

## 基于宫颈上皮与血管特征的阴道镜图像 深度学习模型探索

李燕云<sup>1</sup> 王永明<sup>2</sup> 周 奇<sup>1</sup> 李亦学<sup>3</sup> 王 振<sup>3</sup> 王 珏<sup>4</sup> 孟 妍<sup>4</sup> 蔡青青<sup>5</sup> 隋 龙<sup>1</sup> 华克勤<sup>6△</sup>

(<sup>1</sup>复旦大学附属妇产科医院宫颈科, <sup>4</sup>组织部, <sup>5</sup>女性生殖内分泌相关疾病重点实验室, <sup>6</sup>妇科 上海 200011;

<sup>2</sup>上海长江科技发展有限公司智能医疗业务中心 上海 200233; <sup>3</sup>中国科学院上海生命科学研究院 上海 200031)

**【摘要】 目的** 通过识别阴道镜图像中的上皮与血管特征,探讨深度学习的目标检测技术在宫颈癌前病变定位及分类中的可行性。**方法** 收集2018年3月至2019年7月复旦大学附属妇产科医院病理诊断为宫颈低级别(5 708例)、高级别(2 206例)癌前病变和宫颈癌(514例)患者的28 975张阴道镜图像。依照国际宫颈病理与阴道镜联盟及美国阴道镜与病理协会阴道镜标准化术语,基于16类宫颈上皮与血管征象,对图像进行像素级标注后得到有效标签39 858个。为降低细粒度标注可能存在的误差,进一步将标签归并为低级别、高级别和癌三大类。采用经过二次迁移学习的ResNet101预训练网络作为特征提取器,分别构建基于Faster-RCNN网络结构的高级别病变目标检测和低、高、癌三类目标检测模型。**结果** 基于ImageNet预训练的ResNet101模型,通过宫颈转化区分类的开源阴道镜数据进行第一次迁移学习,再以自有数据的病变分类为目标进行第二次迁移学习得到特征提取器。所构建的高级别和三类病变定位检测模型在测试集上的识别精度均值mAP@IOU=0.5分别为0.82和0.67。**结论** 利用国内最大阴道镜中心的大样本数据,基于上皮与血管特征的精细标注,深度学习模型在宫颈癌前病变检测中取得较好效果。深度学习的目标检测技术在宫颈癌前病变定位及分类中可行,尽管在识别精度上仍有提升空间,但已显示其辅助宫颈筛查尤其是指导定位的可行性。

**【关键词】** 阴道镜; 宫颈癌前病变; 标准化术语; 深度学习; 目标检测

**【中图分类号】** R737.33 **【文献标志码】** A **doi:** 10.3969/j.issn.1672-8467.2021.04.002

## Deep learning model exploration of colposcopy image based on cervical epithelial and vascular features

LI Yan-yun<sup>1</sup>, WANG Yong-ming<sup>2</sup>, ZHOU Qi<sup>1</sup>, LI Yi-xue<sup>3</sup>, WANG Zhen<sup>3</sup>, WANG Jue<sup>4</sup>,  
MENG Yan<sup>4</sup>, CAI Qing-qing<sup>5</sup>, SUI Long<sup>1</sup>, HUA Ke-qin<sup>6△</sup>

(<sup>1</sup>Center for Diagnosis and Treatment of Cervical Diseases, <sup>4</sup>Department of Organization, <sup>5</sup>Shanghai Key Laboratory of Female Reproductive Endocrine Related Diseases, <sup>6</sup>Department of Gynecology, Obstetrics and Gynecology Hospital, Fudan University, Shanghai 200011, China; <sup>2</sup>Intelligent Medical Business Center of Shanghai Changjiang Technology Co., Limited, Shanghai 200233, China; <sup>3</sup>Shanghai Academy of Life Sciences, Chinese Academy of Sciences, Shanghai 200031, China)

**【Abstract】 Objective** To explore the feasibility of object detection and deep learning model applied on the localization and classification of cervical precancerous lesions based on identifying the epithelial and vascular features in colposcopy images. **Methods** A total of 28 975 colposcopic images were collected from Mar 2018 to Jul 2019 in the Obstetrics and Gynecology Hospital of Fudan University, including cervical low-grade lesion (5 708 patients), high-grade lesion (2 206 patients) and cervical cancer (514 patients). According to the colposcopy standardized terminology of the International Federation for Cervical

上海市科委医学引导类科技支撑项目(19411960100);上海市经济和信息化委员会人工智能创新发展专项(2018-RGZN-02041)

<sup>△</sup>Corresponding author E-mail: huakeqin@126.com

网络首发时间:2021-07-06 15:53:04 网络首发地址:https://kns.cnki.net/kcms/detail/31.1885.R.20210706.1347.022.html

