

·论著·

深度学习技术在辅助结直肠腺瘤浸润深度鉴别中的应用研究

许祐铭 姚理文 卢姿桦 于红刚

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

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

【摘要】目的 研究深度学习技术在肠镜电子染色模式下辅助内镜医师鉴别结直肠腺瘤浸润深度的价值。**方法** 回顾性收集2016年11月—2021年6月武汉大学人民医院、南方医科大学深圳医院和宜昌市第一人民医院的3 714个病例的13 246张电子染色图像,构建深度学习模型,对结直肠腺瘤的黏膜下深层浸润和非深层浸润进行鉴别,并在独立测试集与外部测试集中验证模型的性能。使用完整的测试集对比5名内镜医师与深度学习模型的诊断水平。前瞻性收集2021年1—6月来自武汉大学人民医院的35个高清内镜视频,验证在模型辅助下内镜医师的诊断效果。**结果** 该模型在图片测试集中的准确率93.08%(821/882),约登指数0.86,优于内镜医师[最高者分别为91.72%(809/882)和0.78]。在视频中该模型的准确率达97.14%(34/35),约登指数0.94。在模型辅助下,内镜医师的准确率显著提升[最高者97.14%(34/35)]。**结论** 本研究开发的基于深度学习的结直肠腺瘤浸润深度鉴别系统能够准确地识别黏膜下深层浸润病灶,辅助内镜医师提升识别深层浸润病灶的准确率。

【关键词】 结直肠肿瘤; 机器学习; 人工智能; 肿瘤浸润; 诊断, 计算机辅助

基金项目: 国家自然科学基金(81672387); 湖北省重大科技创新项目(2018-916-000-008); 湖北省消化疾病微创诊疗医学临床研究中心项目(2018BCC337)

Application of deep learning to the differentiation of the invasion depth in colorectal adenomas

Xu Youming, Yao Liven, Lu Zihua, 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 evaluate deep learning for differentiating invasion depth of colorectal adenomas under image enhanced endoscopy (IEE). **Methods** A total of 13 246 IEE images from 3 714 lesions acquired from November 2016 to June 2021 were retrospectively collected in Renmin Hospital of Wuhan University, Shenzhen Hospital of Southern Medical University and the First Hospital of Yichang to construct a deep learning model to differentiate submucosal deep invasion and non-submucosal deep invasion lesions of colorectal adenomas. The performance of the deep learning model was validated in an independent test and an external test. The full test was used to compare the diagnostic performance between 5 endoscopists and the deep learning model. A total of 35 videos were collected from January to June 2021 in Renmin Hospital of Wuhan University to validate the diagnostic performance of the endoscopists with the assistance of deep learning model. **Results** The accuracy and Youden index of the deep learning model in image test set were 93.08% (821/882) and 0.86, which were better than those of endoscopists [the highest were 91.72% (809/882) and 0.78]. In video test set, the accuracy and Youden index of the model were 97.14% (34/35) and 0.94. With the assistance of the model, the accuracy of endoscopists was significantly improved [the highest was 97.14% (34/35)]. **Conclusion** The deep learning model obtained in this study

DOI: 10.3760/cma.j.cn321463-20221130-00751

收稿日期 2022-11-30 本文编辑 周昊

引用本文: 许祐铭, 姚理文, 卢姿桦, 等. 深度学习技术在辅助结直肠腺瘤浸润深度鉴别中的应用研究[J].

中华消化内镜杂志, 2023, 40(7): 534-538. DOI: 10.3760/cma.j.cn321463-20221130-00751.



could identify submucosal lesions with deep invasion accurately for colorectal adenomas, and could improve the diagnostic accuracy of endoscopists.

