

基于改进 RetinaNet 的宫颈异常细胞检测算法

刘润坤^{1,2}, 党世杰², 张洪远², 牛银银³, 米贯勋³, 李三华³, 陈振鑫², 赵凌霄^{2*}, 李鹏²¹中国科学技术大学生物医学工程学院(苏州)生命科学与医学部, 安徽 合肥 230026;²中国科学院苏州生物医学工程技术研究所, 江苏 苏州 215162;³河南赛诺特生物技术有限公司, 河南 郑州 450001

摘要 宫颈异常细胞特征细微难以提取、小目标容易漏检、细胞边界回归不准确导致异常细胞检测精度不高, 鉴于此, 本文提出了一种结合注意力的全尺度特征融合 RetinaNet(AFF-RetinaNet)宫颈异常细胞检测算法。首先, 采用 ResNeSt-50 作为特征提取网络提取宫颈异常细胞的细微特征; 其次, 引入平衡特征金字塔(BFP)结构, 对所有特征层进行全尺度融合, 增强小目标的语义信息, 并利用 BFP 中的非局部注意力模块获取图像的全局信息, 以进一步增强特征空间的语义信息; 最后, 采用 CIoU Loss 作为回归分支的损失函数, 以提高对异常细胞边界回归的准确率。另外, 针对实际应用场景, 基于 AFF-RetinaNet 算法实现了全视野宫颈细胞病理学图像(WSI)推理流程, 并基于该推理流程对 WSI 中的异常细胞进行了检测。AFF-RetinaNet 在宫颈异常细胞数据集上的平均精度均值(mAP)为 83.4%, 其中对小目标的 mAP 值(mAP-s)达到了 24.4%, 相较于基准 RetinaNet 算法分别提高了 3.2 个百分点和 10.8 个百分点。基于 AFF-RetinaNet 的 WSI 推理结果在感兴趣区域中的 mAP 为 70.8%。实验结果表明: AFF-RetinaNet 算法可以增强对小尺寸异常细胞的检测能力, 有效提升宫颈异常细胞的检测精度。基于 AFF-RetinaNet 的宫颈 WSI 推理流程可辅助医生快速定位高分辨率宫颈 WSI 中的异常细胞, 有望减轻医生的阅片负担。

关键词 医用光学; 宫颈细胞病理图像; 目标检测; 小目标; 特征融合; 注意力机制

中图分类号 TP391.4

文献标志码 A

DOI: 10.3788/CJL230718

1 引言

宫颈癌是威胁女性健康的第四大常见癌症^[1]。据统计, 宫颈癌现已成为 23 个国家最常见的癌症, 也是 36 个国家癌症病例死亡的主要原因。2020 年, 全世界有超过 60 万宫颈癌新增病例, 34.1 万人因宫颈癌死亡^[2]。研究表明, 宫颈癌的早期筛查可以有效降低其发病率和死亡率。目前, 液基薄层细胞学检查(TCT)以其灵敏度高、可重复性好等特点^[3]成为宫颈癌早期筛查的重要手段^[4]。但是, 传统人工阅片方式依赖于细胞病理医生的专业知识和主观经验, 长时间的眼力、脑力劳动也会影响判读结果, 甚至导致漏判和误判。因此, 宫颈异常细胞自动化检测方法的研究对于提高异常细胞的检测准确率以及减轻医生的工作负荷具有重要意义。

传统方法大多基于图像分割和分类技术实现对宫颈异常细胞的检测。Su 等^[5]先使用自适应阈值法实现对图像中宫颈细胞区域的分割, 并在细胞区域提取出 20 个形态特征和 8 个纹理特征, 然后通过级联 C4.5 算法分类器和逻辑回归分类器识别异常细胞。

Chankong 等^[6]提出了一种针对单细胞图像中宫颈细胞的自动分割和分类方法, 该方法使用模糊 C 均值(CFM)聚类技术和人工神经网络(ANN)实现对宫颈异常细胞的检测。虽然上述研究已经取得了一定成果, 但算法的复杂度普遍较高, 并且对细胞特征的提取依赖于先验知识, 具有一定的局限性。

近年来, 基于卷积神经网络(CNN)的目标检测技术成为计算机视觉领域的研究热点^[7], 该技术允许一次性完成目标的定位和分类, 大大提高了检测效率和准确率, 因此在各种检测任务中得到了广泛应用^[8-10]。Xu 等^[11]将 Faster R-CNN^[12]目标检测算法应用于宫颈异常细胞检测, 证明了目标检测技术在该领域应用的可行性。尹远来等^[13]将特征金字塔网络^[14](FPN)和可变形卷积网络引入 Faster R-CNN 中, 提高了宫颈异常细胞的检测精度。Cao 等^[15]将注意力机制应用到 FPN 中, 提高了算法对异常细胞的特征提取能力。虽然基于目标检测技术的宫颈异常细胞检测方法研究取得了一定进展, 但是整体上宫颈异常细胞检测精度仍受限于以下因素: 1) 宫颈异常细胞特征细微, 不易提取; 2) 细胞重叠、染色不均匀等因素导致小尺寸异常

收稿日期: 2023-02-27; 修回日期: 2023-04-13; 录用日期: 2023-04-25; 网络首发日期: 2023-05-16

基金项目: 苏州市基础研究试点项目(SJC2021022)

通信作者: hitic@sibet.ac.cn

细胞容易漏检;3) 宫颈异常细胞的巨大形态差异导致细胞边界回归不够准确。

针对上述问题,笔者选取检测性能较好的 RetinaNet^[16]作为基线方法,并对其特征提取网络、特征融合网络以及边界回归损失函数进行改进,提出了一种结合注意力机制的全尺度特征融合 RetinaNet (AFF-RetinaNet) 宫颈异常细胞检测算法。首先,采用 ResNeSt-50^[17]作为特征提取网络,以提高模型对宫颈异常细胞的特征提取能力;其次,引入平衡特征金字塔^[18](BFP)进行全尺度特征融合,以增强小目标的语义信息,并通过非局部注意力模块提取图像的全局信息,从而进一步加强特征空间的语义信息;最后,采用 CIoU Loss^[19]训练回归分支,以提高对异常细胞的定位精度。除此之外,基于本文所提 AFF-RetinaNet 宫颈异常细胞检测算法实现了面向实际应用场景的全视野宫颈细胞病理图像(WSI)推理流程,并实现了超高分辨率 WSI 中宫颈异常细胞的自

动检测。

2 AFF-RetinaNet 目标检测算法

2.1 RetinaNet 的基本原理

RetinaNet 是一种单阶段目标检测算法,其结构包含特征提取网络、FPN 和检测头三部分。特征提取网络基于 ResNet 系列网络实现,并输出 3 个不同尺度的特征层,该特征层是 FPN 的输入。经过对上述特征层进行自上而下的特征融合以及对最小尺度的特征层进行两次上采样,得到 5 个尺度的输出特征层,最终通过共享的检测头完成对目标的定位和分类。检测头中的分类分支采用 Focal Loss 缓解难易样本不平衡的问题,回归分支则采用 SmoothL1 Loss 对坐标参数回归进行约束。