Pathology and Colposcopy and American Society for Colposcopy and Cervical Pathology, 39 858 valid labels were obtained after pixel-level labeling based on 16 types of cervical epithelial and vascular signs. In order to reduce the error of fine-grained labeling, labels were further classified into three categories: low-grade, high-grade and cancer. Using ResNet101 pre-training network after secondary transfer learning as feature extractor, the models of high-grade lesion object detection and three categories (low-grade, high-grade and cancer) object detection based on Faster-RCNN network structure were constructed respectively. **Results** Based on the ResNet101 model of ImageNet pre-training, the first transfer learning was performed through the open source colposcopy data of cervical transformation zone classification, and the second transfer learning was carried out based on our own data of lesion classification to obtain the feature extractor. The average recognition accuracy mAP@IOU=0.5 obtained on the high-grade lesion model and the three categories model were 0.82 and 0.67, respectively. **Conclusion** Using the large sample data of the largest colposcopy center in China, the deep learning model can make a good performance in the detection of cervical precancerous lesions based on the fine labeling of epithelial and vascular features. Although there was still room for improvement on accuracy, these models were shown to be potential for cervical cancer screening, especially on guidance for location.

**【Key words】** colposcopy; cervical precancerous lesion; standardized terminology; deep learning; object detection

\* This work was supported by the Science and Technology Commission of Shanghai Municipality in Medical Guidance Science & Technology Pillar Program (19411960100) and Shanghai Municipal Commission of Economy and Informatization in Artificial Intelligence Innovation & Development Project (2018-RGZN-02041).

宫颈癌是最常见的妇科恶性肿瘤<sup>[1]</sup>,有明确的癌前病变过程,即人乳头瘤病毒(human papillomavirus, HPV)感染后出现低度鳞状上皮内病变(low-grade squamous intraepithelial lesion, LSIL)、高度鳞状上皮内病变(high-grade squamous intraepithelial lesion, HSIL),最终进展为浸润癌(invasive carcinoma)。尽管全球范围内宫颈癌筛查已开展数十年,总体效果却远未令人满意<sup>[1-2]</sup>。作为宫颈癌早期诊断三阶梯“细胞学/HPV初筛→阴道镜→组织学确诊”中的主要瓶颈,阴道镜的诊断效率及同质化程度一直是国内外宫颈癌防治中的关键点和难点。为了规范其图像评价标准,国际宫颈病理与阴道镜联盟(International Federation of Cervical Pathology and Colposcopy, IFCPC)于2011年提出了目前全球最全面的阴道镜图像描述体系<sup>[3]</sup>,之后2018年美国阴道镜与病理协会(American Society for Colposcopy and Cervical Pathology, ASCCP)也对其进行了引用<sup>[4]</sup>。该术语体系对宫颈上皮和血管的边界、轮廓、形态等重要解剖部位或组织的图像特征进行了极为详尽的定义及科学化归类,涵盖了醋酸白上皮、镶嵌、点状血管、异形血管等几十种征象的标准化解析<sup>[5-8]</sup>。然

而,作为相对复杂的描述性分类,其临床普及尚有难度。尤其在国内外,高水平的专业阴道镜医生极为短缺,基层医师水平和能力不足,阴道镜诊断效能远不能满足大量的临床需求。

近年,深度学习凭借其强大的特征提取能力在提高医学图像的诊断效率及标准化方面展示出了极佳的应用前景<sup>[9]</sup>。然而,由于较平面图像更为复杂、部位多变、需较强专业背景支撑的跨学科合作等原因,对于阴道镜图像的智能化判别这一细分领域,研究成果却相对较少<sup>[10-14]</sup>。2015—2018年,有学者曾基于传统机器学习算法,通过从醋酸或碘试验图像中提取的颜色和纹理等特征信息来提高识别精度<sup>[15-17]</sup>。近年又有学者尝试利用卷积神经网络训练分类器对宫颈病变进行二分类(LSIL和HSIL)或三分类(HSIL、原位癌、浸润癌)的视觉判别,并取得了一定进展<sup>[10-14]</sup>;基于宫颈照相机摄图、宫颈醋酸前后比值图等,深度学习辅助HSIL识别亦获得尚可的结果<sup>[18-19]</sup>。然而在推动实际临床应用方面,上述成果却远未取得实质性进展。病变的定性可依靠最终组织病理来弥补,而阴道镜定位指导活检或治疗则更为重要且无法取代。此外,上述研究目标多为整张图像的病变分类,无法具体定位病灶区域,

且相关报道极少。曾有个别学者尝试采用分类激活图谱(class activation mapping, CAM)热力图等对宫颈病灶区进行定位识别,但判别效果较差<sup>[20]</sup>。此外,标注特征量少(1~3个特征)导致的信息较为单一,缺乏专业知识背景对图像特征的规范化解读,实验数据量有限(HSIL例数均<500例),难以与临床实际情况相符,亦是目前多数研究的短板<sup>[10-14]</sup>。

本研究借助国际阴道镜术语背景,将不易掌握的复杂描述性分类术语转化为实际且便捷的操作体系,弥补了既往研究中标注特征量少、信息单一、实验数据量有限等短板。同时以临床意义为导向,基于国内最大阴道镜中心的大样本数据,对阴道镜图像进行标准化精细标注,并采用经二次迁移学习的特征提取器构建深度学习目标检测模型,探索智能辅助阴道镜在宫颈病变区域定位及识别中的可行性,研究成果不仅适用于临床筛查,亦可指导活检及后续定位治疗。

