

· 论著 ·

# 基于多模态内镜图像的人工智能模型诊断胃肿瘤性病变研究(含视频)



扫码查看操作视频

陶道 吴练练 杜泓柳 董泽华 于红刚

武汉大学人民医院消化内科 消化系统疾病湖北省重点实验室, 武汉 430060

通信作者: 于红刚, Email: yuhonggang@whu.edu.cn

**【摘要】目的** 构建一种基于多模态内镜图像数据的人工智能胃肿瘤性病变诊断模型,并将其诊断效能与基于单模态图像数据的模型及内镜医师诊断水平进行比较。**方法** 收集武汉大学人民医院内镜中心 2018 年 3 月至 2019 年 12 月期间 463 例患者的 3 267 张白光成像和弱放大成像的胃肿瘤性病变及非肿瘤性病变图像,用于构建基于白光图像和弱放大图像的单模态模型(白光模型和弱放大模型)。将同一病变的白光和弱放大图像组合成图像对,用于构建多模态特征融合模型(多模态模型)。使用 2020 年 3 月至 2021 年 3 月间 97 例患者(102 个病变)的 696 张图像作为图像测试集,比较单模态模型与多模态模型在图像水平和病变水平识别胃肿瘤性病变的诊断效能。收集 2022 年 1 月至 2022 年 6 月期间 80 例患者(80 个病变)的视频片段,用于比较弱放大模型、多模态模型和 7 名内镜医师在病变水平对胃肿瘤性病变的诊断效能。**结果** 在图像测试集中,多模态模型在图像水平诊断胃肿瘤性病变的灵敏度、准确率分别为 84.96%(576/678)、86.89%(1 220/1 289),优于弱放大模型的灵敏度[63.13%(113/179), $\chi^2=42.81$ ,  $P<0.001$ ]和准确率[80.59%(353/438), $\chi^2=10.33$ ,  $P=0.001$ ],同样优于白光模型的灵敏度[70.47%(74/105), $\chi^2=13.52$ ,  $P<0.001$ ]和准确率[67.82%(175/258), $\chi^2=57.27$ ,  $P<0.001$ ]。多模态模型在病变水平诊断的灵敏度、特异度和准确率分别为 87.50%(28/32)、88.57%(62/70)、88.24%(90/102),特异度( $\chi^2=22.99$ ,  $P<0.001$ )和准确率( $\chi^2=19.06$ ,  $P<0.001$ )高于白光模型;但与弱放大模型间差异均无统计学意义( $P>0.05$ )。在视频测试中,多模态模型在病变水平诊断胃肿瘤性病变的灵敏度、特异度和准确率分别为 95.00%(19/20)、93.33%(56/60)和 93.75%(75/80),分别优于内镜医师总体的 77.14%(108/140)、79.29%(333/420)和 78.75%(441/560),差异具有统计学意义( $\chi^2=18.62$ ,  $P<0.001$ ;  $\chi^2=35.07$ ,  $P<0.001$ ;  $\chi^2=53.12$ ,  $P<0.001$ ),且高于资深内镜医师的灵敏度[83.33%(50/60)],差异有统计学意义( $\chi^2=4.23$ ,  $P=0.040$ )。**结论** 构建的基于多模态内镜图像数据的智能胃肿瘤性病变诊断模型,在图片和视频测试集中均有较高表现,且在视频测试中优于内镜医师平均水平。

**【关键词】** 人工智能; 胃肿瘤性病变; 弱放大成像; 多模态**基金项目:** 国家自然科学基金(82202257);武汉市人工智能示范应用场景项目(2022YYCJ01);武汉市知识创新专项-曙光计划项目(2022020801020482);武汉大学深化校企合作项目

## An artificial intelligence system based on multi-modal endoscopic images for the diagnosis of gastric neoplasms (with video)

Tao Xiao, Wu Lianlian, Du Hongliu, Dong Zehua, Yu Honggang

Department of Gastroenterology, Renmin Hospital of Wuhan University, Hubei Key Laboratory of Digestive Diseases, Wuhan 430060, China

Corresponding author: Yu Honggang, Email: yuhonggang@whu.edu.cn

**【Abstract】 Objective** To develop an artificial intelligence model based on multi-modal endoscopic images for identifying gastric neoplasms and to compare its diagnostic efficacy with traditional

DOI: 10.3760/cma.j.cn321463-20230227-00647

收稿日期 2023-02-27 本文编辑 钱程

引用本文:陶道,吴练练,杜泓柳,等.基于多模态内镜图像的人工智能模型诊断胃肿瘤性病变研究(含视频)[J].中华消化内镜杂志,2024,41(9):690-696. DOI: 10.3760/cma.j.cn321463-20230227-00647.



