

DOI:10.19651/j.cnki.emt.2517963

基于 YOLO-BioFusion 的血细胞检测模型^{*}

张 傲 刘 微 刘 阳 杨思瑶 管 勇

(沈阳理工大学信息科学与工程学院 沈阳 110159)

摘要: 血细胞检测是临床诊断中的重要任务,尤其在面对细胞类型多样、尺寸差异显著、目标重叠频繁以及复杂背景时,现有检测模型的精度和鲁棒性仍面临挑战。为解决这些问题,本文提出了一种改进的YOLOv8目标检测模型——YOLO-BioFusion。该模型通过引入ACFN模块,提高了对细小目标和重叠目标的检测能力;应用C2f-DPE和SPPF-LSK模块增强了多尺度特征的融合与提取,提升了模型的鲁棒性和泛化能力;同时,采用Inner-CIoU损失函数加速了模型收敛并提高了定位精度。实验结果表明,在BCCD数据集上,YOLO-BioFusion的mAP@0.5为94.0%,mAP@0.5:0.95为65.2%,分别较YOLOv8-n提高了1.9%和3.2%。与此同时,计算成本仅为6.8 GFLOPs,展示了其在资源受限环境中的应用潜力。该研究为复杂背景下的血细胞检测提供了一种高效且精确的解决方案。

关键词: 血细胞检测;多尺度特征融合;损失函数优化;YOLOv8;重叠目标

中图分类号: TN911.7 文献标识码: A 国家标准学科分类代码: 120.6010

Blood cell detection model based on YOLO-BioFusion

Zhang Ao Liu Wei Liu Yang Yang Siyao Guan Yong

(College of Information Science and Engineering, Shenyang Ligong University, Shenyang 110159, China)

Abstract: Blood cell detection is a crucial task in clinical diagnosis. However, due to the diverse cell types, significant size variations, frequent target overlap, and complex backgrounds, existing detection models still face challenges in terms of accuracy and robustness. To address these issues, this paper proposes an improved YOLOv8-based object detection model, YOLO-BioFusion. The model incorporates the ACFN module to enhance the detection of small and overlapping targets. Additionally, the C2f-DPE and SPPF-LSK modules are introduced to strengthen multi-scale feature fusion and extraction, improving the model's robustness and generalization ability. Meanwhile, the adoption of the Inner-CIoU loss function accelerates model convergence and enhances localization accuracy. Experimental results on the BCCD dataset demonstrate that YOLO-BioFusion achieves an mAP@0.5 of 94.0% and an mAP@0.5:0.95 of 65.2%, outperforming YOLOv8-n by 1.9% and 3.2%, respectively. Moreover, with a computational cost of only 6.8 GFLOPs, YOLO-BioFusion exhibits great potential for applications in resource-constrained environments. This study provides an efficient and accurate solution for blood cell detection in complex backgrounds.

Keywords: blood cell detection; multi-scale feature fusion; optimization of loss function; YOLOv8; overlapping target

0 引言

血液显微镜图像分析是临床检验中常见且重要的技术,通过对血液中主要成分——白细胞(white blood cell,WBC)、红细胞(red blood cell,RBC)和血小板(platelets)数量的分析,帮助医生进行疾病诊断。WBC在抵抗感染方面发挥着关键作用,RBC负责将氧气输送到全身,而 platelets 则与凝血功能密切相关。血细胞数量的变化可以反映多种健康状况。例如,RBC过多可能引起血管阻塞,而数量过

少则可能导致贫血。因此,准确检测血细胞的数量对于疾病的早期诊断至关重要。

传统的血细胞检测方法主要依赖手工设计的特征提取技术。文献[1]采用Otsu阈值法对HSV色彩空间中的图像进行二值化并标记WBC,但该方法未考虑血细胞的结构特征,因此在精度上存在局限性。文献[2]利用分水岭变换分离RBC,并结合K-medoids算法和决策树分类器进行细胞分类。然而,这些方法仍难以处理复杂的细胞特征。文献[3]通过Sobel滤波器检测血细胞边缘,尽管该方法在边

缘检测方面有效,但未能有效区分不同类型的血细胞。这些早期方法结合了先进的图像处理技术(如特征工程)和经典的机器学习模型,显著提升了血细胞检测的精度,但仍面临结构特征提取不足及细胞类型分类能力有限等问题。

随着深度学习技术的快速发展,基于深度学习的目标检测算法已广泛应用于血细胞检测任务。根据检测方式的不同,这些算法大致可分为单阶段检测算法和双阶段检测算法。单阶段算法(如 YOLO^[4] 系列和 SSD^[5] 系列)直接进行分类与回归,具有较高的检测效率,非常适合实时检测任务。相比之下,双阶段算法(如 Faster R-CNN^[6] 和 R-FCN^[7])通过区域提议网络(region proposal network, RPN)生成候选区域,再通过检测网络进行分类与定位,尽管精度较高,但计算复杂度较大。此外,近年来无锚点目标检测算法(如 FCOS^[8] 和 RepPoints^[9])逐渐兴起。这类算法无需预定义锚框,简化了检测流程,然而其计算复杂度较高,且在实时应用中的普及度相对较低。为了进一步优化目标检测算法,研究者们提出了许多创新的方法。例如,文献[10]提出的基于 Transformer 的端到端目标检测算法 DETR,通过消除后处理步骤(如非极大值抑制或锚框),大大简化了检测流程。