## 资料和方法

**数据采集** 回顾性收集2018年3月至2019年7月复旦大学附属妇产科医院宫颈疾病诊治中心9台阴道镜仪器记录的阴道镜数据,每位患者包含3~10张不等的阴道镜图像以及基本信息文件。阴道镜检查仪器包括:3台美国WALLACH公司生产的光电一体数码电子阴道镜(PENTASCOPE)、3台德国Leisegang公司生产的光电一体数码电子阴道镜(BG/LED Y/C)、2台深圳EDAN公司生产的电子数码阴道镜(C6 HD)以及1台江苏TRME同人医疗公司生产的电子数码阴道镜(TR6000G),光电一体数码电子阴道镜均采用Canon EOS600D相机进行拍摄。阴道镜检查原因包括:子宫颈细胞学检查阳性、高危型HPV检测阳性、细胞学检查和高危型HPV检测均阴性但临床症状明显,以及其他原因如可疑罹患宫颈(或阴道、外阴)病变等。阴道镜检查采用常规3%醋酸试验及5%卢戈(Lugol's)液碘染色程序,对阴道镜图像异常区域行多点活检,阴道镜图像未见异常者行常规时钟3、6、9、12点宫颈活检及子宫颈管搔刮术。将活检组织送病理科检查,由本院专业妇产科病理医师阅片后作出诊断。根据2012年版下生殖道和肛门鳞状上皮病变术语(The Lower Anogenital Squamous Terminology,

LAST),病理诊断结果分为:子宫颈正常或炎症、LSIL、HSIL和癌变(包括微浸润癌和浸润癌)<sup>[21]</sup>。

同时,基于患者基本信息文件中的姓名及检查日期,通过与同一时段内组织学病理报告数据中的患者和送检日期进行匹配以获取患者的病理诊断结果。剔除无病理诊断结果以及数据不完整者,共纳入最终病理诊断为宫颈LSIL者5708例、宫颈HSIL者2206例以及宫颈癌者514例,共计8428例患者以及28975张阴道镜图像数据。该回顾性研究符合人体试验伦理标准,并已获得复旦大学附属妇产科医院伦理委员会批准(伦理号:2020-28)。

### 数据标注

**标注规范** 依照2011年国际宫颈病理与阴道镜联盟IFCPC<sup>[3]</sup>及2018年美国阴道镜与病理协会ASCCP阴道镜标准化术语体系<sup>[4]</sup>,宫颈病变异常征象包括:薄醋酸白上皮、细镶嵌及细点状血管,归为1级(LSIL)征象;厚醋酸白上皮、粗镶嵌、粗点状血管、边界锐利、内部边界、隆起,归为2级(HSIL)征象;白斑、侵蚀性病变及碘染不着色,归为非特异性征象;异形血管、脆性血管、外生型病变、坏疽、溃疡等,归为可疑浸润癌征象;湿疣、息肉、明显子宫颈触血等其他征象,归为杂类。本实验基于上述体系,对各类上皮与血管征象制定了20类标注标签,对阴道镜图像的病变区域和级别进行像素级语义标注。包括:(1)薄的醋酸白上皮(含不规则地图样边界);(2)细镶嵌;(3)细点状血管;(4)致密醋酸白上皮(含袖口状腺开口隐窝);(5)粗镶嵌;(6)粗点状血管;(7)边界锐利;(8)内部边界;(9)隆起;(10)异形血管(含非典型血管);(11)脆性血管;(12)表面轮廓不规则;(13)外生型病变;(14)坏疽;(15)溃疡;(16)宫颈有肿块或肿瘤形成。其中,1、4、7、8、9、12、13、14、15和16为上皮特征标签,2、3、5、6、10和11为血管特征标签。0为无效,检查不充分或模糊图像;(17)碘染阴性;(18)阴道图像;(19)外阴图像,该四类标签定义为噪声图像不纳入模型训练样本。

**标注过程** 采用开源标注工具Labelme。候选医师首先通过规范培训并在指定测试图像上的标注准确度不低于65%才可参与数据的标注工作。依据以上标准最终选出了50位熟练掌握2011版IFCPC国际阴道镜术语且具有5年以上阴道镜工作经验的阴道镜专科医师。为了便于在标注过程中相互审核,每2名医师为一小组,按照上述标准规范



对图像进行第一轮标注,标注速度为每人30~50张/天。另由10名具有10年以上阴道镜工作经验的阴道镜医师进行第二轮专家复审,每5名标注医师配备1名专家。审核专家应对不同标签的认定标准达成一致,审核中发现的问题当场修改。最后,由8名算法工程师对标签的规范性进行第三轮复审。整个标注过程严格遵循标注流程,确保图像标注的准确性、规范性和有效性,标注示例见图1。