models and endoscopists. **Methods** A total of 3 267 images of gastric neoplasms and non-neoplastic lesions under white light (WL) endoscopy and weak magnification (WM) endoscopy from 463 patients at the Digestive Endoscopy Center of Renmin Hospital of Wuhan University from March 2018 to December 2019 were utilized. Two single-modal models (WL model and WM model) were constructed based on WL and WM images separately. WL and WM images of corresponding lesions were combined into image pairs for creating a multi-modal (MM) characteristics integration model. A test set consisting of 696 images of 102 lesions from 97 patients from March 2020 to March 2021 was used to compare the diagnostic efficacy of the single-modal models and a multi-modal model for gastric neoplastic lesions at both the image and the lesion levels. Additionally, video clips of 80 lesions from 80 patients from January 2022 to June 2022 were employed to compare diagnostic efficacy of the WM model, the MM model and 7 endoscopists at the lesion level for gastric neoplasms. **Results** In the image test set, the sensitivity and accuracy of MM model were 84.96% (576/678), and 86.89% (1 220/1 289), respectively, for diagnosing gastric neoplasms at the image level, which were superior to 63.13% (113/179) and 80.59% (353/438) of WM model ( $\chi^2=42.81, P<0.001; \chi^2=10.33, P=0.001$ ), and also better than those of WL model [70.47% (74/105),  $\chi^2=13.52, P<0.001; 67.82\%$  (175/258),  $\chi^2=57.27, P<0.001$ ]. The MM model showed a sensitivity of 87.50% (28/32), a specificity of 88.57% (62/70), and an accuracy of 88.24% (90/102) at the lesion level. The specificity ( $\chi^2=22.99, P<0.001$ ) and accuracy ( $\chi^2=19.06, P<0.001$ ) were significantly higher than those of WL model; however, there was no significant difference compared with those of the WM model ( $P>0.05$ ). In the video test, the sensitivity, specificity and accuracy of the MM model at the lesion level were 95.00% (19/20), 93.33% (56/60) and 93.75% (75/80). These results were significantly better than those of endoscopists, who had a sensitivity of 77.14% (108/140), a specificity of 79.29% (333/420), and an accuracy of 78.75% (441/560), with significant differences ( $\chi^2=18.62, P<0.001; \chi^2=35.07, P<0.001; \chi^2=53.12, P<0.001$ ), and was higher than the sensitivity of advanced endoscopists [83.33% (50/60)] with significant difference ( $\chi^2=4.23, P=0.040$ ). **Conclusion** The artificial intelligence model based on multi-modal endoscopic images for the diagnosis of gastric neoplasms shows high efficacy in both image and video test sets, outperforming the average diagnostic performance of endoscopists in the video test.

**【Key words】** Artificial intelligence; Gastric neoplasms; Weak magnifying endoscopy; Multi-modal

**Fund program:** National Natural Science Foundation of China (82202257); Artificial Intelligence Application Demonstration Scenario Project of Wuhan (2022YYCJ01); Knowledge Innovation Program of Wuhan-Shuguang Plan (2022020801020482); College-Enterprise Deepening Reform Project of Wuhan University

胃癌的全球发病率居第 5 位,是癌症相关死亡的主要原因之一<sup>[1-2]</sup>,在中国胃癌相关死亡率居癌症第 2 位<sup>[3]</sup>。然而,早期胃癌症状隐匿,不易察觉,绝大多数患者确诊时,已发展至进展期。进展期胃癌患者的五年生存率仅 40%,而早期胃癌患者的五年生存率高达 90%<sup>[4]</sup>。因此,早期诊断和早期治疗对提高胃癌患者的生存率至关重要。目前,内镜检查是筛查胃癌的主要手段<sup>[3,5]</sup>,然而白光内镜难以辨别早期胃癌的轻微黏膜改变,容易导致漏诊<sup>[6]</sup>。近年来随着染色放大内镜技术的发展,内镜下可清晰显示黏膜微血管和微结构,提高了胃癌的早期诊断率<sup>[7-8]</sup>,但该操作技术难度大,设备成本高昂。弱放大图像增强内镜例如奥林巴斯系列内镜的近聚焦模式,放大倍数约为 45 倍,亦可观察黏膜表面微结构<sup>[9]</sup>,应用广泛,且成本较低,为诊断胃癌等高危病变提供了新的选择。

近年来,人工智能技术广泛应用于医学领域,

在辅助诊断消化道肿瘤方面取得了长足发展<sup>[10-12]</sup>。然而,以往相关研究多聚焦于白光内镜或染色放大内镜下消化道肿瘤的识别<sup>[13-14]</sup>,在弱放大图像增强内镜方面鲜有研究。同时,以往研究多基于单种图像模式数据构建模型,获取的病变特征信息单一,可能导致对病变的识别不完全。本研究通过结合白光内镜和弱放大图像增强内镜特征信息,构建基于多模态数据的人工智能胃肿瘤性病变辅助诊断模型,并与传统的基于单模态数据的模型进行比较,旨在探索一种新方法帮助提高内镜医师对胃肿瘤性病变的诊断水平和检查效率,降低胃肿瘤性病变漏诊率。