针对血细胞检测任务,研究者们提出了多种基于深度学习的优化方法,以提升检测性能和效率。文献[11]引入全局注意力机制和 HorNet-BiFPN 网络结构,增强了对小目标和重叠区域的检测能力,并提升了特征融合效果。文献[12]提出基于 YOLOv5 的改进方法(YOLOv5-ALT),通过优化特征提取模块并引入注意力机制,进一步提升了模型的鲁棒性和检测能力。文献[13]结合卷积块注意力模块和 Transformer 编码器,改进了 YOLOv5(AYOLOv5),增强了模型对关键特征区域的关注能力。此外,文献[14]在 YOLOv5 的不同版本基础上进一步优化了特征提取和检测机制,提升了模型的综合性能。通过这些方法,研究者在血细胞检测任务中提出了高效且可靠的解决方案,充分展示了深度学习在医学图像分析中的应用潜力。

然而,现有的 YOLO 模型在复杂背景下处理多尺度目标检测时仍面临诸多挑战,例如目标重叠、尺寸差异及形态多样性等问题,这些因素容易导致检测精度下降。为了解决这些问题,本文提出了一种改进的 YOLOv8 架构的血细胞检测方法——YOLO-BioFusion。

具体来说,YOLO-BioFusion 引入了多个创新模块,以提升模型在复杂环境中的性能:

1)自适应卷积融合网络(adaptive convolution fusion network, ACFN)模块:该模块通过动态调整卷积操作,使网络能够自适应优化特征提取过程,有效捕捉细小目标和重叠目标的细节信息。

2)C2f-DPE(C2f-dilated path expansion)模块:利用不同扩张速率的深度卷积路径扩展感受野,提升模型对多尺度目标的适应性,并增强对不规则排列目标的检测能力。

3)SPPF-LSK(spatial pyramid pooling fast-large selective kernel)模块:结合 SPPF 结构与大选择性核(large selective kernel, LSK)注意力机制,实现多尺度特征的融合与关键特征的聚焦,增强模型在复杂背景下的鲁棒性,并提高全局信息的捕捉能力。

4)Inner-ClIoU(inner complete IoU)损失函数:通过自适应调整辅助框的大小,提高预测框的定位精度,加速模型的收敛速度。

在实验设计方面,本文将使用 BCCD 数据集和 TXL-PBC 数据集对 YOLO-BioFusion 进行训练与评估,并通过与现有主流模型的对比,验证 YOLO-BioFusion 在复杂背景、目标重叠以及多尺度目标检测中的优势。本研究旨在为血液疾病的早期诊断提供更高效、更可靠的技术支持,同时也为医学影像分析领域提供具有重要理论意义和应用价值的创新成果。

1 YOLO-BioFusion 模型

与血细胞检测面临细胞类型多样性、尺寸差异、目标重叠及复杂背景等挑战,这些因素显著增加了检测难度和精度要求。此外,临床应用还需平衡高精度和计算效率,以支持实时检测需求。

为应对这些挑战,本研究在 YOLOv8 模型基础上提出了轻量化的血细胞检测模型——YOLO-BioFusion。该模型通过多项改进提升了多尺度目标检测能力,增强了细小目标和重叠目标的识别效果,同时保持低计算复杂度,提高了鲁棒性与泛化能力,满足了高效检测的需求。模型结构如图 1 所示。

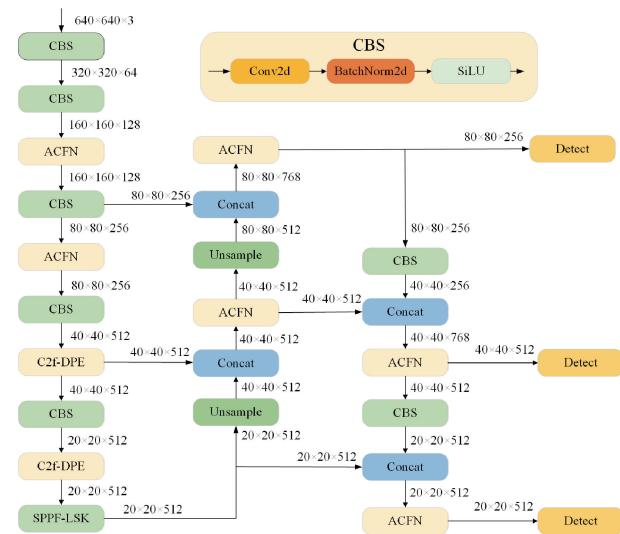


图 1 YOLO-BioFusion 模型结构

Fig. 1 YOLO-BioFusion model architecture

1.1 ACFN 模块

在血细胞检测中,小尺寸血小板和重叠血细胞的特征容易在卷积网络的深层处理中丢失,导致检测精度下降。

这是因为传统卷积在处理细微和复杂背景时难以捕捉到足够的细节。为了应对这些问题,ACFN模块通过自适应卷积层调整特征提取过程,使网络能够根据输入特征的尺度自适应调整卷积操作,从而提升对细小和复杂目标的检测能力。这种自适应性使ACFN模块在处理具有高度变化的图像细节时表现突出,不仅增强了模型的特征表达能力,还有效降低了细小目标被忽略的概率。

如图2所示,ACFN模块的核心在于并行自适应卷积层,能够根据不同尺度的输入特征动态调整卷积操作,从而增强对细小目标的检测能力。该模块首先通过 3×3 卷积层对输入进行初步处理,然后将特征图送入并行的自适应卷积层,分为两个部分进行处理。一部分采用深度卷积(depthwise convolution, DWConv)和逐点卷积(pointwise convolution, PWConv)的组合,独立捕捉每个通道的局部特征,并通过融合这些特征来提升表示的完整性。另一部分则结合了条状卷积和空洞卷积,旨在提高感受野并增强对细节特征的捕捉能力。其中,Conv-sh是在水平方向上进行卷积操作,采用 1×7 的卷积核来捕捉特征图在水平方向上的空间信息,并通过设置dilation=3(空洞卷积)来扩大感受野,从而能够提取更远距离的上下文信息。Conv-sv则是在垂直方向上进行卷积,使用 7×1 的卷积核捕捉垂直方向上的空间特征,并同样结合dilation=6来增强垂直方向的特征捕捉能力。通过这种水平方向和垂直方向分离的卷积操作,模型不仅能够保持较大的感受野,还能在不显著增加参数量的情况下提高计算效率。