【Key words】 Colorectal neoplasms; Machine learning; Artificial intelligence; Tumor infiltrating; Diagnosis, computer-assisted

Fund program: National Natural Science Foundation of China (81672387); Major Science and Technology Innovation Project of Hubei Province (2018-916-000-008); Project of Hubei Provincial Clinical Research Center for Digestive Disease Minimally Invasive Incision (2018BCC337)

2018 年国家癌症中心数据显示,结直肠癌位居恶性肿瘤发病谱第 3 位,死亡谱第 2 位^[1]。由于饮食结构逐渐改变和肠镜检查的普及,近几年我国结直肠癌的发病率呈上升趋势,同时死亡率亦逐步升高,多数患者在确诊时属于中晚期^[2-3]。大部分结直肠癌由腺瘤性息肉演变而来,当怀疑腺瘤性息肉存在癌变时,需要预测癌灶的浸润深度从而选择合适的治疗方式,因此进行准确的光学诊断预测浸润深度至关重要^[4-5]。尽管目前提出了多种结直肠肿瘤内镜下分型,建立了标准化的临床路径,但复杂的理论与漫长的学习周期导致不同内镜医师图像解释的主观差异^[5]。一项前瞻性的大规模临床试验显示,目前 T1 期结直肠癌的光学诊断准确性仍然有限,尤其是发现深层浸润病灶的灵敏度^[6]。

近年来,越来越多消化内镜领域的人工智能研究探索了人工智能赋能医疗的可能性,并在质控、辅助诊断、决策支持等方面改善了胃肠镜的日常实践^[7-10]。在结肠镜息肉检测方面,人工智能已经作为辅助息肉检测的一种方法尝试性地应用于临床,在提升腺瘤检出率方面展现出了良好的效果^[11-14]。在病灶的定性诊断方面,许多研究表明深度学习技术能够帮助医师鉴别增生性息肉与腺瘤性息肉,从而避免不必要的切除^[15-16]。以上研究结果反映出深度学习技术在改进医师检测、认知能力,提升结肠镜质量方面的潜力。若人工智能可以实时辅助内镜医师判断浸润深度,就能进一步观察或活检,发现更多的深层浸润病灶。本研究中,我们构建了一个基于深度学习的结直肠腺瘤浸润深度鉴别系统,辅助内镜医师在内镜电子放大染色成像模式下对结直肠腺瘤的浸润深度进行判断。

资料与方法

一、模型的训练与测试

回顾性收集 2016 年 11 月至 2021 年 6 月于武汉大学人民医院、南方医科大学深圳医院和宜昌市第

一人民医院进行肠镜检查或治疗患者的电子染色图像,所有图片在电子染色模式下由内镜医师选取合适的放大倍数后拍摄,使用日本奥林巴斯公司 260 和 290 系列以及日本富士公司 EC-L590WM 系列内镜。病例入选标准:年龄>18 岁,内镜诊断包含结直肠息肉并且具有病理结果支持。剔除成像质量不好、肠道清洁度不佳、放大倍数不合理导致看不清表面纹理、锯齿状息肉或腺瘤图像。

最终收集病例 3 714 个、图像 13 426 张用于模型的构建与验证,即将数据集分为训练集和测试集。为加强模型性能鲁棒性,南方医科大学深圳医院与宜昌第一医院的图片纳入测试集作为外部验证,训练集图片全部来源于武汉大学人民医院,并且测试集和训练集的图片来源完全不同。训练集和测试集的图像数量比例为 4:1。测试集的图像由一位具有 10 年以上内镜工作经验、肠镜检查超过 3 000 例的专家出具内镜报告并进行病理结果复核。

所有图像以病理结果作为金标准。组织学标本在各中心病理科按照 WHO 标准评估。病灶按照浸润深度被分为 2 组:黏膜下深层浸润癌及以上的结直肠癌组(简称深层浸润癌组),包含浸润深度>SM2 的所有结直肠癌病灶;非黏膜下深层浸润癌组(简称非深层浸润癌组),包含增生性息肉,管状/绒毛状/管状-绒毛状混合腺瘤,低级别上皮内瘤变、高级别上皮内瘤变、黏膜内癌及黏膜下浅层浸润癌(SM1)。当同一病灶包含多种病理性质时,以恶变程度最高的为准。本研究中取得“非深层浸润癌”数据集 7 355 张和“深层浸润癌”数据集 6 071 张,各数据集的图片分布情况见表 1,典型图像见图 1。

结直肠癌浸润深度实时辅助诊断二分类模型

表 1 训练集、测试集的图片分布情况(张)

数据集	合计	非深层浸润癌	深层浸润癌
训练集	10 777	5 494	5 283
内部测试集	1 348	834	514
外部测试集	1 301	1 027	274

