评分系统，采用GAVE、MLP、RM这三个典型的内镜特征，利用深度学习的图片识别技术，开发了一款PHG内镜下的实时辅助识别与分级系统，并评估该系统对内镜医师的辅助效能。

资料与方法

一、研究对象

回顾性收集2015年1月至2023年10月，3家医院消化内镜中心（武汉大学人民医院230例、武汉市中西医结合医院466例、荆州市第二人民医院136例）接受胃镜诊疗的832例肝硬化患者的内镜资料。纳入标准：（1）年龄≥18岁，性别不限；（2）符合2023年《中国肝硬化临床诊治共识意见》[13]的肝硬化诊断标准。排除标准：（1）既往存在肝脏、食管、胃或特殊血管相关手术史；（2）胃镜报告或图片等资料缺失；（3）排除存在活动性出血、模糊、散焦、黏液、反射和泡沫等低质量图片，排除蓝光、放大、染色模式下拍摄的图片，以及食管、十二指肠部位的图片。最终纳入了2 848张胃镜图片。本研究作为回顾性研究，已获得武汉大学人民医院伦理委员会批准豁免知情同意，伦理审查号为WDRY2024-K025。

二、研究方法

1. 金标准：依据Baveno Ⅱ评分系统，2名拥有超过10年肝硬化内镜诊疗经验的专家，对每张图片的GAVE、MLP、RM三个特征分别分类，若两名专家存在分歧，将引入第3名同样资质的专家共同投票决定。

2. 特征定义与分类标准：GAVE指从幽门向胃窦和胃体放射出一段不同距离的平坦或轻微凸起的红色条纹状病变，GAVE0指无GAVE、GAVE1指有GAVE（图1）。MLP指被白黄色凹陷包围的小多边形区域，MLP0指无MLP；MLP1指轻度MLP，其区域中心黏膜呈粉红色；MLP2指重度MLP，其区域中心黏膜呈红色（图2）。RM指扁平或微隆起的红色病变，RM0指无RM，RM1指孤立RM呈离散分布；RM2指融合RM呈融合分布（图3）。

