

文章编号:1007-2780(2023)12-1698-09

基于改进 YOLOv5s 的手术器械检测与分割方法

孟晓亮, 赵吉康, 王晓雨, 张立晔*, 宋 政

(山东理工大学 计算机科学与技术学院, 山东 淄博 255000)

摘要:在内窥镜手术过程中,外科医师需实时掌握手术器械的位置信息。现有目标检测算法受反光和阴影等因素影响,其准确度和漏检率仍有优化的空间。本文提出一种基于改进 YOLOv5s 的手术器械检测与分割方法。首先,通过 Gamma 校正算法校正图像的亮度和对比度,以解决手术器械的反光和阴影遮挡等问题;其次,设计 CBAM 和动态卷积模块,增加重要特征信息的权重,以进一步提高目标检测的准确度并减少模型的漏检率;同时,优化空间金字塔池化模块以扩大感受野,从而更好地识别多尺度目标;最后,设计 FPN 语义分割头,以实现语义分割功能。在内窥镜手术数据集上的实验结果表明,本文目标检测的 mAP@0.5 为 98.2%,语义分割的 mIoU 为 94.0%。所提方法可辅助外科医师快速掌握手术器械的位置和类型,提高手术效率。

关键词:手术器械;目标检测;语义分割;注意力机制

中图分类号:TP391 文献标识码:A doi:10.37188/CJLCD.2023-0025

Detection and segmentation method of surgical instruments based on improved YOLOv5s

MENG Xiao-liang, ZHAO Ji-kang, WANG Xiao-yu, ZHANG Li-ye*, SONG Zheng

(School of Computer Science and Technology, Shandong University of Technology, Zibo 255000, China)

Abstract: In the process of endoscopic surgery, surgeons need to know the position information of surgical instruments in real time. The existing target detection algorithms are affected by factors such as reflection and shadow, and there is still optimization space for the accuracy and missed detection rate. This paper proposes a detection and segmentation method of surgical instruments based on improved YOLOv5s. Firstly, the brightness and contrast of images are corrected by Gamma correction algorithm to solve the problems of reflection and shadow occlusion of surgical instruments. Secondly, convolutional block attention module (CBAM) and dynamic convolution module are designed to increase the weight of important feature information, which further improves the accuracy of target detection and reduces the missed detection rate of the model. At the same time, the spatial pyramid pooling module is optimized to expand the receptive field, so as to better identify multi-scale targets. Finally, the feature pyramid networks (FPN) semantic segmentation head is designed to realize the semantic segmentation. Experimental results on endoscopic surgery dataset show that the mAP@0.5 of target detection in this paper is 98.2%,

收稿日期:2023-01-31;修订日期:2023-03-05.

基金项目:国家自然科学基金(No. 62001272)

Supported by National Natural Science Foundation of China(No. 62001272)

*通信联系人, E-mail: zhangliye@sdut.edu.cn

and the mIoU of semantic segmentation is 94.0%. The proposed method can assist surgeons to quickly grasp the position and type of surgical instruments, and improve the efficiency of surgery.

Key words: surgical instruments; target detection; semantic segmentation; attention mechanism

1 引 言

基于视觉的内窥镜手术作为一种具有代表性的微创手术,在深度学习^[1-2]时代受到越来越多的关注。与传统的手术不同,微创手术在内窥镜相机的辅助下通过一个小的切口进行,组织缝合需要使用专业的手术器械。然而,缺乏对手术技巧、手术质量等有针对性的、客观的反馈是内窥镜手术过程中存在的关键问题之一。手术过程中手术器械检测与分割的准确性对评估外科医师的手术技能具有重要的意义^[3]。

