

基于改进的 Faster R-CNN 的息肉 目标检测和分类方法

杨昆^{1,2,3}, 原嘉成¹, 高聪⁴, 孙宇锋¹, 路宇飞¹, 常世龙¹, 薛林雁^{1,2,3}

(1. 河北大学 质量技术监督学院, 河北 保定 071002; 2. 计量仪器与系统国家地方联合工程研究中心,
河北 保定 071002; 3. 河北省新能源汽车动力系统轻量化技术创新中心, 河北 保定 071002;
4. 河北远东通信系统工程有限公司, 河北 石家庄 050200)

摘要:为了解决结肠镜下腺瘤性息肉和增生性息肉不易分型的问题,提出一种基于改进的 Faster R-CNN 的目标检测及息肉分类模型.在数据预处理阶段,对原有的 2 426 张息肉图像(1 582 张腺瘤性息肉图像,844 张增生性息肉图像)通过 2 种方式进行图像增强,并且通过改进的特征提取、边界框回归以及非极大值抑制的网络,用 602 张图像(386 张腺瘤性图像,216 张增生性息肉图像)进行测试.通过实验证明,在交并比(IoU)取 0.5 时,获得了 86.8% 的平均精度均值,相较于改进之前提升了 2.3%.实验结果验证了该模型的潜在临床应用价值.

关键词:息肉;目标检测;分类;Faster R-CNN

中图分类号:TP192.7

文献标志码:A

文章编号:1000-1565(2023)01-0103-10

Object detection and classification of polyps based on improved Faster R-CNN

YANG Kun^{1,2,3}, YUAN Jiacheng¹, GAO Cong⁴, SUN Yufeng¹, LU Yufei¹,
CHANG Shilong¹, XUE Linyan^{1,2,3}

(1. College of Quality and Technical Supervision, Hebei University, Baoding 071002, China;
2. National & Local Joint Engineering Research Center of Metrology Instrument and System,
Baoding 071002, China; 3. New Energy Vehicle Power System Lightweight Technology
Innovation Center of Hebei Province, Baoding 071002, China; 4. Hebei Far East
Communication System Engineering Co., Ltd, Shijiazhuang 050200, China)

Abstract: In order to solve the problem of difficult classification of adenomatous polyps and hyperplastic polyps under colonoscopy, an object detection and polyp classification model based on improved Faster R-CNN was proposed. Polyp images (1582 adenomatous polyp images, 844 hyperplastic polyp images) were image-enhanced in two ways, and 602 images (386 adenomatous images, 216 hyperplastic

收稿日期:2022-02-12

基金项目:河北省自然科学基金资助项目(A2011201155);河北大学校长科研基金资助项目(XZJJ201914);河北大学多学科交叉研究项目(DXK201914)

第一作者:杨昆(1976—),男,河北保定人,河北大学教授,博士生导师,博士,主要从事生物医学图像处理方向研究.

E-mail: yangkun@hbu.edu.cn

通信作者:薛林雁(1981—),女,河北广平人,河北大学副教授,博士,主要从事生物医学图像处理方向研究.

E-mail: lyxue@hbu.edu.cn

polyp images) were tested. It is proved by experiments that when the intersection-union ratio (IoU) is set to be 0.5, an average precision of 86.8% is obtained, which is 2.3% higher than that before the improvement. The experimental results verified the potential clinical application value of this model.

Key words: polyps; object detection; classification; Faster R-CNN

结直肠癌(CRC)是一种全球性的高发性、高死亡率的疾病,在 2020 年的世界卫生组织癌症报告统计中,结直肠癌在全世界的发病率排名第三,死亡率排名第二^[1]. 大部分结直肠癌由腺瘤性息肉演变而来,这一过程被称作结直肠腺瘤-癌序列^[2]. 结直肠息肉是隆起于结直肠黏膜表面的肿物,其病理类型包括腺瘤性息肉、增生性息肉和其他息肉^[3]. 临床上腺瘤是结直肠癌的癌前状态,因此,早期检测并切除腺瘤是预防结直肠癌的有效手段. 而针对非癌变性的增生性息肉,美国胃肠内窥镜协会^[4]建议保留直肠乙状结肠的增生性息肉,无需采样和内窥镜切除. 此外,研究表明对于微小的非肿瘤性病变,内镜下切除或活检会增加不必要的黏膜损伤和医疗负担^[5].

目前,结直肠息肉的最佳筛查方法仍然是结肠镜检查,但是通过内窥镜成像预测息肉的类型仍然不可靠^[6]. Sharma 等^[7]通过白光成像预测息肉类型的准确率为 45%,虽然作为虚拟色素内窥镜之一的窄带成像(NBI)能够有效地提高结直肠病变组织学诊断的准确性,但是仍然达不到病理诊断的标准,另一项研究中^[8],12 名专业的胃肠病学专家经过培训后进行评估,只有 3 名通过光学活检得到了满意的结果. 息肉体积小,不易识别,且在人工诊断的过程中会受到多种负面因素的影响,例如医生的经验不足、疲劳、诊断时间的限制等,因此在结直肠病变的临床诊断中仍然有较高的漏诊率和误诊率.

随着深度学习的发展,人工智能(AI)在辅助诊断医学图像方面具有很高的潜力^[9]. 在内窥镜图像下应用 AI 算法,可提高息肉的分类和检测精度,早期的机器学习方法通过提取息肉的颜色^[10]、纹理^[11]、形状^[12]、局部二值模式^[13]以及它们的组合特征来进行目标检测,随后,由于受到人工神经网络的启发,深度学习作为机器学习的一个子领域开始兴起,不需要人工设计提取息肉的特征提取器,可以从原始数据中自动学习获得特征. 例如,Paçal 等^[14]采用改进的 YOLO 系列网络对息肉进行目标检测,在公开数据集 ETIS-Larib^[15]中获得了 91.62%的精度. 除此之外,已有利用卷积神经网络(convolutional neural network, CNN)对息肉进行分类的研究,Komeda 等^[16]利用 CNN 分类腺瘤和非腺瘤性息肉,10 次交叉验证的准确率为 75.1%. 在另一项研究中 Byrne 等^[17]同样基于 CNN 模型,以 98%的灵敏度(sensitivity)、83%的特异度(specificity)和 90%的阳性预测值(PPV),高精度的区分腺瘤和增生性息肉. 由于缺少具有息肉类别信息并标注的公共数据集,以上的研究在息肉的检测和分类中虽然取得了突破性的进展,但是缺少在目标检测的同时对息肉进行分类的研究.