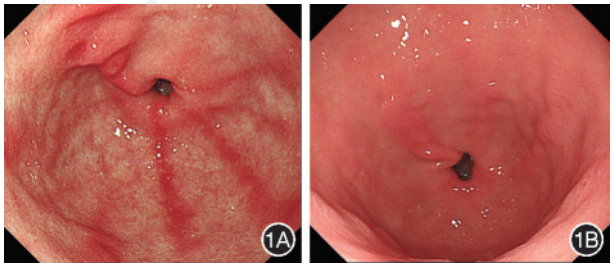

图1　内镜下胃窦毛细血管扩张症　1A：有胃窦毛细血管扩张症；1B：无胃窦毛细血管扩张症



图2　内镜下马赛克图案　2A：重度马赛克图案；2B：轻度马赛克图案；2C：无马赛克图案

www.zhxhnjzz.com

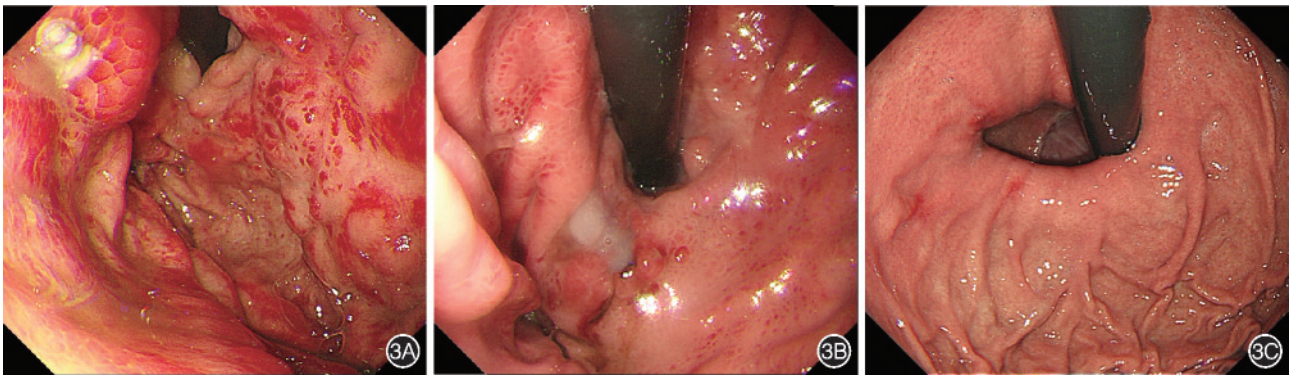

图3　内镜下红色征　3A：融合红色征；3B：孤立红色征；3C：无红色征

3. 数据集分布：专家分类后的金标准图片以8∶1∶1的比例分配到训练集、验证集、测试集，3个数据集间无重叠。训练集用于模型参数的学习，验证集用于超参数调优与模型选择，测试集仅用于最终模型性能的评估，检验其泛化能力。本系统的3个分类模型中，GAVE模型为二分类模型，MLP模型、RM模型为三分类模型。不同模型的数据集间存在交叉，例如无MLP和RM的一张图片可以同时被MLP0和RM0纳入。3个模型的数据集图片分布如表1所示。

表 1　识别门静脉高压性胃病系统的3个模型的训练集、验证集、测试集图片分布（张）

数据集	GAVE模型	MLP模型	RM模型
训练集	420	1 268	1 133
验证集	50	157	140
测试集	50	157	140
合计	520	1 582	1 413

注：GAVE指胃窦毛细血管扩张症；MLP指马赛克图案；RM指红色征

4. 模型训练与验证：本研究使用的深度学习框架为pytorch 1.9.0框架，采用了Ultralytics公司的yolov8-m模型[14]。使用“随机梯度下降”算法作为优化器，其初始学习率设置为0.001，并采用了学习率衰减策略，批次大小设置为64，训练周期数设置为100个。为防止过拟合，在训练中采用了L2权重衰减（权重因子为0.000 1）、丢弃法（丢弃率为0.5）等正则化技术。首先，将图片尺寸标准化（胃镜原始图片统一裁剪到224像素×224像素），并采用数据增强（平移、缩放、随机擦除等）和局部感兴趣区域等数据预处理方法。其次，将裁剪后的训练集和验证集图片输入基于pytorch 1.9.0框架的yolov8-m模型。在每个训练周期中，模型在训练集上进行前向传播，计算损失函数，并通过反向传播算法更新模型权重。每个训练周期结束后，模型在验证集上进行评估，并记录其性能。根据验证集上的性能表现来调整超参数（如学习率）和进行早停决策。最后，将在验证集上性能最佳的模型作为测试模型。

5. 模型测试：在输入模型前，测试集图片同样进行图片尺寸标准化等数据预处理。3名拥有不超过1年肝硬化内镜诊疗经验的初级医师，独立对测试集图片分类；经两周洗脱期之后，他们在3个模型的辅助下分别再次阅片。

三、观察指标

准确率：模型识别正确的图片占所有图片的百分比。某分类的召回率：在真实值为某分类的所有图片中，模型识别正确的图片所占百分比。特异度：针对二分类模型，在真实值为阴性的所有图片中，模型识别正确的图片所占的百分比。某分类的精确率：在模型识别为某分类的所有图片中，模型识别正确的图片所占百分比。

四、统计学方法

利用SPSS 27.0软件进行统计分析，通过准确率、召回率、特异度、精确率评估模型在测试集中的表现。计数资料采用例（%）来表示。McNemar检验用于比较模型与内镜医师的差异。*Kappa*系数用于评估同一名初级医师在模型辅助前后的观察者内一致性（简单*Kappa*系数用于二分类数据；加权*Kappa*系数用于三分类数据，采用线性加权）。*Kappa*系数的大小和一致性强弱的关系为：<0.20表示很差、0.20~0.40表示较差、>0.40~0.60表示中等、>0.60~0.80表示较强、>0.80表示极强。对于3名初级医师整体与金标准间一致性，取3名初级医师分类结果的众数（若分类结果均不相同则统一额外赋值），计算其整体*Kappa*系数。$P<0.05$认为差异有统计学意义。

www.zhxhnjzz.com

结　果

一、GAVE模型

1.GAVE模型的性能：在50张图片的测试集中，模型只识别错了2张图片，其准确率、召回率、特异度为96.0%(48/50)、87.5%(7/8)、97.6%(41/42)。模型的准确率和金标准相比，差异无统计学意义(χ^2=316.226，P=1.000)。GAVE1与GAVE0的精确率分别为87.5%(7/8)和97.6%(41/42)。