针对宫颈异常细胞检测任务,笔者对 RetinaNet 算法进行改进,提出了 AFF-RetinaNet 宫颈异常细胞检测算法,算法的整体结构如图 1 所示。

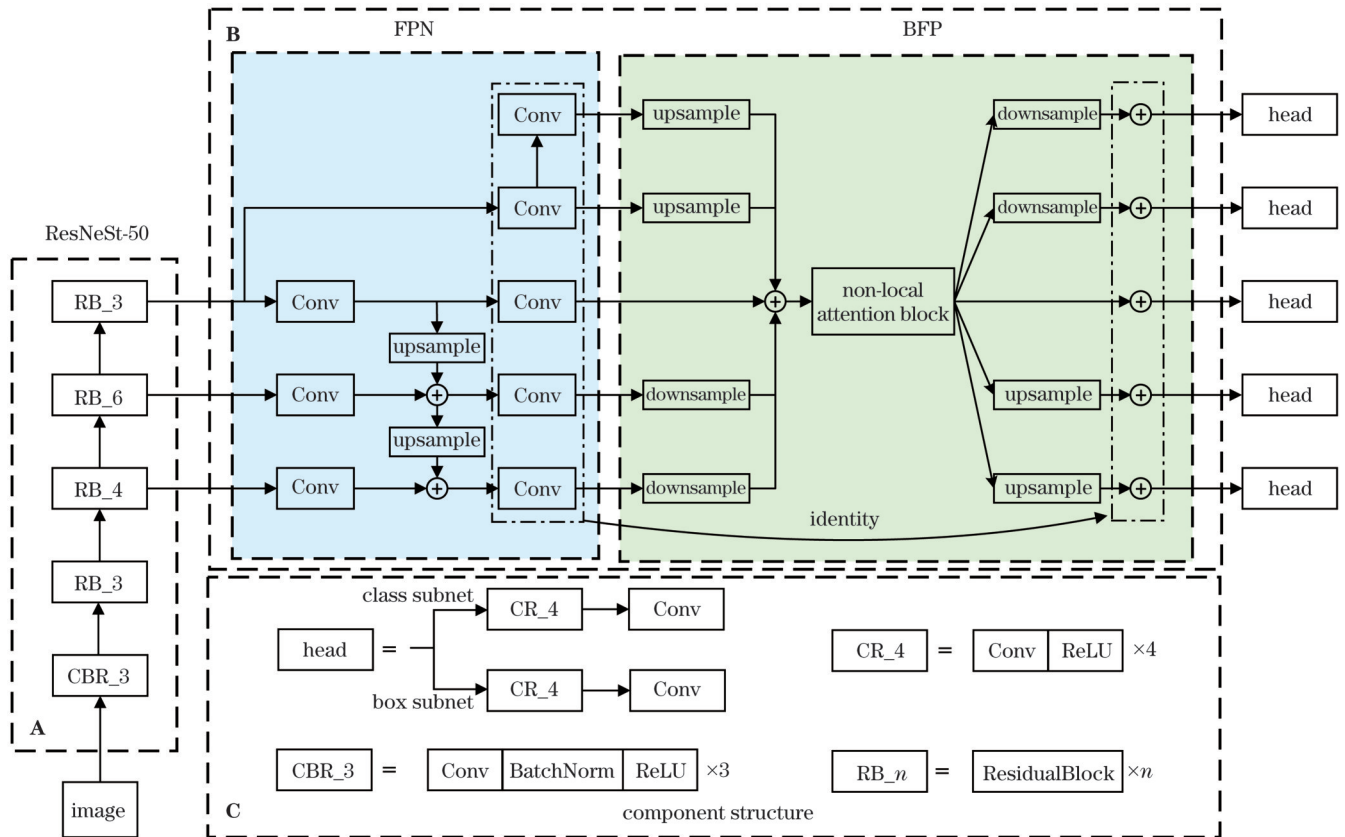


图 1 AFF-RetinaNet 结构图

Fig. 1 Structure of the AFF-RetinaNet

2.2 ResNeSt-50 特征提取网络

用 ResNeSt-50 替代 ResNet 网络作为特征提取网络,ResNeSt-50 负责从复杂的宫颈细胞病理图像中学习异常细胞的细微特征信息,其结构如图 1 中的模块 A 所示。ResNeSt-50 的主要优势在于引入了分组卷积和拆分注意力模块。分组卷积可以对不同通道的特征进行分组计算,精细化提取异常细胞的特征;拆分注

意力模块为分组后的特征产生通道权重,进一步提高关键特征所在通道的比重。ResNeSt-50 的残差模块结构如图 2 所示。

在 ResNeSt-50 残差模块中,输入特征首先在通道维度上被分为 k 组,每组特征在通道维度上被进一步拆分为 r 部分;接着分别通过一个 1×1 和 3×3 卷积后共同进入拆分注意力模块,该模块对输入的 r 部分特征图

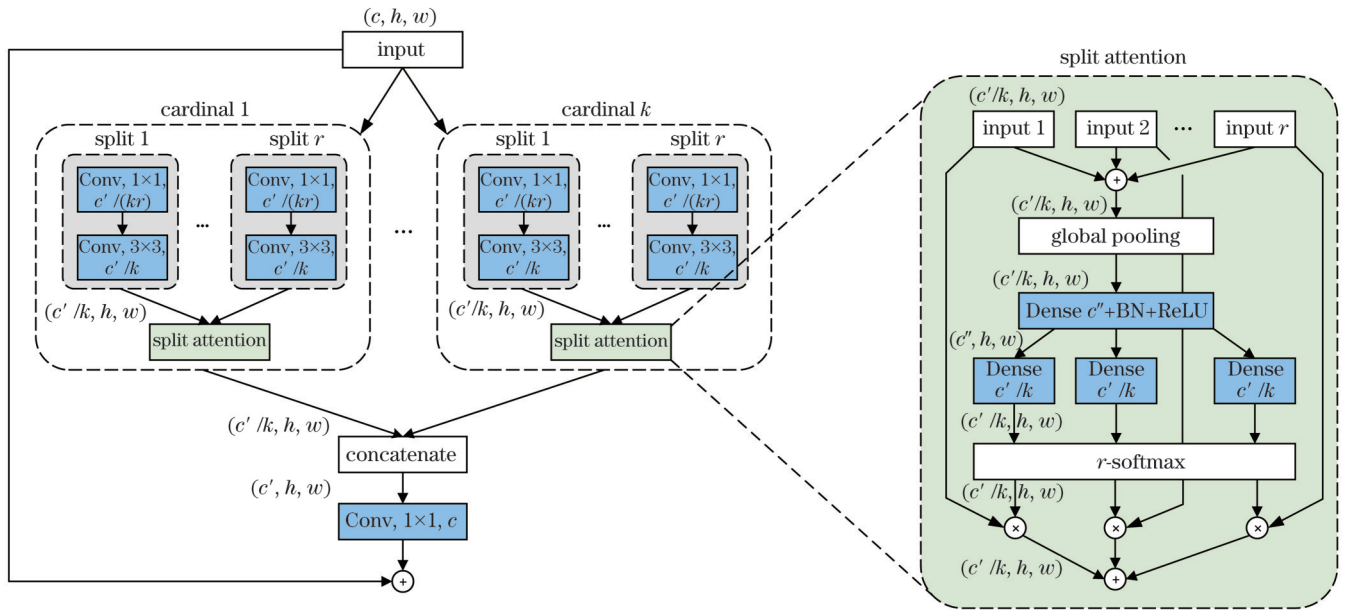


图 2 ResNeSt-50 残差模块结构图

Fig. 2 Residual block structure of ResNeSt-50

在对应通道上进行相加(融合生成一组特征图),同时通过挤压-激励操作为每部分输入特征单独产生并分配通道注意力权重,再将 r 部分特征层在对应通道相加得到一组特征输出;最后,将通过上述方式生成的 k 组特征