本研究构建了从河北大学附属医院和 Colonoscopic Dataset^[18]采集的腺瘤和增生性息肉的数据集,使用了 Faster R-CNN^[19]结合了特征金字塔(feature pyramid networks, FPN)^[20]作为基础网络架构,然后分别对主干网络的特征提取、目标检测框的回归和非极大抑制(non-maximum suppression, NMS)进行了改进,使特征图组间获取了不同权重的特征,丰富了特征图的信息,并且提取目标检测框的每条边的位置信息,使边界框回归更加准确,再通过改变非极大值抑制的方式,较大程度地保留了相邻息肉的目标检测框. 本文提出的改进的 Faster R-CNN 网络有效的提升了 2 种类别息肉的检测精度和查全精度,提升了整体模型的检测性能.

1 实验方法

1.1 Faster R-CNN 网络

Faster R-CNN 是在 R-CNN、Fast R-CNN 基础上加以改进,替换了选择性搜索(selective search, SS)算法,引入了区域候选网络(region proposal network, RPN). RPN 是一个全卷积网络,具体分为 2 个分支,一部分是通过 SoftMax 分类器来判断锚框(anchors)属于正样本或者负样本,另一部分利用边界框回归(bounding box regression)修正 anchors 获得精确的候选区域(region proposal). 原始的 Faster R-CNN 在特征提取阶段运用了 VGGNet^[21],VGGNet 会耗费更多计算资源,并且随着网络深度的增加使得参数量暴涨,

导致更多的内存占用,产生梯度消失现象.在 ResNet^[22] 提出之后,更多地研究采用了 ResNet 作为 Faster R-CNN 的 backbone,并且加入了特征金字塔网络 FPN,来检测不同尺度的目标,使顶层特征通过上采样与低层的特征进行融合,产生了多尺度的特征融合信息,然后分别通过主干网络与 FPN 中提取的原始特征图和 RPN 产生的候选区域后进行感兴趣区域池化(ROI Pooling)操作,再通过全连接层进行分类和边界框回归操作.

1.2 改进的 Faster R-CNN 网络

图 1 显示了改进 Faster R-CNN 的腺瘤与增生性息肉的检测与分类的算法框架,本文改进的 Faster R-CNN 网络使用了 ResNeSt^[23] 作为主干网络,实现了跨通道之间的信息交互,使提取到的息肉特征信息更全面.然后利用 ResNeSt 与 FPN 来提取特征,并且使用了感兴趣区域对齐(ROI Align)替换了原始的 ROI Pooling,在 ROI Align 之后引入边缘感知特征的边界定位(side-aware boundary localization, SABL)模块^[24] 对预测边界框进行回归,传统的边界框回归只是关注于预测的中心点和边界框的尺寸,SABL 模块的回归更关注于边界框的每条边,使边界框的每条边与目标边界对齐,更精准地检测息肉的位置信息.并且针对结肠镜下较多息肉相邻的情况,将 NMS 替换为 Soft-NMS^[25],通过降低不符合规定阈值的被检框的置信度,能较大程度地保留相邻息肉的目标检测信息.

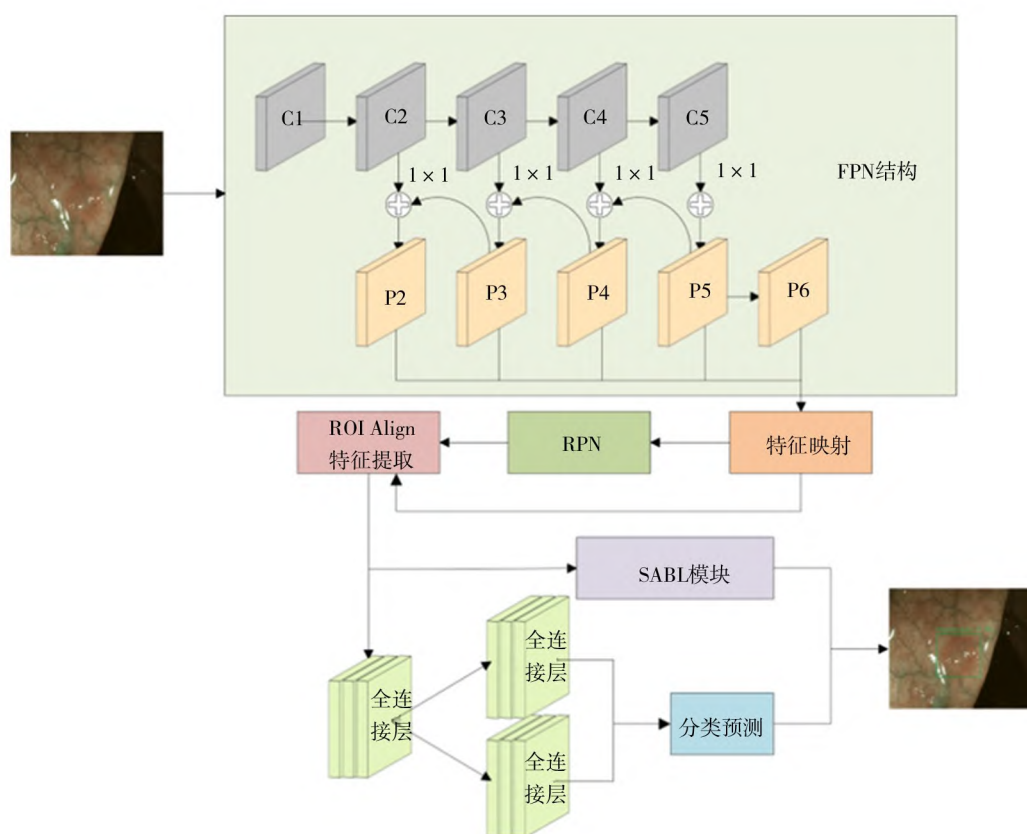


图 1 对于腺瘤与增生性息肉改进的 Faster R-CNN 算法框架

Fig. 1 Improved Faster R-CNN algorithm framework for adenomas and hyperplastic polyps

