doi:10.11659/jjssx.02E024054

·胃肠肝胆专栏·

基于深度学习的人工智能辅助胃镜下实时识别病变及位置模型的建立

郭 宪,吴应洋,江艾芮,樊超强,彭 学,聂绪彪,林 辉,柏健鹰 (陆军军医大学第二附属医院消化内科,重庆 400037)

[摘 要] **目的** 构建一个基于深度学习的人工智能辅助诊断模型,用于实时动态识别胃镜下的胃部病变及位置,并评估其对胃部病变检出和位置识别的能力。**方法** 回顾性分析我院104例患者的胃镜检查视频,对视频帧进行人工标注,将已标注的病变类别图片帧按8∶2的比例划分成训练集与验证集,将已标注的位置类别图片帧根据患者来源按8∶2的比例划分成训练集与验证集,分别用于模型的训练及验证。病变识别部分的训练采用YoloV4模型,位置识别部分的训练采用ResNet152模型。评估辅助诊断模型对病变识别的准确率、敏感度、特异度、阳性预测值、阴性预测值,及对位置识别的准确率。**结果** 共标注图片68 351帧,其中训练集图片54 872帧,包括病变类别41 692帧,位置类别13 180帧;验证集图片13 479帧,包括病变类别10 422帧,位置类别3 057帧。在验证集中,病变识别模型的总体准确率为98.8%,敏感度为96.6%,特异度为99.3%,阳性预测值为96.3%,阴性预测值为99.3%;位置识别模型top5总体准确率为87.1%。**结论** 基于深度学习的人工智能辅助诊断模型用于实时动态识别胃镜下胃部病变及位置有较好的病变检出能力和位置识别能力,临床应用前景巨大。

[关键词] 人工智能;胃部疾病;早期胃癌;深度学习;胃镜检查
[中图分类号] R445.9;TP183 [文献标志码] A [收稿日期] 2024-02-26

Establishment of an artificial intelligence assisted diagnosis model based on deep learning for recognizing gastric lesions and their locations under gastroscopy in real time

GUO Xian, WU Ying-yang, JIANG Ai-rui, FAN Chao-qiang, PENG Xue, NIE Xu-biao, LIN Hui, BAI Jian-ying
(Department of Gastroenterology, the Second Affiliated Hospital of Army Medical University, Chongqing 400037, China)

Abstract: Objective To construct an artificial intelligence assisted diagnosis model based on deep learning for dynamically recognizing gastric lesions and their locations under gastroscopy in real time, and to evaluate its ability to detect and recognize gastric lesions and their locations. **Methods** The gastroscopy videos of 104 patients in our hospital was retrospectively analyzed, and the video frames were manually annotated. The annotated picture frames of lesion category were divided into the training set and the validation set according to the ratio of 8∶2, and the annotated picture frames of location category were divided into the training set and the validation set according to the patient sources at the ratio of 8∶2. These sets were utilized for training and validating the respective models. YoloV4 model was used for the training of lesion recognition, and ResNet152 model was used for the training of location recognition. The accuracy, sensitivity, specificity, positive predictive value, negative predictive value and location recognition accuracy of the auxiliary diagnostic model were evaluated. **Results** A total of 68 351 image frames were annotated, with 54 872 frames used as the training set, including 41 692 frames for lesion categories and 13 180 frames for location categories. The validation set consisted of 13 479 frames, comprising 10 422 frames for lesion categories and 3 057 frames for location categories. The lesion recognition model achieved an overall accuracy of 98.8%, with a sensitivity of 96.6%, specificity of 99.3%, positive predictive value of 96.3%, and negative predictive value of 99.3% in validation set. Meanwhile, the location recognition model demonstrated an top-5 accuracy of 87.1%. **Conclusion** The artificial intelligence assisted diagnosis model based on deep learning for real-time dynamic recognition of gastric lesions and their locations under gastroscopy has good ability in lesion detection and location recognition, and has great clinical application prospects.