国外针对手术器械检测与分割的研究相对较早,且近年来已有很多研究者将深度学习成功应用于实际内窥镜手术中。Jin 等人^[4]对手术过程中的手术器械进行检测,采用 Faster-RCNN 算法框架,主干网络使用 VGG-16 卷积神经网络,在内窥镜手术中实现了对手术器械的实时检测,可以辅助评估外科医师的操作技能。Mahmood 等人^[5]提出一种基于双流残差密集网络(DSRD-Net)的手术器械分割方法,利用残差和密集的空间金字塔池化模块优化算法,精准分割手术器械,减少因视野模糊等导致的人为失误,并在公开可用的胃肠内窥镜数据集上进行测试,取得了较好的实验结果。不过此类算法的参数量和运算量较大,如果再添加额外的优化模块,会导致检测和分割速度下降。基于此,文献^[6]提出一种 Anchor-free 的卷积神经网络,通过沙漏网络对内窥镜手术视频进行逐帧检测并将手术器械建模,以点作为边界框的中心点生成检测框,在保证精度的同时,检测速度超过同时期以双阶段目标检测为代表的 Anchor-based 类算法。但是该网络只能生成手术器械的定位框而未实现对手术器械种类的识别,功能性上略有不足。

国内对于深度学习在微创手术领域的研究起步较晚。2018年,郑腾辉等人^[7]基于全卷积神经网络(FCN)实现了对手术器械和手术器械收纳盒的图像分割,识别目标的准确度可达到像素级,但是检测目标为手术器械收纳盒中的静态手

术器械,因此实用性上有所欠缺。2019年,Ni 等人^[8]提出一种新的网络 Residual Attention U-Net (RAUNet)用于白内障手术中的手术器械语义分割,该网络通过编码-解码结构获取高分辨率的掩膜,同时关注高阶和低阶特征图,从而在保证获取重要特征的基础上有效过滤干扰信息。2022年,Ni 等人^[9]为解决内窥镜手术中照明因素对手术器械的影响,提出了一种新的网络 Surgi-Net,通过添加金字塔注意力模块捕捉多尺度特征,通过自蒸馏消除其他的干扰信息。文献^[10]提出一种改进的 DoubleUNet^[11]分割方法,通过对图像进行预处理,并使用密集连接空洞空间卷积池化金字塔(DenseASPP)模块^[12]和 Focal Tversky Loss 损失函数进一步提取图像特征。该方法可以辅助外科医师切除结肠异常组织,从而降低息肉癌变的概率。

综上所述,基于深度学习的手术器械检测与分割得到很多研究者的关注。虽然 RAUNet 网络已经考虑手术器械反光的问题,但在实际的内窥镜手术中,除手术器械金属材质自身导致的反光问题外,组织器官之间的相互遮挡、光线影响以及阴影遮挡等问题^[13]也需要被考虑。文献^[14]已经证实通过对 YOLO 算法的 Neck 部分的下采样进行优化和添加注意力机制,可以在保证模型轻量化的同时进一步提高检测的精度。但目前 YOLO 算法应用于内窥镜手术的研究相对较少,且目标检测中存在的漏检问题仍然有改进的空间。

针对以上问题,本文提出一种改进的 YOLOv5s 手术器械检测与分割方法。对于内窥镜图像中存在的反光和阴影遮挡等问题,使用 Gamma 校正算法优化图像亮度和对比度。通过添加 CBAM(Convolutional Block Attention Module)^[15]、动态卷积以及空间金字塔池化模块以增强目标特征信息的权重,进一步提高手术器械检测的准确度并降低漏检率。最后,在模型的 Neck 部分添加基于特征金字塔网络(FPN)的语义分割分支,同时实现手术器械的检测和语义分割功能,辅助外科医师提高手术效率,减小手术难度。

2.1 Gamma校正

在基于深度学习的手术器械检测中,图像的亮度会直接影响模型训练的效果。内窥镜视频中出现的手术器械本身多为金属材质,在手术环境下极易出现反光或阴影遮挡等情况,从而影响目标检测的效果。因此,需要在模型训练前对图像进行预处理。本文采用 Gamma 校正算法,对内窥镜图像进行预处理。

Gamma 校正算法的公式如式(1)所示:

$$F = I^r \quad (1)$$

如图 2 所示,红色和蓝色曲线分别代表灰度值过高和过低的情况,当 $r < 1$ 或 $r > 1$ 时, Gamma 算法会对图像灰度值进行调整。对于图像中灰度值较高的区域,经过校正后灰度值会减小,使 r 的值接近于 1;对于图像中灰度值较低的区域,经过校正后灰度值会增大,使 r 的值接近于 1。

