·论著·

基于 YOLO 算法和 ResNet 深度卷积神经网络的结直肠息肉检测(含视频)

扫码查看视频

李素琴[1]　吴练练[1]　宫德馨[1]　胡珊[2]　陈奕云[2]　朱晓芸[1]　李夏[1]　于红刚[1]
[1]武汉大学人民医院消化内科 430060;[2]武汉大学资源与环境学院 430060
李素琴现在中山大学附属第一医院东院消化内科,广州 510700
通信作者:于红刚,Email: yuhonggang@whu.edu.cn

【摘要】 目的 构建一个基于 YOLO 算法和 ResNet 网络的自动检测结直肠息肉的深度卷积神经网络(deep convolutional neural network,DCNN)模型,并测试其功能。**方法** 选取武汉大学人民医院消化内镜中心数据库 2018 年 1 月—2019 年 3 月的肠镜图像及视频并分为 3 个数据集(数据集 1、3、4),另以公共数据集 CVC-ClinicDB(由西班牙巴塞罗那医院提供的 29 个结肠镜检查视频中提取的 612 帧息肉图像组成)作为数据集 2。数据集 1(2018 年 1—11 月的肠息肉图像 3 700 张,无息肉图像 1 000 张)用于 DCNN 模型构建、训练与验证;数据集 2 和数据集 3(2019 年 1—3 月的肠息肉图像 320 张,无息肉图像 400 张)用于 DCNN 模型在图像中的测试;数据集 4(2018 年 12 月肠镜视频 15 个,包含 33 个息肉),用于 DCNN 模型在视频中的测试。主要观察 DCNN 模型检测肠息肉的敏感度、特异度、准确率和假阳性率。**结果** DCNN 模型在数据集 2 中检测肠息肉的敏感度为 93.19%(602/646);在数据集 3 中检测肠息肉的准确率为 95.00%(684/720),敏感度为 98.13%(314/320),特异度为 92.50%(370/400),假阳性率为 7.50%(30/400);在数据集 4 中检测息肉逐息肉个数的敏感度为 100.00%(33/33),逐帧准确率为 96.29%(133 840/138 998),逐帧敏感度为 90.24%(4 066/4 506),逐帧特异度为 96.49%(129 774/134 492),逐帧假阳性率为 3.51%(4 718/134 492)。**结论** 构建的 DCNN模型可用于自动检测结直肠息肉,在静止肠镜图像及肠镜视频中均具有较高的敏感度与特异度,且在视频中测试的假阳性率低,可用于帮助内镜医师检测结直肠息肉。

【关键词】 人工智能;　YOLO 算法;　残差网络;　结直肠息肉

基金项目:中央高校基本科研业务费专项资金(2042018kf1035);湖北省自然科学基金(2016CFA066)

DOI:10.3760/cma.j.cn321463-20200415-00258

A detection model of colorectal polyps based on YOLO and ResNet deep convolutional neural networks (with video)

Li Suqin[1], *Wu Lianlian*[1], *Gong Dexin*[1], *Hu Shan*[2], *Chen Yiyun*[2], *Zhu Xiaoyun*[1], *Li Xia*[1], *Yu Honggang*[1]
[1]*Department of Gastroenterology, Renmin Hospital of Wuhan University, Wuhan 430060, China*; [2]*School of Resources and Environmental Sciences of Wuhan University, Wuhan 430060, China*
Li Suqin is working at Department of Gastroenterology, East Branch of The First Affiliated Hospital, Sun Yat-Sen University, Guangzhou 510700, China
Corresponding author: Yu Honggang, Email: yuhonggang@whu.edu.cn

【Abstract】 Objective To establish a deep convolutional neural network (DCNN) model based on YOLO and ResNet algorithm for automatic detection of colorectal polyps and to test its function. **Methods** Colonoscopy images and videos collected from the database of Digestive Endoscopy Center of Renmin Hospital of Wuhan University from January 2018 to March 2019 were divided into three databases (database 1, 3, 4). The public database CVC-ClinicDB (composed of 612 polyp images extracted from 29 colonoscopy videos provided by Barcelona Hospital, Spain) was used as the database 2. Database 1 (4 700