2.辅助阅片的结果：除初级医师3外，模型和其他医师的召回率、精确率、特异度本身都达到了较高的水平(80%以上)，详见表2。在准确率方面，3名初级医师在模型辅助前后的平均准确率分别为95.3%、99.3%。初级医师1(前χ^2=43.486，后χ^2=0.667)、初级医师2(前χ^2=50.000，后χ^2=3.000)、初级医师3(前χ^2=21.054，后χ^2=1.333)，在模型辅助前、后的准确率，与金标准相比差异均无统计学意义(P>0.05)。3名初级医师与金标准的一致性：初级医师1辅助前(*Kappa*=0.929)、辅助后(*Kappa*=1.000)均极强，初级医师2辅助前后无变化(*Kappa*=1.000)，初级医师3由辅助前的较强(*Kappa*=0.629)升至辅助后的极强(*Kappa*=0.922)。3名医师整体辅助前后与金标准的一致性均极强(整体*Kappa*均为1.000)。

二、MLP模型

1.MLP模型的性能：在157张图片的测试集中，MLP模型的准确率为84.1%(132/157)，和金标准相比，差异无统计学意义(χ^2=3.286，P=0.193)。MLP2的精确率、召回率分别为88.2%(15/17)和75.0%(15/20)，MLP1的精确率和召回率分别为77.9%(60/77)和88.2%(60/68)，MLP0的精确率和召回率分别为90.5%(57/63)和82.6%(57/69)，详见表3。

2.辅助阅片的结果：在准确率方面，3名初级医师在模型辅助前后的平均准确率分别为83.9%、91.9%；初级医师1(χ^2=28.000)、初级医师2(χ^2=55.000)、初级医师3(χ^2=34.084)和金标准相比，模型辅助前差异均有统计学意义(P<0.001)，辅助后则差异均无统计学意义(χ^2=0.667、χ^2=3.000、χ^2=1.333，P>0.05)。3名初级医师与金标准的一致性：初级医师1由辅助前的较强(*Kappa*=0.661)升至辅助后的极强(*Kappa*=0.947)，初级医师2由辅助前的较差(*Kappa*=0.383)升至辅助后的极强(*Kappa*=0.973)，初级医师3由辅助前的较强(*Kappa*=0.640)升至辅助后的极强(*Kappa*=0.964)。3名医师整体与金标准的一致性由辅助前的中等(整体*Kappa*=

表2　GAVE模型与3名初级医师对GAVE的判断结果

观察者	准确率	召回率	特异度	精确率	
				GAVE1	GAVE0
GAVE模型	96.0%(48/50)	87.5%(7/8)	97.6%(41/42)	87.5%(7/8)	97.6%(41/42)
初级医师1	98.0%(49/50)	100.0%(8/8)	97.6%(41/42)	88.9%(8/9)	100.0%(41/41)
辅助后	100.0%(50/50)	100.0%(8/8)	100.0%(42/42)	100.0%(8/8)	100.0%(42/42)
初级医师2	100.0%(50/50)	100.0%(8/8)	100.0%(42/42)	100.0%(8/8)	100.0%(42/42)
辅助后	100.0%(50/50)	100.0%(8/8)	100.0%(42/42)	100.0%(8/8)	100.0%(42/42)
初级医师3	88.0%(44/50)	87.5%(7/8)	88.1%(37/42)	58.3%(7/12)	97.4%(37/38)
辅助后	98.0%(49/50)	87.5%(7/8)	100.0%(42/42)	100.0%(7/7)	97.7%(42/43)