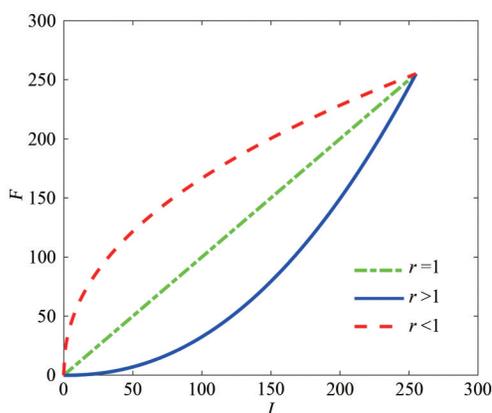


图 2 Gamma 校正算法曲线图

Fig. 2 Graph of Gamma correction algorithm

2.2 CBAM 与动态卷积模块

原有算法特征图信息中每个通道的权重相等,这会导致一些无用信息的比重偏大,而关键信息的比重偏小。因此,我们在 Backbone 底部添加注意力机制模块 CBAM,通过学习的方式获取每个通道的比重并生成权重。CBAM 可同时应用空间和通道两个不同的注意力维度。在语义分割部分,由于上采样会导致部分语义信息丢失,因此,在语义分割头的 C3 部分加入 CBAM,通过通道注意力模块以减少重要信息的丢失,可防止由于单独添加 CBAM 模块导致的网络层数过多的问题。图 3 为 CBAM 的结构图。

另外,将 Neck 部分的 Conv 模块更换为动态

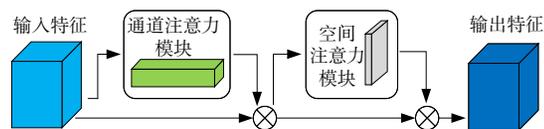


图 3 CBAM 结构

Fig. 3 Structure of CBAM

卷积 ODConv 模块^[19]。注意力机制可赋予权重,通过注意力权重进行学习可以提高 CNN 网络的准确性并减少计算量。普通的动态卷积仅关注卷积核数量,而其他 3 个参数(输入通道数、输出通道数、空间大小)容易被忽略。本文使用的 ODConv 模块可通过对 4 个维度的并行策略实现对内核空间互补注意力的学习,从而进一步提高模型的准确度。

2.3 空间金字塔池化模块优化

YOLOv5s 使用的空间金字塔池化模块是 SPPF (Spatial Pyramid Pooling Fast),在保持 SPP 效果的同时提高其速度。本文使用 SPPC-SPPC 模块^[20](SPP-CSP-Concat),该模块采用 CSP 结构^[21],通过保留 SPP 不同尺度的最大池化层使模型视野进一步扩大,可解决特征重复提取的问题。同时,可建立一条新的卷积分支,将原图通过卷积核后,与通过最大池化层得到的特征图进行二次合并,可有效减少池化后的特征损失,从而进一步提高检测的准确度。

2.4 FPN 语义分割头

图 1 中,在 FPN 和 PAN 的连接位置增加语义分割头,使 80×80 的特征图继续按照 FPN 特征金字塔结构的上采样步骤进行,实现语义分割功能^[22]。在生成 80×80 的特征图后形成两个分支,一个分支进入 PAN 模块,另外一个分支继续采用自顶向下的上采样结构实现语义分割。因此,本文模型可同时实现目标检测和语义分割。

3 实验与结果分析

3.1 数据集

本文数据集采用 Cholec80,该数据集中包含大量的内窥镜手术视频,我们选取 Cholec80 中的 20 个内窥镜手术视频并按照一定的帧频转换成内窥镜图像序列,将其中不存在手术器械或者手术器械较为模糊的图像剔除。数据集包含 4 000 幅图像,其中 3 000 幅图像用于训练,1 000 幅

图像用于测试。本文将数据集中的手术器械分为 7 类,分别标注为 Grasper、Hook、Clipper、Scissors、SpecimenBag、Bipolar 和 Irrigator。

