·论著·

基于深度学习的良恶性胃溃疡人工智能辅助诊断系统研究

黄丽[1,2]　李艳霞[1,2]　吴练练[1,2]　胡珊[3]　陈奕云[4]　张军[1,2]　安萍[1,2]　于红刚[1,2]
[1]武汉大学人民医院消化内科 430060;[2]消化疾病湖北省重点实验室,武汉 430060;
[3]武汉楚精灵医疗科技有限公司 430000;[4]武汉大学资源与环境学院 430079
通信作者:于红刚,Email: yuhonggang1968@163.com

【摘要】　目的　构建人工智能辅助诊断系统,自动发现胃溃疡病灶,鉴别胃良性溃疡与恶性溃疡。**方法**　收集武汉大学人民医院消化内镜中心 2016 年 11 月—2019 年 4 月拍摄的胃镜图片 1 885 张,其中正常胃黏膜图片 636 张、良性胃溃疡图片 630 张、恶性胃溃疡图片 619 张。其中 1 735 张为训练集,150 张为测试集,分别将图片输入基于 fastai 框架的 Res-net50 模型、基于 Keras 框架的 Res-net50 模型和基于 Keras 框架的 VGG-16 模型进行训练。分别构建正常胃黏膜与良性溃疡、正常胃黏膜与恶性溃疡、良性与恶性溃疡 3 个单独的二元分类模型。**结果**　VGG-16 模型表现出了最好的结果,验证集验证模型区分正常黏膜与良性溃疡、正常黏膜与恶性溃疡、良性与恶性溃疡的精确度分别为 98.0%、98.0%和 85.0%。**结论**　本研究获得的模型在发现溃疡病灶上具有较好的能力,有望应用于临床辅助溃疡病灶检出并鉴别良恶性溃疡。

【关键词】　胃溃疡;　内窥镜检查;　人工智能;　诊断,鉴别
基金项目:国家自然科学基金(81672387);中央高校基本科研业务费专项资金(20422018kf1035);湖北省自然科学基金(2016CFA066);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)
DOI:10.3760/cma.j.cn321463-20200105-00547

Artificial intelligence-assisted diagnosis system of benign and malignant gastric ulcer based on deep learning

Huang Li[1,2], *Li Yanxia*[1,2], *Wu Lianlian*[1,2], *Hu Shan*[3], *Chen Yiyun*[4], *Zhang Jun*[1,2], *An Ping*[1,2], *Yu Honggang*[1,2]
[1]*Department of Gastroenterology, Renmin Hospital of Wuhan University, Wuhan 430060, China*; [2]*Key Laboratory of Hubei Province for Digestive Disease, Wuhan 430060, China*; [3]*Wuhan EndoAngel Medical Technology Co., Ltd., Wuhan 430000, China*; [4]*College of Resources and Environment, Wuhan University, Wuhan 430079, China*
Corresponding author: Yu Honggang, Email: yuhonggang1968@163.com

【Abstract】　Objective　To construct an artificial intelligence-assisted diagnosis system to detect gastric ulcer lesions and identify benign and malignant gastric ulcers automatically. **Methods**　A total of 1 885 endoscopy images were collected from November 2016 to April 2019 in the Digestive Endoscopy Center of Renmin Hospital of Wuhan University. Among them, 636 were normal images, 630 were with benign gastric ulcers, and 619 were with malignant gastric ulcers. A total of 1 735 images belonged to training data set and 150 images were used for validation. These images were input into the Res-net50 model based on the fastai framework, the Res-net50 model based on the Keras framework, and the VGG-16 model based on the Keras framework respectively. Three separate binary classification models of normal gastric mucosa and benign ulcers, normal gastric mucosa and malignant ulcers, and benign and malignant ulcers were constructed. **Results**　The VGG-16 model showed the best ability of classification. The accuracy of the

validation set was 98.0%, 98.0% and 85.0%, respectively, for distinguishing normal gastric mucosa from benign ulcers, normal gastric mucosa from malignant ulcers, and benign ulcers from malignant ulcers. **Conclusion** The artificial intelligence-assisted diagnosis system obtained in this study shows noteworthy ability of detection of ulcerous lesions, and is expected to be used in clinical to assist doctors to detect ulcer and identify benign and malignant ulcers.