## 资料与方法

### 一、研究资料

回顾性纳入在武汉大学人民医院内镜中心接

受上消化道白光内镜及弱放大增强内镜的门诊及住院患者的内镜影像资料。胃肿瘤性病变包括胃癌、上皮内瘤变、胃腺瘤、胃神经内分泌瘤等。本研究纳入的胃肿瘤性病变包含早期胃癌(肿瘤细胞局限于黏膜层及黏膜下层的腺癌)、高级别上皮内瘤变和低级别上皮内瘤变;纳入的胃非肿瘤性病变包括肠上皮化生、萎缩、溃疡、糜烂、慢性炎。影像资料中视野质量差、活动性出血、不清晰、泡沫干扰严重的图像被事先剔除。为保证弱放大模型的适用性,将近聚焦模式及相同放大范围(即放大倍数约为 45 倍)的放大窄带光成像和放大蓝激光成像的图像数据结合构建成弱放大成像数据集(图 1)。将同一病变的所有白光图像分别与所有弱放大图像进行任意组合,形成图像对。由两名具有 5 年以上内镜经验的内镜医师对图像和视频进行标记和剪辑,病变均有明确的内镜和组织病理学诊断结果,包含内镜活检和(或)内镜手术病理,以病理诊断结果作为金标准。图像拍摄均使用日本奥林巴斯公司 260 和 290 系列以及日本富士公司 580 和 590 系列内镜。视频均使用录像系统(型号 ELUXEO 7000、LASEREO7000、VP-4450HD, 日本富士公司;型号 EVIS LUCERA CV-260/CLV-260 和 EVIS LUCERA ELITE CV-290/CLV-290SL, 日本奥

林巴斯公司)。本研究通过武汉大学人民医院伦理委员会审核批准,伦理审查号:WDRY2019-K073(X03),审查委员会对回顾性数据豁免知情同意。

1. 数据集 1:来自 2018 年 3 月至 2019 年 12 月 463 例患者(463 个病变)的 3 267 张图片,其中胃肿瘤性病变 143 例(早期癌 82 例、低级别上皮内瘤变 61 例,包含 383 张白光图片、602 张弱放大图片),非肿瘤性病变 320 例(肠上皮化生 86 例、萎缩 68 例、慢性炎 166 例,包含 924 张白光图片、1 358 张弱放大图片)。

2. 数据集 2:来自 2020 年 3 月至 2021 年 3 月 97 例患者(102 个病变)的 696 张图片,其中肿瘤性病变 32 个(早期癌 24 个、低级别上皮内瘤变 8 个,包含 105 张白光图片、179 张弱放大图片,共 678 个图像对),非肿瘤性病变 70 个(肠上皮化生 26 个、萎缩 21 个、慢性炎 23 个,包含 153 张白光图片、259 张弱放大图片,共 611 个图像对)。

3. 数据集 3:来自 2022 年 1 月至 2022 年 6 月在武汉大学人民医院的 80 例患者的 80 个病变视频段,其中包括 20 例肿瘤(早期癌 16 例、低级别上皮内瘤变 4 例)和 60 例非肿瘤视频(肠上皮化生 23 例、萎缩 15 例、慢性炎 22 例)。病变视频均包含白光和弱放