由于模型需要同时实现目标检测和语义分割功能,本文分别通过 LabelImg 和 Labelme 制作两种不同格式的标签。目标检测标签设定为 txt 格式,语义分割标签设定为 png 格式记录掩膜,并对数据集通过 Gamma 校正算法进行调整优化。通过实验发现,如果数据集中不同类别手术器械的数量相对平均时,Grasper 类型的手术器械识别的准确度明显低于其他 6 类手术器械。因此,我们在数据集中提高了 Grasper 类型的比例,将 Grasper 类型的图像数量提高至 1 000 幅,其余 6 类手术器械各为 500 幅图像。

图 4 为数据集样本的部分图像。由图 4(a)可以看出,数据集中手术器械受环境的影响较大,导致图像中亮度分布不均,即使是同一幅图像,手术器械的金属部分反光强烈导致亮度过高,而周围环境和手术器械的其余部分相对较暗。因此,通过 Gamma 校正算法,可使数据集中的图像亮度分布相对均匀,减少外界因素对目标检测的影响,如

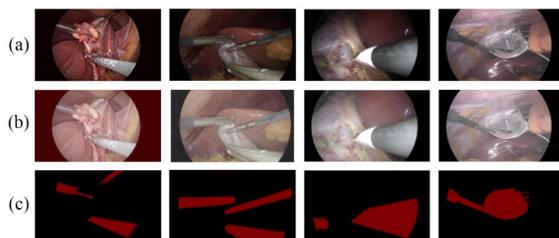


图 4 数据集样本。(a) 原图;(b) Gamma 校正后结果;(c) 分割掩膜。

Fig. 4 Dataset sample. (a) Original image ;(b) Results after Gamma correction ;(c) Segmentation mask.

图 4(b)所示。图 4(c)为分割掩膜结果。

3.2 训练环境

本文的模型训练在 Windows 10 操作系统上进行,模型由 Python 编程语言实现,编程软件采用 PyCharm。采用 CPU 进行模型训练相对缓慢,因此,我们通过 GPU 对网络框架进行模型训练,GPU 的硬件配置如表 1 所示。

表 1 GPU 硬件配置

Tab. 1 Hardware configuration of GPU

名称	配置
处理器	AMD EPYC 7642 Processor
GPU 型号	RTX 3090
显存容量	24 GB
内核数量	24 核

3.3 实验结果分析

3.3.1 模型衡量指标

本文需要同时衡量目标检测和语义分割的效果。对于目标检测,本文选用 mAP@0.5 和损失作为模型的衡量指标;语义分割通过 mIoU 进行衡量。最后,通过测试集验证实际检测和分割的效果。

3.3.2 目标检测效果分析

对于目标检测而言,召回率 (Recall) 和精确度 (Precision) 是衡量目标检测结果的两个非常重要的指标。AP 为通过计算每一个召回率对应精确度的值并进行平均得到的一个衡量标准;mAP 是将目标检测中所有种类的 AP 取平均值的结果。本文将所提方法与其他目标检测算法的 AP 和 mAP 进行比较,结果如表 2 所示。

表 2 不同方法的 AP 和 mAP 比较

Tab. 2 Comparison of AP and mAP with different methods

	本文方法	YOLOv5s	Faster-RCNN	SSD	RetinaNet
Grasper AP	0.963	0.929	0.785	0.906	0.913
Hook AP	0.995	0.995	0.971	0.999	0.999
Clipper AP	0.994	0.975	0.867	0.999	0.972
Scissors AP	0.961	0.947	0.893	0.948	0.948
SpecimenBag AP	0.972	0.953	0.896	0.932	0.971
Bipolar AP	0.995	0.955	0.893	0.957	0.952
Irrigator AP	0.995	0.996	0.771	0.911	0.925
mAP@0.5	0.982	0.964	0.868	0.950	0.954

由表 2 可以看出,无论是在不同手术器械的单项 AP 上还是整体的 mAP 上,本文方法都要优于其他方法。相较于 YOLOv5s,优化后算法在本文手术器械数据集上的 mAP@0.5 提高了 1.8%,达到了 98.2%。在单阶段检测方法中,本文算法在 mAP@0.5 上仍然具有优势,相较于目前流行的 SSD 和 RetinaNet 算法,mAP 分别高出 3.2% 和 2.8%;相较于经典的双阶段检测方法 Faster-RCNN,本文算法的 mAP@0.5 提高了