colonoscopy images from January 2018 to November 2018, including 3 700 intestinal polyp images and 1 000 non-polyp images) was used for establishing training and verifying the DCNN model. Database 2 (CVC-ClinicDB) and database 3 (720 colonoscopy images from January 2019 to March 2019, including 320 intestinal polyp images and 400 non-polyp images) were used for testing the DCNN model on image detection. Database 4 (15 colonoscopy videos in December 2019, containing 33 polyps) was used for testing the DCNN model on video detection. The sensitivity, specificity, accuracy and false positive rate of the DCNN model for detecting intestinal polyps were calculated. **Results** The sensitivity of the DCNN model for detecting intestinal polyps in database 2 was 93.19% (602/646). In database 3, the DCNN model showed the accuracy of 95.00% (684/720), sensitivity of 98.13% (314/320), specificity of 92.50% (370/400), and false positive rate of 7.50% (30/400) for detecting intestinal polyps. In database 4, the DCNN model achieved a per-polyp-sensitivity of 100.00% (33/33), a per-image-accuracy of 96.29% (133 840/138 998), a per-image-sensitivity of 90.24% (4 066/4 506), a per-image-specificity of 96.49% (129 774/134 492), and a per-image-false positive rate of 3.51% (4 718/134 492). **Conclusion** The DCNN model constructed in the study has a high sensitivity and specificity for automatic detection of colorectal polyps both in the colonoscopy images and videos, has a low false positive rate in the videos, and has the potential to assist endoscopists in diagnosis of colorectal polyps.

【**Key words**】 Artificial intelligence; YOLO algorithm; Residual network; Colorectal polyps

Fund program: Fundamental Research Funds for the Central Universities (2042018kf1035); Natural Science Foundation of Hubei Province (2016CFA066)

DOI:10.3760/cma.j.cn321463-20200415-00258

结直肠癌是全球癌症死亡的第三大常见原因[1],早期发现和切除结直肠肿瘤对于降低结直肠癌患者的死亡率至关重要[2]。结直肠镜检查可以直接观察病灶表面及周围情况,被认为是降低结直肠癌发病率和死亡率的金标准[3]。研究显示,结直肠癌一般由腺瘤性息肉发展而来,从腺瘤演变成早期浸润癌通常需 10~15 年[4]。临床上,腺瘤检出率被认为是结肠镜检查质量的评判标准,腺瘤检出率每增加 1%,结直肠癌发病率就降低 3%~6%[5-6]。但由于内镜医师水平参差不齐,结直肠息肉平均漏诊率在 20%[7]。另外,我国各地区消化内镜医师分布比例、内镜资源配置等差异较大,这些现象短时间内解决难度较大。因此,亟需一个能够自动识别结直肠息肉的的客观诊断系统,以减少人为因素的影响。

近年来人工智能技术在多领域迅速发展,深度学习可自动提取相关成像特征[8],避免人类感知偏差,为更详细的图像分析和实时应用打开了新的篇章。深度卷积神经网络(deep convolutional neural network, DCNN)的发展改变了计算机视觉领域,其能从大数据中学习,且具有高精度和高速处理的优点。深度学习的这种特性,使其在医疗领域,尤其是影像学图像应用广泛。目前,深度学习经典的卷积层神经网络架构众多。其中,深度残差网络(residual network, ResNet)的主要作用是图像分类,是在 VGG19 网络的基础上进行的修改,并通过短路机制加入了残差单元,变化主要体现在 ResNet 直接使用步长为 2 的卷积做下采样,并且用全局池化层替换了全连接层。ResNet 的一个重要设计原则是:当特征图大小降低一半时,特征图的数量增加一倍,保持了网络层的复杂度。ResNet 相比普通网络每两层间增加了短路机制,形成残差学习,使得 ResNet 错误率很低,仅为 3.6%。

YOLO 算法采用一个单独的卷积层神经网络模型实现端到端的目标检测[9],是目前为止最先进的目标检测方法之一,可以在一幅图像中同时检测和分类对象。YOLO 的卷积层神经网络将输入的图像分割成 S×S 网格,然后每个单元格负责去检测那些中心点落在该格子内的目标。YOLO 采用卷积网络来提取特征,然后使用全连接层来得到预测值。网络结构参考 GooLeNet 模型,包含 24 个卷积层和 2 个全连接层。YOLO 算法的优点有:一是其采用一个卷积层神经网络来实现检测,其训练与预测都是端到端的,算法简洁且速度快;二是由于 YOLO 算法是对整张图像做卷积,在检测目标时视野更大,可减少背景误判;三是 YOLO 的泛化能力强,在做迁移时,模型鲁棒性高。

在国外,早在 2003 年 Karkanis 等[10]就首次报

道了使用计算机辅助检测系统检测结直肠息肉，检出率>90%，但由于该系统是基于静态图像，研究结果难以用于临床结直肠镜实时检查过程。2016 年 Fernández-Esparrach 等[11]开发的基于卷积三维神经网络的计算机辅助检测系统，通过学习肠镜检查视频，实现视频中肠息肉定位，但由于研究样本数量有限，灵敏度仅为 70% 左右。基于 YOLO 算法和 ResNet 模型个自的优点，我们团队构建了一个 YOLO 算法组合 ResNet 的 DCNN 模型，并用于结直肠息肉检测，现将研究结果总结报道如下。