在通道维度上进行拼接,再通过一个 1×1 卷积与原始特征进行对应元素相加后便得到残差模块的最终输出。以 $r=2$ 为例,拆分注意力机制示意图如图 3 所示,其中 W 、 H 和 C 分别表示特征图的宽、高和通道数。

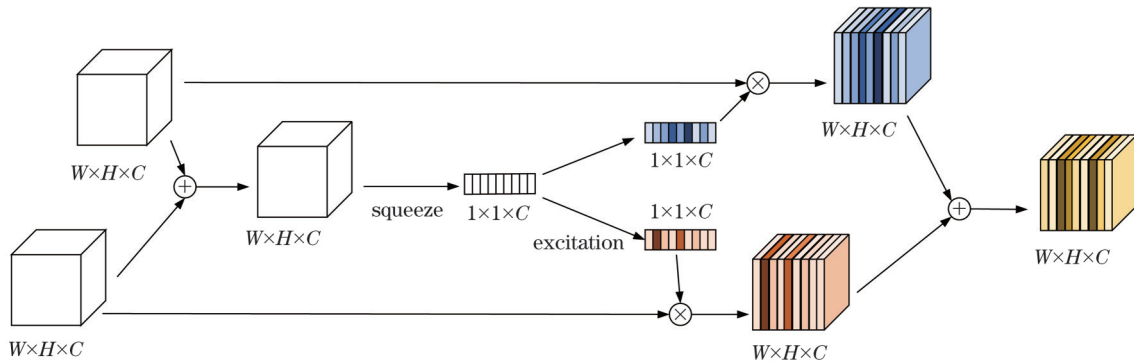


图 3 拆分注意力示意图

Fig. 3 Diagram of split attention

2.3 全尺度特征融合 FPN

小目标的语义信息在特征提取过程中容易丢失,从而造成漏检。虽然基准 RetinaNet 采用 FPN 结构增加检测尺度的策略在一定程度上使得该问题有所改善,但其只对部分尺度特征进行融合的方式使得小目标语义信息加强得不充分,而且忽略了对宫颈异常细胞判别具有重要参考意义的图像全局信息。

为了提高小尺寸异常细胞的检测精度,在 FPN 结构的基础上引入 BFP 构成全尺度特征融合 FPN,其结构如图 1 中的模块 B 所示。首先,将 FPN 输出的所有尺度特征统一缩放到中间尺度,将所有特征进行加权平均,得到包含丰富空间信息和语义信息的单尺度特征;其次,采用非局部注意力模块引入图像全局信息,以增强特征空间的语义信息;最后,对增强后的单尺度

特征分别进行上采样和下采样,生成新的特征金字塔。BFP 整体结构通过残差连接,以避免可能出现的梯度爆炸等问题。

非局部注意力模块结构如图 4 所示,其中 B 表示样本量, C 表示特征通道数, H 和 W 分别表示特征图的高和宽。首先,输入特征 X 分别经过 θ 、 ϕ 、 g 这 3 个 $1 \times 1 \times 1$ 卷积层压缩通道数并调整输出特征的维度;其次,对卷积层 θ 和 ϕ 得到的特征进行矩阵乘法,得到空间注意力权重,并使用 softmax 函数对权重进行归一化;然后,将上述权重与卷积层 g 输出的特征进行矩阵相乘并调整特征维度;最后,通过 1 个 $1 \times 1 \times 1$ 卷积层将加权后的特征恢复至原来的通道数,再将其与 X 进行对应元素相加便可得到最终的输出 Z 。

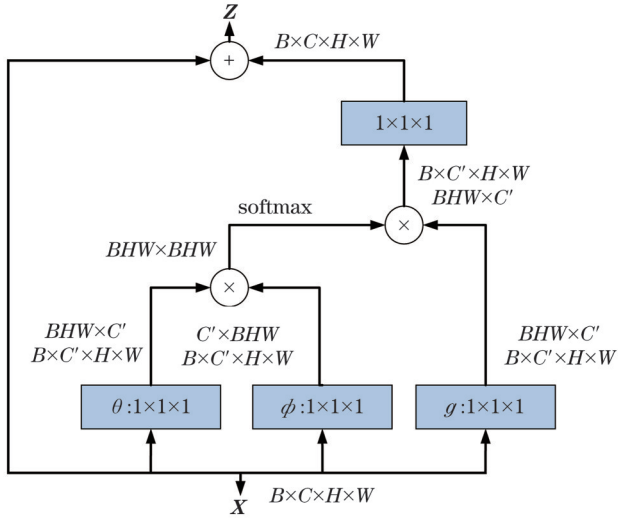


图 4 非局部注意力模块结构图

Fig. 4 Non-local attention block structure

2.4 CIoU Loss

由于受到细胞生长周期和不同致病因素的影响,异常细胞具有形态多样和尺寸差异大的特点,而基准 RetinaNet 所采用的边界回归损失函数(SmoothL1 Loss)没有考虑预测框相对标注框中心点偏移量与高宽偏移量之间的关联性,导致其对异常细胞的边界回归不够准确。因此,本文采用以交并比(IoU,在公式中记为 R_{IoU})为基础的 CIoU Loss 约束回归分支,同时考虑候选框与标注框的 IoU、中心点距离和高宽比等因素,实现对异常细胞边界的准确回归。IoU 的计算公式为

$$R_{IoU} = \frac{A \cap B}{A \cup B}, \quad (1)$$

式中: A 、 B 分别表示两个边界框。CIoU Loss 计算示意图如图 5 所示,计算公式为

$$L_{CIoU} = 1 - R_{IoU} + \frac{d^2}{c^2} + \alpha\nu, \quad (2)$$

式中: d 表示预测框和标注框中心点的距离; c 表示预测框和标注框最小外接矩形对角线的长度; α 为平衡系数; ν 用于衡量两个边界框高宽比的一致性。 α 和 ν 的计算公式分别为

$$\alpha = \frac{\nu}{(1 - R_{IoU}) + \nu}, \quad (3)$$

$$\nu = \frac{4}{\pi^2} \left(\arctan \frac{w^{gt}}{h^{gt}} - \arctan \frac{w}{h} \right)^2, \quad (4)$$

式中: h^{gt} 、 w^{gt} 、 h 和 w 分别表示真实标注框的高和宽以及预测框的高和宽, gt 表示真实标注框(ground truth)。

2.5 WSI 推理流程

在实际应用场景中,病理医生通常基于 WSI 对患者进行诊断。但由于 WSI 的分辨率较高,图像尺寸较大,无法直接使用目标检测算法进行异常细胞检测。因此,笔者对 AFF-RetinaNet 算法进行了进