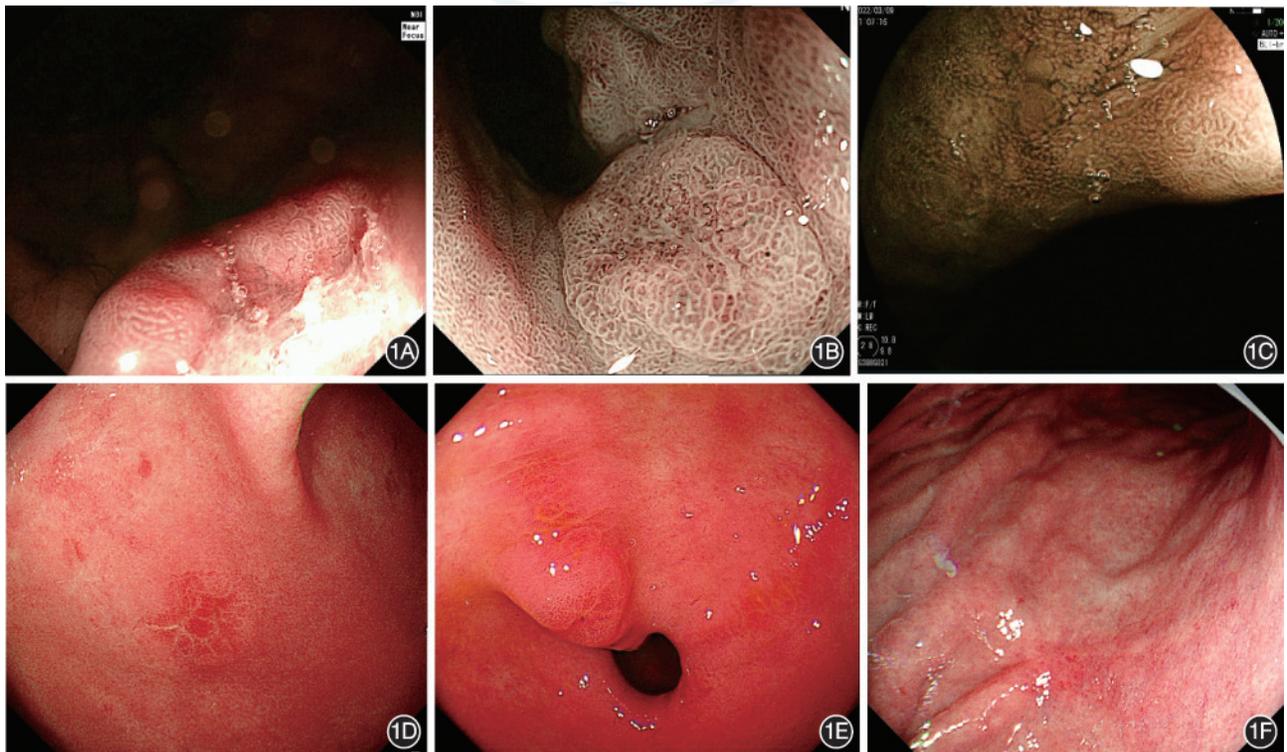


图 1 胃黏膜病变的白光和弱放大典型图像 1A:弱放大成像肿瘤性病变,病理诊断为高级别上皮内瘤变;1B:弱放大成像肿瘤性病变,病理诊断为早期胃癌;1C:弱放大成像非肿瘤性病变,病理诊断为慢性炎;1D:白光内镜肿瘤性病变,病理诊断为低级别上皮内瘤变;1E:白光内镜非肿瘤性病变,病理诊断为慢性炎;1F:白光内镜非肿瘤性病变,病理诊断为萎缩性胃炎

大剪辑片段,白光时长为(10.44±4.67)s,弱放大片段时长(33.29±17.04)s。

## 二、研究方法

### (一)模型构建

1. 白光识别模型(白光模型):将数据集 1 中经过预处理的 1 307 张白光图片输入到基于 keras 2.4.3 框架的 Resnet50 模型中,经过一系列卷积计算后,由池化层转换为包含 2 048 维特征的特征向量,通过分类器计算特征向量,获得“肿瘤性病变与非肿瘤性病变”的二元分类训练模型。

2. 弱放大识别模型(弱放大模型):使用数据集 1 中经过预处理的 1 960 张弱放大成像图片进行训练,训练方法同白光模型。

3. 多模态特征融合模型(多模态模型):双模态神经网络由白光卷积层、弱放大卷积网络层、特征融合卷积层和分类器组成。将数据集 1 中同一病变的任一白光和任一弱放大图片两两组合成图像对输入到双模态网络,白光和放大图像分别进入对应的卷积层,进行特征提取后使用浅层卷积进行特征连接,进行“肿瘤性病变与非肿瘤性病变”的二元分类训练(图 2)。

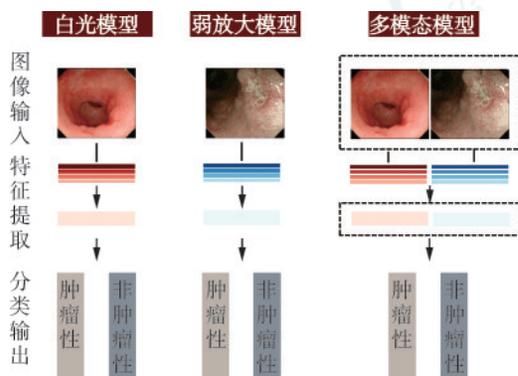


图 2 白光模型、弱放大模型、多模态特征融合模型构建示意图

### (二)模型验证

1. 图像测试:数据集 2 作为内部测试集用于图像测试,白光模型在白光模式下进行测试,弱放大模型在弱放大模式下进行测试,每张图像对应输出一个分类结果,最终在图像水平和病变水平上输出诊断结果。基于病变水平诊断结果的计算逻辑为:当该病变所有图像中至少有一张图像被判断为肿瘤性病变时,该病变诊断为肿瘤性病变,否则,该病变诊断为非肿瘤性病变。多模态模型基于病变的图像对进行测试,在图像对水平和病变水平上进行结果输出。

2. 视频测试:数据集 3 用于视频测试,视频片

段输入到模型后,以 7 帧/s 分解为图像帧,多模态模型存储白光片段的每帧图像,当切换到弱放大模式时,每帧弱放大图像和该病变所有白光图像分别组合成图像对,输出每个图像对的结果,随后根据诊断逻辑输出病变水平结果(视频)。