Keywords: artificial intelligence; gastric diseases; early gastric cancer; deep learning; gastroscopy examination

胃部疾病在我国为常见病、多发病,其中包括慢性胃炎、胃部息肉、胃癌等[1]。2022年我国胃癌新发病例约35.87万[2]。早期胃癌无临床特异性症状,其诊断依赖于完整的胃镜检查及镜下病理活检[3],但早期胃癌多为浅表凹陷或小而平的病变,早期诊治难,发现时多为进展期,手术效果差。据报道,早期胃癌5年生存率可达90%,而进展期胃癌5年生存率仅约20%[4]。由于目前我国胃镜医师数量相对较少,胃镜技术水平参差不齐,部分医生对镜下病变特征认识相对不足,且每日胃镜检查负荷大,容易产生视觉疲劳,导致胃

[基金项目] 重庆市技术创新与应用发展专项重点项目(cstc2019jscx-gksbX0038)

[通信作者] 柏健鹰,E-mail:Drbaijy2015@163.com

镜盲点率、疾病漏诊率较高[5-7]。

随着人工智能高速发展,深度学习、深卷积神经网络在图片识别方面显现出其独特优势[8-11]。深卷积神经网络在心电图识别及糖尿病视网膜病变分级上的应用取得了较好的成绩[12-14]。2015年Miyaki等[15]利用早期胃癌的BLI激光图片训练了一款计算机辅助诊断系统,其在BLI模式下对胃部癌性病变有较好的识别能力。Wu等[16]开发的人工智能辅助诊断系统对早期胃癌识别的准确率为92.5%,敏感度为94.0%,特异度为91.0%;黄丽等[17]研发的基于Keras的VGG-16模型在静态图片上对胃黏膜良恶性溃疡鉴别的精确度达85%。

既往人工智能与胃镜的结合多停留在对静态图片、二分类病变(癌与非癌、良性溃疡与恶性溃疡)的识别,缺乏可以在实时检查中应用的、能够对多病种进行识别的研究,与真实应用场景仍有差距,限制了其在临床上的应用。故本研究拟构建一个基于现实环境的胃镜下实时动态识别多种胃部病变及位置的辅助诊断模型,并评估其对胃部病变检出和位置识别的能力。

1 资料与方法

1.1 临床资料

回顾性分析陆军军医大学第二附属医院消化内镜中心2021年8月至2022年10月收治的104例患者的胃镜检查视频,根据不同病变类型、同一类型病变出现于不同位置及同一位置病变的不同视角对视频进行剪辑,共剪辑成128个视频片段(包括早期胃癌29处、进展期胃癌3处、黏膜下隆起性病变23处、慢性胃炎66处、慢性胃炎伴肠化10处、胃部息肉18处)。被收集视频的胃镜操作者均为副主任医师职称及以上高年资内镜医师,内镜设备为Olympus CV-290、Fuji VP-700hd、Fuji 4450hd,均为主流内镜设备。

纳入标准:①图像清晰,病变显示明显,表面无黏液或附着物;②多角度采集拍摄,尽量使病变位于视野中央,包括远景图(显示病变的位置、形态、颜色等)、中景图(显示病变的大小、边缘、表面隆起或凹陷情况、是否糜烂和出血等)、近景图(显示表面黏膜状态等);③同一病变多张图像(不同角度、染色方法)的充气量、黏膜延展度、病变大小基本保持一致;④图像亮度适中。排除标准:①图像模糊、反光明显;②胃黏膜未充分展开的图像;③手术器械对病变有遮挡。

诊断标准:胃部息肉、黏膜下隆起性病变、早期胃癌、进展期胃癌以病理报告为金标准,慢性胃炎、慢性胃炎伴肠化以胃镜报告为金标准。所有病理、胃镜报告均由高年资副主任医师以上职称医生审核。