【**Key words**】 Stomach ulcer; Endoscopy; Artificial intelligence; Diagnosis, differential

Fund program: National Natural Science Foundation of China (81672387); Fundamental Scientific Research Fund for Central Universities (2042018kf1035); Natural Science Foundation of Hubei Province (2016CFA066); 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)

DOI:10.3760/cma.j.cn321463-20200105-00547

胃溃疡是消化性溃疡中较为常见的一种,近年来幽门螺杆菌感染率迅速下降以及强效抗泌酸药物广泛使用,使得消化性溃疡的流行已大幅度低于 20 年前,但世界范围内抗生素耐药性的增加及老年人群中越来越广泛的抗血栓药物的使用增加了胃溃疡治疗和管理的挑战性[1]。胃溃疡可分为良性溃疡和恶性溃疡。恶性胃溃疡,也可称溃疡型胃癌,其治疗方式及预后与良性胃溃疡大为不同。目前,胃镜检查是发现并诊断胃溃疡的常规检查手段,早期及时发现病灶并进行积极治疗与良好的预后密切相关。通常内镜医师先在普通胃镜下通过形态等经验鉴别胃良恶性溃疡,但胃镜下的病理活检与内镜医师在内镜下的判断依然有不少出入,从而给制定相应的诊疗计划带来较大困难,也极大影响了患者预后[2]。此外,进行胃镜检查的医师由于长时间工作,注意力可能会被削弱,从而导致病灶漏检,医师之间水平的个体化差异对病灶检出率也有很大影响[3]。深度学习的提出推动了人工智能的第 3 次浪潮,近年来,人工智能技术相关的研究在消化内镜领域取得了长足发展,在疾病辅助诊断、质量控制等方面均有极佳表现。本研究通过应用人工智能的深度学习算法构建了一个胃溃疡病灶辅助诊断系统,旨在帮助医师发现溃疡病灶并鉴别良恶性,提高诊断水平和检查效率。

资料与方法

一、研究对象

回顾性收集武汉大学人民医院消化内镜中心 2016 年 11 月—2019 年 4 月进行普通和无痛胃镜检查的患者胃镜图像。入选标准:内镜诊断包含溃疡并且具有活检病理结果的图片作为“良性溃疡”或“恶性溃疡”数据集;不包含溃疡病灶的内镜图像作为“正常胃黏膜”数据集;年龄 18 岁以上。选取各个部位的典型图片。排除标准:不清晰图像;泡沫干扰严重的图像;胃排空不完全患者的胃镜检查图像等。所有图片在白光非放大模式下拍摄,使用是日本 Olympus 公司 260 和 290 系列以及日本 Fujinon 公司 580 和 590 系列内镜。良、恶性胃溃疡图片均有活检病理确诊。

二、研究方法

1.图像预处理:根据入选及排除标准筛去不合格图片后,共纳入 1 885 张胃镜图像,其中包括正常胃黏膜图像 636 张,良性胃溃疡图像 630 张,恶性胃溃疡图像 619 张。为了保证测试集和训练集不来源于同一个患者,按照患者胃镜检查时间顺序将图片大约按 4∶1的比例分割为训练集和测试集,验证集的图片来源与训练集完全不同,确保结果泛化性。详见表 1。通过自有算法对图像进行预处理,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。同时,对图片进行旋转、平移、缩放、随机裁切等操作进行样本增强。

表 1　纳入研究的各类胃镜检查图像数量及数据集分配(张)

数据集	正常胃黏膜	良性溃疡	恶性溃疡
训练集	586	580	569
测试集	50	50	50
合计	636	630	619

2.输入图片进行模型训练获得辅助诊断模型:本研究涉及的框架有 fastai 和 Keras,涉及的模型有

Res-net50 和 VGG-16 两个成熟的卷积神经网络模型。将上一步预处理好的训练集图片分别输入到基于 fastai 框架的 Res-net50 模型、基于 Keras 的 Res-net50 模型和基于 Keras 的 VGG-16 模型，进行正常胃黏膜与良性溃疡、正常胃黏膜与恶性溃疡、良性与恶性溃疡的两两二元分类训练，获得二元分类训练模型。其中训练出的基于 Keras 框架的 Res-net50 模型中间层热图见图 1。