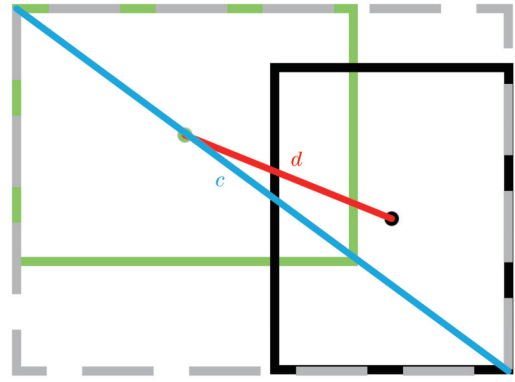


图 5 CIoU Loss 计算示意图

Fig. 5 CIoU Loss calculation diagram

一步拓展,实现了如图 6 所示的 WSI 推理流程。具体步骤为:1) 采用有重叠的滑动窗口对 WSI 进行裁剪,得到若干包含重复区域的子图(patch),裁剪操作的窗口大小及滑动步长设置见 3.1 节;2) 对所有 patch 使用 AFF-RetinaNet 算法进行宫颈异常细胞检测,得到 patch 级别的检测结果;3) 对所有 patch 的预测框坐标信息添加滑动窗口的偏移量,以得到预测框在 WSI 中的坐标信息;4) 对所有的预测框进行表 1 所示的非极大值抑制(NMS)操作,只保留同一位置置信度分数最大的检测框,消除因滑动窗口重叠产生的冗余预测框,得到最终的 WSI 检测结果。

3 实验设计和结果分析

3.1 实验数据获取与处理

本次实验共采集了 245 张宫颈 WSI 数据,所有的切片均采用膜式制片法制备,扫描仪扫描倍率为 20 倍,WSI 分辨率在 35000x35000 左右,以 TIF 格式存储。剔除质量较差的图像,训练和测试数据集纳入 233 例,患者年龄分布范围为 25~60 周岁。细胞病理专家在每张 WSI 上随机选取若干感兴趣区域(ROI),并对 ROI 中包含的所有宫颈异常细胞进行标注,共计选取 1103 个 ROI,标注异常细胞 11188 个,数据标注形式如图 7 所示。

由于 ROI 的尺寸仍然较大,因此采用有重叠的滑动窗口裁剪方式从 ROI 图像上裁剪出包含标注框的 patch,以生成可供模型训练和测试的 TCT 数据集。根据图 8 所示的标注框尺寸分布情况,所有异常细胞的尺寸不超过 700x700,因此本文采用尺寸为 800x800、步距为 600 的滑动窗口对 ROI 进行滑动裁剪,以确保滑动窗口能完整地包含最大标注框,同时使异常细胞在图像中占较大比重,以利于模型对其特征进行学习。对于被滑动窗口截断的标注框,若该标注框在窗口内的面积大于其自身面积的 80%,则保留该标注框位于当前窗口的区域,否则舍弃该标注框并对当前窗口中标注框的区域用白色像素进行填充,以避免引

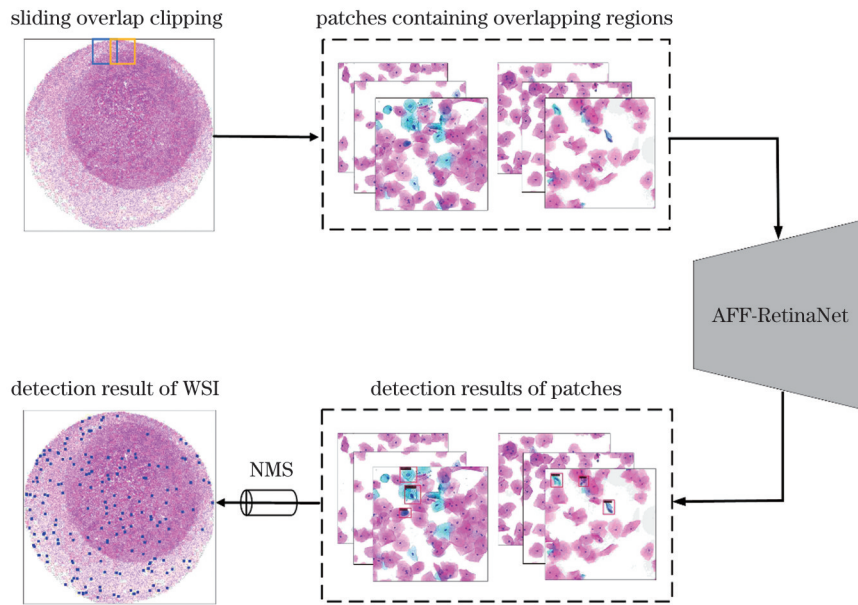


图 6 WSI推理流程示意图

Fig. 6 Diagram of WSI inference process

表 1 NMS算法流程

Table 1 Algorithm process of NMS

NMS algorithm: remove redundant detection boxes	
Input: detection boxes set B , IoU threshold T	
Output: detection boxes set K after NMS	
1:	$K \leftarrow \emptyset$;
2:	all boxes in B are descending ordered by confidence score;
3:	while $B \neq \emptyset$ do
4:	$b_{ms} \leftarrow$ the first detection box in B ;
5:	add b_{ms} to K ;
6:	del b_{ms} from B ;
7:	for b in B do
8:	if IoU of b and $b_{ms} > T$ do
9:	del b from B ;
10:	end if
11:	end for
12:	end while
13:	return K

入无用特征影响模型训练。

数据划分以 ROI 为单位进行。将 233 张 WSI 中

所有的 1103 个 ROI 区域提取出来并按照 4:1 的比例随机划分,得到 883 张用于生成训练集的 ROI 图像和 220 个用于生成测试集的 ROI 图像,对划分好的 ROI 图像采用上述有重叠的滑动窗口裁剪方式进行裁剪,生成最终的 TCT 数据集,其中的训练集包含 8106 张图像,测试集包含 2125 张图像。数据统计信息如表 2 所示。

3.2 实验设置与评价指标

本文的实验环境为: Intel Core i9-10940X, NVIDIA GeForce RTX 3090Ti, ubuntu18.04, 64 位操作系统, MMDetection 深度学习框架^[20]。编程平台为 Pycharm, 编程语言为 Python, 图形加速环境为 CUDA11.1 和 CuDNN8.0.5。实验设置训练轮数为 12 个 epoch, 初始学习率为 0.001, 采用 SGD 优化器和线性预热学习率调整策略, 预热迭代次数为 2000, 批次大小为 8。为了加速模型收敛, ResNeSt-50 采用在 ImageNet 分类任务中得到的预训练权重。其他均为 MMDetection 配置文件中的默认设置。