本研究是一项回顾性单中心诊断性研究,已获得陆军军医大学第二附属医院伦理委员会批准(伦理批号:2023-研第167-01)。

1.2 视频预处理及标注

预处理:将剪辑后的视频片段逐帧进行分割、放缩处理,以保证处理后的视频片段仅包含胃镜检测区域,同时通过几何变换以及颜色变换进行数据增强。

标注:依托医院内部局域网构建CVAT标注平台,将经过预处理的视频上传至标注平台,对视频帧进行位置类别及病变类别标注,其中位置类别分为:食管上段、食管中段、食管下段、贲门(食管侧)、贲门(倒视)、胃底、胃体前壁、胃体后壁、胃体小弯侧、胃体大弯侧、胃窦前壁、胃窦后壁、胃窦大弯侧、胃窦小弯侧、胃角、胃角前壁、胃角后壁、十二指肠球部、十二指肠降部上段、幽门共20个分类;病变类别分为:慢性胃炎、慢性胃炎伴肠化、胃部息肉、黏膜下隆起性病变、早期胃癌、进展期胃癌共6个分类。标注人员将视频帧中出现的主要解剖部位标注于视频空白处(图1),依据病理结果及胃镜检查结果,将病变部分逐帧用矩形包围进行标注(图2),所有标注后的视频由副主任医师以上职称的高年资内镜医师进行审核。

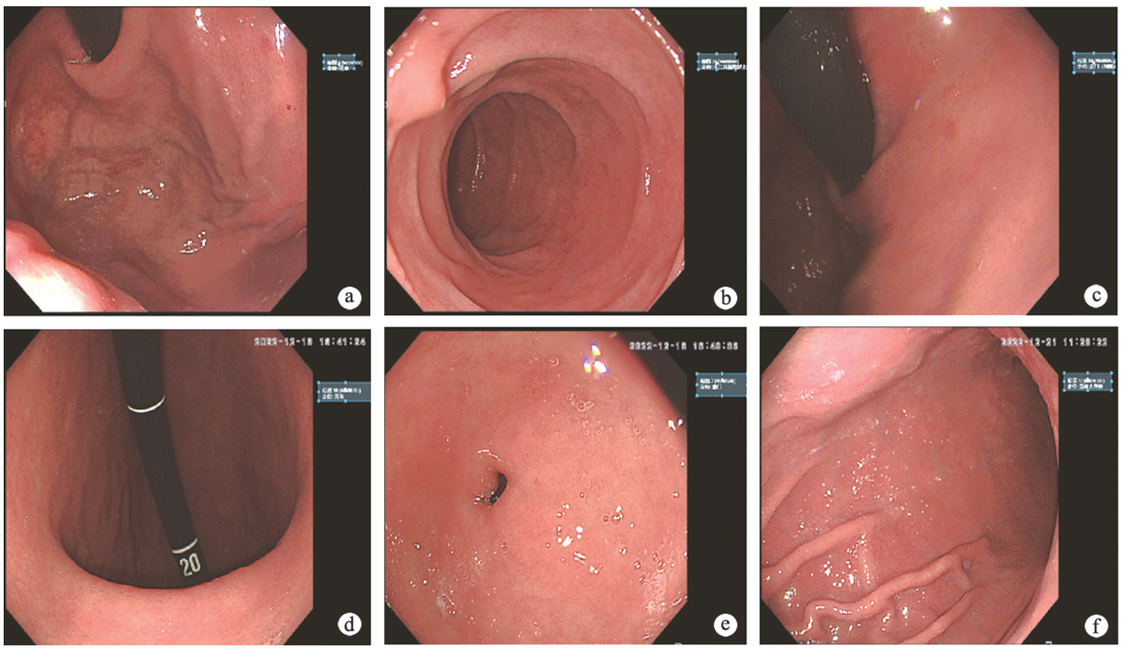
1.3 训练集与验证集的划分

共标注图片68 351帧,其中病变类别图片52 114帧,位置类别图片16 237帧(表1)。为保证训练集与验证集完全独立,将各病变类别图片帧按8:2的比例划分成训练集与验证集(表2);将已标注的各位置类别图片帧,根据患者来源按8:2的比例划分成训练集与验证集(表3)。