1.3 主干网络

原始的 Faster R-CNN 主干网络 VGGNet 逐渐被 ResNet 替代,从 ResNet、ResNeXt^[26] 到 ResNeSt,瓶颈模块作为卷积神经网络的基本构建模块进行了改进. ResNet 瓶颈模块中的 1×1 卷积的作用是对通道先降维再升维,在 3×3 的卷积核提取特征的过程中以较低维度进行卷积运算,减少了计算量,提高了计算效率. ResNeXt 模块采用了分组卷积的思想,将输入通道划分为不同组别,增加了路径数量,分别对不同的路径进行卷积核的运算,最后将输出汇总求和.

如图 2 中, ResNeSt 基本模块架构借鉴了 ResNeXt 的思想, 沿通道维度将输入分为 K 个分支, 第 k 个分支记为 Cardinal k , 然后将每个 Cardinal 更细粒度的分为 R 个分组, 第 r 个子组表示为 Split r , 因此原始的输入特征图沿着通道维度共分为 $G = K \times R$ 组. 利用分割注意力(split-attention)计算每个 Split 的权重, 再融合后作为每个 Cardinal 的输出, 然后将输出的特征图在通道的维度拼接到一起, 最后将 ResNeSt 模块的原始输入特征与其融合作为输出.

每个分支中的分割注意力模块如图 3 所示, 通过计算每一个分支拆分的特征图的权重, 然后再进行融合. 如式(1), 第 k 个分支的组合特征由 r 个子组按元素求和融合得到. 如式(2), 融合后的特征可以通过全局平均池化实现聚合通道维度的全局上下文信息. 如式(3), 然后通过全连接层来自适应地计算每一个子组的权重, 对通道权重向量进行修正. 如式(4), 最后由子组特征图与子组权重进行加权组合, 由此得到每个分支的加权融合特征.

$$\hat{u}^k = \sum_{j=R(k-1)+1}^{Rk} u_j, \quad (1)$$

其中, \hat{u}^k 为第 k 个分支的组合特征, $j \in \{R(k-1)+1, R(k-1)+2, \dots, Rk\}$, $k \in 1, 2, \dots, K$, U_j 为分割注意力中第 j 个输入特征, R 表示子组的个数, K 表示分支的数目, k 表示第 k 个分支.

$$S_c^k = \frac{1}{H \times W} \sum_{i=1}^H \sum_{j=1}^W U_c^k(i, j), \quad (2)$$

其中, H, W 分别代表空间维度的高度和宽度, $U_c^k(i, j)$ 为 \hat{u}^k 的第 c 个通道中位于像素点 (i, j) 的值, s_c^k 为基于全局的池化信息得出的第 c 个通道的每一组权重.

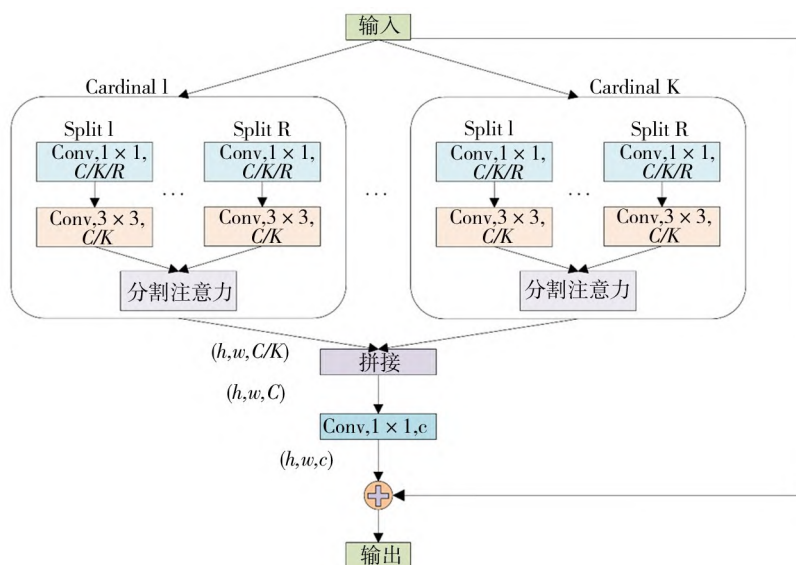


图 2 ResNeSt 模块结构

Fig. 2 ResNeSt module structure

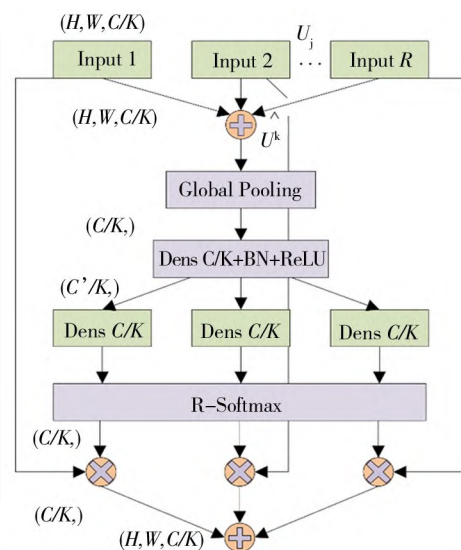


图 3 分割注意力模块结构

Fig. 3 Split-Attention module structure

$$a_i^k(c) = \begin{cases} \frac{\exp(s_i^c(s^k))}{\sum_{j=0}^{RK} \exp(s_i^c(s^k))}, & R > 1, \\ \frac{1}{1 + \exp(-s_i^c(s^k))}, & R = 1, \end{cases} \quad (3)$$