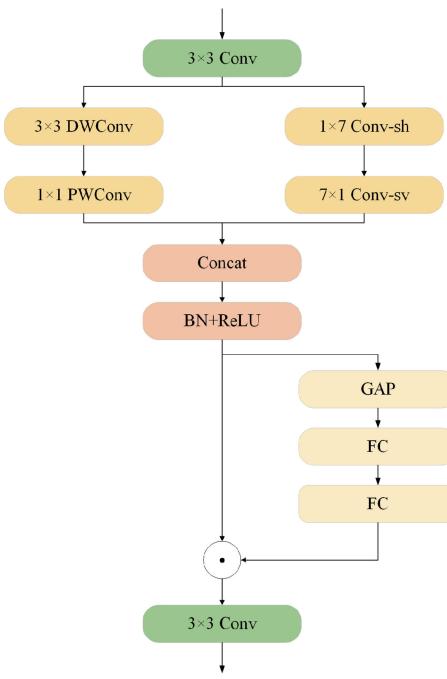


图2 自适应卷积融合网络

Fig. 2 Adaptive convolutional fusion network

一个权重向量。权重向量用于指导联合特征融合,将得到的权重和输入按元素相乘,用于逐通道微调联合特征,强调有用元素,压缩无用元素。该操作的公式为:

$$z_c = \frac{1}{H \times W} \sum_{i=1}^H \sum_{j=1}^W x_c(i, j) \quad (1)$$

其中, z_c 是第 c 个通道的全局池化结果, H 和 W 别是特征图的高度和宽度, $x_c(i, j)$ 表示输入特征图中第 c 通道在位置 (i, j) 的值。

最终,经过一个 3×3 的卷积层对融合后的特征进行处理,生成最终输出特征图。这种设计确保了多路径特征的有效融合,并在复杂背景和高密度重叠目标的检测中表现优异。

1.2 C2f-DPE模块

YOLOv8主干网络在浅层特征与高语义深层特征的融合方面存在不足,导致其对小目标血细胞的检测精度较低。为了解决这一问题,本文提出了深度感知扩展(dilated path expansion, DPE)模块模块,以增强深层特征的提取能力。

如图3所示,在DPE模块中,首先通过一个 3×3 的卷积操作提取初始的基本特征,输出为2C维度的特征图。接下来,特征图通过BN(batch normalization)层和SiLU激活函数进行处理,之后被分成3个并行的分支。每条分支分别使用扩张速率为1、3和5的 3×3 深度卷积,以提取不同层次的语义信息。3个分支的输出特征图再经过BN和SiLU激活函数处理后,汇总成一个4C的特征图。然后,特征图通过 3×3 DWConv和 1×1 PWConv,进一步提取特征。最终,输出特征图与初始输入特征图通过逐元素相加的残差连接操作,生成最终的输出特征图。

本文使用DPE模块替代了原有C2f模块中的Bottleneck结构,从而实现了残差信息的有效融合。改进后的C2f-DPE结构如图4所示。该模块通过采用不同速率的深度卷积,扩大了模型的感受野,使其能够自适应地学习血细胞图像中不规则排列和布局的目标,从而显著提升了模型的泛化能力。

1.3 SPPF-LSK模块

相较于经典的SPP模块,SPPF通过简化操作降低了计算复杂度,同时保持了输出尺寸的自适应性,大幅提升了处理效率。然而,本文发现,SPPF在捕捉小而精细的特征信息方面仍存在不足。尽管池化操作减少了参数量,但这一结构在处理细粒度信息时,容易出现信息丢失的情况,这与本文希望精确捕捉血细胞检测样本复杂性和多样性的目标不完全契合。

为了解决这一问题,本文在SPPF的基础上,引入了LSK注意力机制。LSK注意力机制最初设计用于遥感目标检测任务,能够根据任务需求动态调整空间感受野,捕捉更丰富且动态的上下文信息。LSK注意力机制结构如图5所示。