1.4 研究方法

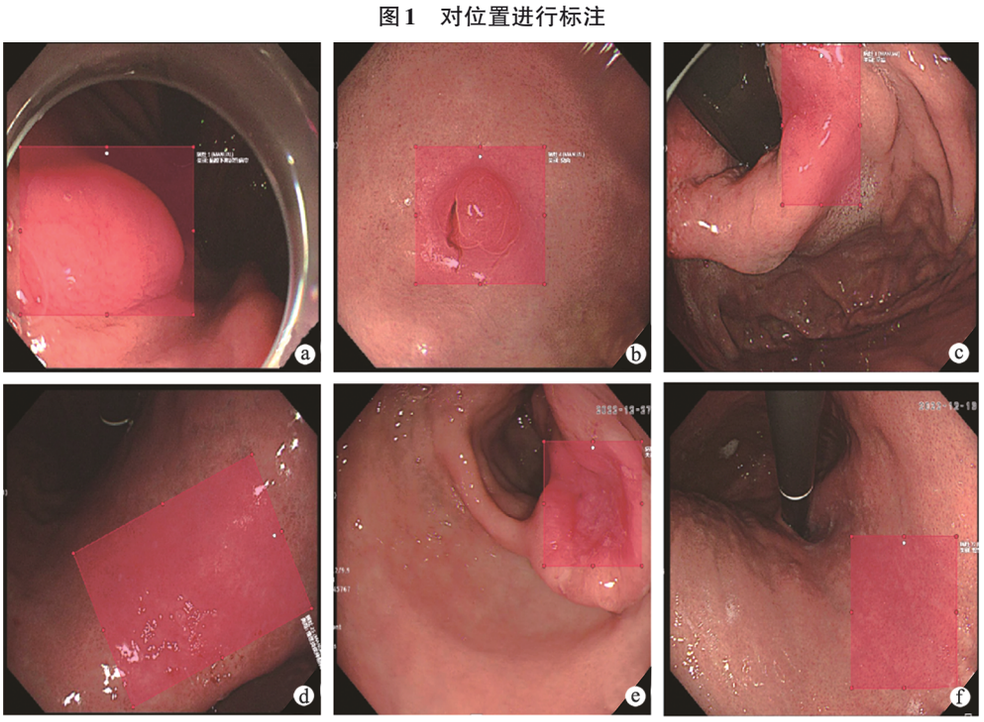
模型的训练:根据病变识别任务和位置识别任务的特点,采用不同的算法进行研究。针对病变识别功能,采用目标检测模型YoloV4对每一帧胃镜图片进行病变的定位与识别,共纳入41 692帧胃镜图像用于YoloV4模型的训练。针对位置识别功能,使用CRNN模型进行视频片段分类,采用ResNet152预训练模型,共纳入13 180帧胃镜图像用于位置识别模型训练。

模型的验证:针对病变识别功能,验证集中的



a:对胃底进行标注;b:对十二指肠降部上段进行标注;c:对贲门(倒视)进行标注;d:对胃角进行标注;e:对幽门进行标注;f:对胃体大弯侧进行标注

图1　对位置进行标注



a:对黏膜下隆起性病变进行标注;b:对胃部息肉进行标注;c:对早期胃癌进行标注;d:对慢性胃炎伴肠化进行标注;e:对进展期胃癌进行标注;f:对慢性胃炎进行标注

图2　对病变进行标注

表1　总标记帧数(帧)