3.模型验证：上述辅助诊断模型训练好之后，将准备好的测试集图片每种各 50 张，输入训练好的模型进行验证并获得模型的精确度。

结　果

通过对训练集的样本进行训练获得模型，再对验证集样本进行验证，可见基于 Keras 的 VGG-16 模型精确度最高，鉴别正常胃黏膜与良性溃疡、正常胃黏膜与恶性溃疡、良性溃疡与恶性溃疡的精确度分别为 98.00%，98.00%，85.00%；基于 fastai 或 Keras 的 Res-net50 模型在正常胃黏膜与良性溃疡和正常胃黏膜与恶性溃疡鉴别方面表现相同，精确度均为 94.00%和 96.00%；在鉴别良性溃疡与恶性溃疡方面，基于 Keras 的 Res-net50 模型比基于 fastai

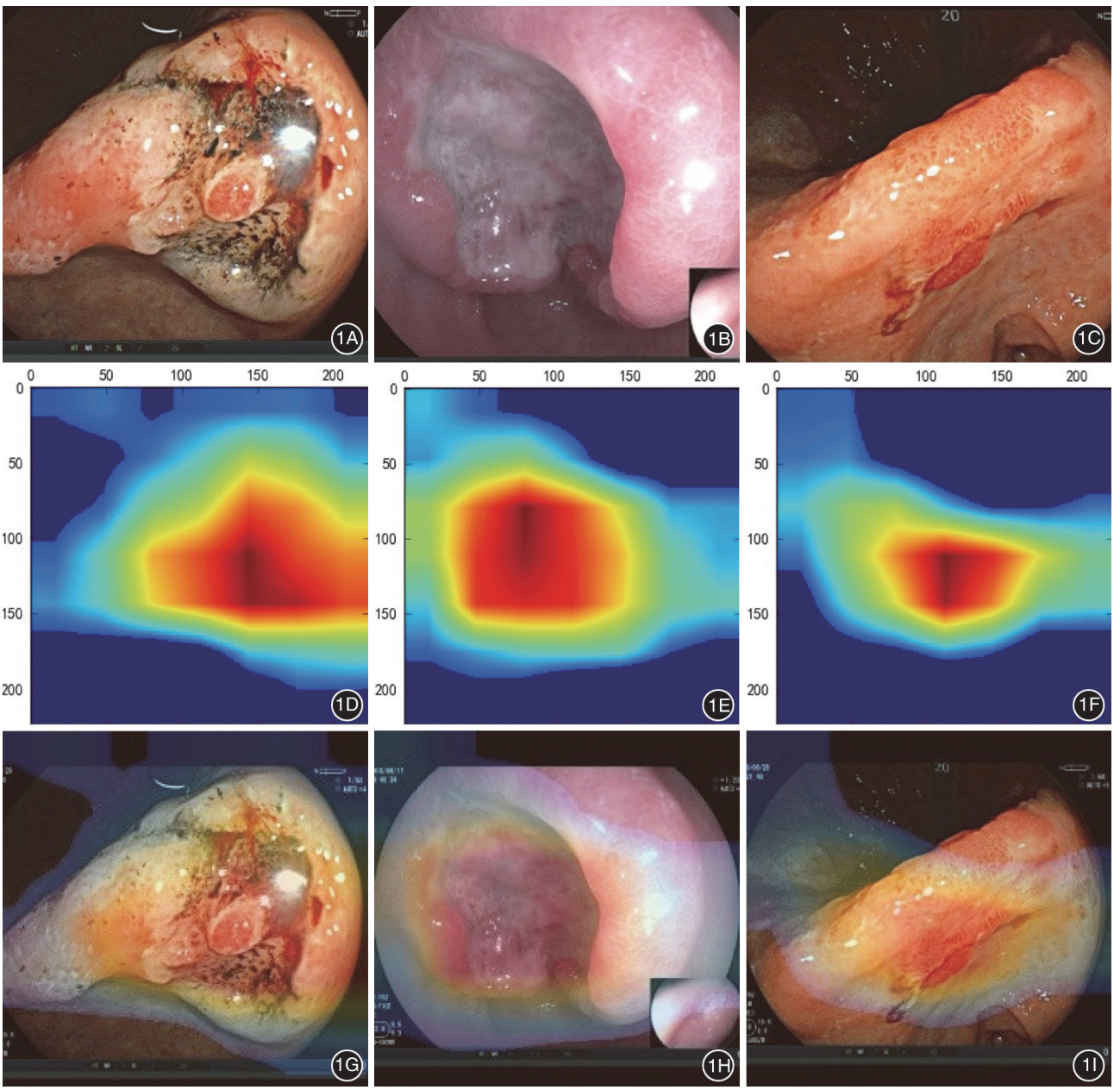

图 1　基于 Keras 框架的 Res-net50 模型中间层热图　1A～1C：输入原图；1D～1F：中间层热图；1G～1I：热图对应到原图上的效果图

框架的 Res-net50 模型精确度高 2%（85.00% 比 83.00%），详见表 2。本研究获得的辅助诊断模型呈现出的辅助诊断结果见图 2。

表 2　胃溃疡病灶辅助诊断系统模型精确度比较（%）

框架	模型	正常胃黏膜与良性溃疡	正常胃黏膜与恶性溃疡	良性溃疡与恶性溃疡
fastai	Res-net50	94.00	96.00	83.00
Keras	Res-net50	94.00	96.00	85.00
Keras	VGG-16	98.00	98.00	85.00

讨　论