此外,数据集 3 还用于 3 种模型与 7 名内镜医师识别胃肿瘤性病变的能力比较。7 名医师均来自武汉大学人民医院,与参与图像标记的 2 名医师互相独立。其中 3 名具有 5 年以上上消化道内镜检查经验的医师被定义为资深医师,2 名具有 3~5 年经验的定义为中级内镜医师,2 名具有 1~2 年经验的定义为初级内镜医师。内镜医师分别独立完成视频测试,2022 年 10 月第一轮测试中,内镜医师仅观看弱放大片段判断病变是否为肿瘤性,诊断结果与弱放大模型诊断结果进行比较。3 周洗脱期后进行第二轮测试,内镜医师观看白光和弱放大片段综合做出判断。测试前及洗脱期内实验相关信息以及所有视频和病变信息均对医师保密。诊断结果与多模态模型结果进行比较。

### 三、统计学方法

应用 SPSS 26.0 软件进行统计学分析,正态分布的计量资料以  $\bar{x} \pm s$  表示,计数资料以例(%)表示。通过计算准确率、灵敏度、特异度、阳性预测值和阴性预测值,评估模型和内镜医师的诊断效能。使用卡方检验比较不同模型之间及模型与内镜医师之间诊断效能的差异, $P < 0.05$  为差异有统计学意义。

## 结 果

### 一、数据集 2 测试结果

在数据集 2 的图像测试中,多模态模型在图像水平的诊断灵敏度、特异度、准确率、阳性预测值和阴性预测值分别为 84.96% (576/678)、89.03% (544/611)、86.89% (1 120/1 289)、89.58% (576/643) 和 84.21% (544/646),见表 1。多模态模型的灵敏度 ( $\chi^2=42.81, P < 0.001$ )、准确率 ( $\chi^2=10.33, P = 0.001$ ) 和阴性预测值 ( $\chi^2=4.77, P = 0.029$ ) 均高于弱放大模型,差异有统计学意义;而特异度 ( $\chi^2=2.69, P = 0.101$ ) 和阳性预测值 ( $\chi^2=1.75, P = 0.185$ ) 与弱放大模型相比,差异无统计学意义。多模态模型灵敏度 ( $\chi^2=13.52, P < 0.001$ )、特异度 ( $\chi^2=49.32, P < 0.001$ )、准确率 ( $\chi^2=57.27, P < 0.001$ )、阳性预测值 ( $\chi^2=76.66, P < 0.001$ ) 和阴性预测值 ( $\chi^2=4.58, P = 0.032$ ) 均高于白光模型,差异有统计学意义。

表1 不同人工智能模型在数据集2图像水平识别胃肿瘤性病变的诊断效能

模型分组	灵敏度	特异度	准确率	阳性预测值	阴性预测值
白光模型	70.47%(74/105)	66.01%(101/153)	67.82%(175/258)	58.73%(74/126)	76.52%(101/132)
弱放大模型	63.13%(113/179)	92.66%(240/259)	80.59%(353/438)	85.61%(113/132)	78.43%(240/306)
多模态模型	84.96%(576/678)	89.03%(544/611)	86.89%(1120/1289)	89.58%(576/643)	84.21%(544/646)

在数据集2病变水平上,多模态模型的灵敏度、特异度、准确率、阳性预测值和阴性预测值分别为 87.50% (28/32)、88.57% (62/70)、88.24% (90/102)、77.78% (28/36) 和 93.94% (62/66), 见表 2。多模态模型的特异度( $\chi^2=22.99, P<0.001$ )、准确率( $\chi^2=19.06, P<0.001$ )和阳性预测值( $\chi^2=10.36, P=0.001$ )均高于白光模型,差异有统计学意义。多模态模型的灵敏度( $\chi^2=0.13, P=0.719$ )、特异度( $\chi^2=3.22, P=0.073$ )、准确率( $\chi^2=2.93, P=0.087$ )、阳性预测值( $\chi^2=2.08, P=0.149$ )和阴性预测值( $\chi^2=0.27, P=0.602$ )均高于弱放大模型,但差异均无统计学意义。

## 二、数据集3测试结果

在数据集3视频测试病变水平上,弱放大模型的特异度和准确率分别为 86.67% (52/60) 和 85.00% (68/80), 见表 3, 高于内镜医师仅观看弱放大视频片段的灵敏度( $\chi^2=18.45, P<0.001$ )和准确率( $\chi^2=21.65, P<0.001$ ), 且与资深内镜医师的结果相当( $\chi^2=0.09, P=0.761; \chi^2=0.02, P=0.899$ )。

多模态模型的灵敏度、特异度和准确率分别为

95.00% (19/20)、93.33% (56/60) 和 93.75% (75/80), 见表 4, 均高于内镜医师依据白光和弱放大视频综合诊断的灵敏度( $\chi^2=18.62, P<0.001$ )、特异度( $\chi^2=35.07, P<0.001$ )和准确率( $\chi^2=53.12, P<0.001$ ), 且灵敏度高于资深内镜医师( $\chi^2=4.23, P=0.040$ ), 差异均有统计学意义。