采用 $R_{IoU}=0.5$ 阈值下的平均精度均值 (mAP, 在

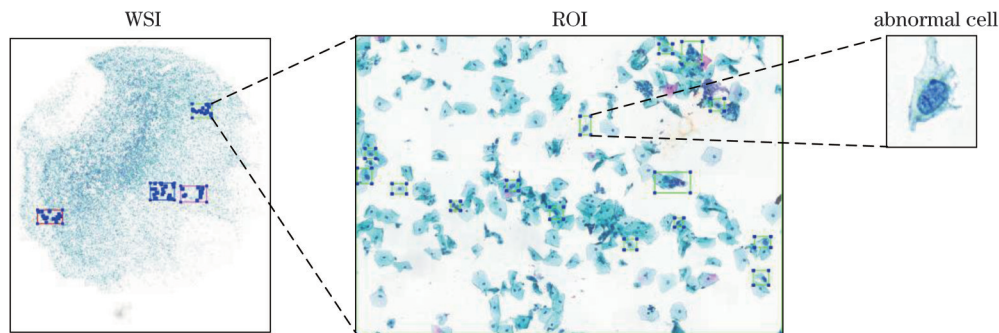


图 7 数据标注示例

Fig. 7 Example of data annotation

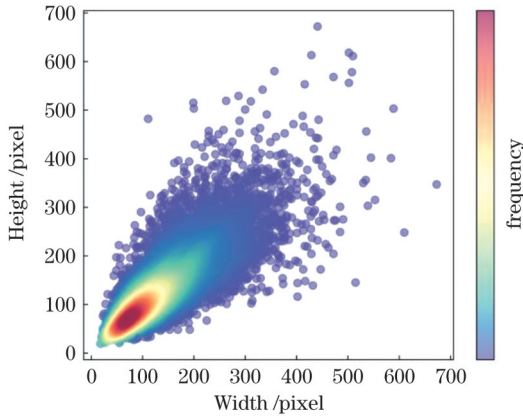


图 8 标注框尺寸分布密度图

Fig. 8 Distribution density graph of label boxes size

公式中记为 R_{mAP}) 对算法整体的检测性能进行评估; 对于 COCO 数据集^[21]中定义的小、中、大目标, 采用 $R_{IoU}=0.50:0.95$ 多阈值下的 mAP 进行评估, 小、中、大目标的 mAP 分别用 mAP-s、mAP-m 和 mAP-l 表示。mAP 的计算公式为

$$R_{mAP} = \frac{1}{C} \sum_{c=c_0}^1 \int P(R) dR, \quad (5)$$

其中,

$$P = \frac{T_p}{T_p + F_p}, \quad (6)$$

表 3 消融实验结果

Table 3 Ablation experimental experiment

Group	ResNeSt-50	BFP	CIoU Loss	mAP / %	mAP-s / %	mAP-m / %	mAP-l / %	Recall %
1	×	×	×	80.2	13.6	38.5	49.1	59.0
2	✓	×	×	81.0	15.0	39.6	50.8	58.9
3	✓	✓	×	81.3	15.9	39.8	50.9	60.7
4	✓	✓	✓	83.4	24.4	41.5	51.8	60.2

组 1 基准 RetinaNet 算法的 mAP 为 80.2%; 组 2 采用 ResNeSt-50 作为特征提取网络, 相比基准算法, mAP 整体提升 0.8 个百分点, 其中小目标 mAP-s 提升 1.4 个百分点, 表明 ResNeSt-50 提高了算法对宫颈异常细胞特征的识别能力; 组 3 进一步引入 BFP 结构来增强小目标的语义信息, 同时引入图像的全局信息辅助细胞分类, 但受限于 SmoothL1 Loss 的缺陷, 算法的整体检测精度相比组 2 仅提升了 0.3 个百分点, 小目标 mAP-s 提升了 0.9 个百分点; 组 4 在以上基础上采用 CIoU Loss 作为回归分支的损失函数, mAP 相比组 3 提升了 2.1 个百分点, 其中小目标 mAP-s 显著提升了 8.5 个百分点, 表明 CIoU Loss 实现了对异常细胞边界更精确的回归预测。相较于基准 RetinaNet 算法, 最终 AFF-RetinaNet 的整体检测精度 mAP 和小目标检测精度 mAP-s 分别提升 3.2 个百分点和 10.8 个百分点。

表 2 实验数据统计信息

Table 2 Statistics of experiment data

Data type	Number
WSI	233
ROI (total)	1103
Abnormal cell	11188
ROI for generating training set	883
Training set	8106
ROI for generating test set	220
Test set	2125

$$R = \frac{T_p}{T_p + F_N}, \quad (7)$$

式中: C 表示检测物体的类别数; P 表示精确率; R 表示召回率; T_p 表示被正确检测到的异常细胞数量; F_p 表示本身为正常细胞但被错检的细胞数量; F_N 表示本身为异常细胞但被算法漏检的细胞数量。

3.3 消融实验

为验证本文所提算法中不同改进策略的有效性, 以 RetinaNet 算法为基准 (组 1), 基于 TCT 数据集对 ResNeSt-50 特征提取网络、BFP 结构以及 CIoU Loss 三部分开展消融实验 (组 2、组 3、组 4)。实验结果如表 3 所示。

消融实验结果表明, 本文提出的三个改进策略均能提升 AFF-RetinaNet 算法对宫颈异常细胞的检测效果。

3.4 与不同检测网络的对比

为验证本文所提 AFF-RetinaNet 整体性能的优越性, 在相同的实验环境下, 选择 YOLOv3^[22]、Faster R-CNN、RetinaNet 等在宫颈异常细胞检测领域表现突出的先进算法以及 Cascade R-CNN^[23]、FCOS^[24] 等其他主流目标检测算法为基线, 对比以上算法与 AFF-RetinaNet 模型的参数数量以及在 TCT 数据集上的 mAP、mAP-s 和每秒检测帧数 (FPS)。由表 4 可知, 相比基准 RetinaNet, AFF-RetinaNet 虽然在一定程度上损失了检测速度, 但其在模型参数数量仅增加 0.3×10^6 的条件下, 在 TCT 数据集上的 mAP 和 mAP-s 分别达到 83.4% 和 24.4%, 优于表 4 中的其他算法。

表 4 主流算法在 TCT 数据集上的性能对比

Table 4 Performance comparison of mainstream algorithms on TCT dataset

Algorithm	mAP / %	mAP-s / %	FPS	Parameter quantity / 10 ⁶
YOLOv3	70.7	14.9	101.0	61.5
Cascade R-CNN	75.1	15.4	31.2	68.9
Faster R-CNN	77.4	14.3	37.3	41.1
LAD ^[25]	77.6	14.9	21.6	31.8
FCOS	79.1	16.3	44.8	31.8
TOOD ^[26]	79.9	15.5	20.7	31.8
RetinaNet	80.2	13.6	49.2	36.1
DDOD ^[27]	80.7	14.3	39.6	31.9
AFF-RetinaNet	83.4	24.4	25.4	36.4

3.5 AFF-RetinaNet 检测结果分析

为了更直观地展示 AFF-RetinaNet 相比于基准

RetinaNet 的优势, 随机选取测试集中的 4 张图像, 对检测结果进行可视化。真实标注、RetinaNet 以及 AFF-RetinaNet 检测结果如图 9(a)~(c) 所示, 其中从左到右每一列依次为示例 1、示例 2、示例 3 和示例 4。示例 1 中存在两个中等尺寸的异常细胞, 此时两种算法均可准确检测到异常细胞; 示例 2 中存在多个小目标, 此时 RetinaNet 出现较为严重的小目标漏检情况, 而 AFF-RetinaNet 仍可准确地检测到大部分小目标; 示例 3 中包含 5 个异常细胞, 且不同细胞的高宽比存在一定差距, RetinaNet 除了检测到以上细胞外, 还误检了 2 个目标, 而 AFF-RetinaNet 则进一步降低了误检的个数; 示例 4 为重叠细胞检测场景, RetinaNet 漏检了 1 个目标, AFF-RetinaNet 可以准确检测到所有目标。

通过对上述检测结果进行对比分析可知, AFF-RetinaNet 可以较为准确地在 TCT 图像中检测宫颈异常细胞, 虽然存在少数误检和漏检的情况, 但相比于基准 RetinaNet 算法检测效果显著提升。