本研究中，我们初步探索了基于深度学习的图像分类技术对内镜下正常胃黏膜、良性胃溃疡和恶性胃溃疡的鉴别。通过回顾性收集武汉大学人民医院胃镜检查图片资料，训练和验证模型，获得胃溃疡检出和良恶性鉴别的辅助诊断模型，其发现病灶的精确度高达 98%，鉴别良恶性溃疡的精确度达 85%。

胃癌作为常见的癌症之一，严重影响着患者生活质量，早期发现、早期处理能获得更好的预后。消化内镜检查作为早期消化道癌筛查的常规检查，在全国各地广泛应用。然而，地区之间医疗资源的差异、医师之间水平的差异等，造成了消化内镜检查质量参差不齐。Menon 和 Trudgill[4] 的一项荟萃分析表明，在确诊前 3 年内，约有 11.3%的上消化道癌被漏诊，这是一个不容小觑的数字，而机器在辅助诊断方面表现出的客观性和非疲劳性为这一问题带来极具前景的解决方案。本研究中胃溃疡病灶辅助诊断系统对正常胃黏膜与良性胃溃疡、恶性胃溃疡的鉴别精确度高达 98%，这极大减少了溃疡病灶的漏诊，提高胃镜检查质量。

此前有不少传统研究比较良性溃疡与恶性溃疡在形态方面的差别，内镜医师在普通胃镜下诊断恶性胃溃疡的主要依据是病灶的形态特征，比如底部凹凸不平、不规则形状、污秽厚苔、粗糙边缘、局部皱襞肥厚僵硬或呈结节状等。但对于病史较长的部分良性溃疡，在内镜大体观上也常具有恶性溃疡的特征，需要病理活检来明确诊断结果[5]。良恶性溃疡的鉴别需要内镜医师具有丰富的经验和较多的理论知识。本研究获得的模型在良恶性溃疡的鉴别上具有 85%的精确度，对于经验稍薄弱的内镜医师或鉴别起来较困难的溃疡病灶具有重大意义，可有效提高基层医院的诊疗水平。