## 讨 论

本研究中,我们探索了一种基于多模态数据的深度学习的方法,同时融合了白光成像和弱放大成像中的病变特征,能够对胃黏膜病变是否为肿瘤进行判断。在图片测试集中,该模型诊断胃肿瘤的灵敏度、特异度、准确率分别为 84.96%、89.03%、86.89%。在视频测试集中,该模型的灵敏度和特异度分别为 95.00% 和 93.33%, 显著优于 7 名内镜医师平均水平,且与资深内镜医师表现持平。

胃癌仍然是世界上常见的癌症之一,严重影响患者生命健康及生活质量,胃癌的早发现、早治疗能改善预后,提高患者生存率<sup>[3]</sup>。消化内镜检查是

表2 不同人工智能模型在数据集2病变水平识别胃肿瘤性病变的诊断效能

模型分组	灵敏度	特异度	准确率	阳性预测值	阴性预测值
白光模型	84.38%(27/32)	51.43%(36/70)	61.76%(63/102)	44.26%(27/61)	87.80%(36/41)
弱放大模型	84.38%(27/32)	77.14%(54/70)	79.41%(81/102)	62.79%(27/43)	91.53%(54/59)
多模态模型	87.50%(28/32)	88.57%(62/70)	88.24%(90/102)	77.78%(28/36)	93.94%(62/66)

表3 弱放大模型和内镜医师在视频集中基于病变水平识别胃肿瘤性病变的诊断效能(第一轮测试)

诊断者分组	灵敏度	特异度	准确率
弱放大模型	80.00%(16/20)	86.67%(52/60)	85.00%(68/80)
内镜医师(n=7)	70.00%(98/140)	75.00%(315/420)	73.75%(413/560)
资深内镜医师(n=3)	81.67%(49/60)	85.56%(154/180)	84.58%(203/240)
中级内镜医师(n=2)	70.00%(28/40)	74.17%(89/120)	73.13%(117/160)
初级内镜医师(n=2)	52.50%(21/40)	60.00%(72/120)	58.13%(93/160)

表4 多模态模型和内镜医师在视频集中基于病变水平识别胃肿瘤性病变的诊断效能(第二轮测试)

诊断者分组	灵敏度	特异度	准确率
多模态模型	95.00%(19/20)	93.33%(56/60)	93.75%(75/80)
内镜医师(n=7)	77.14%(108/140)	79.29%(333/420)	78.75%(441/560)
资深内镜医师(n=3)	83.33%(50/60)	93.33%(168/180)	90.83%(218/240)
中级内镜医师(n=2)	72.50%(29/40)	79.17%(95/120)	77.50%(124/160)
初级内镜医师(n=2)	72.50%(29/40)	58.33%(70/120)	61.88%(99/160)

消化道肿瘤筛查的重要工具,然而,早期肿瘤轻微病变难以识别,以及地区间医疗资源的差异、医师间水平的差异,使得早期胃肿瘤的漏诊率仍然居高不下<sup>[15]</sup>。近年来内镜技术的高速发展,窄带光成像内镜、放大内镜的出现,突出显示了消化道黏膜微血管和微结构的改变,为内镜医师辨别早期胃肿瘤提供了帮助<sup>[7]</sup>。但染色放大内镜对设备要求高,且技术难度大,学习周期长,需要内镜医师具有丰富的理论及操作经验,仍然不能广泛用于消化道早期肿瘤的筛查。新近出现的弱放大成像模式既可以实现黏膜细微血管和结构的观察,又能在常规筛查中应用,为诊断胃肿瘤提供了新的机会<sup>[9]</sup>。

近年来,基于深度学习的人工智能在消化内镜领域掀起了应用热潮,在消化内镜质量控制和病变识别方面取得了可观的进展<sup>[10-11]</sup>。人工智能具有可提取人眼无法识别的细微黏膜特征、不受其他因素干扰等多种优势<sup>[16]</sup>,目前人工智能对早期胃肿瘤的诊断仍是研究的热点<sup>[17]</sup>。Hirasawa 等<sup>[18]</sup>利用 13 584 张白光内镜图像训练的人工智能模型,对胃癌的总体诊断灵敏度为 92.2%。Wu 等<sup>[19]</sup>进行了一项随机对照试验,验证了人工智能辅助白光内镜检查,可显著降低胃肿瘤的漏诊率。He 等<sup>[20]</sup>开发了一个在放大染色内镜下辅助诊断早期胃癌的系统,并在真实的临床实践中达到了 92.59% 的灵敏度和 83.67% 的准确率。

然而,尚未有研究探索过基于弱放大成像模式的人工智能模型。此外,在真实的临床环境中,当白光内镜下发现可疑病变时,设备允许时内镜医师通常会进一步使用弱放大成像或放大染色内镜进一步观察,综合判断病变性质。而以往的研究主要使用白光或放大染色单模态数据构建模型,将病变的诊断过程和逻辑分割成了多个独立的环节,忽略了不同成像模式下的病变特征差异,可能导致对病变的识别不全。最近的一项研究证实,与单独使用白光内镜和放大染色相比,人工智能联合使用白光和放大染色在诊断结肠癌浸润深度方面具有更好的性能<sup>[21]</sup>。本研究创新性地探索了基于多模态数据的人工智能模型在诊断胃肿瘤方面的能力,开发并验证了该模型的性能,为人工智能辅助诊断胃肿瘤提供了一个新思路。