对象与方法

一、研究对象

选取武汉大学人民医院消化内镜中心数据库 2018 年 1 月—2019 年 3 月的肠镜图像及视频，按排除标准去除不合格的图像及视频，剩余图像由 4 名医学博士进行分类，分为无息肉的正常图像和有息肉（不区分腺瘤性息肉与非腺瘤性）图像，再由 2 名高年资（>8 年肠镜检查经验）内镜医师对两类图像进行确认。每个肠镜视频都记录了从找到回盲瓣到肛门口的完整的退镜过程（退镜时间≥6 min），且每个视频至少包含 1 个肠息肉，视频按 25 帧/s 展开，上述 2 名高年资医师回顾性地记录每个视频每帧图像中息肉的存在及每个视频息肉个数，并以 2 名高年资医师的诊断为金标准。将最后筛选出的图像及视频分为 3 个数据集：数据集 1（2018 年 1—11 月的肠息肉图像 3 700 张，无息肉图像 1 000 张），DCNN 模型训练/验证集；数据集 4（2018 年 12 月肠镜视频 15 个，包含 33 个息肉），视频测试集；数据集 3（2019 年 1—3 月的肠息肉图像 320 张，无息肉图像 400 张），图像测试集。数据集 2 为公共数据集 CVC-ClinicDB（https：// polyp. grand-challenge. org/site/Polyp/CVCClinicDB），由西班牙巴塞罗那医院提供的 29 个结肠镜检查视频中提取的 612 帧息肉图像组成。排除标准：年龄<18 岁；肠道清洁度差；结直肠部分切除者；进展期癌；炎症性肠病。所用肠镜型号：日本富士 EC-530WM/W，EC-L590WM，EC-600WM；日本奥林巴斯 CF-240I，CF-Q260AI，CF-H290I，CF-HQ290I。图像在白光非放大模式下拍摄。

二、模型构建

1.图像标记：由 2 名高年资医师对数据集 1 的 3 700张息肉图像进行息肉标注（图 1），图像标注工具使用剑桥大学 Visual Geometry Group 开发的 VGG Image Annotator，该标注工具能够在线和离线使用，可标注矩形、圆、椭圆、多边形、点和线。

2.构建模型：由计算机算法工程师用 YOLO 算法及 ResNet 两种网络模型组合搭建成一个 DCNN 模型。DCNN 模型工作步骤（图 2）：第一步将图像输入系统，YOLO 在低阈值（阈值为 0.2）下抓取疑似息肉病灶图像，将图像分类为无息肉组和疑似息肉组；第二步，系统将疑似息肉图片进行剪切，剪切成统一大小 360 像素×360 像素，再统一将疑似病灶放大 1.5 倍，并使其处于图像正中间；第三步，用 ResNet 分类检测（阈值为 0.1），将疑似息肉图片分为非息肉组和息肉组。DCNN 模型处理图像的速度为25 帧/s。数据集 1 中 3 700 张息肉图像（已行息肉标注）用于 YOLO 模型训练/验证，预设进行 100 轮，实际在第 62 轮时模型偏差最低，且在之后连续的 10 轮中偏差不再降低，认为第 62 轮得到的模型为最佳模型，因此设置了提前结束机制，在第 72 轮后结束，见图 3。再从数据集 1 的 3 700 张息肉图像