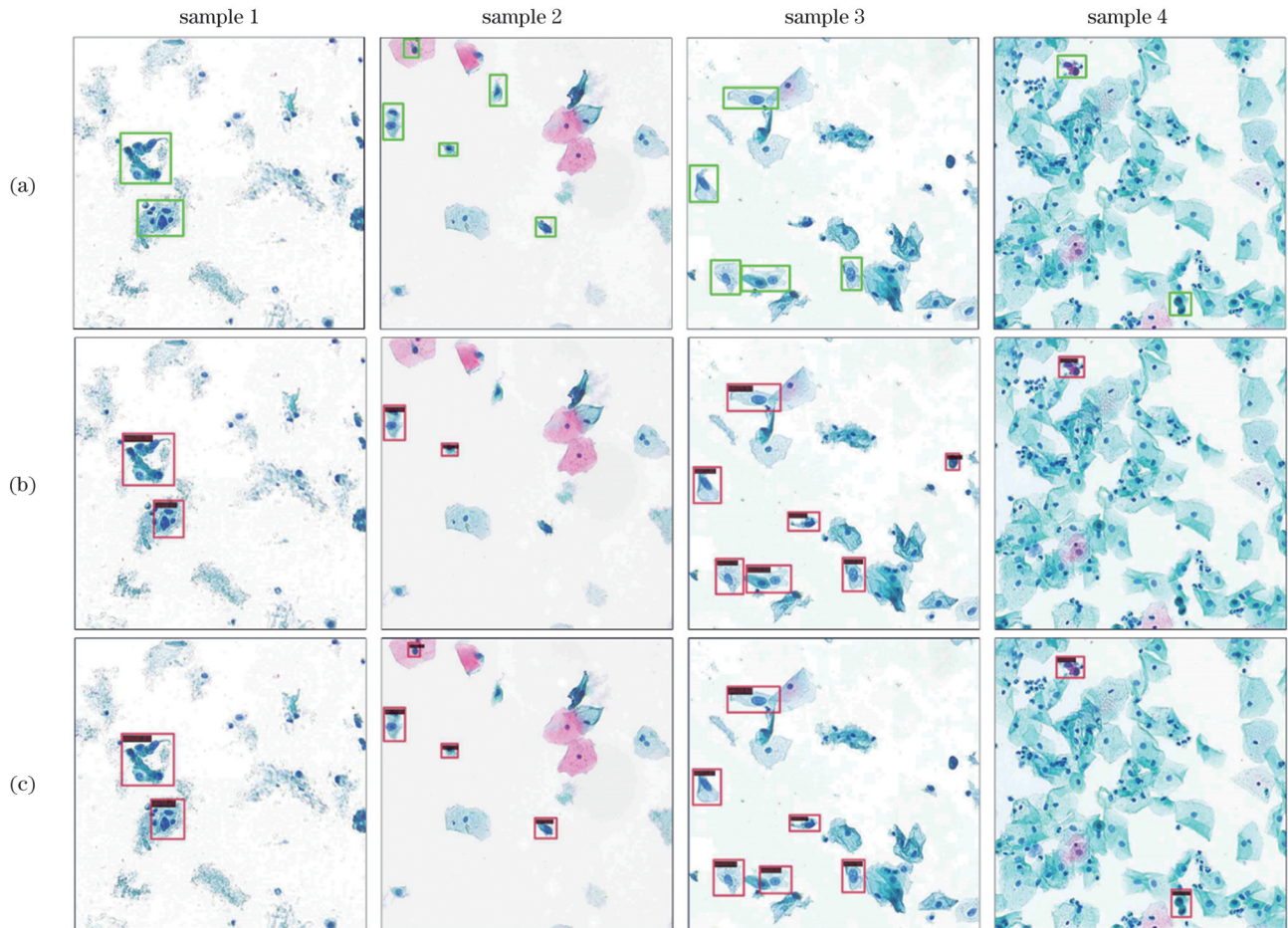


图 9 AFF-RetinaNet 和 RetinaNet 的检测结果对比。(a) 真实标注; (b) RetinaNet; (c) AFF-RetinaNet

Fig. 9 Comparison of AFF-RetinaNet and RetinaNet detection results. (a) Ground truth; (b) RetinaNet; (c) AFF-RetinaNet

3.6 WSI 推理结果分析

如图 10 所示, WSI 推理流程可检测 WSI 中所有区域的疑似异常细胞, 但由于标注只存在于 ROI 中, 因此只有 ROI 中的检测结果才有评价意义。本节针对

生成测试集的 220 个 ROI 中的检测结果计算 mAP, 以定量评估 WSI 推理流程的结果。基于不同目标检测算法的 WSI 推理结果的 mAP 如表 5 所示, 其中基于 AFF-RetinaNet 的 WSI 推理结果的 mAP 最优, 达到

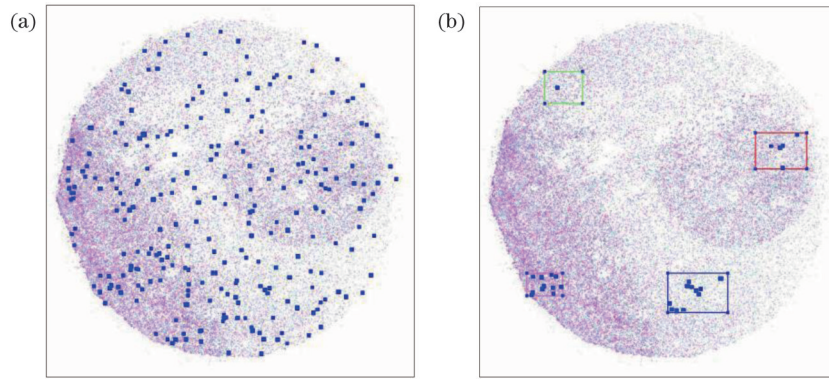


图 10 WSI推理结果和真实标注对比。(a)WSI推理结果;(b)真实标注

Fig. 10 Comparison of WSI inference results and ground truth. (a) WSI inference results; (b) ground truth

表 5 基于不同算法的 WSI推理结果对比

Table 5 Comparison of WSI inference results based on different algorithms

Algorithm	mAP / %
YOLOv3	56.5
LAD	59.0
FCOS	60.6
DDOD	62.0
TOOD	64.6
Cascade R-CNN	66.3
Faster R-CNN	67.4
RetinaNet	70.2
AFF-RetinaNet	70.8

70.8%。为进一步直观地展示基于 AFF-RetinaNet 的 WSI推理结果的准确性,随机选取 2 个 ROI 图像检测结果进行可视化,如图 11 所示。从图中可以看出基于