然后通过全局池化层和两个全连接层抽取特征,得到

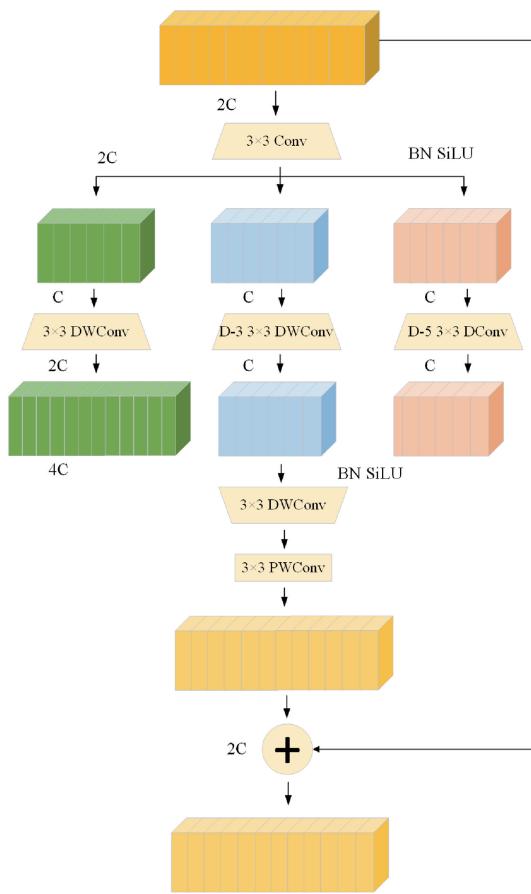


图 3 DPE 模块

Fig. 3 Depth sense expansion module

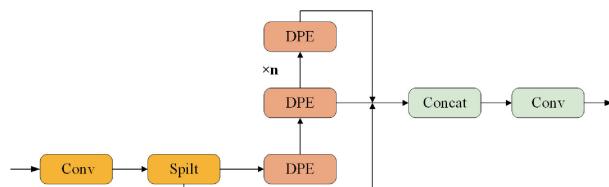


图 4 C2f-DPE 模块

Fig. 4 C2f-DPE module

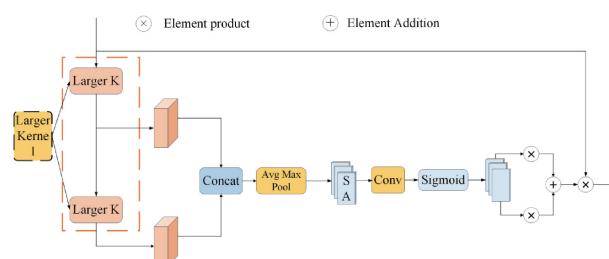


图 5 LSK 注意力机制

Fig. 5 LSK attention mechanism

LSK 注意力机制首先将一个大卷积核分解为两个深度卷积, 获取更多的上下文信息, 同时可以减少参数更加灵活, 并在每个深度卷积后加上一个 1×1 的卷积进行通道融合。

合, 即:

$$U_0 = X, U_{i+1} = F_i^{dw}(U_i) \quad (2)$$

$$U_i = F_i^{1 \times 1}(U_i), i \in [1, N] \quad (3)$$

其中, F_i^{dw} 代表深度卷积的卷积核大小; $F_i^{1 \times 1}$ 代表 1×1 卷积。

将具有不同感受野的特征进行拼接操作, 即:

$$U = \text{Concat}(U, \dots, U_i) \quad (4)$$

然后在的通道方向上应用全局平均池化和全局最大池化压缩全局通道信息:

$$\text{SA}_{\text{avg}} = \text{AvgPool}(U) \quad (5)$$

$$\text{SA}_{\text{max}} = \text{MaxPool}(U) \quad (6)$$

为了获取更丰富的特征信息, 将经过平均池化和最大池化的特征图通过卷积操作 $F^{2 \rightarrow N}$ 进行拼接, 并通过 sigmoid 激活函数得到空间注意力特征图。将大卷积核的特征与空间注意力特征图进行加权处理, 得到注意力特征 S , 公式如下:

$$S = F^{2 \rightarrow N}([\text{SA}_{\text{avg}}, \text{SA}_{\text{max}}]) \quad (7)$$

$$S_i = \text{sigmoid}(S_i) \quad (8)$$

$$S = F\left(\sum_{i=1}^N (S_i \cdot U_i)\right) \quad (9)$$

LSK 注意力最终的输出为输入特征 X 和注意力特征 S 的逐元素乘积, 公式如下:

$$Y = X \cdot S \quad (10)$$

如图 6 所示, SPPF-LSK 结构结合了 SPPF 和 LSK 注意力机制。通过这一结合, SPPF-LSK 模块不仅继承了 SPPF 在效率上的优势, 还通过 LSK 注意力机制增强了多尺度特征的捕捉能力, 尤其在细粒度信息的提取方面表现突出。该模块在血细胞检测任务中, 能够在保持全局感知的同时, 聚焦于关键区域, 从而显著提高了对重要特征的识别能力。

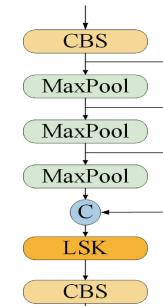


图 6 SPPF-LSK 模块

Fig. 6 SPPF-LSK module

1.4 Inner-CIOU 损失函数

在传统的目标检测任务中, 常见的做法是通过增加新的损失项来加速模型的收敛。然而, IOU(intersection over union)损失在自适应优化方面存在局限性, 难以针对不同的检测器和任务进行有效调整, 限制了模型的泛化能力。为了解决这一问题, 本文引用了 Inner-CIOU 损失函数, 该

方法基于CIOU(complete IOU)损失^[15],并引入了一个ratio尺度因子,用于动态调整辅助框的大小,从而加速模型收敛并提高预测精度。

Inner-CIOU根据检测目标的IOU和尺寸差异,生成不同大小的辅助框来计算损失。对于高IOU的大目标,生成的辅助框会比真实框略小,利用较小的辅助框来计算损失可以增强梯度,进而加快高IOU样本的回归速度。对于低IOU的小目标,生成的辅助框会大于真实框,这样可以扩大回归的有效范围,提高模型对低IOU样本的敏感性,增强其回归效果。最终,将Inner-IOU融入CIOU边框回归损失函数的公式如下定义:

$$L_{\text{inner-cio}} = L_{\text{ciov}} + IOU_{\text{inner}} \quad (11)$$

$$IOU_{\text{inner}} = \frac{\text{inter}}{\text{union}} \quad (12)$$

$$\text{union} = (w^g \times h^g + (\text{ratio})^2) + (w \times h \times (\text{ratio})^2) - \text{inter} \quad (13)$$

$$\text{inter} = (\min(b_x^g, b_x) - \max(b_x^g, b_x)) \times (\min(b_y^g, b_y) - \max(b_y^g, b_y)) \quad (14)$$

$$b_x = \frac{x_b^g + w_b^g}{2} - \frac{x_b + w_b}{2} + \text{ratio} \cdot b_w \quad (15)$$

$$b_y = \frac{y_b^g + h_b^g}{2} - \frac{y_b + h_b}{2} + \text{ratio} \cdot b_h \quad (16)$$

在Inner-CIOU损失函数中,union和inter分别表示辅助预测框与辅助真实框的并集和交集。针对血细胞检测中小目标密集的场景,使用较大尺寸的辅助框可以有效增强目标的回归效果。该方法特别关注生成过程中框的重叠和密集度,即使在低IOU的情况下,也能取得更好的回归效果。