注：GAVE指胃窦毛细血管扩张症；GAVE1指有GAVE；GAVE0指无GAVE

表3　MLP模型与3名初级医师对MLP的判断结果

观察者	准确率	MLP2		MLP1		MLP0	
		精确率	召回率	精确率	召回率	精确率	召回率
MLP模型	84.1%(132/157)	88.2%(15/17)	75.0%(15/20)	77.9%(60/77)	88.2%(60/68)	90.5%(57/63)	82.6%(57/69)
初级医师1	82.2%(129/157)	62.5%(10/16)	50.0%(10/20)	100.0%(50/50)	73.5%(50/68)	75.8%(69/91)	100.0%(69/69)
辅助后	96.2%(151/157)	81.8%(18/22)	90.0%(18/20)	97.0%(64/66)	94.1%(64/68)	100.0%(69/69)	100.0%(69/69)
初级医师2	65.0%(154/157)	100.0%(3/3)	15.0%(3/20)	81.1%(30/37)	44.1%(30/68)	59.0%(69/117)	100.0%(69/69)
辅助后	98.1%(129/157)	100.0%(18/18)	90.0%(18/20)	97.1%(67/69)	98.5%(67/68)	98.6%(69/70)	100.0%(69/69)
初级医师3	71.3%(112/157)	54.5%(18/33)	90.0%(18/20)	89.7%(26/29)	38.2%(26/68)	71.6%(68/95)	98.6%(68/69)
辅助后	97.5%(153/157)	94.7%(18/19)	90.0%(18/20)	97.1%(66/68)	97.1%(66/68)	98.6%(69/70)	100.0%(69/69)

注：MLP指马赛克图案；MLP2指重度MLP；MLP1指轻度MLP；MLP0指无MLP

www.zhxhnjzz.com

0.601），升至辅助后的极强（整体*Kappa*=0.964）。

三、RM模型

1.RM模型的性能：在140张图片的测试集中，RM模型的准确率为87.9%（123/140），和金标准相比，差异无统计学意义（χ^2=2.891，*P*=0.409）。RM2的精确率和召回率分别为94.7%（18/19）和78.3%（18/23），RM1的精确率和召回率分别为72.2%（26/36）和81.3%（26/32），RM0的精确率和召回率分别为92.9%（79/85）和92.9%（79/85），详见表4。

2.辅助阅片的结果：在准确率方面，3名初级医师在模型辅助前后的平均准确率分别为81.9%和83.1%；和金标准相比，初级医师1在模型辅助前（χ^2=3.143，*P*=0.370）、后（χ^2=2.000，*P*=0.572）差异均无统计学意义。初级医师2（前χ^2=9.714，*P*=0.021；后χ^2=8.000，*P*=0.018）、初级医师3（前χ^2=28.235，*P*<0.001；后χ^2=29.286，*P*<0.001）在模型辅助前后，与金标准相比，差异均有统计学意义。3名初级医师与金标准的一致性：初级医师1由辅助前的较强（*Kappa*=0.797）升至辅助后的极强（*Kappa*=0.807）。初级医师2辅助前（*Kappa*=0.679）、后（*Kappa*=0.798）均较强。初级医师3在辅助前（*Kappa*=0.652）、后（*Kappa*=0.619）均较强。3名医师整体辅助前（整体*Kappa*=0.792）、后（整体*Kappa*=0.798）与金标准的一致性均较强。

讨　论

近年来，PHG受到愈来愈多的关注，2023年美国肝病学会指南首次建议合并轻度以上的PHG应被认为存在代偿性晚期慢性肝病[15]。然而，分级标准的不统一，极大阻碍了PHG的研究进展。本研究参考Baveno Ⅱ评分系统，首次开发了一款对GAVE、MLP及RM这3个内镜特征实时识别并分级的人工智能辅助诊断系统，3个模型的准确率为84.1%~96.0%，均达到较高水平。

随着技术进步，深度学习已应用消化内镜的盲区检测、巴雷特食管的范围测量、早期胃癌诊断等诸多方面[16]。本研究采用的yolov8-m模型是一种单独的卷积层神经网络模型，能实现端到端的目标检测[17]，相对于传统的YOLO算法（如yolov7或更早的版本），拥有更轻量级的网络结构、更强的特征提取网络、更好的稳健性等诸多优势，相对于其他尺寸的yolov8模型（yolov8-n/l/x），在准确率和推理速度之间达到了更好的平衡。考虑到Baveno Ⅱ评分系统的可重复性与出血预测能力已经得到了验证[8]，且包含的内镜特征更易于模型提取，本研究以此建立了GAVE、MLP及RM三个分类模型。