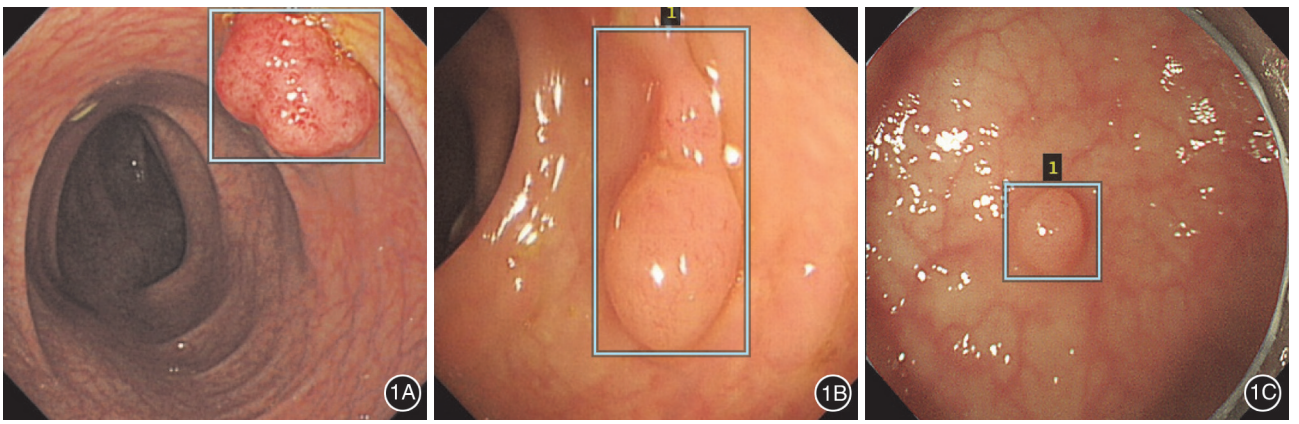


图 1　用矩形框标注每张图像中的息肉

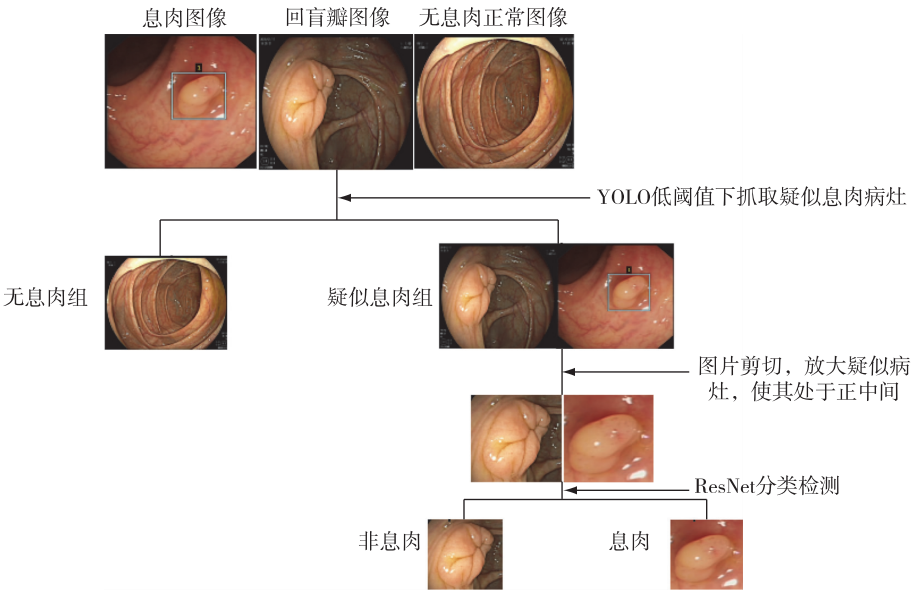


图 2　深度卷积神经网络模型工作流程图

(已行息肉标注)中利用计算机随机数法随机选取 1 000张,并和数据集 1 的 1 000 张无息肉图像用于 ResNet 训练/验证,其中 800 张息肉图像和 800 张无息肉图像用于 ResNet 模型训练/验证,共进行 5 轮,在第 3 轮时偏差达到最小,准确率达到最高,在之后连续的 2 轮中偏差没有继续下降,准确率未再提高,认为第 3 轮得到的模型是最佳模型,见图 4;剩余的 200 张有息肉图像和 200 张无息肉图像用于 ResNet 模型测试,模型的准确率为 96%(384/400)。对于同一例患者的同一个息肉在不同角度或不同距离拍摄的图像,我们会特别关注以保证其同在训练集或测试集中。

三、测试模型

数据集 2 和数据集 3 用于图像测试,数据集 4 用于视频测试。计算 DCNN 模型检测结直肠息肉的敏感度、特异度和准确率。

结　果

1.数据集 2 的 DCNN 模型检测结果:612 帧息肉图像中,部分图像中包含多处息肉,总计 646 处息肉。DCNN 模型检测出 620 处息肉,其中真阳性 602 处、假阳性 18 处,另有 44 处未检出,DCNN 在数据集 2 中检测结直肠息肉的敏感度为 93.19%(602/646)。

2.数据集 3 的 DCNN 模型检测结果:数据集 3 中,含息肉图像 320 张,无息肉图像 400 张;息肉位于回盲部 7 个,升结肠 41 个,横结肠 50 个,降结肠 35 个,乙状结 84 个,直乙交界 11 个,直肠 92 个;息肉≤5 mm 273 个,>5~10 mm 36 个,>10 mm 11 个;息肉山田分型Ⅰ型 120 个,Ⅱ型 170 个,Ⅲ型 20 个,Ⅳ型 10 个。DCNN 的检测结果见表 1。DCNN 检测肠息肉的准确率为 95.00%(684/720),敏感度为