2 实验

2.1 数据集

在本研究中,提出的网络模型在扩展后的BCCD数据集上进行了训练。原始BCCD数据集^[16]包含364张血液涂片图像,涵盖3种不同类型的细胞,样本如图7所示。为了提高模型的泛化能力和鲁棒性,采用了数据增强技术,包括随机裁剪、翻转和颜色调整,将数据集扩展至900张分辨率率为640×640的图像。增强后的数据集按7:2:1的比例划分为训练集、验证集和测试集,以优化模型的训练过程和性能评估。

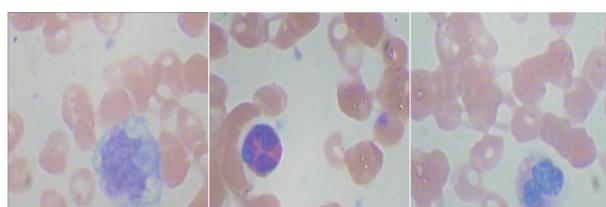


图7 BCCD数据集样本

Fig. 7 BCCD dataset samples

2.2 实验平台与参数设置

本研究的实验在配置了PyTorch 1.11.0和CUDA 11.3的环境中进行,编程语言为Python 3.8,操作系统为Ubuntu 20.04。实验平台的硬件配置包括18核vCPU的AMD EPYC 9754处理器、60GB内存,以及配备24GB显存的NVIDIA GeForce RTX 4090D显卡。

在实验中,本文通过多轮超参数调优,以确保模型在性能上的最佳表现。学习率的设置对模型训练具有重要影响。从初始学习率0.001开始,随着训练的推进逐步降低学习率。这一策略旨在训练初期快速探索参数空间,而在接近收敛时,通过细致调整学习率,避免模型跳过最优解,从而实现更加精确的收敛。最终,经过多次实验验证,选择了0.0001作为学习率的最终设定,以确保模型能够稳定且高效地收敛。

批次大小的选择对训练的稳定性和GPU内存的有效利用起着至关重要的作用。通过多轮实验验证,最终确定批次大小为16为最优设置。这一选择能够保证训练过程的稳定性,并充分发挥GPU的并行计算能力,避免内存不足或资源浪费的问题。

在优化器选择方面,本文对比了多种优化方法,包括Adam、AdamW和随机梯度下降(stochastic gradient descent,SGD)。最终,结合动量(0.9)和权重衰减(0.0005)的SGD优化器表现最佳。动量的引入使得模型在更新参数时保持一定的惯性,进而加快了收敛速度;而权重衰减则通过正则化权重,有效避免了模型陷入局部最优解。此外,本文还采用了数据增强技术,进一步提升了模型的泛化能力。通过这些精心设计的超参数调优策略,确保了实验的高效性和结果的可靠性。

2.3 实验结果分析

1) 评价指标

本研究采用精确率(precision)、召回率(recall)、mAP@0.5、mAP@0.5:0.95和GFLOPs作为评估模型性能的关键指标。精确率用于衡量模型在预测为阳性样本中,实际阳性样本所占的比例,反映了模型的预测准确性;召回率则评估模型对实际阳性样本的检出能力,即能够正确识别的阳性样本占所有阳性样本的比例。平均精度mAP(mean average precision,mAP)作为目标检测任务中常用的综合评估指标,结合了精确率与召回率,其中mAP@0.5和mAP@0.5:0.95分别表示在不同IoU阈值下计算出的平均精度,能够全面衡量模型在不同置信度下的检测表现。GFLOPs则用来衡量模型的计算复杂度,表示每次前向传播所需的浮点运算量,GFLOPs越低,说明模型在计算资源和实时检测任务中的适应性越强。

2) YOLO-BioFusion模型性能评估

YOLO-BioFusion模型在测试集上的评估结果如表1所示。模型在检测和识别3种血细胞类型时表现优异,mAP@0.5达到了94.0%,mAP@0.5:0.95为65.2%,精

度为 87.5%，召回率为 90.9%。从表 1 可以看出，模型在 WBC 和血小板检测中的 mAP@0.5 均超过了 90%，表明其在识别这两类细胞时具有较高的准确性。相比之下，RBC 的数量庞大且常存在重叠现象，检测中偶尔会出现漏检，但其 mAP@0.5 仍达到 89.3%。

表 1 BCCD 数据集的性能结果

Table 1 Performance results for the BCCD dataset

Class	Precision	Recall	mAP@0.5	mAP@0.5:0.9
RBC	78.2	84.6	89.3	63.5
WBC	96.6	100	98.9	80.7
Platelets	87.7	88.2	93.8	51.4
All	87.5	90.9	94.0	65.2

图 8 展示了模型在 mAP@0.5 下的性能表现。可以看到，随着训练迭代次数的增加，mAP@0.5 呈现出稳定上升的趋势，最终达到了 94.0%。这表明在设定的 IoU 阈值为 0.5 时，模型能够有效捕获并精准定位目标区域。图 8 中的曲线也反映出模型在不同训练阶段的收敛情况，显示出良好的学习能力和鲁棒性。图 9 则展示了模型在 mAP@0.5:0.95 下的表现。该指标涵盖了不同 IoU 阈值下的检测性能，从 0.5~0.95 逐步递增。模型最终在 mAP@0.5:0.95 下达到 65.2%，表明它在多种复杂场景下都能保持较高的检测精度。曲线的走势表明，模型不仅在低 IoU 阈值下表现出色，在更高要求的 IoU 标准下依然能够保持较好的性能。这说明模型在面对多样化的目标形态和重叠度要求时，依旧能够实现准确、稳定的检测能力。