项目	病变类别	位置类别	合计
训练集帧数	41 692	13 180	54 872
验证集帧数	10 422	3 057	13 479
总标注帧数	52 114	16 237	68 351

10 422帧病变类别胃镜图像用于模型的验证。针对位置识别功能,验证集中的3 057帧位置类别胃镜图像用于模型的验证。

1.5　统计学方法

数据采用SPSS 26.0统计学软件进行分析,计数资料以率(%)表示。使用病变识别的准确率、敏感度、

表2 病变类别的训练集与验证集(帧)

项目	早期胃癌	进展期胃癌	黏膜下隆起性病变	慢性胃炎	慢性胃炎伴肠化	胃部息肉	合计
训练集帧数	21 922	660	4 673	8 308	3 545	2 584	41 692
验证集帧数	5 480	165	1 168	2 077	886	646	10 422
总标记帧	27 402	825	5 841	10 385	4 431	3 230	52 114

表3 位置类别的训练集与验证集(帧)

项目	十二指肠球部	十二指肠降部上段	幽门	胃窦前壁	胃窦后壁	胃窦大弯侧	胃窦小弯侧	胃角	胃角前壁	胃角后壁
训练集帧数	667	1 441	558	850	105	447	656	905	399	26
验证集帧数	195	532	138	51	70	37	136	206	108	90
总标记帧	862	1 973	696	901	175	484	792	1 111	507	116

项目	胃体前壁	胃体后壁	胃体小弯侧	胃体大弯侧	胃底	食管下段	食管中段	食管上段	贲门(倒视)	贲门(食管侧)	总计
训练集帧数	589	409	862	659	1 826	390	433	832	152	974	13 180
验证集帧数	158	116	111	156	185	118	168	201	55	226	3 057
总标记帧	747	525	973	815	2011	508	601	1 033	207	1 200	16 237

特异度、阳性预测值、阴性预测值以及位置识别的top5准确率来评价模型对胃部病变检出和位置识别的能力。

2 结果

2.1 病变识别功能在验证集中的表现

病变识别的总体准确率为98.8%,敏感度为96.6%,特异度为99.3%,阳性预测值为96.3%,阴性预测值为99.3%。各病变类别识别的准确率、敏感度、特异度、阳性预测值及阴性预测值见表4。

2.2 位置识别功能在验证集中的表现

位置识别任务模型top1总体准确率为60.1%,top3总体准确率为82.3%,top5总体准确率为87.1%。各位置类别识别的top1、3、5准确率见表5。

表4 模型在验证集中对各类型病变识别的结果(%)

项目	早期胃癌	进展期胃癌	黏膜下隆起性病变	慢性胃炎	慢性胃炎伴肠化	胃部息肉
准确率	98.2	99.5	99.7	97.1	99.6	98.8
敏感度	98.1	98.2	99.4	92.8	96.7	90.6
特异度	98.2	99.5	99.7	98.1	99.9	99.4
阳性预测值	98.4	76.8	97.9	92.5	98.8	90.6
阴性预测值	97.9	100	99.9	98.2	99.7	99.4

表5 模型在验证集中对位置识别的结果(%)

准确率	十二指肠球部	十二指肠降部上段	幽门	胃窦前壁	胃窦后壁	胃窦大弯侧	胃窦小弯侧	胃角	胃角前壁	胃角后壁
top1	65.6	92.5	72.5	11.8	38.6	64.9	74.3	68.4	63.9	3.3
top3	89.7	98.3	89.1	37.3	64.3	89.2	82.4	82.0	90.7	63.3
top5	94.4	98.9	95.7	45.1	78.6	94.6	85.3	89.8	94.4	75.6

准确率	胃体前壁	胃体后壁	胃体小弯侧	胃体大弯侧	胃底	食管下段	食管中段	食管上段	贲门(倒视)	贲门(食管侧)
top1	34.8	32.8	49.5	10.3	88.6	30.5	47.0	59.2	16.4	77.4
top3	53.2	49.1	64.9	42.9	94.6	100	97.0	99.0	34.5	91.6
top5	68.4	59.5	79.3	52.6	95.1	100	99.4	99.5	34.5	92.5