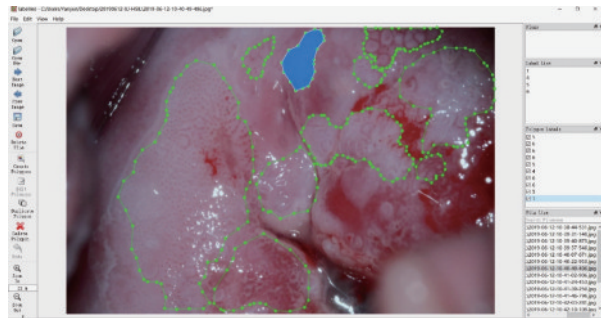


图1 基于Labelme软件对20类宫颈上皮与血管征象标签及区域标准化标注示例

Fig 1 Example of standardized labels based on 20 types of cervical epithelial and vascular signs in Labelme

**标注后处理** 将20类标签归并为五大类:低级别征象标签(1、2和3)、高级别征象标签(4、5、6、7、8和9)、浸润癌征象标签(10、11、12、13、14、15和16)、无法分类标签(17)以及无效噪声标签(0、18和19)。无效标签噪声数据标签包括质量较差图像以及外阴、阴道图像。碘染图像由于对上皮和血管的分辨率较差,则归为无法分类标签。最终,基于28 975张阴道镜图像,共获得标签57 618个。剔除无法分类及无效噪声标签后,共得到三分类有效标注图像

19 607张,包括LSIL图像11 984张、HSIL图像6 871张和癌变图像752张;共获取有效标签39 858个,包括低级别征象标签24 262个、高级别征象标签13 199个以及癌变征象标签2 397个。

**建模任务** 从临床需求看,宫颈HSIL是目前临床治疗的分界点,国内外均以HSIL识别准确率作为阴道镜检查的质控标准<sup>[4]</sup>。从病变区域特征看,LSIL病变区域特征不明显,识别难度较大;癌变区域特征明显,但数据量偏少;而HSIL病变区域的上皮与血管特征较LSIL显著,且样本量介于LSIL和癌变之间。因此,本研究将HSIL阴道镜图像的病变区域检测作为主要建模任务。在上述研究的基础上,进一步对LSIL、HSIL、癌三类别病变区域检测和类型识别进行了研究。

**实验方法** 本实验的建模任务可定义为计算机视觉(computer vision, CV)领域的目标识别和检测。自2014年以来,基于深度学习的目标检测框架分为two-stage和one-stage两大类,前者以经典方法Faster R-CNN<sup>[22]</sup>为代表,后者以YOLO<sup>[23]</sup>和SSD<sup>[24]</sup>为主要框架。由于Faster R-CNN具有更好的检测精度和鲁棒性,本文选择Faster-RCNN网络结构来检测宫颈病变区域。同时,用ResNet101<sup>[25]</sup>网络结构作为提取目标特征的主干网络(backbone)。具体实验方法如图2所示。

