

改进 YOLOv5 框架的血细胞检测算法^①



张 昊, 郑广海, 张 鑫, 吕 娜

(大连交通大学 软件学院, 大连 116052)

通信作者: 郑广海, E-mail: zgh@djtu.edu.cn

摘 要: 对于血液中红细胞、白细胞、血小板等成分的观察和计数是临床医学诊断的重要依据. 血细胞的异常意味着可能存在凝血异常、感染、炎症等与血液相关的问题. 人工检测血细胞不仅耗费人力, 且容易出现误检、漏检的情况. 因此, 针对上述情况, 提出一种新颖的血细胞检测算法—YOLOv5-CBF. 该算法在 YOLOv5 框架的基础上, 通过在主干网络中加入坐标注意力 (coordinate attention, CA) 机制, 提高检测精度; 将颈部网络中的 FPN+PAN 结构中改为结合了跨尺度特征融合方法 (bidirectional feature pyramid network, BiFPN) 思想的特征融合结构, 使目标多尺度特征有效融合; 在三尺度检测的基础上增加了一个小目标检测层, 提高对数据集中小目标血小板的识别精度. 通过在数据集 BCCD 上进行的大量的实验结果表明: 与传统的 YOLOv5 算法相比较, 该算法在 3 类血细胞检测的平均精度提升 2.7%, 试验效果良好, 该算法对血细胞检测具有很高的实用性.

关键词: 深度学习; 血细胞检测; YOLOv5; 多尺度特征融合; 注意力机制

引用格式: 张昊, 郑广海, 张鑫, 吕娜. 改进 YOLOv5 框架的血细胞检测算法. 计算机系统应用, 2023, 32(5): 123-131. <http://www.c-s-a.org.cn/1003-3254/9082.html>

Improved YOLOv5 Algorithm for Blood Cell Detection

ZHANG Hao, ZHENG Guang-Hai, ZHANG Xin, LYU Na

(Software Technology Institute, Dalian Jiaotong University, Dalian 116052, China)

Abstract: The observation and counting of red blood cells, white blood cells, and platelets in the blood are an important basis for clinical medical diagnosis. Abnormal blood cells mean that there may be blood-related problems such as clotting abnormalities, infections, and inflammation. As artificial blood cell detection is not only labor-intensive but also prone to false detection and misses, a novel blood cell detection algorithm YOLOv5-CBF is proposed to address the above problem. On the basis of the YOLOv5 framework, the algorithm improves detection accuracy by adding a coordinate attention (CA) mechanism to the backbone network. The FPN+PAN structure in the neck network is changed to the feature fusion structure combining the idea of the bidirectional feature pyramid network (BiFPN), a cross-scale feature fusion method; in this way, the multi-scale features of the target can be effectively fused. In addition to the three-scale detection, a small target detection layer is added to improve the identification accuracy of small target platelets in the dataset. The results of a large number of experiments conducted on the dataset BCCD show that the algorithm presents an average accuracy improvement of 2.7% in the detection of the three blood cells compared to the conventional YOLOv5 algorithm, demonstrating good performance. The algorithm is highly practical for blood cell detection.

Key words: deep learning; blood cell detection; YOLOv5; multi-scale feature fusion; attention mechanism

^① 基金项目: 辽宁省教育厅服务地方项目 (JDL2020004)

收稿时间: 2022-10-27; 修改时间: 2022-11-29; 采用时间: 2022-12-11; csa 在线出版时间: 2023-02-28

CNKI 网络首发时间: 2023-03-01

1 引言

根据显微图像进行血细胞检测在疾病识别领域发挥着重要作用. 血细胞在医学上分为3种类型: 白细胞(WBC)、血小板(platelets)和红细胞(RBC)^[1]. 成年男性的正常血细胞数目约为每毫升4.7~610万, 女性的正常血细胞数目约为4.2~540万, 这些血细胞的数目和形态对临床诊断而言是至关重要的, 当人体发生感染或者其他疾病时, 医生可以通过血细胞的检查来辅助诊断^[2]. 血细胞的检测早期普遍采用人工镜检的方式, 即先手工制片、染色, 再通过人工分类来完成. 人工检测血细胞的方法容易受到主观因素的影响, 且由于数据量大且目标小, 容易导致准确率不高且耗费人力^[3,4]而使用基于深度卷积神经网络的自动检测算法来准确高效地检测血细胞已经取得了较好的检测效果, 提高医疗系统的检测效率^[5-7].

随着深度学习领域的发展, 越来越多的目标检测方法被提出, 目标检测技术广泛应用于医疗影像、缺陷检测、自动驾驶等领域中. 主流的目标检测法大致分为两类: 一类是两阶段检测, 比如R-CNN算法^[8], 该算法在第1阶段生成候选对象区域放入深度网络中提取特征, 在第2阶段将候选区域进行分类, 这样的两阶段检测虽然运行速度缓慢但精度很高. 另一类是单阶段检测: 比如YOLO系列^[9-12]和SSD算法^[13], 其相较两阶段检测的检测速度有所提升, 准确率也在逐渐提高.

综上, 本文选择YOLOv5-6.0算法作为基础网络, 在复杂度低的同时有较高的识别精度, 对检测精度以及速度要求较高的血细胞检测上, 具有较高的实用性. 本文在YOLOv5的骨干网络中, 加入了坐标注意力^[14], 增强特征提取的能力, 将YOLOv5模型的FPN^[15]+PAN^[16]结构引入BiFPN^[17]的思想进行多尺度特征融合, 在三尺度检测的基础上增加了一个小目标检测层. 通过在公共血细胞检测数据集上的对比实验验证, 本文方法的检测效果显著提升, 可以很好地应用于血细胞检测任务中.