其中, $a_i^k(c)$ 是第 i 个子组的权重, S_i^c 是由 2 个全连接层加上 ReLU 激活函数构成的注意力权重函数.

$$V_c^k = \sum_{i=1}^R a_i^k(c) U_{R(k-1)+i}, \quad (4)$$

其中 V_c^k 表示每个分支第 c 个通道的加权融合特征, $U_{R(k-1)+i}$ 表示第 $R(k-1)+i$ 个子组的特征图.

1.4 特征金字塔

通过 ResNeSt 的最后一层特征进行预测并不能够提取到目标的多尺度信息, 因为低层的特征语义信息

较少,高层的特征语义信息较丰富,但是目标的位置信息较粗略.因此结合 FPN 结构来解决目标检测中多尺度变化问题,如图 1 所示,通过将 ResNeSt 不同层的特征图进行了自上而下(top-down)、横向连接(lateral connection)的形式的融合.首先分别建立在 ResNeSt 的 conv2, conv3, conv4 和 conv5 的输出上,构建了 $\{C2, C3, C4, C5\}$ 4 个尺度的特征,而 4 个对应融合后的特征是 $\{P2, P3, P4, P5\}$. 自上而下路线采用上采样(upsampling)进行,前一层的特征经过 1×1 的卷积核改变通道数,使前一层的特征能够与本层特征进行信息融合,因此,融合后的卷积层既得到了低层特征较高的分辨率,也具有了较高的语义信息,提高了对腺瘤性和增生性息肉的目标检测的准确性.

1.5 边缘感知特征的边界定位

边界框回归用于将附近的候选区域与目标对齐,精确的定位需要对边界框的每条边都精准地放置,然而,原始的 Faster R-CNN 的边界框回归方法只关注于预测中心以及尺寸,忽视了预测边界框的每条边.尤其在锚(anchor)与目标之间存在多种可能的位移时,相较于整体移动框以及修正尺寸,使框的每条边与目标边界对齐更简便.因此,本研究使用基于其周围的上下文信息分别定位的 SABL 模块以取代传统的边界框回归.

SABL 模块的整体框架如图 4 所示,通过 ROI Align 得到的特征图通过边缘感知特征提取方法来提取具有边缘感知的特征,将目标空间分为多个桶区域(buckets),确定边框与 buckets 的位置关系,然后使用分桶(bucketing)边缘定位做初步的粗略估计,根据每个 bucketing 的置信度取平均与先前的分类得分相乘,计算得到 NMS 得分,再通过回归精细化偏移量得到准确的位置信息以及最终的 NMS 得分.

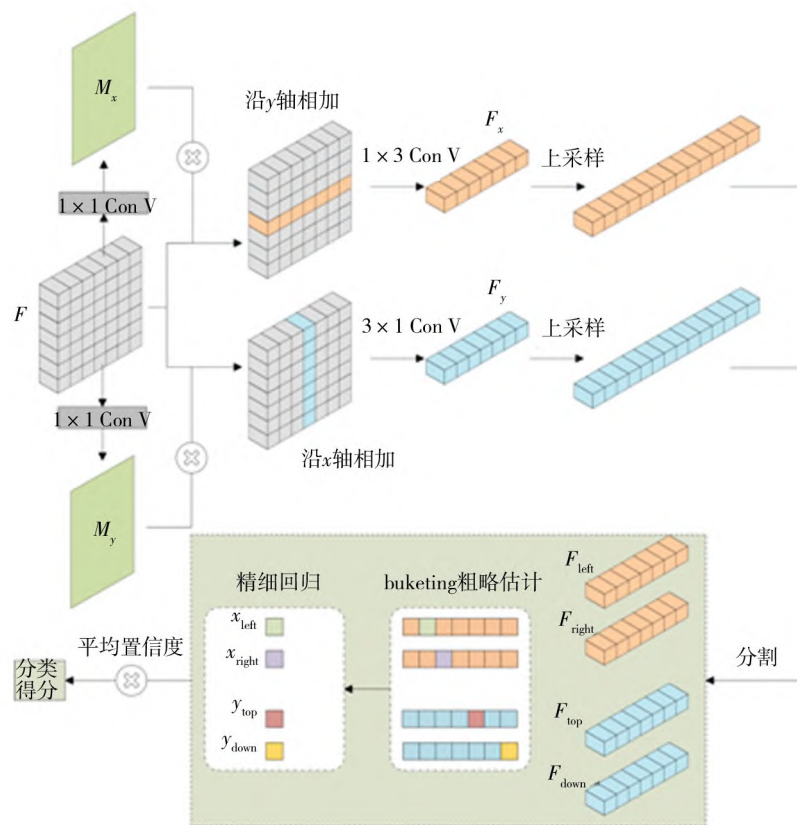


图 4 SABL 模块内部结构

Fig. 4 Internal structure diagram of SABL module

边缘感知特征提取是通过 RPN 提取、ROI Align 操作的特征图来提取边界框的上下左右 4 个位置的预测值,预测左右边界时将特征沿 y 轴叠加,预测上下边界时将特征沿 x 轴叠加,首先将特征图通过 2 个 3×3 的卷积层得到特征 F ,再分别通过 2 个 1×1 卷积进行自注意力信息提取,得到 x 、 y 方向上的特征映射 M_x 、 M_y ,分别将 M_x 、 M_y 作用于 F ,如式(5),得到 x 、 y 轴上的一维特征图 F_x 、 F_y ,经过上采样以及切割后得到