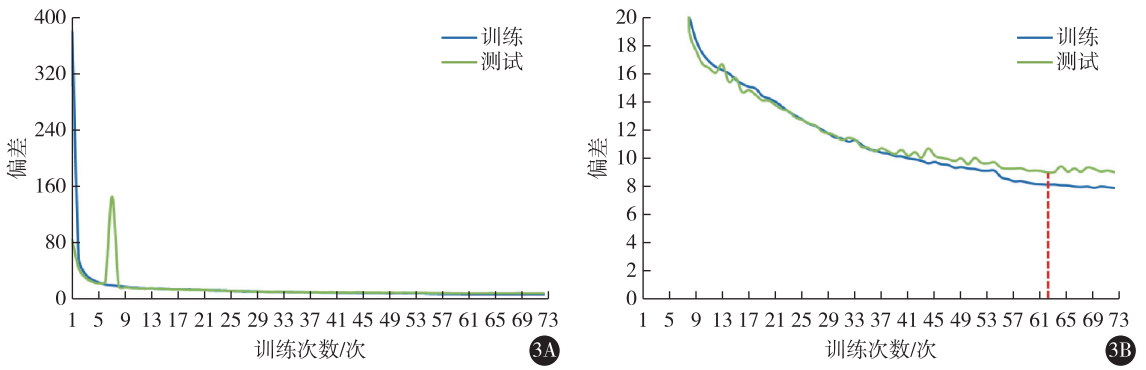


图 3　YOLO 模型训练/验证曲线　3A:72 轮的整体曲线图;3B:第 8~72 轮的局部曲线图

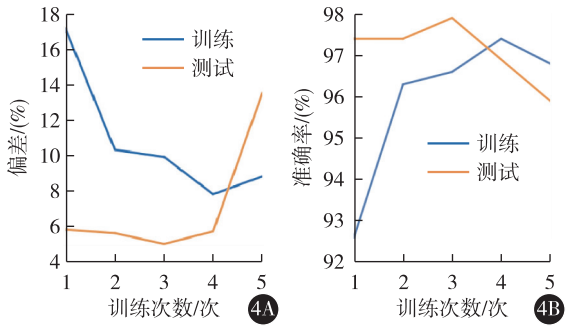


图 4　ResNet 模型训练/验证曲线　4A：模型在第 3 次验证时偏差率达到最小；4B：模型在第 3 次验证时准确率达到最高

98.13%（314/320），特异度为 92.50%（370/400）。未检测出的 6 个息肉均为<5 mm 的息肉（山田Ⅰ型）。DCNN 检测直径<5 mm 肠息肉的准确率为 97.80%（267/273）。

表 1　数据集 3 的深度卷积神经网络模型检测结果（张）

检测结果	图像数	金标准	
		息肉阳性	息肉阴性
息肉阳性	344	314	30
息肉阴性	376	6	370
合计	700	320	400

3.数据集 4 的 DCNN 模型检测结果：数据集 4 的 15 个肠镜视频共包含 33 个息肉，息肉位于回盲部1 个、升结肠 2 个、横结肠 5 个、降结肠 2 个、乙状结肠 10 个、直肠 13 个；息肉≤5 mm 29 个，>5～10 mm 3 个，>10 mm 1 个；息肉山田分型Ⅰ型 19 个，Ⅱ型11 个，Ⅲ型 2 个，Ⅳ型 1 个。15 个肠镜视频总时间约94 min，按 25 帧/s 展开，共 138 998 帧图像，DCNN 的检测结果见表 2，33 个息肉均被 DCNN 模型检出。检测息肉逐息肉个数的敏感度为100.00%（33/33），逐帧准确率为 96.29%（133 840/138 998），逐帧敏感度为 90.24%（4 066/4 506），逐帧特异度为 96.49%（129 774/134 492）。

表 2　数据集 4 的深度卷积神经网络模型检测结果（帧）

检测结果	图像数	金标准	
		息肉阳性	息肉阴性
息肉阳性	8 784	4 066	4 718
息肉阴性	130 214	440	129 774
合计	138 998	4 506	134 492

4.假阳性分析：数据集 3 中，有 30 张［7.50%（30/400）］图像误判为息肉，包括黏膜表面反光 10 张、回盲瓣 1 张、肠黏膜堆积形成皱褶 5 张、粪渣/粪水 6 张、光影 2 张、泡沫 4 张、血管网圈 2 张，以黏膜表面反光、粪渣/粪水、肠黏膜堆积形成皱褶为主［70.00（21/30）］；数据集 4 中，有 4 718 帧［3.51%（4 718/134 492）］视频图像误判为息肉，包括黏膜表面反光 1 674 帧、回盲瓣 664 帧、肠黏膜堆积形成皱褶 1 255 帧、粪渣/粪水 402 帧、光影 461 帧、泡沫 136 帧、血管网圈 58 帧、局部黏膜炎症充血 40 帧、肠道空腔 28 帧，以黏膜表面反光、肠黏膜堆积形成皱褶、回盲瓣为主［76.16（3 593/4 718）］。假阳性图像见图 5。