本研究结果显示,在相同的图像和视频测试集中,多模态模型的灵敏度、特异度均优于白光模型和弱放大模型,表明较之以往基于单一成像模式数据的训练模型,基于多模态数据的人工智能模型具

有更高的优越性和突破性。将同一病变的不同成像模式下的特征进行提取,融合了不同模式下的特征优势,实现病变特征信息的最大化利用,体现了对病变识别的综合性和全面性。另一方面,以往的研究对于模型的训练,多依赖于大量样本的标注,本研究将同一病变的白光图像和弱放大图像任意两两组合构成图像对作为样本单位,增加了样本的多样性,有效扩充了训练集样本数量,有望实现小样本量基础上开发人工智能模型。此外,单张图像受拍摄角度和光源等影响,往往会丢失部分病变特征,使判断出现误差,且临床实践中,内镜医师往往通过多角度全方位地观察病变综合得出判断,因此本研究验证模型性能时,不局限于在图像层次上测试,还在包含多张图像的病变水平上测试性能,并进一步在视频中对胃黏膜病变进行识别,可实现在临床环境中动态和实时地辅助内镜医师,实现人机交互。

本研究仍然存在一定的局限性。首先,本研究仅回顾性地纳入了武汉大学人民医院的单中心数据,未来还需进一步在多中心数据基础上进行训练和前瞻性验证,减少偏倚风险,增加模型泛化性。其次,尽管在视频数据集上比较了模型和内镜医师的能力,未来仍然需要开展临床试验验证该模型在真实临床环境的有效性和实用性。

总之,本研究构建了一种基于多模态数据的人工智能模型辅助诊断胃肿瘤性病变,在图像和视频测试集中均有较高的准确率,且在视频测试中优于内镜医师平均水平,具有很好的临床应用前景。

**利益冲突** 所有作者声明不存在利益冲突

**作者贡献声明** 陶道、杜泓柳、董泽华:实验操作、论文撰写、数据整理、统计学分析;吴练练、于红刚:研究指导、论文修改、经费支持

## 参 考 文 献

- [1] Sung H, Ferlay J, Siegel RL, et al. Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries[J]. *CA Cancer J Clin*, 2021,71(3):209-249. DOI: 10.3322/caac.21660.
- [2] Smyth EC, Nilsson M, Grabsch HI, et al. Gastric cancer[J]. *Lancet*, 2020, 396(10251): 635-648. DOI: 10.1016/S0140-6736(20)31288-5.
- [3] Wang FH, Zhang XT, Li YF, et al. The Chinese Society of Clinical Oncology (CSCO): clinical guidelines for the diagnosis and treatment of gastric cancer, 2021[J]. *Cancer Commun (Lond)*, 2021, 41(8): 747-795. DOI: 10.1002/cac2.12193.
- [4] Katai H, Ishikawa T, Akazawa K, et al. Five-year survival analysis of surgically resected gastric cancer cases in Japan: a