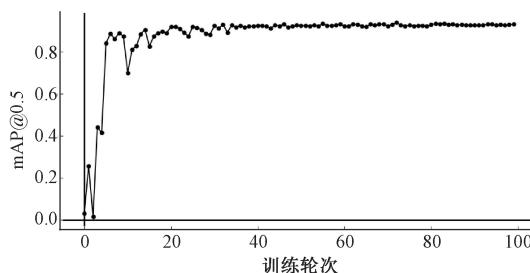


图 8 mAP@0.5 模型性能曲线

Fig. 8 mAP@0.5 model performance curve

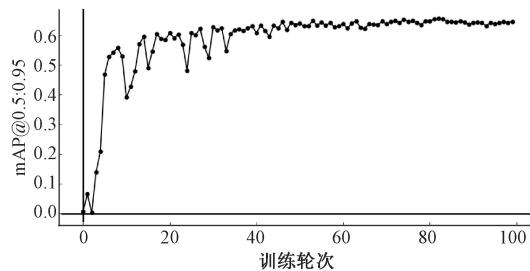


图 9 mAP@0.5:0.95 模型性能曲线

Fig. 9 mAP@0.5:0.95 model performance curve

图 10 展示了模型对 3 种血细胞类型的混淆矩阵，结果表明改进后的网络在测试集中能够很好地区分不同血细胞类型。图 11 则展示了 YOLO-BioFusion 模型在血细胞检测任务中的精度-召回率(precision-recall, PR)曲线，显示该模型在所有类别(红细胞、白细胞、血小板)上的 mAP@0.5 为 94.0%。其中，白细胞的 mAP@0.5 高达 98.9%，PR 曲线几乎贴近右上角，说明白细胞的检测极为准确，漏检和误检极少。血小板的 PR 曲线稍低于白细胞，但仍表现出很高的精度，尤其在重叠的小目标检测上表现稳定。红细胞的检测精度相对略低，这可能是由于其数量多且存在重叠，影响了检测表现。总体来看，各条 PR 曲线在高召回率下仍保持了较高的精度，说明该模型在全面检测时，能够在保持高精度的同时提高召回率。

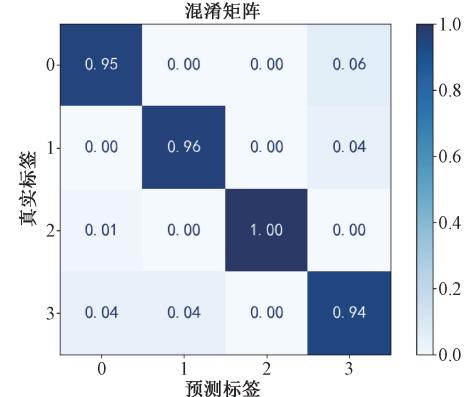


图 10 血细胞检测混淆矩阵

Fig. 10 Confusion matrix for blood cell detection

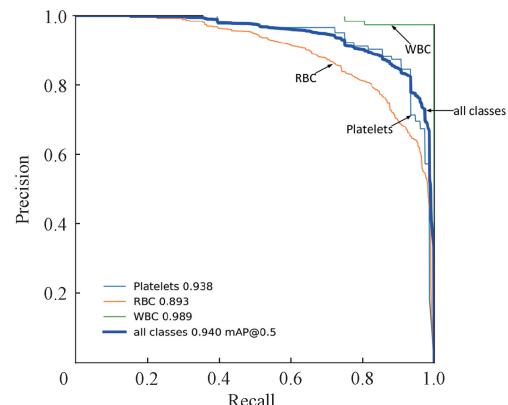


图 11 血细胞检测 PR 曲线

Fig. 11 PR curve of blood cells

图 12 展示了 YOLOv8-n 与 YOLO-BioFusion 在同一复杂背景下的检测结果对比。可以观察到，YOLOv8-n(左侧)在检测某些目标时出现了漏检，特别是在血细胞重叠较为严重的区域(箭头所指)，部分目标未被正确识别。相比之下，YOLO-BioFusion(右侧)在处理相同图像时表现出了更为精准的检测结果。该模型能够更有效地识别复杂背景中的重叠目标，漏检和误检现象明显减少，并且能够成功捕

捉到更多细节,尤其是在血细胞密集的区域。这表明,YOLO-BioFusion在处理复杂背景时表现出了显著的优势。

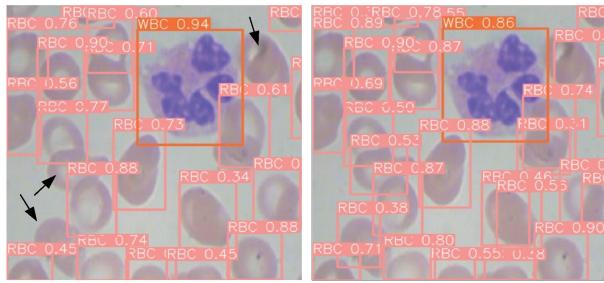


图 12 模型预测结果对比

Fig. 12 Comparison of the model prediction results

3)不同模型性能对比

为了评估所提出模型的性能,进行了充分的对比实验,本文首先比较了多种先进的目标检测模型,包括 Faster R-CNN、YOLOv3、YOLOv5、YOLOv8、YOLOv9^[17]、YOLOv10^[18]和 RT-DETR 模型。此外,本文还将所本文模型与血细胞任务中最新的检测方法进行了对比,包括

Random forests^[19]、SVM^[20]、Fuzzy classifier^[21]、ISE-YOLO^[22]、DCBC DEEPL^[23]、TE-YOLOF B3^[24]、YOLOV5-CW^[25]和 AYOLOv5。