**图像预处理** 通过剪切、模糊处理、旋转、翻转、亮度调节、缩小、放大、不规则变形、对比度调节等操作,进行数据增强,使实验数据多元化,模拟真实数据,并解决不同标签类别之间的样本数量不均问题。保留真实的病灶标注,并将标注区域转化

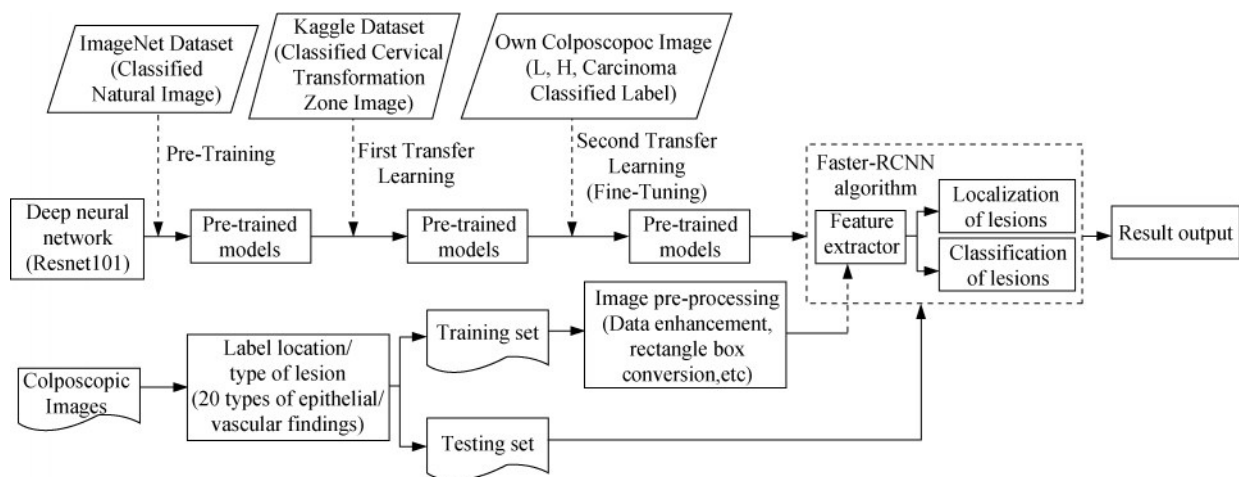


图2 深度学习模型构建技术路线图

Fig 2 The technical routes of deep learning models construction

为矩形框,以适应目标检测任务。将数据按照8:2分为训练集和测试集。

**基于迁移学习的特征提取器** 第一次迁移:基于开源自然图像 ImageNet 数据集 (<http://www.image-net.org>) 训练的 Resnet101 模型,利用开源宫颈阴道镜图像宫颈转化区分类数据(<https://www.kaggle.com/c/intel-mobileodt-cervical-cancer-screening/data>)训练,进行第一次迁移学习。第二次迁移:基于上述模型,从自有阴道镜图像数据中挑选出单一分类(LSIL、HSIL 或癌)标签图像,以病变类型(LSIL、HSIL、癌)识别为目标继续对网络最后几层进行微调(fine-tuning),为第二次迁移学习。

**高级别病变区域检测** 将上述预训练的 Resnet101 作为特征提取器作为 Faster-RCNN 的主干网络,用于病变区域特征提取。以开源 Faster-RCNN 代码进行实验,参数采用源代码的默认参数。以嫁接方式进行精调,替换上面数层卷积层,目标输出为病变区域及类型(参数: BATCH=4, EPOCH=50, RESNET101 主干网参数训练)。用随机数生成方式,从含有有效标签的 HSIL 图像数据中任意抽取数据,进行算法训练。用训练得到的参数对未标注的测试集阴道镜图像预测病变区域,获得算法预测结果图,与医生标注的病变区域进行比对,计算出真阳性样本量(true positive, TP)、假阳性样本量(false positive, FP)、假阴性样本量(false negative, FN)、召回率(recall, R)、精确度(precision, P)、交并比(intersection over union, IOU)、平均识别精度均值(mean average precision, mAP)及受试者操作特征曲线(receiver operator characteristic curve, ROC curve)。

**三分类区域检测与病变类型识别** 用随机数生成方式,从含有有效标签的不同病变图片中任意抽取数据,进行算法训练。用训练集得到的参数对测试集图片进行预测病变区域及病变类型,同上述方法获得 mAP@IOU=0.5 值。

结 果

本实验采用的硬件平台为 HPE DL380 gen10 服务器。其中,配置 2 个 xeon gold 6148 (2.4G 20Core) CPU, 2 个 Nvidia v100 32G GPU 显卡, 512GB 内存。软件环境为 Ubuntu16.04 操作系统, CUDA9.0, cuDNN7.5 for CUDA9.0, Python3.7, 深

度学习框架 PyTorch 1.0 版本的 Faster-RCNN (<https://github.com/jwyang/faster-rcnn.pytorch.git>)。

**迁移学习结果** 第一次迁移学习:1 997 张训练;371 张测试,测试集上的分类准确率(分类正确图片数/总图片数)为 75%(表 1)。

表 1 基于 ImageNet 宫颈转化区分类的第一次迁移学习数据分布

Tab 1 Data distribution of the first transfer learning based on ImageNet cervical transformation zone classification (images)

Set	Cervical transformation zone			Total (n)
	Type 1	Type 2	Type 3	
Training set	330	1 064	603	1 997
Testing set	59	202	110	371
Total (n)	389	1 266	713	2 368

第二次迁移:从自有阴道镜图像训练集中筛选出单分类标签图像,即图片仅包含低级别、高级别或癌变三种标签中的单一类标签图像,共计 17 876 张高级别图像。其中,随机选择 16 108 张图片作为训练集, 1 768 张图片作为测试集,各类别标签数据分布如表 2 所示。测试集上的模型区分 LSIL、HSIL、癌变的准确率(分类正确图片数/总图片数)为 70%。

表 2 基于自有单分类标签阴道镜图像的第二次迁移学习数据分布

Tab 2 Data distribution of the second transfer learning based on colposcopy images of single-lesion labelled classification (images)

Set	Colposcopy images of single-lesion labelled classification			Total (n)
	LSIL	HSIL	Invasive carcinoma	
Training set	10 611	4 822	675	16 108
Testing set	1 128	557	83	1 768
Total (n)	11 739	5 379	758	17 876

**高级别病变区域检测结果** 从只包含单一 HSIL 标签的阴道镜图像中随机抽取出 1 790 张图片作为训练集;61 张图片为测试集。将测试集图片的检测结果与医生标注区域进行比对,获得真阳性样本量 TP 为 44;假阳性样本量 FP 为 17、假阴性样本量 FN 为 5。计算得到召回率(R)为 89.8%,精确率(P)为 72.1%,并根据高级别病变区域检测的结果绘制 R-P 图例和 ROC 曲线(图 3A、B)。当交并比 IOU 为 0.5 时,精度均值 AP@IOU=0.5 为 0.82。预测模型最终输出结果示例见图 4。