3 讨论

胃镜检查是多种胃部疾病(尤其是早期胃癌)的重要检查手段[18-19]。由于胃是空腔脏器,易变形,多褶皱,加上黏膜表面常有黏液附着、黏膜反光、气泡形成等情况,易对镜下视野形成极大干扰,且我国胃镜医师数量相对较少[5],胃镜技术水平参差不齐[6],而患者胃镜检查需求量大,医师在高负荷情况下易产生视觉疲劳,影响诊断准确性,因此,在多种因素共同作用下,胃部疾病在镜下容易漏诊。

人工智能的高速发展以及其与医疗领域结合的巨大潜力和成果让我们看到了解决这一困境的可能。为了贴近临床应用场景,我们研究开发了这款可以在现实环境下实时对多种胃部病变及位置进行识别的辅助诊断模型。该辅助诊断模型旨在白光胃镜下实时辅助医师对包括早癌在内的多种胃部疾病进行诊断,亦可以加深初级胃镜医师对镜下病变的学习。模型主要包含病变识别和位置识别两个功能,当实时视频流输入到模型后,若发现可疑病变,将以不同颜色矩形框对病变进行标记,以提示医师在胃镜检查中高度关注该病灶,及时予以进一步观察,必要时进行染色及活检等进一步操作。在检查过程中,该辅助诊断

模型会自动记录可疑病变的图片帧及诊断,后台自动记录已检查部位,并实时显示所有未检查部位。

黄丽等[17]研发的基于Keras的VGG-16模型,可在静态图片上对胃黏膜良恶性溃疡进行鉴别,准确率达85%;Xie等[20]改进注意特征融合的胃息肉检测模块,使其可以在更加复杂的胃部背景下识别胃息肉,将CSP-Darknet53网络嵌入该注意特征融合模块后,其在数据集中对胃息肉检出的平均精密度、F1得分、精密度和召回率分别达92.33%、0.89、95.92%和83.63%。既往研究多针对静态图片进行分类鉴别,无法在实时动态下进行分类鉴别,与真实应用场景仍有一定距离,而我们研发的辅助诊断模型可在胃镜检查过程中以实时视频流输入的方式进行实时辅助诊断,对病变进行识别,更符合真实胃镜检查场景需要,有较好的临床应用前景。

赵曲川等[21]基于U-Net构建了慢性萎缩性胃炎诊断模型,其在验证集中对慢性萎缩性胃炎识别的敏感度、特异度、准确率分别达92.73%、92.24%、92.63%;Hu等[22]开发了一款辅助诊断模型,可以对NBI图像进行早期胃癌的识别,该模型在内部验证集中的*AUC*为0.808。既往人工智能与胃镜的研究多为二分类疾病的鉴别,而往往忽略了如慢性胃炎、慢性胃炎伴肠化、胃部息肉、黏膜下隆起性病变等均有恶变为胃癌的可能[23]。本研究研发的辅助诊断模型可对慢性胃炎、慢性胃炎伴肠化、胃部息肉、黏膜下隆起性病变、早期胃癌、进展期胃癌共6种病变进行识别,覆盖了临床上更多的疾病种类,病变识别总体准确率达98.8%,敏感度为96.6%,特异度为99.3%,对各病变种类识别的准确率、敏感度、特异度均在90%以上,尤其对早期胃癌识别的准确率达到了98.2%,敏感度为98.1%,特异度为98.2%,阳性预测值为98.4%,阴性预测值为97.9%,该辅助诊断模型有助于胃镜医生更好地检出早期胃癌及多种常见胃部病变并及时进行相应治疗,从而降低胃癌病死率,在实际应用中可以覆盖到更多的患者,有更广阔的临床应用前景。