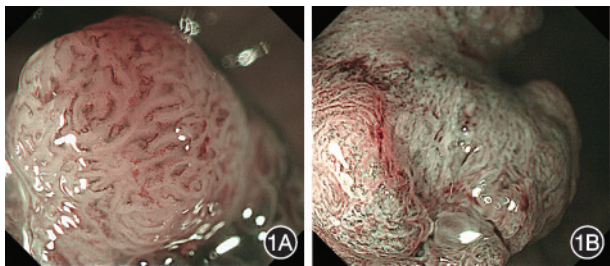


图1 结直肠肿瘤内镜下电子染色典型图像 1A:非深层浸润癌;1B:深层浸润癌

采用成熟的深层卷积神经网络模型 Resnet50 算法构建,能够随着内镜医师移动内镜镜头,实时判断内镜电子染色后的病灶是侵袭性或非侵袭性。Resnet50 算法解决了深层神经网络随着深度增加性能却明显降低的退化问题,极大地拓展了神经网络的深度,能够更有效地提取图像特征。图像预处理采用自有算法,裁切掉图片不必要的边框以减少图片噪声,降低图中不必要的信息对训练结果精度的影响。对图片进行旋转、平移、缩放、随机裁切等操作进行数据增强。输入图片的大小为 224 像素×224 像素,采用基于 Tensor Flow 1.12.2 的 Keras 深度学习框架对模型进行训练和调优。同时使用 Dropout、Early stopping 等方法避免过拟合。采用迁移学习进行预训练。

二、图片测试集人机对比实验

模型构建完成后,在武汉大学人民医院消化内镜中心邀请 2 名高年资内镜医师(肠镜经验超过 5 年)与 3 名内镜初学者(肠镜经验未满 2 年)独立对测试集进行评估,并与模型的性能进行对比。在人机对比中,以病灶为单位统计诊断效果,每个病灶选取 3~5 张高清图像交给内镜医师判断浸润深度。在剔除患者姓名等无关诊断信息的基础上,为贴合临床诊断逻辑,将测试集图像与对应病灶的白光图像一同交与医师判断。

三、模型辅助效果验证

1. 建立视频数据集:与图片测试集的病例不同,另收集 2021 年 1—6 月来自武汉大学人民医院的 35 例患者的高清内镜视频,其中深层浸润及以上 18 例,包括黏膜下深层浸润癌 10 例,进展期结肠癌 8 例;非深层浸润癌 17 例。视频病例的纳入标准与图片一致,也以病理结果作为金标准。每个病灶的视频均由同一名博士生剪辑,包含电子染色模式下观察整体形态的远观片段和观察表面结构的近观片段,在剪辑时隐去患者信息。剪辑后的视频

与对应患者的病理资料一起交给专家审核。人工智能模型仅对近观片段进行判断,远观片段作为附加资料提供给参与测试的内镜医师进行辅助判断。受试者在测试开始后才能看到视频,无法提前观看视频。

2. 深度学习模型辅助下医师的诊断表现测试:参与人机对比实验的 2 名高年资内镜医师与 3 名内镜初学者参与了本次测试,参与者均有染色放大内镜独立操作经验。参与医师在同一时间被邀请至同一教室中,在规定时间内(2 h)内用统一的电脑播放测试视频。测试视频中,原内镜图像的边框会替换为深度学习模型的诊断结果,以便隐去患者信息与提供实时提示效果。医师判断每个病例是否为深层浸润癌后,还会在备注中写下对结果的判断依据与置信度(很有把握、比较有把握、不确定)。将内镜医师的诊断准确度与静态测试集进行对比。

本研究通过武汉大学人民医院伦理委员会审查(批件号 WDRY2019-K091)。

四、数据分析

采用 SPSS 26.0 统计软件处理数据,计算准确率、灵敏度、特异度与约登指数等指标,分别评估深度学习模型和内镜医师的诊断能力,并利用视频模拟临床环境,观察在深度学习模型辅助下内镜医师诊断能力的变化。