从表 2 的对比结果可见,本文提出的 YOLO-BioFusion 模型在多个关键性能指标上表现优异,明显超过了当前主流目标检测算法。首先,在精度和召回率方面,模型分别达到了 87.5% 和 90.9%,相较于 YOLOv5 和 YOLOv8 等经典模型显著提升,表明其在精确识别和定位血细胞方面具备更高的鲁棒性与可靠性,有效减少了误检和漏检现象。其次,模型在 mAP@0.5 和 mAP@0.5:0.9 上分别取得 94.0% 和 65.2% 的成绩,尤其是在多尺度、复杂背景和目标重叠场景下展现出卓越的检测性能,反映出模型在应对多样化检测挑战时的强大适应性。与之相比,模型的计算效率同样值得关注,GFLOPs 为 6.8,显著低于其他对比模型,显示出其在高效性和资源友好性上的优势,使其更适合应用于计算资源受限的场景。此外,模型参数量仅为 2.9 M,远小于 YOLOv3-l(103.7 M)和 RT-DETR-r18(19.8 M),在维持高检测性能的同时大幅降低了模型复杂度,显著提升了实际应用的可行性。

表 2 不同目标检测模型性能对比

Table 2 Performance comparison of different target detection models

Model	Precision	Recall	mAP@0.5	mAP@0.5:0.9	parameters	GFLOPs
Faster R-CNN	75.5	86.6	85.7	59.8	41.5	128.9
RT-DETR	79.0	88.3	88.5	62.0	19.8	56.9
YOLOV3-ting	84.4	89.0	91.1	62.9	18.9	12.1
YOLOv3-l	85.2	88.9	91.8	63.8	103.7	282.2
YOLOV3-ting	84.4	89.0	91.1	62.9	18.9	12.1
YOLOv5-n	84.9	91.0	92.3	63.1	2.5	7.1
YOLOv5-s	85.8	90.5	92.0	63.2	9.1	24.0
YOLOv8-n	84.2	91.6	92.1	62.0	3.0	8.1
YOLOv8-s	86.0	90.0	92.5	63.4	11.1	28.7
YOLOv9-n	79.0	82.7	85.9	59.6	2.7	8.4
YOLOv10-n	81.2	86.7	88.5	62.0	2.2	6.5
YOLO-BioFusion	87.5	90.9	94.0	65.2	2.9	6.8

从表 3 的对比结果来看,本文提出的 YOLO-BioFusion 模型和其他应用去血细胞检测的模型相比见也表现出明显的优势,在精度(87.5%)、召回率(90.9%)和 mAP@0.5(94.0%)等关键指标上均超越现有模型,与传统机器学习方法(Random Forests、SVM、Fuzzy Classifier)相比,本文模型在精确度和鲁棒性上大幅提升,而与最新的深度学习方法(ISE-YOLO、DCBC DEEPL、TE-YOLOF B3、AYOLOv5)相比同样取得更好的表现。

4)主干网络对比

为了更好地展示改进后主干网络的优越性和有效性,本文将 YOLO-BioFusion 模型的主干网络与当前流行的主干网

络(包括 Swin Transformer^[26]、EfficientFormer^[27]、efficientViT^[28]、LSKNet^[29]和 FasterNet^[30])进行了对比。从表 4 中的结果可以看出,YOLO-BioFusion 模型在精度(Precision(87.5)、mAP@0.5(94.0)和 mAP@0.5:0.9(65.2))关键指标上均优于其他主干网络模型。这些结果充分证明了改进后的主干网络在提升检测精度方面的显著有效果。

5)注意力机制对比

为验证 LSK 注意力机制与 SPPF 结构相结合在血细胞检测任务中的有效性,本文选择了 Squeeze-and-Excitation Attention^[31]、TripletAttention^[32]、Efficient Multi-Scale Attention^[33]和 Convolutional Block Attention^[34]机制

表 3 对比不同文献中血细胞检测模型的对比分析

Table 3 Comparative analysis of blood cell detection models in different literature

Model	时间	文献	Precision	Recall	mAP@0.5
Random forests	2018	[19]	—	—	74.3
SVM	2019	[20]	—	—	79.7
Fuzzy classifier	2020	[21]	—	—	76.0
ISE-YOLO	2021	[22]	—	—	85.7
DCBC DEEPL	2022	[23]	79.9	80.3	82.4
TE-YOLOF B3	2022	[24]	—	—	91.9
YOLOV5-CW	2023	[25]	84.0	85.7	89.9
AYOLOv5	2023	[13]	86.2	91.5	93.3
YOLO-BioFusion	—	—	87.5	90.9	94.0

表 4 主干网络优化对比

Table 4 Comparison of backbone network optimization

模型主干	Precision	Recall	mAP@0.5	mAP@0.5:0.9
Swin Transformer	81.4	90.4	90.0	62.2
EfficientFormer	82.6	87.8	91.2	63.1
EfficientViT	80.8	91.3	88.9	63.7
LSKNet	79.1	91.2	90.3	63.1
FasterNet	75.9	88.8	89.3	59.1
YOLO-BioFusion	87.5	90.9	94.0	65.2

进行对比实验,所有注意力机制均在相同的位置进行添加,结果如表 5 所示。实验表明,SPPF-LSK 组合取得了最

佳 mAP@0.5 和 mAP@0.5:0.9,充分展示了 LSK 与 SPPF 结合在血细胞检测性能上的显著优势。

表 5 注意力机制对比

Table 5 Contrast of attention mechanisms

注意力机制	Precision	Recall	mAP@0.5	mAP@0.5:0.9
Squeeze-and-Excitation Attention	88.5	88.4	93.4	62.0
Triplet Attention	86.2	91.7	92.8	64.5
Efficient Multi-Scale Attention	86.5	91.5	92.7	64.8
Convolutional Block Attention	85.2	89.9	93.3	64.8
LSK Attention	87.5	90.9	94.0	65.2