AFF-RetinaNet 的宫颈 WSI推理流程可有效检测超大尺寸 WSI 中的大部分异常细胞,有望减轻医生的阅片负担。

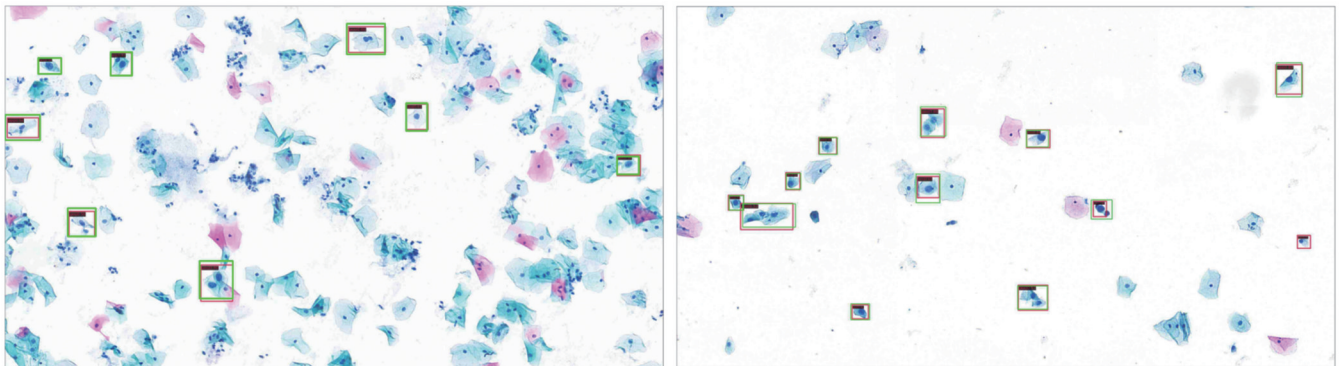


图 11 基于 AFF-RetinaNet 的 WSI推理结果在 ROI 中的可视化示例

Fig. 11 Visualization examples of AFF-RetinaNet-based WSI inference results in ROI

4 结 论

本文提出了一种 AFF-RetinaNet 目标检测算法,用于检测宫颈细胞病理图像中的异常细胞。首先,算法采用 ResNeSt-50 作为特征提取网络,其中的分组卷积和拆分注意力机制提高了算法对宫颈异常细胞的特征提取能力;其次,引入 BFP 结构实现全尺度的特征融合,保留了更多小目标的语义信息,提高了小目标的

检测精度,同时,通过 BFP 中的非局部注意力模块引入图像全局信息辅助细胞分类,以提高异常细胞的检测准确率;最后,采用 CIoU Loss 约束边界回归分支,并额外考虑了中心距离和形状关系等因素,提高了对不同形态异常细胞的定位精度。此外,根据实际应用场景,基于 AFF-RetinaNet 实现了 WSI推理流程,可自动检测超高分辨率 WSI 中的宫颈异常细胞。实验结果表明,AFF-RetinaNet 的 mAP 和 mAP-s 分别达到了

83.4% 和 24.4%，相较于基准 RetinaNet 分别提高 3.2 个百分点和 10.8 个百分点，增强了对小尺寸异常细胞的检测能力，有效提升了异常细胞的检测精度。基于 AFF-RetinaNet 的 WSI 推理流程检测结果显示该算法的 mAP 为 70.8%，说明所提算法可有效检测超高分辨率 WSI 中的大部分异常细胞，辅助细胞病理医生开展宫颈癌细胞学筛查工作。

综上，所提 AFF-RetinaNet 算法提高了宫颈异常细胞的整体检测精度，显著提升了小尺寸异常细胞的检测效果，基于该算法的 WSI 推理流程可以实现超高分辨率宫颈 WSI 中异常细胞的自动检测。在未来的研究工作中，笔者将进一步收集多类别宫颈异常细胞标签数据，以实现异常细胞的精细化识别，为医生提供更加详细的辅助诊断信息。

参 考 文 献

- [1] World Health Organization. Cervical cancer[EB/OL]. (2022-02-22) [2022-08-24]. <https://www.who.int/zh/news-room/fact-sheets/detail/cervical-cancer>.
- [2] 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: a Cancer Journal for Clinicians, 2021, 71(3): 209-249.
- [3] Bedell S L, Goldstein L S, Goldstein A R, et al. Cervical cancer screening: past, present, and future[J]. Sexual Medicine Reviews, 2020, 8(1): 28-37.
- [4] Adweb K M A, Cavus N, Sekeroglu B. Cervical cancer diagnosis using very deep networks over different activation functions[J]. IEEE Access, 2021, 9: 46612-46625.
- [5] Su J, Xu X, He Y J, et al. Automatic detection of cervical cancer cells by a two-level cascade classification system[J]. Analytical Cellular Pathology, 2016, 2016: 9535027.
- [6] Chankong T, Theera-Umpon N, Auephanwiriyakul S. Automatic cervical cell segmentation and classification in Pap smears[J]. Computer Methods and Programs in Biomedicine, 2014, 113(2): 539-556.
- [7] 段仲静, 李少波, 胡建军, 等. 深度学习目标检测方法及其主流框架综述[J]. 激光与光电子学进展, 2020, 57(12): 120005.
- [8] Duan Z J, Li S B, Hu J J, et al. Review of deep learning based object detection methods and their mainstream frameworks[J]. Laser & Optoelectronics Progress, 2020, 57(12): 120005.
- [8] 何自芬, 陈光晨, 陈俊松, 等. 多尺度特征融合轻量化夜间红外行人实时检测[J]. 中国激光, 2022, 49(17): 1709002.
- [8] He Z F, Chen G C, Chen J S, et al. Multi-scale feature fusion lightweight real-time infrared pedestrian detection at night[J]. Chinese Journal of Lasers, 2022, 49(17): 1709002.
- [9] 刘丰, 韩民, 万军, 等. 基于深度学习的牙齿病变自动检测算法[J]. 中国激光, 2022, 49(20): 2007207.
- [9] Liu F, Han M, Wan J, et al. Automatic detection of dental lesions based on deep learning[J]. Chinese Journal of Lasers, 2022, 49(20): 2007207.
- [10] 徐志京, 柏雪. 基于双重特征增强的遥感舰船小目标检测[J]. 光学学报, 2022, 42(18): 1828002.
- [10] Xu Z J, Bai X. Small ship target detection method for remote sensing images based on dual feature enhancement[J]. Acta Optica Sinica, 2022, 42(18): 1828002.
- [11] Xu M Q, Zeng W X, Sun Y H, et al. Cervical cytology intelligent diagnosis based on object detection technology[EB/OL]. [2022-12-05]. <https://openreview.net/pdf?id=r14Qkwsif>.
- [12] Ren S Q, He K M, Girshick R, et al. Faster R-CNN: towards real-time object detection with region proposal networks[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2017, 39(6): 1137-1149.
- [13] 尹远来, 赵磊. 基于改进的 Faster R-CNN 模型的异常鳞状上皮细胞检测[J]. 智能计算机与应用, 2021, 11(2): 7-13.
- [13] Yin Y L, Zhao L. Atypical squamous cells detection based on improved Faster R-CNN[J]. Intelligent Computer and Applications, 2021, 11(2): 7-13.
- [14] Lin T Y, Dollár P, Girshick R, et al. Feature pyramid networks for object detection[C]//2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), July 21-26, 2017, Honolulu, HI, USA. New York: IEEE Press, 2017: 936-944.
- [15] Cao L, Yang J Y, Rong Z W, et al. A novel attention-guided convolutional network for the detection of abnormal cervical cells in cervical cancer screening[J]. Medical Image Analysis, 2021, 73: 102197.
- [16] Lin T Y, Goyal P, Girshick R, et al. Focal loss for dense object detection[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2020, 42(2): 318-327.
- [17] 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), June 19-20, 2022, New Orleans, LA, USA. New York: IEEE Press, 2022: 2735-2745.
- [18] Pang J M, Chen K, Shi J P, et al. Libra R-CNN: towards balanced learning for object detection[C]//2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), June 15-20, 2019, Long Beach, CA, USA. New York: IEEE Press, 2020: 821-830.
- [19] Zheng Z H, Wang P, Liu W, et al. Distance-IoU loss: faster and better learning for bounding box regression[J]. Proceedings of the AAAI Conference on Artificial Intelligence, 2020, 34(7): 12993-13000.
- [20] Chen K, Wang J Q, Pang J M, et al. MMDetection: open MMLab detection toolbox and benchmark[EB/OL]. (2019-06-17) [2023-02-13]. <https://arxiv.org/abs/1906.07155>.
- [21] Lin T Y, Maire M, Belongie S, et al. Microsoft COCO: common objects in context[M]//Fleet D, Pajdla T, Schiele B, et al. Computer vision-ECCV 2014. Lecture notes in computer science. Cham: Springer, 2014, 8693: 740-755.
- [22] Redmon J, Farhadi A. YOLOv3: an incremental improvement[EB/OL]. (2018-04-08) [2023-02-13]. <https://arxiv.org/abs/1804.02767>.
- [23] Cai Z W, Vasconcelos N. Cascade R-CNN: delving into high quality object detection[C]//2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition, June 18-23, 2018, Salt Lake City, UT, USA. New York: IEEE Press, 2018: 6154-6162.
- [24] Tian Z, Shen C H, Chen H, et al. FCOS: a simple and strong anchor-free object detector[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2022, 44(4): 1922-1933.
- [25] Nguyen C H, Nguyen T C, Tang T N, et al. Improving object detection by label assignment distillation[C]//2022 IEEE/CVF Winter Conference on Applications of Computer Vision (WACV), January 3-8, 2022, Waikoloa, HI, USA. New York: IEEE Press, 2022: 1322-1331.
- [26] Feng C J, Zhong Y J, Gao Y, et al. TOOD: task-aligned one-stage object detection[C]//2021 IEEE/CVF International Conference on Computer Vision (ICCV), October 10-17, 2021, Montreal, QC, Canada. New York: IEEE Press, 2022: 3490-3499.
- [27] Chen Z H, Yang C, Li Q F, et al. Disentangle your dense object detector[C]//Proceedings of the 29th ACM International Conference on Multimedia, October 20-24, 2021, Virtual Event, China. New York: ACM Press, 2021: 4939-4948.