结 果

一、模型及内镜医师在图片测试集中的表现

在全体测试集的 882 个病例中,模型判断浸润深度的准确率、灵敏度、特异度、约登指数分别为 93.08% (821/882)、93.20% (192/206)、93.05% (629/676)、0.86。在外部验证的 361 个病例中,该模型的准确率为 89.20% (322/361)、灵敏度为 91.40% (85/93)、特异度为 88.43% (237/268)、约登指数为 0.80。该模型在内、外部验证中均表现出良好的性能。2 名高年资内镜医师的准确率、灵敏度、特异度、约登指数分别为 91.72% (809/882)、83.98% (173/206)、94.08% (636/676)、0.78 和 89.12% (786/882)、78.64% (162/206)、92.31% (624/676)、0.71。3 名低年资内镜医师的平均准确率、平均灵敏度、平均特异度、约登指数为 83.22%、66.50%、92.01%、0.59。在图片测试集中,模型的表现优于高年资内镜医师与低年资内镜医师。

二、视频验证结果

模型在视频验证集中的准确率达到 97.14%。2 名高年资医师在模型的辅助下,准确率提升至 97.14% 和 94.29%;3 名低年资医师在模型辅助下,准确率分别为 94.29%, 91.43%, 85.57%。深度学习模型辅助内镜医师,准确率和灵敏度有明显提升,特异度无明显变化,详见表 2。

讨 论

本研究利用 10 777 张内镜放大染色模式下结直肠息肉的窄带光成像图像,构建了一个基于深度学习技术的内镜放大染色模式下的结直肠癌辅助诊断系统。并在图像、视频两种类型的数据中验证了系统判断是否存在黏膜下深层浸润的能力。内镜医师发现息肉后,切换电子染色模式即可触发系统,系统将在内镜医师利用放大电子染色观察病灶表面腺管和微血管结构时,实时给出模型判断的病灶性质,帮助内镜医师发现黏膜下深层浸润的区域。在图片测试集中,模型的表现优于高年资内镜医师,并且在外部测试集中表现出了良好的鲁棒性。在视频验证集中,模型的灵敏度达到 100%。充分验证表明,在临床环境中,深度学习模型能够稳定可靠地给出病灶是否存在黏膜下深层浸润的判断,并且在模型的辅助下,内镜医师的诊断能力显著上升。

早期结直肠癌的患者预后情况明显优于晚期结直肠癌,所以尽早筛查、诊断早期结直肠癌对延长患者生存期有着至关重要的意义^[4,17-18]。随着微创手术的发展,对结直肠癌患者的根治除经典外科手术外,内镜手术因其出血少、并发症少等一系列优点,应用范围也越来越广^[19]。若病灶发生了深层次的癌变,则应由外科切除避免复发。结直肠癌的预后与早期筛查、诊断、治疗有着密切的关系,早期结直肠癌因淋巴结转移率低而预后良好,内镜手术

治疗后五年生存率可超过 90%;而中晚期结直肠癌需外科手术干预,五年生存率仅为 10%。研究证实,病变浸润深度是淋巴结转移的重要危险因素,因此评估浸润深度对治疗方案的选择起着决定性作用^[17-18]。合理选择治疗方案可以避免患者遭受二次手术的痛苦,提高患者的生存质量,同时有助于医疗资源的合理分配^[5]。

近年来,人工智能在包括消化内镜在内的各个医学领域取得了巨大的进展。目前的研究已初步证实人工智能在结直肠癌浸润深度诊断中的可行性,但现有的研究仍有不足。Tokunaga 等^[20]开发了一个使用白光图像预测浸润深度的人工智能模型,但白光下肠道黏膜色泽单一,腺管开口及微血管无法清晰辨别,该系统仍需结合放大染色等高级成像提供的腺管与微血管结构进行诊断。Luo 等^[21]开发了一个基于深度学习算法的模型来预测深部黏膜下浸润性癌。然而,内镜图像中可能存在深层浸润的区域需要由经验丰富的内镜专家手动标记,这与临床实践中减轻医师负担的需求不符。医师通过放大染色肠镜来回、反复地观察病灶,通过诊断分型理论寻找息肉的表面异型结构,这个过程具有实时性的特点。但目前的研究仅限于人工提取特征,能够在视频中应用、实时辅助内镜医师观察某一视角下的腺管结构的研究仍然缺乏。本研究使用了更接近实际临床环境的视频对模型的性能进行验证,视频相比于图像包含了更多的息肉角度和更复杂的诊断策略,能够验证系统的实时动态识别能力。并且在视频验证中,模型达到了 97.14% 的准确率,在该模型的辅助下,内镜医师的诊断准确率和灵敏度均有显著提高,反应灵敏度与特异度的综合指标约登指数也有相应提高。