本研究的GAVE模型表现出了极高的准确率、精确率、召回率和特异度。在模型辅助后，3名医师的准确率都达到了98%及以上，只有1名初级医师分错了1张图片。既往研究中，内镜医师识别GAVE时，同样表现出了极高水平（特异度100%和假阳性率0%）[18]。此外，GAVE的低患病率限制了数据量，GAVE的“西瓜胃”外观容易识别无需判断严重程度，因此我们认为本研究的GAVE模型作为二分类模型的高效能表现是值得信赖的，而非模型过拟合与数据多样性不足的结果。

MLP与RM都属于三分类特征。有研究发现，倘若MLP、RM仅判断“有无”，其观察者间一致性较强（*Kappa*>0.75）；但判断严重程度时（如无MLP、轻度MLP、重度MLP），*Kappa*系数却会明显降低[8, 18]。在本研究中，3名初级医师独立分类MLP、RM时，错误的图片多集中于相邻的级别，如MLP1与MLP2、RM0与RM1。在模型辅助后，MLP只剩相邻级别的图片被分类错误，并且错误数量从6~45张减至3~4张；而RM非相邻级别（RM0与RM2）图片

表4　RM模型与3名初级医师对RM的判断结果

观察者	准确率	RM2		RM1		RM0	
		精确率	召回率	精确率	召回率	精确率	召回率
RM模型	87.9%（123/140）	94.7%（18/19）	78.3%（18/23）	72.2%（26/36）	81.3%（26/32）	92.9%（79/85）	92.9%（79/85）
初级医师1	87.1%（122/140）	81.0%（17/21）	73.9%（17/23）	77.4%（24/31）	75.0%（24/32）	92.0%（81/88）	95.3%（81/85）
辅助后	87.1%（122/140）	81.0%（17/21）	73.9%（17//23）	75.0%（24/32）	75.0%（24/32）	93.1%（81/87）	95.3%（81/85）
初级医师2	81.4%（114/140）	92.3%（12/13）	52.2%（12/23）	66.7%（22/33）	68.8%（22/32）	85.1%（80/94）	94.1%（80/85）
辅助后	85.7%（120/140）	93.3%（14/15）	60.9%（14/23）	66.7%（24/36）	75.0%（24/32）	92.1%（82/89）	96.5%（82/85）
初级医师3	77.1%（108/140）	87.5%（7/8）	30.4%（7/23）	50.0%（16/32）	50.0%（16/32）	85.0%（85/100）	100.0%（100/100）
辅助后	76.4%（107/140）	88.9%（8/9）	34.8%（8/23）	51.9%（14/27）	43.8%（14/32）	81.7%（85/104）	100.0%（85/85）

注：RM指红色征；RM2指融合RM呈融合分布；RM1指孤立RM呈离散分布；RM0指无RM分布

www.zhxhnjzz.com

分类错误的数量从3~5张减至0~2张。因此，本研究认为模型对MLP的严重程度、RM有无的判断有较好的辅助效果。

MLP与RM模型在单个分类中同时具有较高的精确率和召回率，而初级医师大多无法兼顾二者。例如就本研究的RM2而言，模型的精确率和召回率都高于3名初级医师，分别为94.7%（18/19）、78.3%（18/23）；而初级医师1在初级医师中精确率最低[81.0%（17/21）]，召回率最高[73.9%（17/23）]，初级医师2精确率较高[92.3%（12/13）]，召回率较低[52.2%（12/23）]。在模型辅助后，初级医师的精确率、召回率各有所提升，其中MLP模型的辅助效果比RM模型更为明显。此外，由于RM的严重程度主要体现在RM的分布状态，需要胃黏膜的整体评估，未来有望通过增加图片量和勾画病灶，进一步提升RM模型的辅助效果。