11.4%,验证了本文方法在 AP 上的优势。

将本文所提方法和 YOLOv5s 方法的损失曲线进行对比分析,结果分别如图 5 和图 6 所示。

从图 5 和图 6 可以看出,对于 bounding box 检测框损失,本文所提方法和 YOLOv5s 方法都接近 0.02。对于 obj 目标检测损失,最终两者在数值上比较接近。但在验证集 val 的 obj 损失曲线上,YOLOv5s 在降低到 0.012 后缓慢回升,并在

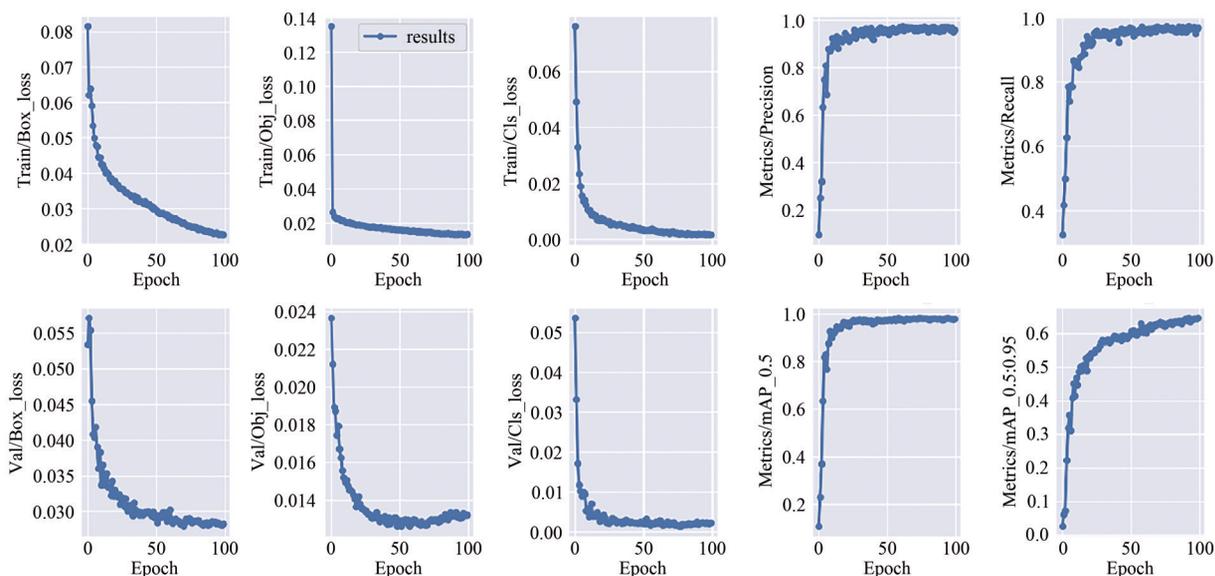


图 5 本文所提方法的损失曲线

Fig. 5 Loss curve of the proposed method

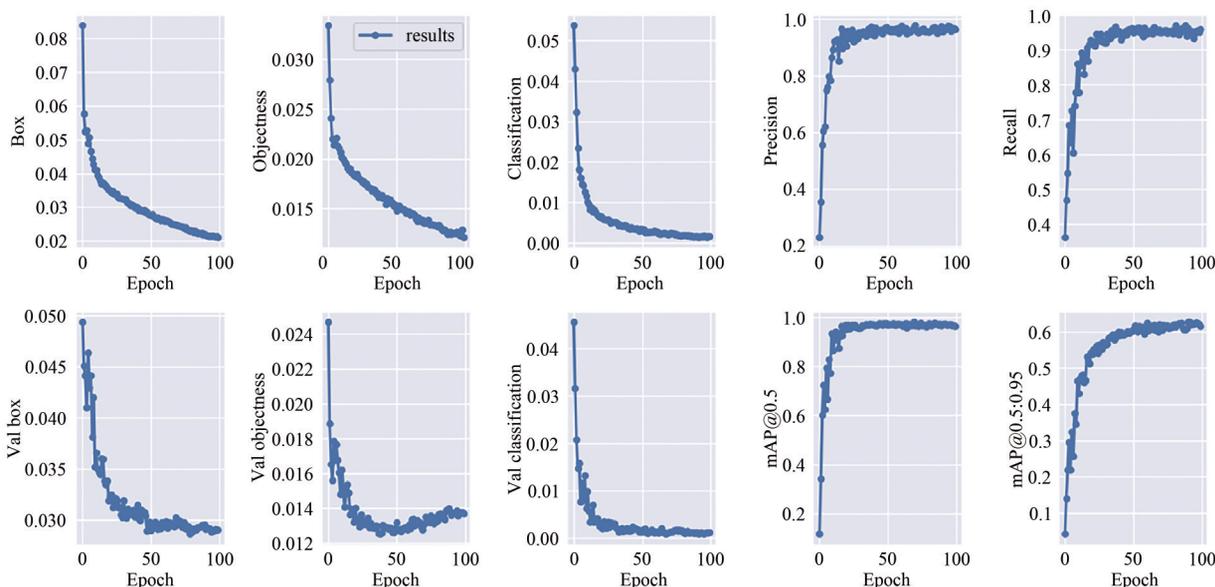


图 6 YOLOv5s 方法的损失曲线

Fig. 6 Loss curve of the YOLOv5s method

0.014附近上下波动;而本文所提方法在降低到0.012后,一直较为平稳的维持在0.014以下。从曲线和数据上可以看出,YOLOv5s在测试集val上的损失曲线更加平滑,当epoch为100时,训练集损失曲线也趋于平稳。本文所提方法由于引入语义分割模块,训练集train的obj损失曲线和测试集val的损失曲线几乎趋近于 90° ,在训练集train最终损失值相差不大的情况下,本文所提方法的损失曲线下降速度最快,在进入平稳期后波动也最小,不会像YOLOv5s一样出现明显的上下波动。在验证集val上,本文所提方法的损失曲线无论是在训练的波动幅度还是最终损失值上都优于YOLOv5s方法,进一步验证了本文