讨　论

目前深度学习在消化内镜领域的研究多集中在国外。Kumagai 等[12]应用深度学习人工智能来分析食管的细胞内镜图像，其鉴别恶性及非恶性病变的总体准确率为 90.9%，敏感度为 92.6%。Lee 等[13]使用 ResNet 与 VGGNet 卷积层网络模型对比检测胃溃疡、胃癌，结果显示 ResNet 有更高的鉴别水平，区分正常与溃疡图像、正常与胃癌图像的准确率超过 90%。Gross 等[14]对 434 个小结直肠息肉（≤10 mm）通过计算机辅助检测区分非腺瘤性和腺瘤性肠息肉，准确率达到了 93.1%。

在公共数据集 CVC-ClinicDB（数据集 2）的图像检测中，检测肠息肉的敏感度为 93.19%，高于同样使用该数据集测试的报道结果（88.24%）[15]。在数据集 3 的图像检测中，检测息肉的敏感度为 98.13%，特异度为 92.50%；对<5 mm 息肉检测的准确率为 97.80%，高于类似文献报道[15-17]；未检测出的 6 个息肉均为<5 mm 的息肉（山田Ⅰ型）。在数据集 4 的 15 个肠镜视频检测中，当息肉出现在视野后模型会用蓝色矩形框标出息肉，检测肠息肉的敏感度为 90.24%，特异度为 96.49%。DCNN 模型在视频中检测息肉的敏感度较图像中低（90.24%比 98.13%），主要考虑为视频中画面不稳，且息肉刚开始出现在视野时一般距离较远，导致目标过小；而在图像中检测息肉的特异度较视频中低（92.50%比 96.49%），主要考虑为数据集 3 中息肉阴性图像数据过少，可能存在选择偏倚。

肠镜检查过程中存在镜身不稳等导致画面模糊不清的视频节段，这部分的视频逐帧展开后形成大量的模糊图像，无法观察。在实际肠镜检查过程中，医生也无法鉴别模糊视频节段的是否存在息