- retrospective analysis of more than 100,000 patients from the nationwide registry of the Japanese Gastric Cancer Association (2001-2007) [J]. *Gastric Cancer*, 2018, 21(1): 144-154. DOI: 10.1007/s10120-017-0716-7.
- [5] Yao K, Uedo N, Kamada T, et al. Guidelines for endoscopic diagnosis of early gastric cancer[J]. *Dig Endosc*, 2020, 32(5): 663-698. DOI: 10.1111/den.13684.
- [6] Zhang X, Li M, Chen S, et al. Endoscopic screening in Asian countries is associated with reduced gastric cancer mortality: a meta-analysis and systematic review[J]. *Gastroenterology*, 2018, 155(2):347-354.e9. DOI: 10.1053/j.gastro.2018.04.026.
- [7] Hu YY, Lian QW, Lin ZH, et al. Diagnostic performance of magnifying narrow-band imaging for early gastric cancer: a meta-analysis[J]. *World J Gastroenterol*, 2015, 21(25): 7884-7894. DOI: 10.3748/wjg.v21.i25.7884.
- [8] Zhang Q, Wang F, Chen ZY, et al. Comparison of the diagnostic efficacy of white light endoscopy and magnifying endoscopy with narrow band imaging for early gastric cancer: a meta-analysis[J]. *Gastric Cancer*, 2016, 19(2):543-552. DOI: 10.1007/s10120-015-0500-5.
- [9] Kakushima N, Yoshida N, Doyama H, et al. Near-focus magnification and second-generation narrow-band imaging for early gastric cancer in a randomized trial[J]. *J Gastroenterol*, 2020, 55(12):1127-1137. DOI: 10.1007/s00535-020-01734-3.
- [10] Goyal H, Sherazi S, Mann R, et al. Scope of artificial intelligence in gastrointestinal oncology[J]. *Cancers (Basel)*, 2021, 13(21):5494. DOI: 10.3390/cancers13215494.
- [11] Niu PH, Zhao LL, Wu HL, et al. Artificial intelligence in gastric cancer: application and future perspectives[J]. *World J Gastroenterol*, 2020, 26(36):5408-5419. DOI: 10.3748/wjg.v26.i36.5408.
- [12] Zhang C, Wu L. The application of artificial intelligence in gastrointestinal endoscopy: a state-of-the-art review[J]. *JDH*, 2022, 1(1): 3-18. DOI: 10.55976/jdh.120221423-18.
- [13] Wu L, Xu M, Jiang X, et al. Real-time artificial intelligence for detecting focal lesions and diagnosing neoplasms of the stomach by white-light endoscopy (with videos) [J]. *Gastrointest Endosc*, 2022, 95(2):269-280.e6. DOI: 10.1016/j.gie.2021.09.017.
- [14] 朱益洁, 吴练练, 何鑫琦, 等. 基于不同训练方法的 2 种人工智能系统在染色放大内镜下识别早期胃癌的能力比较[J]. *中华消化杂志*, 2022, 42(7):433-438. DOI: 10.3760/cma.j.cn311367-20211214-00680.
- [15] Pimenta-Melo AR, Monteiro-Soares M, Libânio D, et al. Missing rate for gastric cancer during upper gastrointestinal endoscopy: a systematic review and meta-analysis[J]. *Eur J Gastroenterol Hepatol*, 2016, 28(9):1041-1049. DOI: 10.1097/MEG.0000000000000657.
- [16] Bi WL, Hosny A, Schabath MB, et al. Artificial intelligence in cancer imaging: clinical challenges and applications[J]. *CA Cancer J Clin*, 2019, 69(2):127-157. DOI: 10.3322/caac.21552.
- [17] Kim JH, Nam SJ, Park SC. Usefulness of artificial intelligence in gastric neoplasms[J]. *World J Gastroenterol*, 2021, 27(24): 3543-3555. DOI: 10.3748/wjg.v27.i24.3543.
- [18] Hirasawa T, Aoyama K, Tanimoto T, et al. Application of artificial intelligence using a convolutional neural network for detecting gastric cancer in endoscopic images[J]. *Gastric Cancer*, 2018, 21(4): 653-660. DOI: 10.1007/s10120-018-0793-2.
- [19] Wu L, Shang R, Sharma P, et al. Effect of a deep learning-based system on the miss rate of gastric neoplasms during upper gastrointestinal endoscopy: a single-centre, tandem, randomised controlled trial[J]. *Lancet Gastroenterol Hepatol*, 2021, 6(9): 700-708. DOI: 10.1016/S2468-1253(21)00216-8.
- [20] He X, Wu L, Dong Z, et al. Real-time use of artificial intelligence for diagnosing early gastric cancer by magnifying image-enhanced endoscopy: a multicenter diagnostic study (with videos)[J]. *Gastrointest Endosc*, 2022, 95(4):671-678.e4. DOI: 10.1016/j.gie.2021.11.040.
- [21] Lu Z, Xu Y, Yao L, et al. Real-time automated diagnosis of colorectal cancer invasion depth using a deep learning model with multimodal data (with video) [J]. *Gastrointest Endosc*, 2022, 95(6):1186-1194.e3. DOI: 10.1016/j.gie.2021.11.049.

## · 读者 · 作者 · 编者 ·

### 《中华消化内镜杂志》2024 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)	POEM(经口内镜食管下括约肌切开术)	Hb(血红蛋白)
EST(经内镜乳头括约肌切开术)	NOTES(经自然腔道内镜手术)	PaO <sub>2</sub> (动脉血氧分压)
EUS(超声内镜检查术)	MRCP(磁共振胰胆管成像)	PaCO <sub>2</sub> (动脉血二氧化碳分压)
EUS-FNA(超声内镜引导细针穿刺抽吸术)	GERD(胃食管反流病)	ALT(丙氨酸转氨酶)
EMR(内镜黏膜切除术)	RE(反流性食管炎)	AST(天冬氨酸转氨酶)
ESD(内镜黏膜下剥离术)	IBD(炎症性肠病)	AKP(碱性磷酸酶)
ENBD(经内镜鼻胆管引流术)	UC(溃疡性结肠炎)	IL(白细胞介素)
ERBD(经内镜胆道内支架放置术)	NSAIDs(非甾体抗炎药)	TNF(肿瘤坏死因子)
APC(氩离子凝固术)	PPI(质子泵抑制剂)	VEGF(血管内皮生长因子)
EVL(内镜下静脉曲张套扎术)	HBV(乙型肝炎病毒)	ELISA(酶联免疫吸附测定)
EIS(内镜下硬化剂注射术)	HBsAg(乙型肝炎病毒表面抗原)	RT-PCR(逆转录-聚合酶链反应)