$F_{\text{left}}, F_{\text{right}}, F_{\text{top}}, F_{\text{down}}$.

$$\begin{cases} F_x = \sum_y F(y, :) * M_x(y, :), \\ F_y = \sum_x F(:, x) * M_y(:, x), \end{cases} \quad (5)$$

得到的 4 个边界特征大小是 $1 * 7$ 维, 7 个维度代表 7 个 bucket, 采取二分类的方式判断边界与 bucket 之间的位置关系, 并基于 bucket 的中心回归边界的精细位置, 最终通过调整 NMS 得分使回归边界框能够拥有高精度的定位、高类别的置信度。

如式(6), 通过训练交叉熵损失函数(Cross_entropy loss)用于寻找 bucket 的位置损失 $L_{\text{bucketing}}$.

$$\text{Cross_entropy} = - \sum_{k=1}^N (P_k * \log q_k), \quad (6)$$

其中 p_k 表示真实值, q_k 表示预测值。

如式(7), 并且用 Smooth L1 loss 计算 bucket 和真实边界框位置的损失 L_{reg} .

$$\text{smooth } l_1(x) = \begin{cases} 0.5x^2, & |x| < 1, \\ |x| - 0.5, & \text{otherwise}, \end{cases} \quad (7)$$

其中 x 表示为真实值和预测值的差值。

1.6 改进的非极大值抑制

NMS 在目标检测中来提取置信度高的目标检测框, 抑制置信度较低的误检框, 在对腺瘤性和增生性息肉进行模型输出目标框时, 会在同一目标处出现大量重复的检测框, NMS 针对重叠的候选框, 若大于规定的某一阈值则删除, 低于阈值的将保留, 但是如果目标周围出现了大于规定阈值的同样类型的目标检测框, NMS 也会将其删除, 由于在结肠直肠中会出现多个息肉聚集的情况, 本研究用 Soft-NMS 代替 NMS, 如式(8), 当出现 IoU 大于阈值对应的检测框时, 置信度分数乘上一个小于 1 的系数, 能够通过权重来降低检测框原有的置信度, 重叠区域的增大会导致置信度更快的衰减。

$$S_i = \begin{cases} S_i, & IoU(M, b_i) < N_t \\ S_i(1 - IoU(M, b_i)), & IoU(M, b_i) \geq N_t \end{cases}, \quad (8)$$

其中, M 为当前得分的最高检测框, b_i 为待处理的检测框, N_t 为设定的阈值, S_i 设定为当前处理框的预测分数。

2 实验研究

2.1 实验环境及流程

本实验通过构建腺瘤与增生性息肉的数据集, 在 Faster R-CNN 网络下训练数据并测试, 再对模型进行改进, 运用评价指标来验证网络模型效果并调整到最优。所有网络模型均在 Ubuntu 18.04 系统下, 核心配置为 Inter Core i7-7800 * 1 以及 128 G 的 DDR4 运行内存, 采用 GPU 为 Nvidia GeForce GTX 2070 * 1、Pytorch 1.3.1、Cuda 10.2。训练中的参数设置: 批处理大小(Batchsize)设置为 2, 输入图片尺寸统一为 $1\,333 * 800$, 学习率设置为 0.01, 权重衰减设置为 0.000 1, 动量设置为 0.9, 用随机梯度下降法(SGD)优化损失函数, 共进行了 20 个周期(epoch)的训练。

2.2 数据收集

在目前可公开的数据集中, 例如 CVC-ClinicDB^[27]、Kvasir-SEG^[28]等, 都只收集了息肉的图片或视频帧, 并没有注明息肉的类型, Colonoscopic Dataset 数据集虽然收集并注明了腺瘤性、增生性和锯齿状息肉的视频, 但是由于病人样本数较少, 会造成过拟合的现象。因此, 在本研究中, 第 1 部分数据来源于 2019 年 6 月至 2020 年 9 月, 在河北大学附属医院进行的结肠镜检查的 654 个病人的图像和病理报告, 并且挑选出其中至少含有一个息肉的图片, 经过了结肠镜医师的证实, 同时所有息肉的结肠镜图像均由结肠镜医师精确标注。在收集并标注的图片中, 医师根据病理报告证实为腺瘤性息肉的图像有 1 928 张, 证实为增生性息肉的图像有 234 张。第 2 部分数据来源于公开数据集 Colonoscopic Dataset 中 60 个病人的结肠镜视频, 将其转化为视频帧并手动标注, 其中增生性息肉图像 826 张, 腺瘤性息肉图像 40 张。一共收集了腺瘤性息肉图像 1 968 张(64.99%), 增生性息肉图像 1 060 张(35.01%), 按照训练集: 测试集 = 4 : 1 的比例将收集的图像分为 2 部

分:1)训练集:2 426 张息肉图像,其中包括 1 582 张腺瘤性息肉图像和 844 张增生性息肉图像,通过训练集的图像拟合模型、调整参数,达到最优效果.2)测试集:602 张图像,其中包括 386 张腺瘤性图像,216 张增生性息肉图像,通过测试集的图像评估训练模型.

2.3 图像增广

图像增广技术是通过图像进行旋转、放缩等一系列的处理后,产生了多倍与原图相似但不相同的图像.为了扩充数据集,并且使模型有泛化能力,对训练集的 2 426 张图片进行了图像增广处理.针对数据集的图像,制定了 2 种类型的增广方式,第 1 种:对图像进行了放缩、旋转、错切等操作,此种方式保留了图像自身的像素,只改变了图像中感兴趣区域的位置信息.第 2 种:对图像进行了锐化、仿射变换等一系列改变像素值的图像增强方式.采取的 2 种不同方式的图像增广处理使图片扩充了 13 倍.