张希钢等[24]基于Transformer模型构建了胃镜视频部位实时识别的模型,其在测试集中的准确率为84.3%;Wu等[16]开发的人工智能辅助诊断系统在10分类位置识别的准确率为90.0%,在26分类位置识别的准确率为65.9%。本研究研发的辅助诊断模型可实时对胃内20个解剖位置进行识别,在20分类位置识别的top5总体准确率为87.1%,与其他研究结果接近,可见该辅助诊断系统在实时胃镜检查中有较好的位置识别能力。

本研究中,该辅助诊断系统对多种胃部疾病均有较高的诊断效能,但位置top1识别率偏低,主要原因在于同一视野中存在多个解剖位置,而我们只标注了主要的解剖位置,这导致模型单一位置输出结果常与标注结果不一致。将位置识别由单结果输出调整为多结果输出后,位置识别率显著提高。在接下来的研究中,我们拟对胃镜图片按解剖部位进行切割,针对性训练,可能会取得更好的检验效果。

综上,基于深度学习的人工智能实时动态辅助诊断系统在胃镜检查中对胃部疾病有较好的实时检出能力。但本研究为回顾性研究,尚未进行前瞻性的临床验证。同时,本研究中所有的数据均来自陆军军医大学第二附属医院消化内镜中心,为单中心研究,尚未在其他医院进行推广应用,模型的训练过程中不可避免地存在一定的过拟合。在下一步的研究中,我们将在周边医院进行推广,纳入多中心数据进一步训练验证,并在临床上前瞻性应用。随着多中心数据的纳入、模型的进一步升级优化、硬件算力的提高,该系统的诊断效能会进一步提高。

[参考文献]

[1] 何敏均,季璐玓,连 力,等. 2000-2019年中国早发型胃癌和晚发型胃癌流行病学趋势分析[J]. 中华流行病学杂志,2023,44(8):1198-1202. doi:10.3760/cma.j.cn112338-20230302-00117.

[2] 郑荣寿,陈 茹,韩冰峰,等. 2022年中国恶性肿瘤流行情况分析[J]. 中华肿瘤杂志,2024,46(3):221-231. doi:10.3760/cma.j.cn112152-20240119-00035.

[3] 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.

[4] 张梦娇,徐 铭,吴练练,等. 内镜人工智能诊断辅助系统对胃局灶性病变检出的应用(含视频)[J]. 中华消化内镜杂志,2023,40(5):372-378. doi:10.3760/cma.j.cn321463-20221022-00575.

[5] 王洛伟,马旭东,李兆申. 加强消化内镜诊疗质控 推动消化内镜高质量发展[J]. 中华消化内镜杂志,2022,39(11):857-859. doi:10.3760/cma.j.cn321463-20220411-00043.

[6] 高 野,冯拥璞,刘 雨,等. 新时代消化内镜技术培训和内镜医师培养[J]. 中华消化内镜杂志,2020,37(1):3-10. doi:10.3760/cma.j.issn.1007-5232.2020.01.002.

[7] 徐 铭,姚理文,于红刚. 智能消化内镜质控系统在胃镜检查中的应用价值研究[J]. 中华消化内镜杂志,2022,39(2):133-138. doi:10.3760/cma.j.cn321463-20200426-00358.

[8] LeCun Y, Bengio Y, Hinton G. Deep learning[J]. Nature, 2015, 521(7553):436-444. doi:10.1038/nature14539.

[9] 王 瑜,吴 江,伍东升. 卷积神经网络在职业性尘肺病影像学诊断中的应用研究进展[J]. 生物医学工程学杂志,2024,41(2):413-420. doi:10.7507/1001-5515.202309079.

[10] Khozin S. From organs to algorithms: redefining cancer classification in the age of artificial intelligence[J]. Clin Transl Sci, 2024, 17(9):e70001. doi:10.1111/cts.70001.