本研究还有一定的局限性。（1）作为一项回顾性研究，样本量相对较少，需要大样本的前瞻性研究进一步开发及验证；（2）评价模型辅助效果时，只利用图片验证，未进行视频验证；（3）只以图片为单位，未对单个患者的所有胃镜图片完整评估，因此未将3个特征按Baveno Ⅱ评分系统计算总分。

综上，这款PHG实时识别与分级系统的3个模型都达到了较高的准确率、精确率和召回率，尤其是GAVE模型，而MLP模型对初级医师的辅助效果最佳，有望提高内镜医师的诊断效率，降低其漏诊率和误诊率，为PHG的后续研究提供参考。未来期待能与食管静脉曲张的人工智能模型结合，开发出一款肝硬化胃镜下全面评估出血风险的预测系统。

利益冲突　所有作者声明不存在利益冲突

作者贡献声明　古皓文：设计与实施研究、采集数据、统计分析、撰写文章；杨杰：设计与实施研究、采集数据、技术支持；肖勇、万新月、胡伟、谢先木、黄定鹏、姚承明、施信良：指导与实施研究、采集数据；刘仕倩：设计研究、对文章的知识性内容做批评性审阅；黄丽：对文章的知识性内容做批评性审阅；张弛：实施研究、采集数据；郑碧清：研究指导、技术支持；陈明锴：研究指导、论文修改

参 考 文 献

[1] Yi Z, Zhang M, Ma Z, et al. Role of the posterior mucosal defense barrier in portal hypertensive gastropathy[J]. Biomed Pharmacother, 2021, 144: 112258. DOI: 10.1016/j.biopha.2021.112258.

[2] Merli M, Nicolini G, Angeloni S, et al. The natural history of portal hypertensive gastropathy in patients with liver cirrhosis and mild portal hypertension[J]. Am J Gastroenterol, 2004, 99(10):1959-1965. DOI: 10.1111/j.1572-0241.2004.40246.x.

[3] Cubillas R, Rockey DC. Portal hypertensive gastropathy: a review[J]. Liver Int, 2010, 30(8): 1094-1102. DOI: 10.1111/j.1478-3231.2010.02286.x.

[4] Sarin SK. Diagnostic issues: portal hypertensive gastropathy and gastric varices[C]// Portal hypertension Ⅱ. Proceedings of the second Baveno international consensus workshop on definitions, methodology and therapeutic strategies. Oxford: Blackwell Science, 1996.

[5] McCormack TT, Sims J, Eyre-Brook I, et al. Gastric lesions in portal hypertension: inflammatory gastritis or congestive gastropathy?[J]. Gut, 1985, 26(11): 1226-1232. DOI: 10.1136/gut.26.11.1226.

[6] Spina GP, Arcidiacono R, Bosch J, et al. Gastric endoscopic features in portal hypertension: final report of a consensus conference, Milan, Italy, September 19, 1992[J]. J Hepatol, 1994, 21(3):461-467. DOI: 10.1016/s0168-8278(05)80329-0.

[7] Tanoue K, Hashizume M, Wada H, et al. Effects of endoscopic injection sclerotherapy on portal hypertensive gastropathy: a prospective study[J]. Gastrointest Endosc, 1992, 38(5): 582-585. DOI: 10.1016/s0016-5107(92)70522-7.

[8] Stewart CA, Sanyal AJ. Grading portal gastropathy: validation of a gastropathy scoring system[J]. Am J Gastroenterol, 2003, 98(8):1758-1765. DOI: 10.1111/j.1572-0241.2003.07595.x.

[9] Cucchiarelli S, Santopaolo F, Lamazza A, et al. Pitfalls in the reporting of upper endoscopy features in cirrhotic patients[J]. Dig Liver Dis, 2019, 51(3): 382-385. DOI: 10.1016/j.dld.2018.08.014.

[10] Primignani M, Tosetti G. Portal hypertensive gastropathy after variceal eradication: more bleeding risk or just more reddening? [J]. Hepatol Int, 2016, 10(6): 847-850. DOI: 10.1007/s12072-016-9750-5.