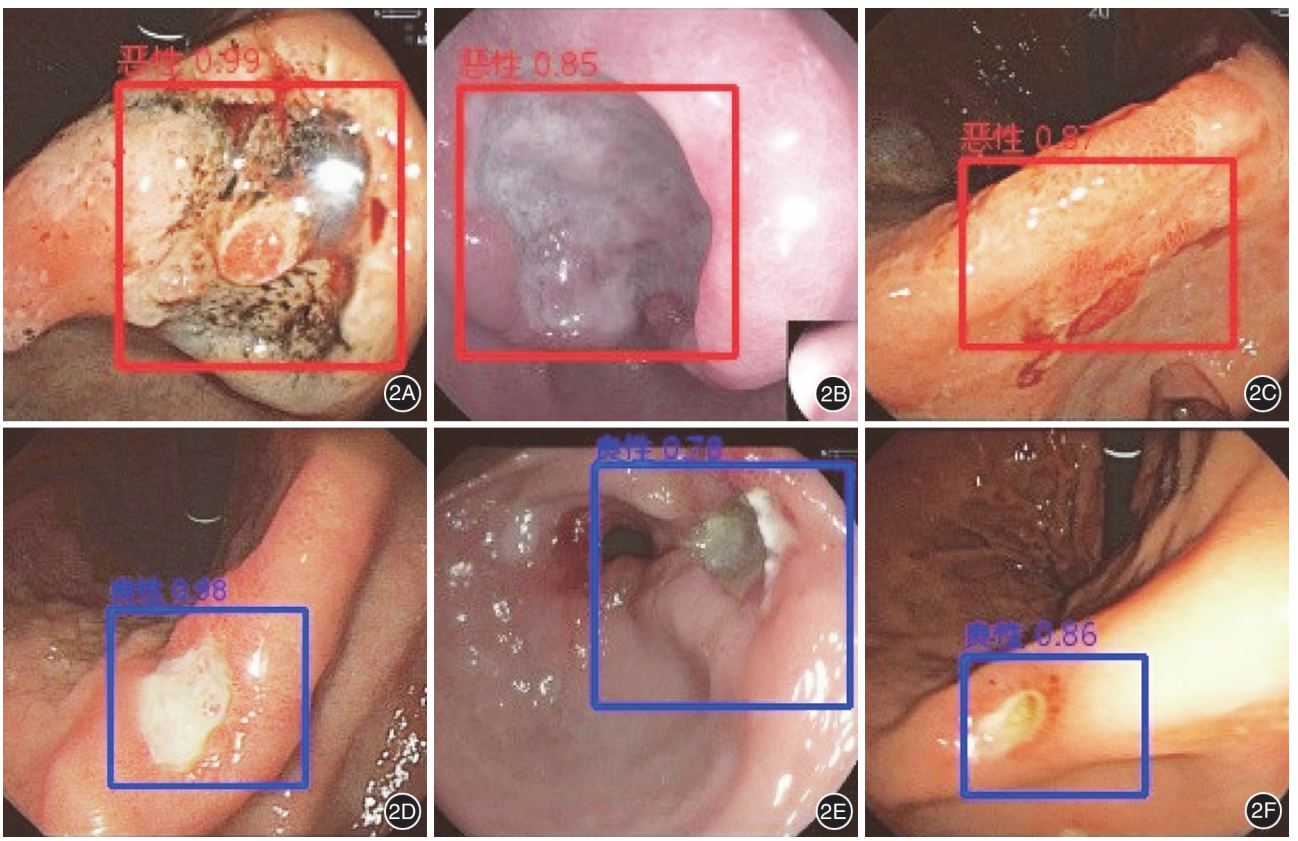

图 2　模型辅助诊断效果展示图　2A～2C：模型预测为恶性溃疡，用红色方框框出，“恶性”后面的数字表示认为溃疡为恶性的置信度；2D～2F：模型预测为良性溃疡，用蓝色方框框出，“良性”后面的数字表示认为溃疡为良性的置信度

近年来基于深度学习的人工智能辅助诊断系统在医疗领域表现极佳。2017 年, Esteva 等[6] 利用 12 945张临床皮肤病影像训练出的皮肤癌辅助诊断系统在诊断能力上与临床医师旗鼓相当。近两年来人工智能在消化内镜领域也有广泛的应用, Wang 等[7] 利用人工智能系统自动识别息肉,在结肠镜检查过程中实时提醒医师; Wu 等[8] 开发的人工智能质量控制系统在胃镜检查过程中实时监测医师检查时是否覆盖 26 个部位,以达到减少胃镜检查盲区的目标。两者的临床试验均证明人工智能产品具有提高胃肠镜检查质量的作用。不难预见,未来人工智能在医疗领域可以在辅助诊断、质量控制、医院管理方面大展身手,利用飞速发展的科技建立信息化的智慧医院。

本研究结果表明,在图像分类方面,不同框架和不同模型组合之间差距较小,当样本量足够大时,甚至可能消除差异。本研究中表现最好的是基于 Keras 框架的 VGG-16 模型,在发现溃疡病灶方面精确度高达 98%,同时在良恶性溃疡区分方面也有较好的结果,精确度为 85%,在样本量相对较小的情况下有这样高的精确度也预示着深度学习在良恶性溃疡病灶发现和鉴别上具有极高的应用前景。同时,模型可实现对病灶的定位。值得注意的是,本研究最初并未标记图片上病灶的位置以供机器学习,算法在训练过程中自己总结特征并定位病灶,这说明深度学习模型具有较好的特征提取和推演能力。