尽管近年来我国微创手术蓬勃发展,但受限于高昂的培养成本与漫长的培养周期,我国内镜医师的数量还远远不能满足我国庞大的人口基数的需求,且不同地区的医疗资源分布不均,部分内镜中

表 2 视频测试集中深度学习模型和内镜医师诊断黏膜下深层浸润病变的表现

判断者	准确率(%)	灵敏度(%)	特异度(%)	约登指数
深度学习模型	97.14(34/35)	100.00(18/18)	94.12(16/17)	0.94
高年资内镜医师 1	97.14(34/35)	100.00(18/18)	94.12(16/17)	0.94
高年资内镜医师 2	94.29(33/35)	94.44(17/18)	94.12(16/17)	0.89
低年资内镜医师 1	94.29(33/35)	88.89(16/18)	100.00(17/17)	0.89
低年资内镜医师 2	91.43(32/35)	94.44(17/18)	88.24(15/17)	0.83
低年资内镜医师 3	85.57(31/35)	88.89(16/18)	88.24(15/17)	0.77

心仅能进行常规的内镜检查。目前结直肠主流的分型有 pit pattern 分型、JNET 分型,均需要对表面腺管结构与微血管进行长时间的仔细观察,这对操作肠镜的稳定性和镜头的细微移动提出了很高的要求。因此,浸润深度的判断需要内镜医师具有丰富的经验以及扎实的理论基础,不同内镜医师之间的主观差异较大。在本研究中,利用大量数据构建的人工智能模型鉴别浸润深度的准确率优于内镜医师,有助于提升内镜质量与诊断准确度,有望弥补不同水平医师之间诊断质量的差距。

本研究存在的局限性。首先,尽管本研究在多中心的测试集上进行了验证,但训练集均收集自武汉大学人民医院,若能排除入院偏倚因素,收集多家中心的数据将有助于模型获取更丰富的图像特征,从而提升模型的泛用性,有利于在不同的环境中表现出稳定的结果。其次,本研究仅纳入日本奥林巴斯与日本富士这两家内镜厂商的内镜图片,当模型应用于其他厂商的内镜图像时性能可能会下降。

综上,本研究利用多中心、大样本的图像构建并验证了深度学习模型,且采用前瞻性视频验证的方式检验了该模型在实时环境中的性能,该模型表现整体优于高年资与低年资内镜医师,并能够显著提升医师发现深层浸润病灶的灵敏度。