Abnormal Cervical Cell Detection Algorithm Based on Improved RetinaNet

Liu Runkun^{1,2}, Dang Shijie², Zhang Hongyuan², Niu Yinyin³, Mi Guanxun³, Li Sanhua³, Chen Zhenxin², Zhao Lingxiao^{2*}, Li Peng²

¹*Division of Life Sciences and Medicine, School of Biomedical Engineering (Suzhou), University of Science and Technology of China, Hefei 230026, Anhui, China;*

²*Suzhou Institute of Biomedical Engineering and Technology, Chinese Academy of Science, Suzhou 215162, Jiangsu, China;*

³*Henan Celnovte Biotechnology Co., Ltd., Zhengzhou 450001, Henan, China*

Abstract

Objective With the development of digital pathology and artificial intelligence technology, research on the automatic detection of abnormal cervical cells has made great progress. Of the different technologies, object detection technology based on deep learning can simultaneously locate and classify an object, making it a promising application in the field of abnormal cervical cell detection. However, the detection accuracy of abnormal cervical cells still has room for improvement because of the subtle features of abnormal cells that are difficult to extract, small targets that are easily missed, and inaccurate boundary regression. Therefore, in this study, a full-scale feature fusion RetinaNet algorithm combined with attention (AFF-RetinaNet) is proposed for abnormal cervical cell detection to improve the detection accuracy of abnormal cervical cells. In addition, for practical application scenarios, a whole slide image (WSI) inference process based on the AFF-RetinaNet algorithm was implemented for detecting abnormal cells in WSI, which can help pathologists quickly locate abnormal cervical cells in high-resolution WSI and reduce the burden of having to read cervical cytology images.

Methods To improve the accuracy of abnormal cervical cell detection, AFF-RetinaNet is proposed. First, ResNeSt-50 was used as the feature extraction network to extract the fine features of abnormal cervical cells. The structure of balanced feature pyramid (BFP) was then applied to integrate all feature layers at full scale and obtain the global information of the image, which can enhance the semantic information of small targets. Finally, CIoU loss was used as the loss function of the regression branch to improve the accuracy of abnormal cell boundary regression. In the WSI inference process based on AFF-RetinaNet, the WSI was first divided into several patches, and AFF-RetinaNet was then used to obtain the detection results of each patch. Finally, all detection results of patches were integrated, and the non-maximum suppression (NMS) algorithm was used as a post-process to obtain the final WSI detection results.

Results and Discussions Figure 1 shows the proposed AFF-RetinaNet for abnormal cervical cell detection. The ResNeSt-50 network was used as the feature extraction network (Fig. 2). Results show that the overall mean average precision (mAP) of the AFF-ResNeSt-50 increases by 0.8 percentage points compared with the RetinaNet, whereas the mAP of small objects (mAP-s) increases by 1.4 percentage points (Table 3), indicating that the ResNeSt-50 network can improve the ability to identify the features of abnormal cells. BFP was added to integrate all feature layers at full scale and introduce the global information of the image. However, due to the defects of SmoothL1 Loss, the mAP and mAP-s increase by only 0.3 and 0.9 percentage points, respectively (Table 3). Finally, the CIoU loss was used as the loss function of the regression branch to obtain a more accurate boundary regression. Accordingly, the mAP and mAP-s increase by 2.1 and 8.5 percentage points, respectively (Table 3). In a comparison experiment with other mainstream target detection algorithms, AFF-RetinaNet exhibits optimal detection performance (Table 4), further verifying the superiority of AFF-RetinaNet. The WSI inference process based on the AFF-RetinaNet algorithm (Fig. 6) can automatically detect abnormal cells in a high-resolution WSI (Fig. 10) with an mAP of 70.8% in the regions of interest (Table 5).

Conclusions In this study, a detection algorithm called AFF-RetinaNet was proposed for detecting abnormal cervical cells. ResNeSt-50 was used as the feature extraction network to improve the ability to identify the features of abnormal cells. BFP realizes full-scale feature fusion and introduces the global information of an image to retain more semantic information about small objects. CIoU loss considers additional factors, such as the central distance and shape relationship, which improve the accuracy of abnormal cell boundary regression. In addition, the WSI inference process based on the AFF-RetinaNet can automatically detect abnormal cells in a high-resolution WSI. Experimental results show that the mAP of AFF-RetinaNet on the abnormal cervical cell dataset is 83.4%, and the mAP-s reaches 24.4%, which are 3.2 and 10.8 percentage points higher, respectively, than those of the baseline RetinaNet and better than those of the other mainstream object detection algorithms. The detection results of the cervical WSI inference process based on AFF-RetinaNet reveal an mAP of 70.8% in the regions of interest. Our study shows that AFF-RetinaNet can improve the detection accuracy of abnormal cervical cells and that the WSI inference process based on AFF-RetinaNet can assist pathologists in cervical cancer screening, giving the proposed algorithm significant clinical application value.

Key words medical optics; cervical cytopathological image; object detection; small object; feature fusion; attention mechanism