窄带光成像和电子染色等技术可以通过相对简便的操作获得更加清晰的胃溃疡轮廓,并可以更加清晰地观察其周围的微血管形态、腺体结构等在周边黏膜的表现,帮助鉴别良恶性溃疡,也提高胃癌的活检准确率[9-10]。但是考虑到中国医疗资源发展不均衡,部分基层医院较少能够提供染色放大内镜检查,所以白光下人工智能辅助诊断更具有普适性,因此本研究选择白光下图片进行训练,所训练模型更适合推广使用。

然而,本研究还存在一些局限性。本研究中纳入的样本量较少,需要后期增加样本量进行训练,或许能获得更高的精确度。同时,本研究所获取的内镜图像资料只局限于 Olympus 公司的 260 和 290 系列以及 Fujinon公司的 580 和 590 系列,我们下一步将对其他型号内镜所拍摄的图像进行训练,提高模型使用的普适性。在本研究中,良性溃疡与恶性溃疡的区分精确度相对较低,但是本研究对良性和恶性溃疡与正常胃黏膜区分的结果较好,精确度可达 98%。说明本研究所训练的模型发现病灶的能力较强。鉴于胃癌相对于十二指肠癌更高发[11],目前对于胃部的溃疡病灶医师会常规取活检,因此,发现病灶的意义可能强于良恶性的鉴别,良性溃疡与恶性溃疡的鉴别精确度虽没有达到较高水平,但是在临床实际应用过程中影响相对较小,较高的发现病灶能力保证了检查的质量。

总的来说,本研究提出一种胃溃疡病灶辅助诊断、良恶性胃溃疡鉴别系统,训练模型精确度较高,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

参 考 文 献

[1] Lanas A, Chan FKL. Peptic ulcer disease[J]. Lancet, 2017, 390(10094): 613-624. DOI: 10.1016/s0140-6736(16)32404-7.

[2] 赵聪, 农春燕, 李慧, 等. 内镜超声检查术在良恶性胃溃疡鉴别诊断中的价值探讨[J]. 四川医学, 2001, 22(1): 34-35. DOI: 10.3969/j. issn. 1004-0501. 2001. 01. 022.

[3] Lee JH, Kim YJ, Kim YW, et al. Spotting malignancies from gastric endoscopic images using deep learning[J]. Surg Endosc, 2019, 33(11): 3790-3797. DOI: 10.1007/s00464-019-06677-2.

[4] Menon S, Trudgill N. How commonly is upper gastrointestinal cancer missed at endoscopy? A meta-analysis[J]. Endosc Int Open, 2014, 2(2): E46-50. DOI: 10.1055/s-0034-1365524.

[5] 李菁, 洪阳, 李雪梅, 等. 62 例胃巨大溃疡内镜与临床特点分析[J]. 检验医学与临床, 2010, 7(19): 2165-2166. DOI: 10.3969/j. issn. 1672-9455. 2010. 19. 073.

[6] Esteva A, Kuprel B, Novoa RA, et al. Dermatologist-level classification of skin cancer with deep neural networks[J]. Nature, 2017, 542(7639): 115-118. DOI: 10.1038/nature21056.

[7] Wang P, Berzin TM, Glissen BJR, et al. Real-time automatic detection system increases colonoscopic polyp and adenoma detection rates: a prospective randomised controlled study[J]. Gut, 2019, 68(10): 1813-1819. DOI: 10.1136/gutjnl-2018-317500.

[8] Wu L, Zhang J, Zhou W, et al. Randomised controlled trial of WISENSE, a real-time quality improving system for monitoring blind spots during esophagogastroduodenoscopy[J]. Gut, 2019, 68(12): 2161-2169. DOI: 10.1136/gutjnl-2018-317366.

[9] 褚衍六, 高孝忠, 乔秀丽, 等. 窄带成像内镜对胃良恶性溃疡的诊断价值研究[J]. 中国内镜杂志, 2009, 15(6): 580-583, 587.

[10] 丁百静, 刘云燕, 李生, 等. 放大窄带成像内镜对不明性质胃溃疡的诊断价值[J]. 胃肠病学和肝病学杂志, 2015, 24(7): 815-818. DOI: 10.3969/j. issn. 1006-5709. 2015. 07. 014.

[11] Bray F, Ferlay J, Soerjomataram I, et al. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries[J]. CA Cancer J Clin, 2018, 68(6): 394-424. DOI: 10.3322/caac. 21492.

（收稿日期：2020-01-05）

（本文编辑：朱悦）