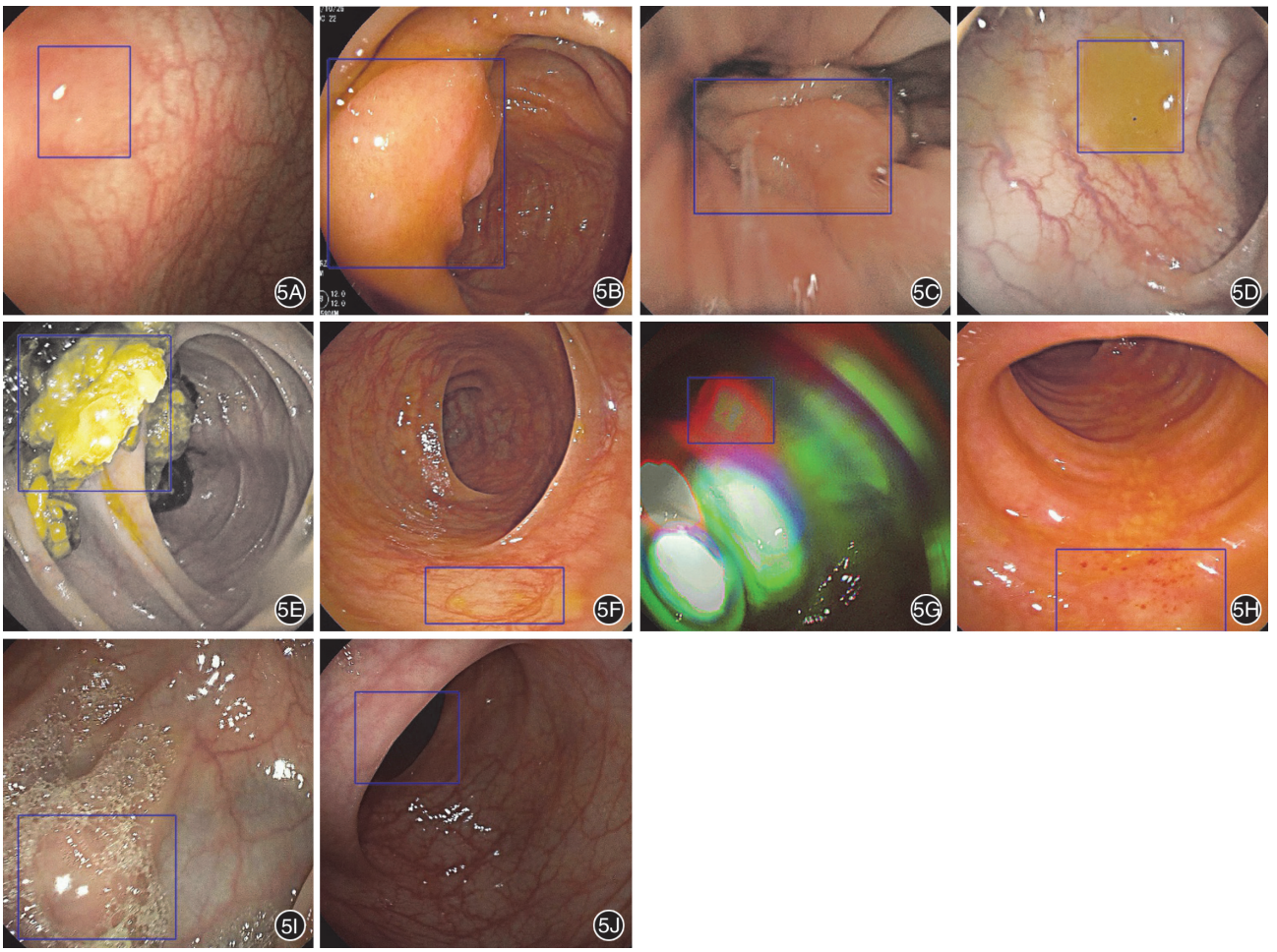


图 5　深度卷积神经网络模型误判成息肉的图像　5A：黏膜表面反光；5B：回盲瓣；5C：肠黏膜堆积形成皱褶；5D：粪水；5E：粪渣；5F：血管网圈；5G：光影；5H：局部黏膜炎症充血；5I：泡沫；5J：肠道空腔

肉，需要等待视野清晰后再次观察。因此，本研究用于 ResNet 模型训练及测试的 1 000 张无息肉图像中，包含无息肉清晰肠黏膜图像及模糊不清无法观察的肠镜图像，目的是让 ResNet 模型能够分类检出模糊不清图像，将其归类于息肉阴性图像，这降低了视频检测中的假阳性。采用数据集 4 的 15 个肠镜视频进行检测时，息肉假阳性率为 3.51%，低于相关文献报道[9,18-19]。

从人的视觉理论分析，一般认为图像处理系统处理速度大于 25 帧/s 为实时分析。本研究 DCNN 模型处理速度为 25 帧/s，基本达到实时分析的效果。本研究选取的肠镜视频为从回盲瓣开始到肛门口结束的完整退镜过程，而非部分文献报道中的从息肉在视野开始出现到息肉在视野消失的节段肠镜视频[15,19]，更接近临床实际情况。本研究的所有肠镜图像及视频是在普通白光下、非放大模式下拍摄的。因为目前我国要提高结直肠腺瘤、早期结直肠癌的检出率，最重要的方法是肠镜筛查的普及。虽然电子染色内镜、显微内镜及光学相干断层成像术等可以更好地观察息肉表面微血管网形态，甚至细胞结构，实现在内镜检查过程中区分肿瘤性和非肿瘤性病变，但其操作过程复杂，技术要求较高，普及度不够，不适合用于筛查，因此我们选择了普通白光下、非放大模式下拍摄的肠镜图像作为研究对象，认为其适用范围更广，更有临床意义。

我们构建的模型使用的是 YOLO 及 ResNet，两者都是目前运用广泛且成熟的网络模型。但目前深度学习的卷积层神经网络构架不断优化，在 ResNet 之后还出现了 DenseNet、ResNeXt、DPN 模型。DenseNet 建立的是前面所有层与后面层的密集连接，通过特征在通道上的连接来实现特征重用，与 ResNet 相比在参数和计算成本更少的情形下就能达到更优的性能。ResNeXt 是把 ResNet 的单个卷积改成了多支路的卷积，保持 ResNet 的高可移

植性优点，同时又提高精度。DPN 模型是在 ResNeXt 的基础上引入了 DenseNet 的核心内容，既可以共享网络特征，又拥有探索新特征的能力，结合了两者的优势，性能更优化。未来我们将进行不同卷基层网络构架单一或组合成的 DCNN 模型对比，找出检测肠息肉准确率最高的模型。

本研究还存在一些不足之处。首先，本研究是一个单中心研究，图像数据相对较少，且回顾性地收集图像，排除了大量干扰图像（如含有泡沫、粪水等），而这些在临床实际中是不可避免的，会影响模型在临床中的实用性；其次，模型训练的肠息肉图像需人工标记，这需要花费大量的时间；再者，未设置对照组，如 DCNN 模型与不同年资内镜医师对比检测息肉的敏感度及诊断时间对比。之后的研究中，我们将从多中心收集图像数据，并前瞻性收集图像，扩大训练集；加入干扰图像，并训练 DCNN 识别干扰图像，提高其临床实用性；研发相关的配套软件，将 DCNN 模型用于真实临床检查过程中。

总之，本研究构建的 DCNN 模型在静止图像及肠镜视频中检测肠息肉均具有较高敏感度与特异度，处理速度快，将来可能实现在实时肠镜检查过程中帮助内镜医师进行自动检测及定位息肉；也可能实现远程会诊，帮助解决偏远及基层地区内镜医师水平不足的问题，提高结直肠息肉检出率。同时，也需要进行更多的研究和临床试验来进一步证实 DCNN 模型用于检测结直肠息肉的有效性与实用性。

利益冲突　所有作者声明不存在利益冲突

（本文视频地址：http://www.xhnj.com/video/1008384.htm）

参 考 文 献

[1] Torre LA, Bray F, Siegel RL, et al. Global cancer statistics, 2012[J]. CA Cancer J Clin, 2015, 65(2): 87-108. DOI: 10.3322/caac.21262.