方法的有效性和优势。

3.3.3 语义分割效果分析

mIoU是评价语义分割效果的重要指标^[23]。本文通过添加基于FPN网络的语义分割头实现了语义分割功能,目标检测算法的mIoU通过矩形框计算。为更好地比较分析语义分割的效果,本文在相同条件下训练其他语义分割网络并与本文方法进行比较。由于目标检测部分已经实现对于手术器械的分类功能,语义分割只需将手术器械和背景进行区分即可,因此语义分割的类别这里只定义Instruments和Background两类,将所提方法的mIoU值与其他语义分割方法进行比较,结果如表3所示。

表3 不同方法的mIoU值比较

Tab.3 Comparison of mIoU with different methods

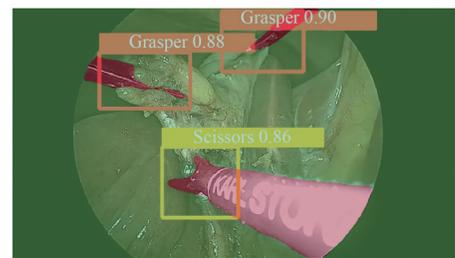
	本文方法/%	PSPNet/%	FCN/%	Deeplabv3/%	UNet/%
mIoU	94.0	93.1	74.1	91.8	92.7
Background	99.3	98.5	94.5	98.4	98.3
Instruments	88.7	87.7	53.6	85.2	87.0

由表3可以看出,本文方法在语义分割方面有着良好的表现,mIoU值达到了94.0%,远高于经典分割网络FCN。相较于Deeplabv3和UNet图像分割网络,本文方法的mIoU分别高出2.2%和1.3%。相较于目前流行的语义分割网络PSPNet,mIoU提高了0.9%,进一步验证了本文方法的有效性和优势。