3 结果及分析

3.1 评价指标

本文采用精确率(Precision, P)、召回率(Recall, R)、平均精度(Average Precision, AP)、平均精度的均值(mean Average Precision, mAP)对腺瘤、增生性息肉检测和分类的算法进行评价,如式(9)~(12).由 P 、 R 值得到每个类别的 P - R 曲线,曲线下的面积则为 AP , mAP 反映了在所有类别的精度表现,在目标检测中通常用 mAP 作为评估算法精度的最佳指标.本文通过设定不同的阈值(IoU)分别对算法的性能进行评价.其中 IoU 表示 2 个框(A 、 B)之间的重合度,如式(13)所示.因此 $mAP_{0.5}$ 和 $mAP_{0.75}$ 分别表示:当设定 IoU 为 0.5 和 0.75 时 mAP 的取值.

$$P = \frac{T_p}{T_p + F_p}, \quad (9)$$

$$R = \frac{T_p}{T_p + F_N}, \quad (10)$$

其中, T_p 、 F_p 和 F_N 分别表示正确地识别出正样本的数量、错误地识别成正样本的数量和错误地识别成负样本的数量.

$$AP = \int_0^1 P(R) dR, \quad (11)$$

$$mAP = \frac{1}{classes} \sum_{i=1}^{classes} \int_0^1 P(R) d(R), \quad (12)$$

$$IoU = \frac{A \cap B}{A \cup B}. \quad (13)$$

3.2 测试结果

本研究将 FPN 结构融合到 Faster R-CNN 的网络作为基准(baseline),并且以 ResNet50 作为 backbone.表 1 统计了改进 Faster R-CNN 的每个模块所产生的对 mAP 、 mAR 以及 IoU 取 0.5 时的 mAP 值的影响,改变了特征提取的 ResNet50,替换为 ResNeSt50,提升了平均精度的均值.并且用到了边缘感知的边界定位(SABL)改变了边框回归的原始方法,无论是精度与召回率都有较为明显的提升,而通过 Soft-NMS 更加关注息肉的召回率,虽然在 IoU 取 0.5 时比 NMS 低 0.1%,但是提升了 5.5%的平均召回率的均值,基于对息肉图像的考量,查全能够降低漏检的概率,作为医学辅助诊断的手段,更多的需起到提示医生的作用,并且在 10 个阈值上的平均精度均值较之前提升了 1.4%.

表 2 的统计结果是在 IoU 为 0.5 时,测试的图像在不同的网络模型下,腺瘤性息肉和增生性息肉的 AP 值,其 SSD 和 RetinaNet 为单阶段目标检测网络, Cascade R-CNN 为双阶段的级联型目标检测网络模型.由表可见 Faster R-CNN 的 AP 表现结果优于其他网络,改进后的 Faster R-CNN 相较于之前分别提升了 1.8%和 2.8%,并且基于精确率与召回率绘制了每个类别的 P - R 曲线,如图 5 所示,改进后的 Faster R-CNN 在每个类别的 P - R 曲线最终都优于其他网络模型.

表 1 模块结构对于测试结果的影响

Tab. 1 Effect of the network module structure on the test results

改进结构	$mAP_{0.5}/\%$	$mAP/\%$	$mAR/\%$
ResNet50	84.5	61.3	70.8
ResNeSt50	85.9	61.7	70.5
SABL-ResNeSt50	86.9	62.3	71.1
SABL-ResNeSt50-SoftNMS	86.8	63.7	76.6

表 2 不同类别的测试结果与其他网络比较

Tab. 2 Comparison of test results of different categories with those of other networks

网络	$AP_{0.5}(\text{Adenoma})/\%$	$AP_{0.5}(\text{Hyperplastic})/\%$
Faster R-CNN	82.9	86.1
Cascade R-CNN	79.4	83.5
SSD	77.0	83.8
RetinaNet	82.4	84.1
改进的 Faster R-CNN	84.7	88.9

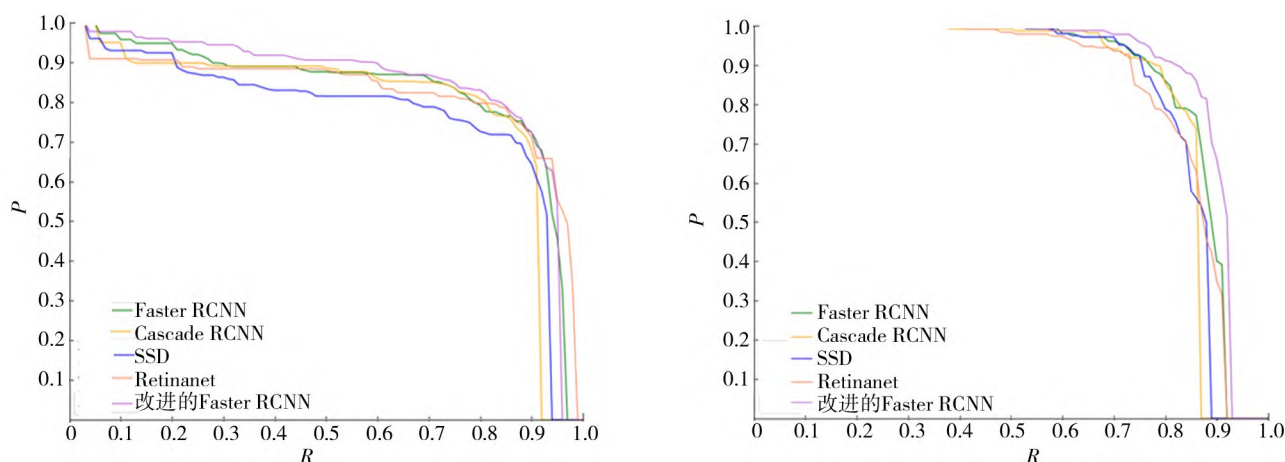
图 5 不同网络模型对于不同类别息肉的 P - R 曲线Fig. 5 P - R curves of different types of polyps under different network models

表 3 统计结果的 mAP 和 mAR 是在 IoU 为 $0.5 \sim 0.95$ 的所有阈值下的结果取平均值,其中每隔 0.05 取一个 IoU 值,分别代表所有类别、阈值的平均精度的均值和平均召回率的均值,从表中可以看出,在腺瘤和增生性息肉的检测和分类中,改进的 Faster R-CNN 网络相较于原网络和其他网络, $AP_{0.5}$ 、 $mAP_{0.75}$ 、 mAP 和 mAR 都有显著提升.