2 算法原理

YOLOv5是YOLO系列的第5代目标检测算法, 包含了1.0、5.0、6.0等多种版本, 包含了YOLOv5s、YOLOv5m、YOLOv5l、YOLOv5x等多种规模的网

络, 模型的大小依次增加、精度也依次增加. YOLOv5-6.0沿用v3、v4的整体布局, 具有更高的检测精度和速度, 与YOLOv5差距不大. YOLOv5-6.0由输入端、主干网络、颈部网络、预测分类器4个部分组成. 本文选择的基础网络YOLOv5-6.0的整体结构如图1所示.

2.1 输入端

输入端将尺寸不同的输入样本采用自适应图片缩放的方式将图像尺寸缩放到一个固定的尺寸, 进行自适应填充, 再送入主干网络中进行训练. YOLOv5的数据增强采用了和YOLOv4一样的Mosaic数据增强的方式, 用于解决数据集样本不足以及尺寸分布不均衡的问题, Mosaic数据增强将4张图片通过随机裁剪、随机亮度调整、翻转等方式进行拼接, 减少了GPU的使用, 增强了网络的鲁棒性.

2.2 主干网络

YOLOv5-6.0的主干网络将卷积结构中的激活函数改为SiLU函数, 并将Focus模块换成了一个卷积核(kernel)为6, 步长(stride)为2, 填充(padding)为2的卷积层, 对于现有的一些GPU设备而言, 使用卷积层比使用Focus模块更高效. YOLOv5-6.0将空间金字塔池化结构(SPP)改为SPPF结构, 并且将SPPF的结构放在主干网络的最后一层. SPPF的结构如图1所示, 将输入串行通过卷积层和多个5×5大小的最大池化层(MaxPool), 提高了计算速度.

2.3 颈部网络

YOLOv5-6.0的Neck部分采用特征金字塔网络(FPN)+路径聚合网络(PAN)的结构, 特征金字塔结构通过上采样的方式将深层特征图中较强的语义信息传到浅层. 路径聚合网络通过下采样的方式将浅层的位置信息传递到深层, 进行多尺度的特征融合.

2.4 预测分类器

YOLOv5的head用3个1×1的卷积层替代全连接层(FC)进行预测和分类, 分别生成20×20, 40×40, 80×80这3个不同尺度的特征图, 输出结果为4+1+C, 分别为预测框坐标, 目标置信度、类别概率值.

3 改进的YOLOv5算法

本实验基于YOLOv5-6.0算法提出了改进网络YOLOv5-CBF. 在主干网络中引入CA模块, 增强模型

对3类血细胞目标的定位能力,其参数如表1所示,Input size表示该层的输入尺寸,Number表示该层的数量,Module表示模块名,Args表示模块的初始化参数,分别为:输入通道数,输出通道数,卷积核尺寸.在颈部网络上结合BiFPN的思想,引入AWFF结构进行加权

特征融合,加强特征的多尺度表达.在三尺度检测的基础上增加了一个小目标检测层,提高对数据集中小目标血小板的识别精度,其检测速度能够很好地满足实时检测的要求,改进后的YOLOv5-CBF整体结构如图2所示.