3.3.4 检测与分割结果

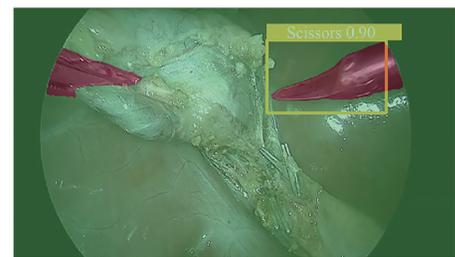
本文选取数据集中具有代表性的两帧图像进行检测与分割结果分析,结果如图7所示。

由图7(a)可以看出,本文方法能够较好地实现手术器械的检测和语义分割,在多个手术器械同时存在的情况下也能够进行准确识别。但也会存在漏检的情况,如图7(b)所示,相较于被检测出的Scissors类型的手术器械,漏检的手术器械在图像中所占比例相对较小,且受到手术环境的遮挡,导致模型提取的特征信息不足,被检测器误认为是背景信息,因此出现了漏检现象。但是语义分割模块却能够很好地识别该手术器械并进行准确地分割,在一定程度上弥补了目标检测的漏检问题。



(a) 多器械检测与分割结果

(a) Results of multi-instruments detection and segmentation



(b) 漏检结果

(b) Missed detection result

图7 手术器械检测与分割结果

Fig.7 Detection and segmentation results of surgical instruments

4 结 论

本文采用改进的YOLOv5s结合FPN的方

法实现了手术器械的检测与语义分割功能。通过 Gamma 校正算法,减轻反光和阴影等因素的影响;通过 CBAM、动态卷积模块、空间金字塔池化模块和 FPN 语义分割模块,提高检测与分割的准确度,并降低漏检率。实验结

果表明,手术器械的检测与语义分割效果良好,手术器械检测的 mAP@0.5 为 98.2%,语义分割的 mIoU 为 94.0%。所提方法可辅助外科医师提高手术效率,减小手术难度,实现术中精准导航。

参 考 文 献:

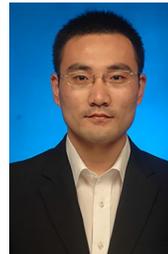
- [1] RIVENSON Y, WU Y C, OZCAN A. Deep learning in holography and coherent imaging [J]. *Light: Science & Applications*, 2019, 8: 85.
- [2] ZUO C, QIAN J M, FENG S J, *et al.* Deep learning in optical metrology: a review [J]. *Light: Science & Applications*, 2022, 11(1): 39.
- [3] FU Z M, JIN Z Y, ZHANG C G, *et al.* The future of endoscopic navigation: a review of advanced endoscopic vision technology [J]. *IEEE Access*, 2021, 9: 41144-41167.
- [4] JIN A, YEUNG S, JOPLING J, *et al.* Tool detection and operative skill assessment in surgical videos using region-based convolutional neural networks [C]//*Proceedings of 2018 IEEE Winter Conference on Applications of Computer Vision*. Lake Tahoe, USA: IEEE, 2018: 691-699.
- [5] MAHMOOD T, CHO S W, PARK K R. DSRD-Net: dual-stream residual dense network for semantic segmentation of instruments in robot-assisted surgery [J]. *Expert Systems with Applications*, 2022, 202: 117420.
- [6] LIU Y Y, ZHAO Z J, CHANG F L, *et al.* An anchor-free convolutional neural network for real-time surgical tool detection in robot-assisted surgery [J]. *IEEE Access*, 2020, 8: 78193-78201.
- [7] 郑腾辉,陶青川. 基于全卷积神经网络的手术器械图像语义分割算法[J]. *现代计算机*, 2019(9):80-84.
ZHENG T H, TAO Q C. Semantic segmentation algorithm for images of surgical instruments based on fully convolutional networks [J]. *Modern Computer*, 2019(9): 80-84. (in Chinese)
- [8] NI Z L, BIAN G B, ZHOU X H, *et al.* RAUNet: residual attention U-Net for semantic segmentation of cataract surgical instruments [C]//*Proceedings of the 26th International Conference on Neural Information Processing*. Sydney, Australia: Springer, 2019: 139-149.
- [9] NI Z L, ZHOU X H, WANG G A, *et al.* SurgiNet: pyramid attention aggregation and class-wise self-distillation for surgical instrument segmentation [J]. *Medical Image Analysis*, 2022, 76: 102310.
- [10] 徐昌佳,易见兵,曹锋,等. 采用 DoubleUNet 网络的结直肠息肉分割算法[J]. *光学精密工程*, 2022, 30(8): 970-983.
XU C J, YI J B, CAO F, *et al.* Colorectal polyp segmentation algorithm using DoubleUNet network [J]. *Optics and Precision Engineering*, 2022, 30(8): 970-983. (in Chinese)
- [11] JHA D, RIEGLER M A, JOHANSEN D, *et al.* DoubleU-Net: a deep convolutional neural network for medical image segmentation [C]//*Proceedings of the 2020 IEEE 33rd International Symposium on Computer-Based Medical Systems*. Rochester, USA: IEEE, 2020: 558-564.
- [12] YANG M K, YU K, ZHANG C, *et al.* DenseASPP for semantic segmentation in street scenes [C]//*Proceedings of 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition*. Salt Lake City, USA: IEEE, 2018: 3684-3692.
- [13] PAN J J, LIU W M, GE P, *et al.* Real-time segmentation and tracking of excised corneal contour by deep neural networks for DALK surgical navigation [J]. *Computer Methods and Programs in Biomedicine*, 2020, 197: 105679.
- [14] 樊新川,陈春梅. 基于 YOLO 框架的轻量化高精度目标检测算法[J]. *液晶与显示*, 2023, 38(7):945-954.
FAN X C, CHEN C M. Lightweight and high-precision object detection algorithm based on YOLO framework [J]. *Chinese Journal of Liquid Crystals and Displays*, 2023, 38(7): 945-954. (in Chinese)
- [15] WOO S, PARK J, LEE J Y, *et al.* CBAM: convolutional block attention module [C]//*Proceedings of the 15th European Conference on Computer Vision*. Munich, Germany: Springer, 2018: 3-19.