利益冲突 所有作者声明不存在利益冲突

作者贡献声明 许佑铭:实施研究、分析数据、撰写论文;姚理文:采集数据、统计分析;卢姿桦:采集数据;于红刚:研究指导、论文修改、经费支持

参 考 文 献

- [1] 吴春晓,顾凯,龚杨明,等. 2015 年中国结直肠癌发病和死亡情况分析[J]. 中国癌症杂志, 2020, 30(4):241-245. DOI: 10.19401/j.cnki.1007-3639.2020.04.001.
- [2] 孙可欣,郑荣寿,张思维,等. 2015 年中国分地区恶性肿瘤发病和死亡分析[J]. 中国肿瘤, 2019, 28(1):1-11. DOI: 10.11735/j.issn.1004-0242.2019.01.A001.
- [3] 王维琼. 2016 年中国恶性肿瘤发病和死亡分析[J]. 临床医药文献电子杂志, 2017, 4(19): 3604. DOI: 10.3877/j.issn.2095-8242.2017.19.030.
- [4] 崔子璐,李鹏. 消化内镜技术在判断早期结肠癌浸润深度中的应用[J]. 临床消化病杂志, 2020, 32(1):1-4. DOI: 10.3870/lxh.j.issn.1005-541X.2020.01.001.
- [5] 中华人民共和国国家卫生健康委员会. 中国结直肠癌诊疗规范(2020 版)[J]. 中华消化外科杂志, 2020, 19(6):563-588. DOI: 10.3760/cma.j.cn115610-20200504-00348.
- [6] Backes Y, Schwartz MP, Ter Borg F, et al. Multicentre prospective evaluation of real-time optical diagnosis of T1 colorectal cancer in large non-pedunculated colorectal polyps using narrow band imaging (the OPTICAL study)[J]. Gut, 2019, 68(2):271-279. DOI: 10.1136/gutjnl-2017-314723.
- [7] Le Berre C, Sandborn WJ, Aridhi S, et al. Application of artificial intelligence to gastroenterology and hepatology[J]. Gastroenterology, 2020, 158(1): 76-94. e2. DOI: 10.1053/j.gastro.2019.08.058.
- [8] Min JK, Kwak MS, Cha JM. Overview of deep learning in gastrointestinal endoscopy[J]. Gut Liver, 2019, 13(4):388-393. DOI: 10.5009/gnl18384.
- [9] Sánchez-Peralta LF, Bote-Curiel L, Picón A, et al. Deep learning to find colorectal polyps in colonoscopy: a systematic literature review[J]. Artif Intell Med, 2020, 108:101923. DOI: 10.1016/j.artmed.2020.101923.
- [10] Tanwar S, Vijayalakshmi S, Sabharwal M, et al. Detection and classification of colorectal polyp using deep learning[J]. Biomed Res Int, 2022, 2022:2805607. DOI: 10.1155/2022/2805607.
- [11] Hassan C, Spadaccini M, Iannone A, et al. Performance of artificial intelligence in colonoscopy for adenoma and polyp detection: a systematic review and meta-analysis[J]. Gastrointest Endosc, 2021, 93(1): 77-85. e6. DOI: 10.1016/j.gie.2020.06.059.
- [12] Tamai N, Saito Y, Sakamoto T, et al. Effectiveness of computer-aided diagnosis of colorectal lesions using novel software for magnifying narrow-band imaging: a pilot study[J]. Endosc Int Open, 2017, 5(8): E690-E694. DOI: 10.1055/s-0043-105490.
- [13] Repici A, Badalamenti M, Maselli R, et al. Efficacy of real-time computer-aided detection of colorectal neoplasia in a randomized trial[J]. Gastroenterology, 2020, 159(2): 512-520. e7. DOI: 10.1053/j.gastro.2020.04.062.
- [14] 李素琴,吴练练,宫德馨,等. 基于 YOLO 算法和 ResNet 深度卷积神经网络的结直肠息肉肉检测(含视频). 中华消化内镜杂志, 2020, 37(8): 584-590. DOI: 10.3760/cma.j.cn321463-20200415-00258.
- [15] Yoda Y, Ikematsu H, Matsuda T, et al. A large-scale multicenter study of long-term outcomes after endoscopic resection for submucosal invasive colorectal cancer[J]. Endoscopy, 2013, 45(9): 718-724. DOI: 10.1055/s-0033-1344234.
- [16] 宫德馨,张军,周巍,等. 深度学习技术在提升结直肠息肉性质鉴别准确率中的应用. 中华消化内镜杂志,2021,38(10): 801-805. DOI:10.3760/cma.j.cn321463-20210109-00018.
- [17] 黄小杏,黄世勇,谢伟. 150 例结直肠癌浸润深度与淋巴结转移情况的临床病理研究[J]. 中国当代医药, 2013, 20(33): 4-6,10.
- [18] 李红,胡林莉,贾环. 浸润深度与结直肠癌患者临床特征及预后的关系[J]. 癌症进展, 2021, 19(3): 290-293. DOI: 10.11877/j.issn.1672-1535.2021.19.03.20.
- [19] 吴涵,余志龙,黄陈. 人工智能技术在结直肠癌中的应用与展望[J]. 中华结直肠疾病电子杂志, 2019, 8(2):115-119. DOI: 10.3877/cma.j.issn.2095-3224.2019.02.002.
- [20] Tokunaga M, Matsumura T, Nankinzan R, et al. Computer-aided diagnosis system using only white-light endoscopy for the prediction of invasion depth in colorectal cancer[J]. Gastrointest Endosc, 2021, 93(3): 647-653. DOI: 10.1016/j.gie.2020.07.053.
- [21] Luo X, Wang J, Han Z, et al. Artificial intelligence-enhanced white-light colonoscopy with attention guidance predicts colorectal cancer invasion depth[J]. Gastrointest Endosc, 2021, 94(3):627-638.e1. DOI: 10.1016/j.gie.2021.03.936.