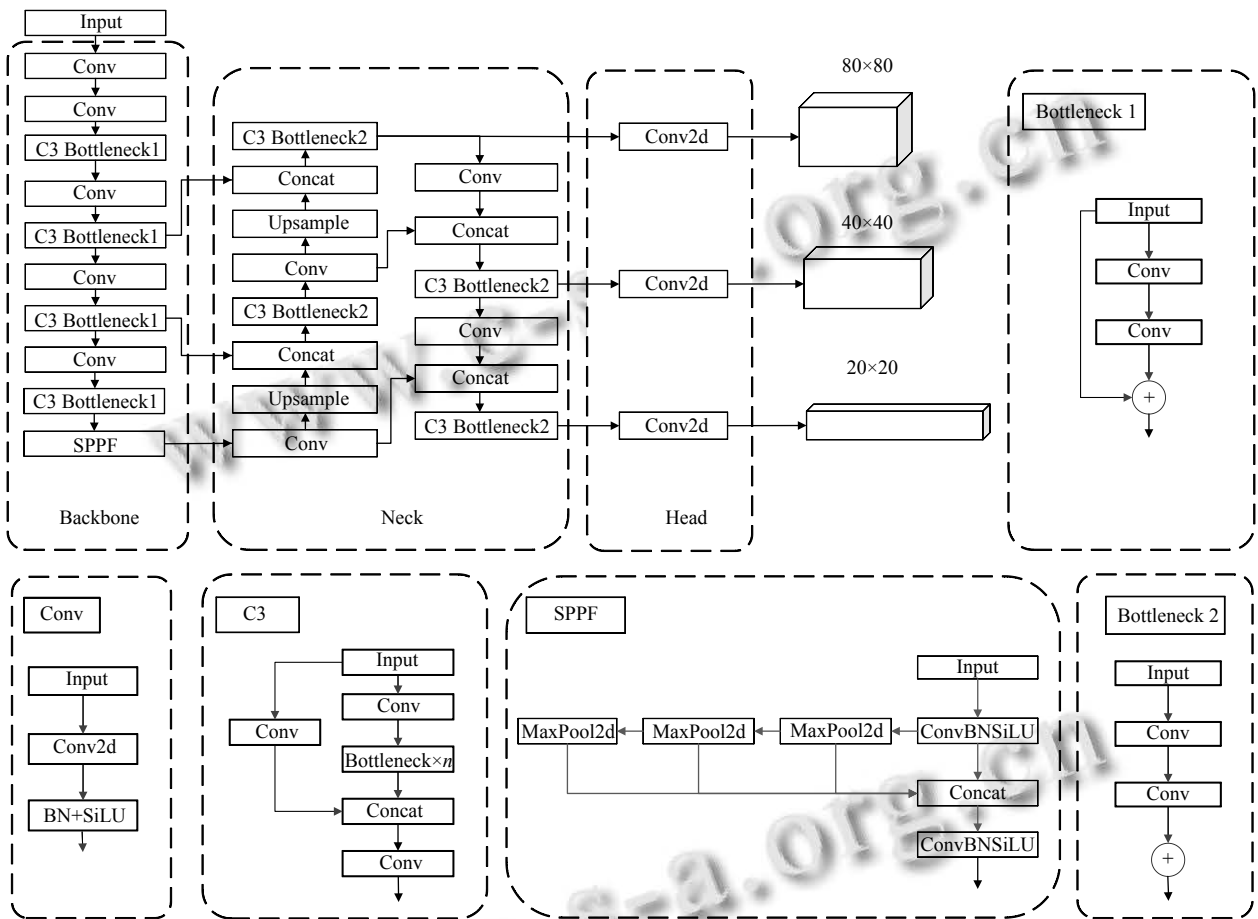


图1 YOLOv5-6.0网络结构

表1 主干网络参数

Layers	Module	Input size	Number	Args
1	Conv	640×640×3	1	[3, 32, 6, 2, 2]
2	Conv	320×320×32	1	[32, 64, 3, 2]
3	CA结构	160×160×64	1	[64, 64]
4	C3结构	160×160×64	1	[64, 64, 1]
5	Conv	160×160×64	1	[64, 128, 3, 2]
6	C3结构	80×80×128	2	[128, 128, 2]
7	Conv	80×80×128	1	[128, 256, 3, 2]
8	C3结构	40×40×256	3	[256, 256, 3]
9	Conv	40×40×256	1	[256, 512, 3, 2]
10	C3结构	20×20×512	1	[512, 512, 1]
11	CA结构	20×20×512	1	[512, 512]
12	SPPF	20×20×512	1	[512, 512, 5]

3.1 注意力机制模块

在深度学习领域,注意力机制常用来帮助模型更准确地定位和识别感兴趣的对象.神经网络中的注意力机制有通道注意^[18]、空间注意^[19]和混合注意^[20]等多种类型.Hou等的研究成果^[14]表明,通道注意力(例如,SE注意力)虽然可以提升模型的性能,但它们通常会忽略位置信息,而CA模块将通道注意力按维度分解,并分别沿不同的空间方向聚合特征,捕获跨通道信息和位置信息.

CA模块可以帮助轻量级的模型更加精确的定位和识别感兴趣的目标,增强YOLOv5-6.0网络的特征

表达能力, 进而提高血细胞的检测精度, CA 结构如图 3 所示.

CA 对输入的特征图 X . 首先使用两个一维的池化核分别沿水平与垂直方向对每个通道进行编码, 第 c 个通道在其对应的高度和宽度输出可以分别表示为:

$$z_c^h(h) = \frac{1}{W} \sum_{0 \leq i < W} X_c(h, i) \quad (1)$$

$$z_c^w(w) = \frac{1}{H} \sum_{0 \leq j < H} X_c(j, w) \quad (2)$$

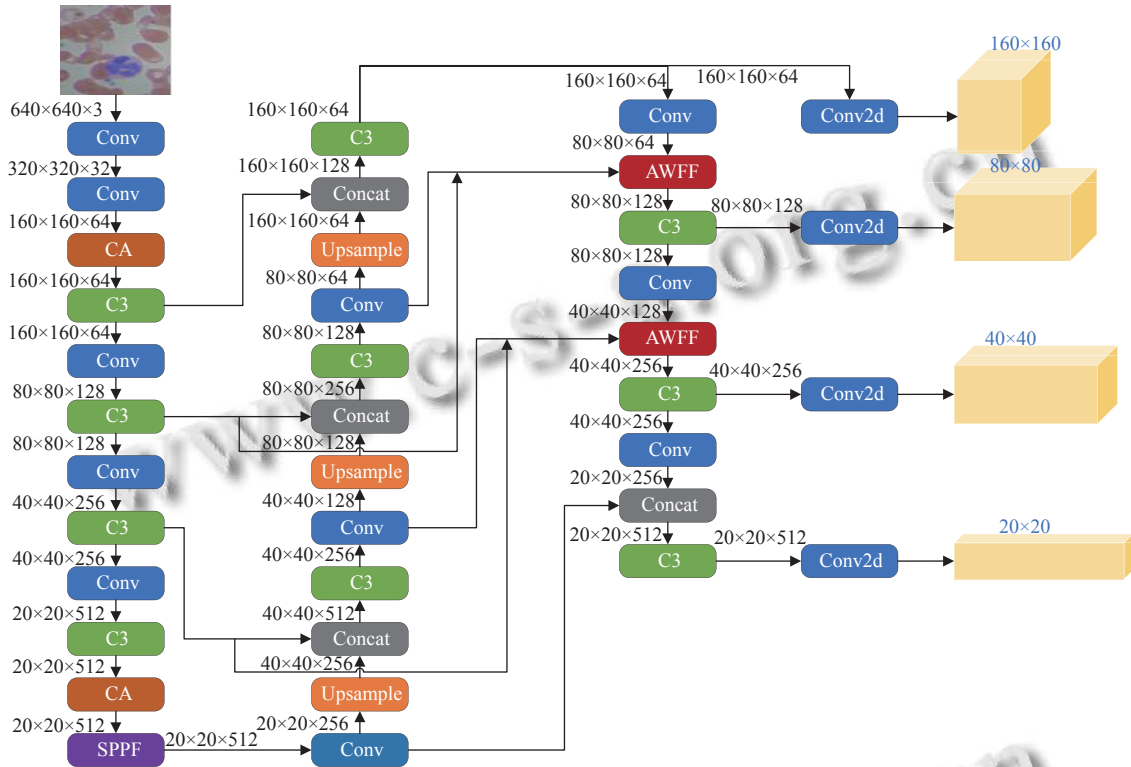


图 2 YOLOv5-CBF 网络结构

$$f = \delta(F_1([Z^h, Z^w])) \quad (3)$$

利用激活函数将 f 分解为 2 个单独的张量 f^h 和 f^w . 用 2 个 1×1 卷积变换分别将 f^h 和 f^w 变换为与输入 X 具有相同通道数的张量 $g^h = \sigma(F_h(f^h))$ 和 $g^w = \sigma(F_w(f^w))$. 最后, 输出为特征图与两个权重图相乘:

$$y_c(i, j) = x_c(i, j) \times g_c^h(i) \times g_c^w(j) \quad (4)$$

CA 模块增强了特征图的识别能力, 可以帮助轻量级的 YOLOv5 模型更加精确的定位和识别感兴趣的目标. YOLOv5-CBF 在 Backbone 中加入了两个 CA 模块, 改进后的 Backbone 网络详细参数如表 1 所示.

3.2 跨尺度特征融合方法 BiFPN

YOLOv5 的 Neck 部分采用 FPN+PAN 的结构. FPN 结构将语义特征自顶向下传递, 路径聚合结构将定位特征自底向上传递. 这种多尺度特征的融合将特征图变换成同样的尺寸进行连接, 通常对融合输出特

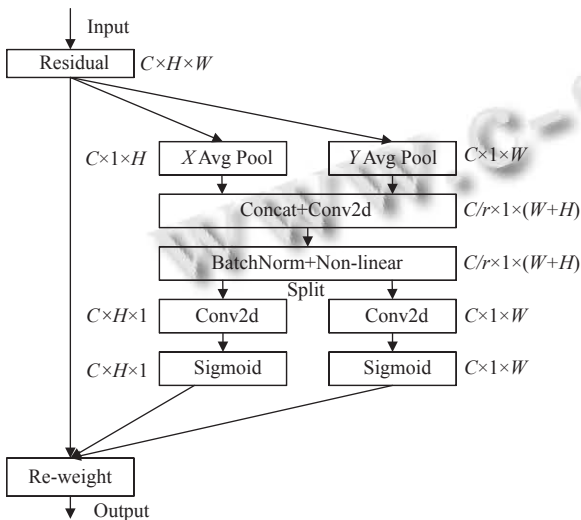


图 3 CA 结构

其次, 对两个带有特定方向信息的特征图进行拼接, 得到:

征的贡献是均衡的,因此无法充分利用不同尺度间的特征.为更为充分地进行多尺度特征融合,提高检测精度,本算法的 Neck 部分借鉴了一种具有跨尺度连接和加权特征融合的 BiFPN 结构^[17]的思想.

BiFPN 结构第 1 步删除只有一条输入边的节点.原因是只有一条输入边且不参与特征融合的特征相较于其他集成不同特征的节点的贡献较小.其次,如果它们处于同一级别,我们将从原始输入额外添加一条边到输出节点进行跳跃连接. BiFPN 结构将每个双向路径看作一个特征网络层,并多次复制同一层,以实现更优的特征融合结构.上述优化所产生的 BiFPN 结构如图 4 所示,下箭头为上采样,上箭头为下采样.

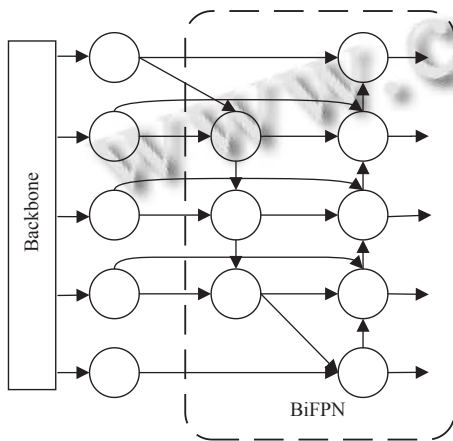


图 4 BiFPN 特征融合结构

为解决在不同深度下特征难以有效融合的问题,本文提出了注意力加权特征融合 (attention-weighted feature fusion, AWFF) 结构,为每一个输入特征设置了一个可学习的特征融合因子,进行特征加权融合,获取更多上下层高效的语义特征信息,进而提高目标检测算法的准确率. YOLOv5-CBF 中使用了两个 AWFF 模块, AWFF 网络结构如图 5 所示.