[11] Wang Y, Hong Y, Wang Y, et al. Automated multimodal machine learning for esophageal variceal bleeding prediction based on endoscopy and structured data[J]. J Digit Imaging, 2023, 36(1):326-338. DOI: 10.1007/s10278-022-00724-6.

[12] Chen M, Wang J, Xiao Y, et al. Automated and real-time validation of gastroesophageal varices under esophagogastroduodenoscopy using a deep convolutional neural network: a multicenter retrospective study (with video)[J]. Gastrointest Endosc, 2021, 93(2): 422-432. e3. DOI: 10.1016/j.gie.2020.06.058.

[13] 中华医学会消化病学分会. 中国肝硬化临床诊治共识意见[J]. 中华消化杂志, 2023, 43(4): 227-247. DOI: 10.3760/cma.j.cn311367-20230228-00093.

[14] Jocher G, Chaurasia A, Qiu J. Ultralytics YOLO[Z]. 2023.

[15] Kaplan DE, Ripoll C, Thiele M, et al. AASLD practice guidance on risk stratification and management of portal hypertension and varices in cirrhosis[J]. Hepatology, 2024, 79(5):1180-1211. DOI: 10.1097/HEP.0000000000000647.

[16] 中华医学会消化内镜学分会大数据协作组. 上消化道内镜人工智能系统临床应用专家共识(2023,武汉)[J]. 中华消化内镜杂志, 2024, 41(2): 85-96. DOI: 10.3760/cma.j.cn321463-20231007-00399.

[17] 李茂晖, 吴传平, 鲍艳, 等. 论YOLO算法在机器视觉中应用原理[J]. 教育现代化, 2018, 5(41):174-176. DOI: 10.16541/j.cnki.2095-8420.2018.41.077.

[18] de Macedo GF, Ferreira FG, Ribeiro MA, et al. Reliability in endoscopic diagnosis of portal hypertensive gastropathy[J]. World J Gastrointest Endosc, 2013, 5(7): 323-331. DOI: 10.4253/wjge.v5.i7.323.

www.zhxhnjzz.com

VedVision唯视
胆道子镜直视系统

Vedkang唯德康

一次性使用胆胰管成像导管



电子内窥镜图像处理器

观入微，术无限，应于手

开启胆胰疾病诊治的直视操作时代

江苏唯德康医疗科技有限公司
Jiangsu Vedkang Medical Science and Technology Co., Ltd.

A 地址：江苏省武进经济开发区果香路52号
T 电话：0519-69877755, 69877756
F 传真：0519-69877753
E 邮箱：sales@vedkang.com

生产企业：江苏图云医疗科技有限公司
产品注册证及名称：
苏械注准 20222061594（电子内窥镜图像处理器）
苏械注准 20222061739（一次性使用胆胰管成像导管）
苏械广审（文）第 270803-07238 号

⚠ 禁忌内容或注意事项详见说明书
以上仅指本公司产品

广告

健可诺®
国药准字H20213838

磷酸钠盐散
Sodium Phosphates Powder

广告

清肠新选择 舒适又方便



独特包装
控制用药风险

适 应 症 用于患者结肠X-光线及肠道内窥镜检查前或手术前清理肠道。

用法用量 本品用于肠道准备时服药一般分两次，每次服药1袋。
第一次服药时间在手术或检查前一天晚上7点，用法为用800ml以上温凉水溶解后服用。
第二次服药时间在手术或检查当天早上7点(或在操作或检查前至少3个小时)，或遵医嘱，用法同第一次。
为获得良好肠道准备效果，建议患者在可承受范围内多饮用水。

不良反应 常见的不良反应为腹胀、恶心、腹痛、呕吐，还可能会出现用药期间和用药后的短暂的电解质紊乱、乏力、眩晕、过敏反应、肝功能检查ALT，AST升高、肛门刺激症状。其他详见说明书。

禁 忌 1.禁用于先天性巨结肠、肠梗阻、腹水、充血性心脏病或肾功能衰竭患者。2.使用本品禁止联合使用其他缓泻药物。3.对本品磷酸钠盐和其他成分过敏者禁用。

健能药业 JEWELLAND

川药广审(文)第261102-04383号　生产企业:四川健能制药有限公司　本广告仅供医学药学专业人士阅读