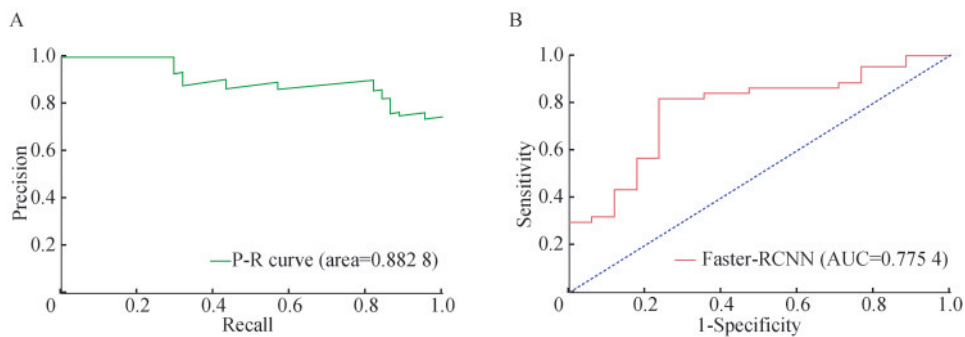
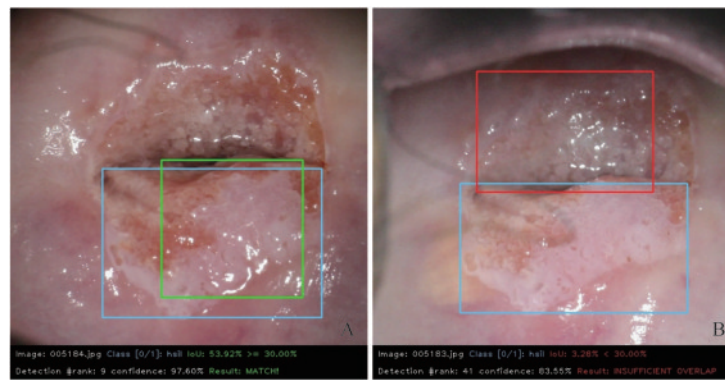


图3 深度学习模型对61张高级别病变测试图像识别的P-R图例(A)和ROC曲线(B)

Fig 3 P-R image recognition (A) and ROC curve (B) of 61 images for high-grade lesion recognized by deep learning model



A: Correct, results match; B: Wrong, results insufficiently overlap. Blue box: The lesion area which doctors marked; Green box: The lesion area which was determined as correct model output when IOU is 0.5; Red box: The lesion area which was determined as wrong model output when IOU is 0.5.

图4 IOU为0.5时模型在测试集上输出结果(A)与医师标注区域(B)对比示例

Fig 4 Example of comparison between model's result (A) and doctor's result (B) on the test data when IOU is 0.5

**三分类区域检测与病变类型识别结果** 用随机数生成方式,任意抽取5 000张图片作为训练集,其中HSIL图片1 752张,LSIL图片3 056张,癌图片192张,另抽取200张图片作为测试集。将训练模型对测试集图片预测病变区域与医生标注区域进行对比,每个类别的平均识别精度均值mAP@IOU=0.5为0.67。

## 讨 论

本研究借助国际阴道镜标准化术语中多种宫颈上皮及血管征象特征对图像进行像素级语义标注,采用二次迁移学习的ResNet101预训练网络为特征提取器构建Faster-RCNN目标检测模型,分别实现了对宫颈阴道镜图像HSIL单一目标和LSIL、HSIL和癌变三类目标的病变区域定位及识别。证实了基于专业医学背景和大样本数据,深度学习技术不仅可辅助阴道镜诊断分类,且在病灶定位上亦可取得较好效果。

自20世纪60年代阴道镜在全球广泛应用至今,影响其准确性和可重复性的主要原因之一是缺乏统一规范的图像评价标准。尽管多种阴道镜评分系统曾被运用,如Reid法、改良Reid法以及Swede法等,但其诊断效能却一直不理想。为此,IFCPC于2011年发布了最新阴道镜术语系统,对阴道镜下上皮和血管的边界、轮廓、形态等重要解剖部位或组织的图像特征进行了全面解析和科学化归类。本小组前期曾对IFCPC阴道镜术语体系进行了前瞻性大样本临床比较研究以及多项回顾性队列研究,证实了其较好的临床诊断效能,为其实际运用奠定了重要的循证基础<sup>[7-8]</sup>。本研究借助该专业阴道镜背景优势,对阴道镜下上皮及血管特征进行统一认知和解读,基于20类宫颈上皮与血管征象对图像进行标准化精细标注,归并形成分类映射,获得高质量标注数据,辅助计算机识别最具影响的特征群,细化训练目标,从而构建更为稳定可靠的机器学习模型。本研究将不易掌握的复杂描述性分类术语转化为实际且便捷的操作体系,弥补