在 AWFF 中先用 1×1 的 Conv 操作将 C3 特征图的通道数压缩,再将输出的特征图 F_1 与 Neck 部分输出的特征图进行拼接.再通过 Conv 操作进行加权融合,融合公式为:

$$F_i = \text{Conv}(\text{Concat}(\beta_1 \times M_i, \alpha_1 \times \text{Resize}(\text{Conv}(X_i)))) \quad (5)$$

$$O_i = \text{Concat}(\beta_2 \times N_i, \alpha_2 \times \text{Resize}(\text{Conv}(F_i))) \quad (6)$$

其中, α_i, β_i 是可学习的特征融合因子,反映了对应特征的重要程度,特征融合因子值越大,在特征融合中参与

的比重越大,且 $\alpha_i + \beta_i = 1$,最后再加入 CA 模块,帮助模型更加精确的定位和识别感兴趣的目标.

本文借鉴了 BiFPN 的思想,采用加权的方式对输入特征设置了可学习的特征融合权重,改进后的 YOLOv5-CBF 整体结构如图 2 所示.

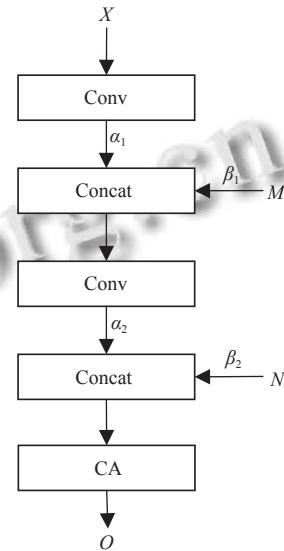


图 5 AWFF 网络模块

3.3 多尺度检测的改进

通常目标检测算法对中大型目标的数据集(例如 COCO 数据集)都有较好的检测效果,而较深的特征图很难学习到血细胞数据集中小目标血小板的特征.体积小的血细胞的目标区域信息可能在多次卷积后被丢失,从而导致检测精度降低.为了改进此问题, YOLOv5-CBF 将输入图片的像素调整到 $640 \times 640 \times 3$,分别在 20×20 、 40×40 、 80×80 以及 160×160 四种不同大小的检测尺度上进行预测,改进后的模型如图 2 所示.

YOLOv5-CBF 在增加了一个检测尺度之后,在 Neck 部分输出为 $80 \times 80 \times 128$ 的特征图的 C3 层后增加 Conv 层、二倍上采样层和 C3 层.再将二倍上采样层特征层与主干网络 C3 层的特征进行融合,通过 Conv 层进行下采样,进行加权特征融合,并生成 160×160 的检测尺度.

改进后的模型通过实验检测了 3 种目标血细胞.与传统的 YOLOv5 算法相比,改进算法 YOLOv5-CBF 更充分地提取了浅层的特征信息,提高了检测的精度.每个网格中会预测 3 个不同尺寸的边界框,每个边界框会产生一个 $4+1+C$ 维的向量,分别为边框坐标、类别置信度、预测类别的置信度(BCCD 中为类

别) 3) 计算方法如下:

$$Confidence = Pr(object) \times IoU_{pred}^{truth} \quad (7)$$

4 实验分析

本实验的处理器为 Intel Core i7 9750H, GPU 为 NVIDIA GeForce RTX 2060, 深度学习框架为 PyTorch, 开发环境为 PyTorch 1.8.1, Python 3.8, Cuda 11.1, 优化器采用 Adam 优化器.

4.1 数据集处理

为了评估本文所提出的 YOLOv5-CBF 网络, 本实验使用的训练集和测试集均采用公开数据集 BCCD, 该数据集包括 364 张图像, 比例为 7:2:1, 即 70% 的数据用于训练, 20% 的数据用于调整模型的超参数, 10% 的数据用于测试模型. BCCD 数据集包含少量的白细胞 (WBC)、血小板 (platelets) 以及大量的红细胞 (RBC), 其中, 检测重叠、粘连的红细胞和小目标血小板是检测的难点.

模型训练时, 需要针对血细胞数据集样本不足的情况进行数据增强, 通过在图像的 0 到 15% 之间随机裁剪、随机亮度调整、翻转等操作后对 4 张图片进行拼接, 丰富了血细胞数据集, 进而间接增大了每批数据量 (batch_size) 的大小.

所有输入血细胞图像尺寸统一调整为 $640 \times 640 \times 3$, batch size 设置为 16, 初始学习率设置为 0.001, 动量项取 0.9, epoch 设置为 100, 进行训练, 得到最佳的模型权重. 并与 Faster-R-CNN^[21]、YOLOv4、YOLOv3 和 YOLOv5 网络进行对比实验.

4.2 评价指标

本实验对于 BCCD 检测采用查准率 (Precision)、查全率 (Recall)、检测精度 (AP)、类平均精度 (mAP) 作为主要的评价指标. 改进后 YOLOv5-CBF 算法的 3 类血细胞的 P-R 曲线如图 6 所示, 通过计算 P-R 曲线下方的面积即可获得该类血细胞的 mAP 值. 红细胞、白细胞和血小板 3 种类别的 mAP 值依次为 88.8%、98.6%、89.5%.