表 3 所有类别的测试结果与其他网络比较

Tab. 3 Comparison of test results of all categories with those of other networks

网络	$mAP_{0.5}/\%$	$mAP_{0.75}/\%$	$mAP/\%$	$mAR/\%$
Faster R-CNN	84.5	73.4	61.3	70.8
Cascade R-CNN	81.5	71.6	59.9	69.4
SSD	80.4	57.7	51.6	62.6
RetinaNet	83.3	66.4	58.3	70.6
改进的 Faster R-CNN	86.8	75.6	63.7	76.6

如图 6 展示了腺瘤性和增生性息肉检测和分类的结果对比,蓝色框和文字为测试结果的金标准,分别能表示出息肉的位置和息肉的类别,绿色框和文字代表检测到的增生性息肉和网络判断的类别与置信度,黄色框和文字代表检测到的腺瘤性息肉和网络判断的类别与置信度.第 1 组漏检的增生性息肉在改进的网络测

试中提取到了更深的特征,由此检测到并准确地分类.针对第 2、3 组测试图像,提升了 2 种类型息肉的置信度,有更大的概率确认此息肉的类型,且有效提升了平均准确度的均值.

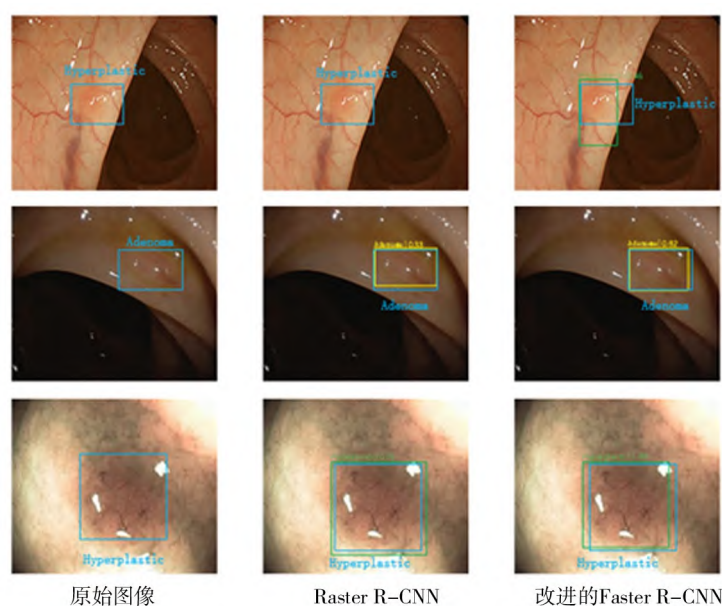


图 6 改进前后的网络模型的测试结果对比

Fig. 6 Comparison of test results of the network model before and after the improvement

4 结论

本文针对区分腺瘤和增生性息肉的难题,对 Faster R-CNN 的特征提取、边界框回归和非极大值抑制分别进行了改进,提取了更深入的特征的同时获得了较高的分类置信度,并且 Soft-NMS 的操作相较于改进之前在保证精度的同时获得了更优的查全率.通过实验调整出最优的参数,结果显示,改进后的 Faster R-CNN 网络获得了 86.8% 的 $mAP_{0.5}$ 值,相比改进之前提升 2.3%,能够更加准确地检测出息肉,同时也能获得更高的分类精度,对于 AI 未来的临床辅助诊断提供了一个潜在的方向.

参 考 文 献:

- [1] SUNG H, FERLAY J, SIEGEL R L, et al. Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries[J]. Ca-cancer J Clin, 2021, 71(3): 209-249. DOI:10.3322/caac.21660.
- [2] CARBALLAL S, MOREIRA L, BALAGUER F. Serrated polyps and serrated polyposis syndrome[J]. Cirugía Española Engl Ed, 2013, 91(3):141-148. DOI:10.1016/j.cireng.2013.07.002.
- [3] WU X, HE X, LI S, et al. Long non-coding RNA uc002kmd.1 regulates CD44-dependent cell growth by competing for miR-211-3p in colorectal cancer[J]. PLoS One, 2016, 11(3): e0151287. DOI:10.1371/journal.pone.0151287.
- [4] REX D K, KAHN C, O'BRIEN M, et al. The american society for gastrointestinal endoscopy PIVI (preservation and incorporation of valuable endoscopic innovations) on real-time endoscopic assessment of the histology of diminutive colorectal polyps[J]. Gastrointest Endosc, 2011, 73(3): 419-422. DOI:10.1016/j.gie.2011.01.023.
- [5] KESWANI R N, LAW R, CIOLINO J D, et al. Adverse events after surgery for nonmalignant colon polyps are common and associated with increased length of stay and costs[J]. Gastrointest Endosc, 2016, 84(2): 296-303. e1. DOI:10.1016/j.gie.2016.01.048.
- [6] PATINO-BARRIENTOS S, SIERRA-SOSA D, GARCIA-ZAPIRAIN B, et al. Kudo's classification for colon polyps assessment using a deep learning approach[J]. Appl Sci, 2020, 10(2): 501. DOI:10.3390/app10020501.
- [7] SHARMA P, FRYE J, FRIZELLE F. Response to Re: accuracy of visual prediction of pathology of colorectal polyps: how accurate are we? [J]. ANZ J Surg, 2014, 84(10): 797-798. DOI:10.1111/ans.12790.
- [8] LADABAUM U, FIORITTO A, MITANI A, et al. Real-time optical biopsy of colon polyps with narrow band imaging