了既往研究中标注特征量少、信息单一、实验数据量有限等短板。

尽管国内外开展宫颈癌智能化检测的研究已经有十余年,但由于较平面图像更为复杂、部位多变、需较强专业背景支撑的跨学科合作等原因,对于宫颈癌前病灶阴道镜图像的智能化判别,研究成果相对较少。2015年Song等<sup>[15]</sup>结合临床诊断结果和5%醋酸试验后的阴道镜图像特征等多模态数据,提出了基于数据驱动病变特征的提取方法,并针对HSIL和LSIL二分类得到了74%的识别精度。之后有学者对上述特征提取或传统机器学习方法进行了改良,提高识别精度(80.87%和81.3%)<sup>[16-17]</sup>。也有学者尝试探索深度学习技术运用于宫颈病变视觉判别的可行性,Xu等<sup>[10]</sup>利用卷积神经网络对1 000张阴道镜宫颈图像进行训练,结合宫颈细胞学和HPV检测结果,最终模型识别LSIL+的准确性为88.91%。也有学者分别基于485张宫颈阴道镜图像(142张HSIL、257张原位癌、86张浸润癌)和330名患者(97例LSIL、213例HSIL),采用深度学习算法构建了三分类(HSIL、原位癌、浸润癌)和二分类(HSIL和LSIL)模型,分别获得了平均50%的识别精度和82.3%的准确度<sup>[12-14]</sup>。近期,Zhang等<sup>[11]</sup>基于1 709例患者阴道镜宫颈图像,利用预训练DenseNet卷积神经网络对ImageNet和Kaggle数据集的各层参数进行微调,二分类诊断CIN2+准确率为73.08%(AUC $\approx$ 0.75)。除阴道镜图像之外,亦有学者对宫颈的其他图像进行过类似研究。2019年通过对哥斯达黎加宫颈癌筛查人群长达7年的随访,利用固定焦距摄像机进行数字化宫颈摄影,以深度学习技术辅助识别CIN2+病变<sup>[18]</sup>。陆晗<sup>[19]</sup>则提出了一种用醋酸实验前后的宫颈比值图像代替原图的方法,利用VGGNet-16模型对宫颈癌前病变进行分类识别;但均并未取得比醋酸后阴道镜图像明显优越的识别精度。上述研究均在机器辅助宫颈病变识别的方法探索上取得了一定进展。本研究以临床意义为导向,基于国内最大阴道镜中心的大样本数据,通过跨学科合作,综合对比了多种目标检测的深度学习网络架构,并在数据预处理、数据增强、主干网络选择与训练、模型超参数优化、算法改进等几个方面对模型进行优化,构建了HSIL单分类模型和三分类(LSIL、HSIL、癌)Faster-RCNN模型,对HSIL及以上病变的识别获得了

89.8%的召回率(即敏感度)和72.1%的精确度,该模型的诊断效能与本研究中心既往资料中,经IFCPC术语培训的5~10年工作经验的阴道镜专科医师水平相当(敏感度约65%~90%)<sup>[7-8]</sup>。本研究通过选择和输出针对临床实际数据和应用场景的最优模型,获得稳定的分类预测效果,更加贴合国内临床筛查需求。

本研究的另一特点在于对阴道镜下病变的定位识别。既往报道的研究目标多为图像级的病变分类,在病变区域的定位识别方面相关报道极少。曾有个别学者尝试CAM热力图等对宫颈病灶区进行定位识别,但效果十分有限,仅能对白色肿胀且伴粗点状血管的区域作出HSIL判别<sup>[20]</sup>。然而,定位指导活检或指导治疗是阴道镜的主要功能之一,有着无法取代的临床需求。本研究基于对28 975幅不同程度宫颈病变阴道镜图像的精准定位标注,结合深度学习中的目标检测技术,分别实现了对宫颈阴道镜图像HSIL单一目标和LSIL、HSIL和癌变三类目标的病变区域定位识别,当目标检测交并比IOU $\geq$ 0.5的情况下,两种模型mAP分别达到0.82和0.67。结果证实了模型用于辅助阴道镜指导病变区域定位的可行性,不仅适用于临床筛查,亦可指导活检及后续定位治疗。

本研究模型在阴道镜图像的分类和病变定位上均取得了较好成绩。尽管在识别精度上仍有进一步提升的空间,但已显示出其实际临床应用潜能。本研究不足之处是针对三分类区域检测只进行了初步的探索,在数据预处理、数据增强、模型调参、迁移学策略等方面还有较大发挥空间。此外,研究推广仍然面临瓶颈和挑战,如不同医疗机构的阴道镜设备非标准化、计算机硬件及网络支撑不足等,都将影响未来智能阴道镜的实际临床应用。