查准率是衡量在血细胞检测任务中是所有预测为正样本的结果中, 预测某细胞正确的个数 (TP) 占正确总个数 (TP+FP) 的比值.

查全率指的是表示血细胞的真实类别为正, 最后预测得到的结果也为正的个数 (TP) 占所有应该被检

测到的细胞 (TP+FN) 的比例.

AP 用来计算单类别模型平均准确度, N 代表类别总数, 血细胞检测任务中的类别总数为 3, 各指标具体计算公式如下:

$$Precision = \frac{TP}{TP + FP} \quad (8)$$

$$Recall = \frac{TP}{TP + FN} \quad (9)$$

$$AP = \int_0^1 PRdR \quad (10)$$

$$mAP = \frac{\sum_{i=0}^n AP_i}{n} \quad (11)$$

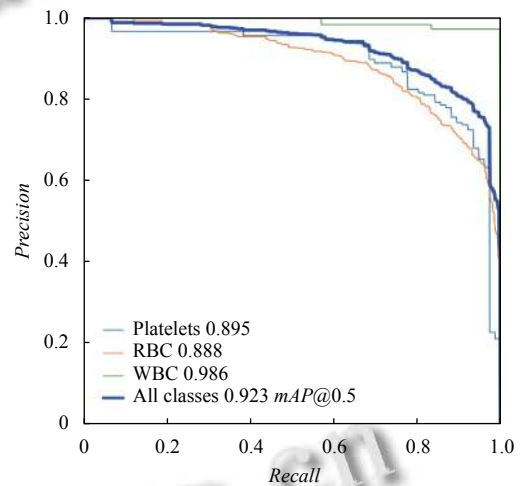


图 6 YOLOv5-CBF 的 P-R 曲线

4.3 实验结果分析

本文将 YOLOv5-CBF 模型与现有的主流的单阶段、两阶段目标检测模型通过同一组血细胞检测图片进行对比分析. 对比结果表 2 所示.

表 2 中, FPS (frame per second) 为检测速度, 即算法每秒可以检测到的图像数量. 分析表 2 中的数据可以看出, YOLOv5-CBF 相比大部分目标检测网络不仅复杂度较小, 而且精度上也有着明显优势. 相较于 Faster-R-CNN^[21]、YOLOv4、YOLOv3、TE-YOLOF-B3、ISE-YOLO 和 YOLOv5-6.0 网络模型在 mAP@0.5 指标上分别高出 15.8 个百分点, 7.6 个百分点, 8.4 个百分点, 0.4 个百分点、6.6 个百分点和 2.7 个百分点, 对比分析实验表明该算法在 mAP@0.5 指标上是所有对比模型中最优的. 在检测速率上, YOLOv5-CBF 相较于同是用于在公开血细胞数据集 BCCD 上进行检测的 TE-

YOLOF-B3 和 ISE-YOLO 算法, FPS 分别增加了 8.9 和 17.4, YOLO-CBF 在相同的血细胞数据集上的检测速度和精度相较于前两个网络模型都有所提升. 本算法速度不及原始的 YOLOv5-6.0 网络, 检测单张图片的时间增加了 0.03 s, FPS 为 51.9, 可满足在硬件终端上检测任务的实时性需求, 并有效改善原始的网络误检、漏检的问题. 实验表明本算法在保证检测速度的基础上, 提升了 mAP , 在所有对比模型中综合性能最好.

表2 对比分析实验

Model	Input size	Backbone	Class	AP (%)	$mAP@0.5$ (%)	FPS
Faster-R-CNN	1000×600	ResNet	RBC	0.722	0.765	9.2
			WBC	0.803		
			Platelets	0.770		
YOLOv3	608×608	Darknet-53	RBC	0.829	0.839	34.5
			WBC	0.914		
			Platelets	0.774		
YOLOv4	640×640	CSPDarkNet53	RBC	0.798	0.847	36.1
			WBC	0.930		
			Platelets	0.813		
TE-YOLOF-B3 ^[22]	416×416	EfficientNet-B3	RBC	0.873	0.919	43
			WBC	0.987		
			Platelets	0.898		
ISE-YOLO ^[23]	416×416	Darknet-53+SE	RBC	0.927	0.857	34.5
			WBC	0.965		
			Platelets	0.896		
YOLOv5-6.0	640×640	CSPDarkNet53	RBC	0.838	0.896	56.7
			WBC	0.977		
			Platelets	0.873		
YOLOv5-CBF	640×640	CSPDarkNet53+CA	RBC	0.888	0.923	51.9
			WBC	0.986		
			Platelets	0.895		

使用原始的 YOLOv5-6.0、Faster-R-CNN、YOLOv3、YOLOv4 与改进 YOLOv5-CBF 目标检测算法从 BCCD 数据集中随机选取 3 张图片进行检测, 检测效果如图 7 所示.