- [16] OZAWA T, HAYASHI Y, ODA H, *et al.* Synthetic laparoscopic video generation for machine learning-based surgical instrument segmentation from real laparoscopic video and virtual surgical instruments [J]. *Computer Methods in Biomechanics and Biomedical Engineering: Imaging & Visualization*, 2021, 9(3): 225-232.
- [17] ELFWING S, UCHIBE E, DOYA K. Sigmoid-weighted linear units for neural network function approximation in reinforcement learning [J]. *Neural Networks*, 2018, 107: 3-11.
- [18] ELHASSAN M A, YANG C H, HUANG C X, *et al.* SPFNet: subspace pyramid fusion network for semantic segmentation [J/OL]. *arXiv*, 2022: 2204.01278.
- [19] LI C, ZHOU A J, YAO A B. Omni-dimensional dynamic convolution [C]//*Proceedings of the 10th International Conference on Learning Representations*. Madrid, Spain: OpenReview. net, 2022, DOI: 10.48550.
- [20] WANG C Y, BOCHKOVSKIY A, LIAO H Y M. YOLOv7: trainable bag-of-freebies sets new state-of-the-art for real-time object detectors [C]//*Proceedings of 2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition*. Vancouver, Canada: IEEE, 2023: 7464-7475.
- [21] WANG C Y, LIAO M H Y, WU Y H, *et al.* CSPNet: a new backbone that can enhance learning capability of CNN [C]//*Proceedings of 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition*. Seattle, USA: IEEE, 2020: 1571-1580.
- [22] YU L T, WANG P C, YU X Y, *et al.* A holistically-nested U-Net: surgical instrument segmentation based on convolutional neural network [J]. *Journal of Digital Imaging*, 2020, 33(2): 341-347.
- [23] SUN Y W, PAN B, FU Y L. Lightweight deep neural network for real-time instrument semantic segmentation in robot assisted minimally invasive surgery [J]. *IEEE Robotics and Automation Letters*, 2021, 6(2): 3870-3877.

作者简介:



孟晓亮(1988—),男,山东潍坊人,博士,讲师,2018年于哈尔滨理工大学获得博士学位,主要从事视觉检测与图像处理、深度学习方面的研究。E-mail: xiaoliang@sdut.edu.cn



张立晔(1986—),男,山东济南人,博士,副教授,2018年于哈尔滨工业大学获得博士学位,主要从事计算机视觉、机器学习方面的研究。E-mail: zhangliye@sdut.edu.cn