**作者贡献声明** 李燕云 数据采集,论文构思、撰写和修订。王永明,王振,蔡青青 数据统计。周奇,王珏,孟妍 数据采集。李亦学,隋龙 论文构思和指导。华克勤 论文综合指导。

**利益冲突声明** 所有作者均声明不存在利益冲突。

## 参 考 文 献

[1] 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.
- [2] MCGUIRE S. World Cancer Report 2014. Geneva, Switzerland: World Health Organization, International Agency for Research on Cancer, WHO Press, 2015 [J]. *Adv Nutr*, 2016, 7(2):418-419.
- [3] BORNSTEIN J, BENTLEY J, BÖSZE P, *et al.* 2011 colposcopic terminology of the International Federation for Cervical Pathology and Colposcopy [J]. *Obstet Gynecol*, 2012, 120(1):166-172.
- [4] KHAN MJ, WERNER CL, DARRAGH TM, *et al.* ASCCP colposcopy standards: role of colposcopy, benefits, potential harms, and terminology for colposcopic practice [J]. *J Low Genit Tract Dis*, 2017, 21(4):223-229.
- [5] NAM K. Colposcopy at a turning point [J]. *Obstet Gynecol Sci*, 2018, 61(1):1-6.
- [6] REMA PN, MATHEW A, THOMAS S. Performance of colposcopic scoring by modified International Federation of Cervical Pathology and Colposcopy Terminology for diagnosing cervical intraepithelial neoplasia in a low-resource setting [J]. *South Asian J Cancer*, 2019, 8(4):218-220.
- [7] LI YY, DUAN XL, SUI L, *et al.* Closer to a uniform language in colposcopy: study on the potential application of 2011 International Federation for Cervical Pathology and Colposcopy Terminology in clinical practice [J]. *Biomed Res Int*, 2017, 2017(1):1-10.
- [8] 李燕云, 张宏伟, 郑瑞莲, 等. 应用2011年版阴道镜术语对子宫颈病变行阴道镜诊断与病理诊断的一致性分析 [J]. *中华妇产科杂志*, 2015, 50(5):361-366.
- [9] KAJI S, KIDA S. Overview of image-to-image translation by use of deep neural networks: denoising, super-resolution, modality conversion, and reconstruction in medical imaging [J]. *Radiol Phys Technol*, 2019, 12(3):235-248.
- [10] XU T. Multimodal deep learning for cervical dysplasia diagnosis [C]. *Proceedings of International Conference on Medical Image Computing & Computer-assisted Intervention*, 2016, Athens: MICCAI.
- [11] ZHANG T, LUO YM, LI P, *et al.* Cervical precancerous lesions classification using pre-trained densely connected convolutional networks with colposcopy images [J]. *Biomed Signal Process Control*, 2020, 55(1):1-11.
- [12] SATO M, HORIE K, HARA A, *et al.* Application of deep learning to the classification of images from colposcopy [J]. *Oncol Lett*, 2018, 15(3):3518-3523.
- [13] MIYAGI Y, TAKEHARA K, MIYAKE T. Application of deep learning to the classification of uterine cervical squamous epithelial lesion from colposcopy images [J]. *Mol Clin Oncol*, 2019, 11(6):583-589.
- [14] MIYAGI Y, TAKEHARA K, NAGAYASU Y, *et al.* Application of deep learning to the classification of uterine cervical squamous epithelial lesion from colposcopy images combined with HPV types [J]. *Oncol Lett*, 2020, 19(2):1602-1610.
- [15] SONG D, KIM E, HUANG X, *et al.* Multimodal entity coreference for cervical dysplasia diagnosis [J]. *IEEE Trans Med Imaging*, 2015, 34(1):229-245.
- [16] XU T, ZHANG H, XIN C, *et al.* Multi-feature based benchmark for cervical dysplasia classification evaluation [J]. *Pattern Recognit*, 2017, 63(1):468-475.
- [17] ASIEDU MN, SIMHAL A, CHAUDHARY U, *et al.* Development of algorithms for automated detection of cervical pre-cancers with a low-cost, point-of-care, pocket colposcope [J]. *IEEE Trans Biomed Eng*, 2019, 66(8):2306-2318.
- [18] HU L, BELL D, ANTANI S, *et al.* An observational study of deep learning and automated evaluation of cervical images for cancer screening [J]. *J Natl Cancer Inst*, 2019, 111(9):923-932.
- [19] 陆晗. 基于深度学习和宫颈图像的癌前病变识别 [D]. 南昌航空大学, 2019.
- [20] 张丹. 基于深度学习的阴道镜 HSIL 检测 [D]. 浙江大学, 2018.
- [21] DARRAGH TM, COLGAN TJ, THOMAS COX J, *et al.* The lower anogenital squamous terminology standardization project for HPV-associated lesions: background and consensus recommendations from the College of American Pathologists and the American Society for Colposcopy and Cervical Pathology [J]. *Int J Gynecol Pathol*, 2013, 32(1):76-115.
- [22] SHAOQING R, KAIMING H, ROSS G, *et al.* Faster R-CNN: towards real-time object detection with region proposal networks [C]. *Montreal: Neural Information Processing Systems*, 2015.
- [23] REDMON J, DIVVALA S, GIRSHICK R, *et al.* You only look once: unified, real-time object detection [C]. *Las Vegas: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2016.
- [24] LIU W, ANGUELOV D, ERHAN D, *et al.* Ssd: single shot multibox detector [C]. *Amsterdam: European Conference on Computer Vision*. Springer, Cham, 2016: 21-37.
- [25] K. HE, X. ZHANG, S. REN, *et al.* Deep residual learning for image recognition [C]. *Las Vegas: CVPR*, 2016.

(收稿日期:2020-07-23; 编辑:沈玲)