从图 7(b) 的血细胞检测效果可以看出 Faster-R-CNN 算法在大目标 (WBC) 中的检测效果较好, 预测的置信度普遍较高, 但图 7 (b2) 中血小板存在一个漏检, 左侧重叠与粘连的红细胞部分检测效果很差, 且 Faster R-CNN 标注框的大小不够精准. 从图 7(c) 的检测效果可以看出, 相较于 Faster-R-CNN 算法, YOLOv3 算法

对于红细胞粘连部分的检测效果有所改善, 但仍存在少量的误检与漏检问题, 如图 7(c1) 右下角的小目标血小板检测中出现了一个漏检. YOLOv4 算法与 YOLOv3 检测效果差距不大, 对于红细胞粘连部分的检测也存在少量的错检与漏检问题. 如图 7(e) 所示, YOLOv5-6.0 对粘连红细胞的错检与漏检问题明显改善, 但图 7(e2) 右侧对小目标血小板的检测仍存在一个漏检. 如图 7(f) 所示, 采用 YOLOv5-CBF 算法检测左侧重叠的红细胞虽然也存在部分漏检问题, 但相较于其他算法有很大改善. 如图 7(f2) 所示, YOLOv5-CBF 准确检测出 4 种对比算法均未能检测出的小目标血小板.

4.4 消融实验

为了验证各个模块的有效性, 分析各个模块对 YOLOv5-CBF 算法精度影响, 本文进行了消融实验, 以 YOLOv5-6.0 为基础网络逐步增加改进点, 实验结果如表 3 所示. 根据表 3 消融实验的数据显示, 在主干网络中加入 CA 模块的网络相较于原网络的 mAP 值提升了 1.4%, 说明注意力机制网络能抑制无用特征, 有效提高血细胞检测的性能. 在三尺度检测的基础上增加了一个小目标检测层, 相较于原网络的 mAP 值提升了 2.1%. 嵌入 AWF 模块后, mAP 提高至 92.1%. YOLOv5-CBF 相较于原 YOLOv5-6.0 网络在 3 类血细胞检测的平均精度提升 2.7%, 提升了检测准确率, 并保持了较好的检测速率, 具有很高的实用性.

5 结论与展望

为解决血细胞尺度变化大, 红细胞粘连以及数据样本少导致的检测精度低等问题, 本文提出了一个实时高效的 YOLOv5-CBF 算法. 针对数据集目标尺度变化大的问题, 本文在主干网络中添加 CA 模块, 增强模型对目标的精确定位能力, 从而提高检测精度. 为解决颈部网络的特征融合不充分, 导致多特征表达能力下降的问题, 提出了结合了 BiFPN 思想的特征融合结构, 使目标特征更有效的表达, 在一定程度上减少了红细胞在堆叠情况下由于特征表达能力不足而造成的漏检和误检, 提高了红细胞检测的准确性, 进而提高血细胞的检测精度. 在三尺度检测的基础上增加了一个小目标检测层, 提高血小板的识别精度. 实验结果证明 YOLOv5-CBF 算法具有优秀的检测精度和实时检测效果, 能很好地应用于血细胞检测这样的密集小目标检测任务中. 当利用本文算法对严重堆叠的红细胞进行

目标检测时, 会存在不同程度的漏检和误检, 下一步工作将针对严重堆叠问题展开进一步研究, 提升检测精

度, 进一步优化模型结构与损失函数, 加快模型的收敛速度.

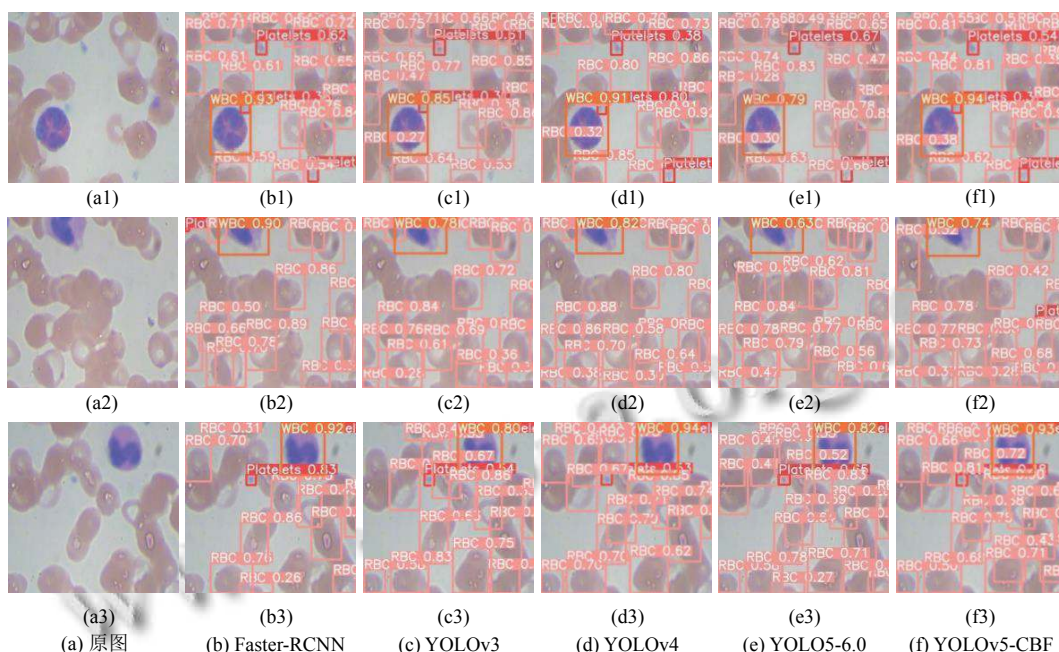


图7 可视化对比图

表3 消融实验

Model	Precision	Recall	mAP@0.5	mAP@0.5:0.95
YOLOv5-6.0	0.857	0.863	0.896	0.581
+CA	0.832	0.922	0.91	0.622
+BiFPN	0.815	0.91	0.904	0.59
+检测尺度	0.851	0.909	0.917	0.625
+BiFPN-AWFF	0.884	0.902	0.921	0.631
YOLOv5-CBF	0.886	0.904	0.923	0.637