- in community practice does not yet meet key thresholds for clinical decisions[J]. *Gastroenterology*, 2013, 144(1): 81-91. DOI:10.1053/j.gastro.2012.09.054.
- [9] 蒋西然, 蒋韬, 孙嘉瑶, 等. 深度学习人工智能技术在医学影像辅助分析中的应用[J]. *中国医疗设备*, 2021, 36(6): 164-171. DOI:10.3969/j.issn.1674-1633.2021.06.040.
- [10] KARKANIS S A, IAKOVIDIS D K, MAROULIS D E, 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] AMELING S, WIRTH S, PAULUS D, et al. Texture-based polyp detection in colonoscopy[C]// *Bildverarbeitung für die Medizin 2009: Algorithmen - Systeme - Anwendungen, Proceedings des Workshops vom 22. bis 25. März 2009 in Heidelberg*. DBLP, 2009. DOI:10.1007/978-3-540-93860-6_70.
- [12] HWANG S, OH J, TAVANAPONG W, et al. Polyp detection in colonoscopy video using elliptical shape feature[C]// *2007 IEEE International Conference on Image Processing, San Antonio, TX, USA, IEEE*, : II-465. DOI:10.1109/ICIP.2007.4379193.7.
- [13] NAWARATHNA R, OH J, MUTHUKUDAGE J, et al. Abnormal image detection in endoscopy videos using a filter bank and local binary patterns[J]. *Neurocomputing*, 2014, 144: 70-91. DOI:10.1016/j.neucom.2014.02.064.
- [14] PACAL I, KARABOGA D. A robust real-time deep learning based automatic polyp detection system[J]. *Comput Biol Med*, 2021, 134: 104519. DOI:10.1016/j.combiomed.2021.104519.
- [15] SILVA J, HISTACE A, ROMAIN O, et al. Toward embedded detection of polyps in WCE images for early diagnosis of colorectal cancer[J]. *Int J Comput Assist Radiol Surg*, 2014, 9(2): 283-293. DOI:10.1007/s11548-013-0926-3.
- [16] KOMEDA Y, HANDA H, WATANABE T, et al. Computer-aided diagnosis based on convolutional neural network system for colorectal polyp classification: preliminary experience[J]. *Oncology (Suppl 1)*, 2017, 93(1): 30-34. DOI:10.1159/000481227.
- [17] BYRNE M F, CHAPADOS N, SOUDAN F, et al. Real-time differentiation of adenomatous and hyperplastic diminutive colorectal polyps during analysis of unaltered videos of standard colonoscopy using a deep learning model[J]. *Gut*, 2019, 68(1): 94-100. DOI:10.1136/gutjnl-2017-314547.
- [18] MESEJO P, PIZARRO D, ABERGEL A, et al. Computer-aided classification of gastrointestinal lesions in regular colonoscopy[J]. *IEEE Trans Med Imaging*, 2016, 35(9): 2051-2063. DOI:10.1109/TMI.2016.2547947.
- [19] REN S Q, HE K M, GIRSHICK R, et al. Faster R-CNN: towards real-time object detection with region proposal networks[J]. *IEEE Trans Pattern Anal Mach Intell*, 2017, 39(6): 1137-1149. DOI:10.1109/TPAMI.2016.2577031. DOI:10.1109/TPAMI.2016.2577031.
- [20] LIN T Y, DOLLÁR P, GIRSHICK R, et al. Feature pyramid networks for object detection[J]. *2017 IEEE Conf Comput Vis Pattern Recognit CVPR*, 2017: 936-944. DOI:10.1109/CVPR.2017.106.
- [21] SIMONYAN K, ZISSERMAN A. Very deep convolutional networks for large-scale image recognition[J]. *Computer Science*, 2014. DOI:10.48550/arXiv.1409.1556.
- [22] HE K M, ZHANG X Y, REN S Q, et al. Deep residual learning for image recognition[C]// *2016 IEEE Conference on Computer Vision and Pattern Recognition, Las Vegas, NV, USA, IEEE*, : 770-778. DOI:10.1109/CVPR.2016.90.
- [23] ZHANG H, WU C R, ZHANG Z Y, et al. ResNeSt: split-attention networks[C]// *2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*. New Orleans, LA, USA, IEEE: 2735-2745. DOI:10.1109/CVPRW56347.2022.00309.
- [24] WANG J Q, ZHANG W W, CAO Y H, et al. Side-aware boundary localization for more precise object detection[C]// *Comput Vis-ECCV 2020*, 2020: 403-419. DOI:10.1007/978-3-030-58548-8_24.
- [25] BODLA N, SINGH B, CHELLAPPA R, et al. Soft-NMS—improving object detection with one line of code[C]// *2017 IEEE International Conference on Computer Vision, Venice, Italy, IEEE*: 5562-5570. DOI:10.1109/ICCV.2017.593.
- [26] XIE S N, GIRSHICK R, DOLLÁR P, et al. Aggregated residual transformations for deep neural networks[C]// *2017 IEEE Conference on Computer Vision and Pattern Recognition*. Honolulu, HI, USA, IEEE, : 5987-5995. DOI:10.1109/CVPR.2017.634.
- [27] BERNAL J, SÁNCHEZ F J, FERNÁNDEZ-ESPARRACH G, et al. WM-DOVA maps for accurate polyp highlighting in colonoscopy: validation vs. saliency maps from physicians[J]. *Comput Med Imaging Graph*, 2015, 43: 99-111. DOI:10.1016/j.compmedimag.2015.02.007.
- [28] POGORELOV K, RANDEL K R, GRIWODZ C, et al. KVASIR: A multi-class image dataset for computer aided gastrointestinal disease detection[C]// *Proceedings of the 8th ACM on Multimedia Systems Conference*. Taipei Taiwan, New York, NY, USA: ACM, 2017: 164 - 169. DOI:10.1145/3083187.3083212.

(责任编辑:孟素兰)