[2] Shaukat A, Mongin SJ, Geisser MS, et al. Long-term mortality after screening for colorectal cancer[J]. N Engl J Med, 2013, 369(12): 1106-1114. DOI: 10.1056/NEJMoa1300720.

[3] Nishihara R, Wu K, Lochhead P, et al. Long-term colorectal-cancer incidence and mortality after lower endoscopy[J]. N Engl J Med, 2013, 369(12): 1095-1105. DOI: 10.1056/NEJMoa1301969.

[4] 王鹏，贾淑娟，冯晓宏. 2015 年与 2005 年大肠息肉检出率及病理类型对比分析[J]. 解放军医学院学报，2017，38(6)：524-526，530. DOI：10.3969/j.issn.2095-5227.2017.06.009.

[5] Corley DA, Jensen CD, Marks AR, et al. Adenoma detection rate and risk of colorectal cancer and death[J]. N Engl J Med, 2014, 370(14): 1298-1306. DOI: 10.1056/NEJMoa1309086.

[6] Urban G, Tripathi P, Alkayali T, et al. Deep Learning Localizes and Identifies Polyps in Real Time With 96% Accuracy in Screening Colonoscopy[J]. Gastroenterology, 2018, 155(4): 1069-1078.e8. DOI: 10.1053/j.gastro.2018.06.037.

[7] Rustagi T. Quality indicators for colonoscopy and the risk of interval cancer[J]. N Engl J Med, 2010, 363(14): 1372; author reply 1373. DOI: 10.1056/NEJMc1006842.

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

[9] 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.

[10] Karkanis SA, Iakovidis DK, Maroulis DE, et al. Computer-aided tumor detection in endoscopic video using color wavelet features[J]. IEEE Trans Inf Technol Biomed, 2003, 7(3): 141-152. DOI: 10.1109/titb.2003.813794.

[11] Fernández-Esparrach G, Bernal J, López-Cerón M, et al. Exploring the clinical potential of an automatic colonic polyp detection method based on the creation of energy maps[J]. Endoscopy, 2016, 48(9): 837-842. DOI: 10.1055/s-0042-108434.

[12] Kumagai Y, Takubo K, Kawada K, et al. Diagnosis using deep-learning artificial intelligence based on the endocytoscopic observation of the esophagus[J]. Esophagus, 2019, 16(2): 180-187. DOI: 10.1007/s10388-018-0651-7.

[13] 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.

[14] Gross S, Trautwein C, Behrens A, et al. Computer-based classification of small colorectal polyps by using narrow-band imaging with optical magnification[J]. Gastrointest Endosc, 2011, 74(6): 1354-1359. DOI: 10.1016/j.gie.2011.08.001.

[15] Wang P, Xiao X, Glissen Brown JR, et al. Development and validation of a deep-learning algorithm for the detection of polyps during colonoscopy[J]. Nat Biomed Eng, 2018, 2(10): 741-748. DOI: 10.1038/s41551-018-0301-3.

[16] Kominami Y, Yoshida S, Tanaka S, et al. Computer-aided diagnosis of colorectal polyp histology by using a real-time image recognition system and narrow-band imaging magnifying colonoscopy[J]. Gastrointest Endosc, 2016, 83(3): 643-649. DOI: 10.1016/j.gie.2015.08.004.

[17] Shin Y, Balasingham I. Comparison of hand-craft feature based SVM and CNN based deep learning framework for automatic polyp classification[J]. Conf Proc IEEE Eng Med Biol Soc, 2017, 2017: 3277-3280. DOI: 10.1109/EMBC.2017.8037556.

[18] 王智杰，高杰，孟茜茜，等. 基于深度学习的人工智能技术在早期胃癌诊断中的应用[J]. 中华消化内镜杂志，2018，35(8)：551-556. DOI：10.3760/cma.j.issn.1007-5232.2018.08.004.

[19] Misawa M, Kudo SE, Mori Y, et al. Artificial Intelligence-Assisted Polyp Detection for Colonoscopy: Initial Experience[J]. Gastroenterology, 2018, 154(8): 2027-2029.e3. DOI: 10.1053/j.gastro.2018.04.003.

（收稿日期：2020-04-15）
（本文编辑：顾文景）