参考文献

- Atkins CG, Buckley K, Blades MW, *et al.* Raman spectroscopy of blood and blood components. *Applied Spectroscopy*, 2017, 71(5): 767–793. [doi: 10.1177/0003702816686593]
- Garraud O, Tissot JD. Blood and blood components: From similarities to differences. *Frontiers in Medicine*, 2018, 5: 84. [doi: 10.3389/fmed.2018.00084]
- Biswas S, Ghoshal D. Blood cell detection using thresholding estimation based watershed transformation with Sobel filter in frequency domain. *Procedia Computer Science*, 2016, 89: 651–657. [doi: 10.1016/j.procs.2016.06.029]
- Tiwari P, Qian J, Li QC, *et al.* Detection of subtype blood cells using deep learning. *Cognitive Systems Research*, 2018, 52: 1036–1044. [doi: 10.1016/j.cogsys.2018.08.022]

- Shahin AI, Guo YH, Amin KM, *et al.* White blood cells identification system based on convolutional deep neural learning networks. *Computer Methods and Programs in Biomedicine*, 2019, 168: 69–80. [doi: 10.1016/j.cmpb.2017.11.015]
- Liang GB, Hong HC, Xie WF, *et al.* Combining convolutional neural network with recursive neural network for blood cell image classification. *IEEE Access*, 2018, 6: 36188–36197. [doi: 10.1109/ACCESS.2018.2846685]
- Zhang DD, Zhang PF, Wang LS. Cell counting algorithm based on YOLOv3 and image density estimation. *Proceedings of the 2019 IEEE 4th International Conference on Signal and Image Processing*. Wuxi: IEEE, 2019. 920–924.
- Girshick R, Donahue J, Darrell T, *et al.* Rich feature hierarchies for accurate object detection and semantic segmentation. *Proceedings of the 2014 IEEE Conference on Computer Vision and Pattern Recognition*. Columbus: IEEE, 2014. 580–587.
- Redmon J, Divvala S, Girshick R, *et al.* You only look once: Unified, real-time object detection. *Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition*. Las Vegas: IEEE, 2016. 779–788.
- Redmon J, Farhadi A. YOLO9000: Better, faster, stronger.

- Proceedings of the 2017 IEEE Conference on Computer Vision and Pattern Recognition. Honolulu: IEEE, 2017. 6517–6525.
- 11 Redmon J, Farhadi A. YOLOv3: An incremental improvement. arXiv:1804.02767, 2018.
 - 12 Bochkovskiy A, Wang CY, Liao HYM. YOLOv4: Optimal speed and accuracy of object detection. arXiv:2004.10934, 2020.
 - 13 Liu W, Anguelov D, Erhan D, *et al.* SSD: Single shot multibox detector. Proceedings of the 14th European Conference on Computer Vision. Amsterdam: Springer, 2016. 21–37.
 - 14 Hou QB, Zhou DQ, Feng JS. Coordinate attention for efficient mobile network design. Proceedings of the 2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Nashville: IEEE, 2021. 13713–13722.
 - 15 Lin TY, Dollár P, Girshick R, *et al.* Feature pyramid networks for object detection. Proceedings of the 2017 IEEE Conference on Computer Vision and Pattern Recognition. Honolulu: IEEE, 2017. 2117–2125.
 - 16 Wang WH, Xie EZ, Song XG, *et al.* Efficient and accurate arbitrary-shaped text detection with pixel aggregation network. Proceedings of the 2019 IEEE/CVF International Conference on Computer Vision. Seoul: IEEE, 2019. 8439–8448.
 - 17 Tan MX, Pang RM, Le QV. EfficientDet: Scalable and efficient object detection. Proceedings of the 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Seattle: IEEE, 2020. 10778–10787.
 - 18 Hu J, Shen L, Albanie S, *et al.* Squeeze-and-excitation networks. Proceedings of the 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Salt Lake City: IEEE, 2018. 7132–7141.
 - 19 Jaderberg M, Simonyan K, Zisserman A. Spatial transformer networks. Proceedings of the 28th International Conference on Neural Information Processing Systems. Montreal: MIT Press, 2015. 2017–2025.
 - 20 Woo S, Park J, Lee JY, *et al.* CBAM: Convolutional block attention module. Proceedings of the 15th European Conference on Computer Vision. Munich: Springer, 2018. 3–19.
 - 21 Ren SQ, He KM, Girshick RB, *et al.* Faster R-CNN: Towards real-time object detection with region proposal networks. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2017, 39(6): 1137–1149. [doi: [10.1109/TPAMI.2016.2577031](https://doi.org/10.1109/TPAMI.2016.2577031)]
 - 22 Xu FX, Li XK, Yang H, *et al.* TE-YOLOF: Tiny and efficient YOLOF for blood cell detection. Biomedical Signal Processing and Control, 2022, 73: 103416. [doi: [10.1016/j.bspc.2021.103416](https://doi.org/10.1016/j.bspc.2021.103416)]
 - 23 Liu C, Li DW, Huang P. ISE-YOLO: Improved squeeze-and-excitation attention module based YOLO for blood cells detection. Proceedings of the 2021 IEEE International Conference on Big Data (Big Data). Orlando: IEEE, 2021. 3911–3916.

(校对责编: 孙君艳)