6) 损失函数对比

为了评估 Inner-CIoU 中 ratio 系数对模型性能的影响,本文进行了相关对比实验,实验结果如表 6 所示。研究表明,随着 ratio 从 0.7 逐步增加至 1.1,模型在 ratio=0.8 时表现最佳。此外,本文还对比了多种主流损失函数,包括 EIoU、DIoU、SIoU、Inner-EIoU 和 Inner-CIoU,实验结果如表 7 所示。结果表明,Inner-CIoU 在 mAP@0.5 和 mAP@0.5:0.9 两项关键指标上均取得最佳表现,进一步验证了其在血细胞检测任务中的卓越性能。这一发现充分说明 Inner-CIoU 在提升检测精度方面的有效性和优势。

7) 消融实验

为了评估所提出网络中各模块的效果,进行了消融实

表 6 Inner-CIoU 不同 Ratio 系数对比

Table 6 Comparison of the different Ratio coefficients of Inner-CIoU

Ratio	Precision	Recall	mAP@0.5	mAP@0.5:0.9
0.7	86.2	90.5	93.3	64.0
0.8	87.5	90.9	94.0	65.2
0.9	85.4	90.2	93.0	64.6
1.0	86.0	88.9	92.4	63.8

验,验证每个改进模块对模型整体性能的影响。表 8 展示了消融实验的结果,涵盖了标准 YOLOv8-n 及其与各改进模块(ACFN、C2f-DPE、SPPF-LSK 和 Inner-CIOU)的组合。

表7 损失函数对比

Table 7 The comparison of the loss functions

损失函数	Precision	Recall	mAP@0.5	mAP@0.5:0.9
EIOU	84.5	92.4	93.3	65.0
DIOU	86.6	90.5	90.3	63.6
SIOU	85.4	89.2	92.7	62.8
Inner-EIOU	86.2	89.7	93.1	64.2
Inner-CIOU	87.5	90.9	94.0	65.2

实验结果表明,加入不同的改进模块显著提升了模型的各项性能指标。引入ACFN模块后,精度(从84.0%提高至85.2%,mAP@0.5从91.9%升至92.5%,mAP@0.5:0.95从62.0%增加到64.4%)。ACFN模块通过自适应卷积操作,特别优化了对小目标的检测能力,显著提高了对细小和重叠目标的检测精度。同时,GFLOPs从8.1降低到6.5,表明该模块在提高检测精度的同时有效减轻了计算复杂度。引入C2f-DPE模块后,召回率提升至92.5%,mAP@0.5增加到92.8%,mAP@0.5:0.95提升

表8 消融实验结果

Table 8 Results of the ablation experiments

Model	Precision	Recall	mAP@0.5	mAP@0.5:0.95	GFLOPs
YOLOv8-n	84.0	91.9	92.4	62.0	8.1
YOLOv8+ACFN	85.2	90.7	92.5	64.4	6.5
YOLOv8+C2f-DPE	86.3	92.5	92.8	62.6	8.0
YOLOv8+SPPF-LSK	82.5	91.9	92.7	62.8	8.4
YOLOv8+Inner-CIOU	85.3	90.8	92.9	63.0	8.1
YOLOv8+ACFN+C2f-DPE	83.9	91.1	92.3	63.6	6.3
YOLOv8+ACFN+SPPF-LSK	84.1	91.6	92.6	63.4	6.9
YOLOv8+C2f-DPE+SPPF-LSK	83.4	90.5	93.0	64.5	8.3
YOLO-BioFusion	87.5	90.9	94.0	65.2	6.8

至62.6%。C2f-DPE模块通过多尺度的感受野扩展,增强了对不同尺寸目标的适应性,特别是对于血细胞图像中多尺度的目标,提升了模型对不同尺度特征的融合能力。GFLOPs略微下降,进一步证明该模块在实际应用中不仅提升了检测能力,还优化了计算效率。当引入SPPF-LSK模块时,虽然精度略微下降,但mAP@0.5达到了92.7%,mAP@0.5:0.95提升至62.8%。SPPF-LSK模块通过融合空间金字塔池化和大尺度卷积注意力机制,特别优化了多尺度特征的捕捉,尤其在复杂背景中,能够提升模型对细粒度信息的提取,稳定且精准地处理目标重叠或密集区域。结合ACFN和C2f-DPE后,虽然精度略有波动,但mAP@0.5:0.95提升至63.6%,GFLOPs降至6.3,表明这两个模块的组合在减轻计算负担的同时,仍能保持较高的检测精度,进一步提升了模型在复杂任务中的泛化能力。

最终,当ACFN、C2f-DPE、SPPF-LSK和Inner-CIOU四个模块结合使用时,模型的mAP@0.5达到了94.0%,

mAP@0.5:0.95提升至65.2%,显示出这些模块的综合作用,成功解决了小目标检测、多尺度特征融合和复杂背景下目标检测的挑战,验证了该模型在复杂检测任务中的强大效果。

8)添加噪声和不同光照对模型的影响

本文对数据集中的图像进行了高斯噪声处理,以模拟实际临床环境中可能出现的图像干扰。噪声通常会在图像采集、传输或处理过程中引入,影响图像质量。通过在血液涂片图像中加入这些噪声,能够评估模型在噪声干扰下的鲁棒性。实验结果(表9)表明,尽管噪声的加入导致模型在mAP@0.5上出现轻微下降,但引入ACFN、C2f-DPE、SPPF-LSK和Inner-CIOU后,模型的mAP@0.5相较于原始YOLOv8-n模型有所提升,表现出更强的能力来捕捉重叠和小尺寸目标。这表明,ACFN、C2f-DPE、SPPF-LSK和Inner-CIOU能够有效提高模型在噪声环境中的检测精度和鲁棒性。

表9 添加噪声后模型性能对比

Table 9 Compares the model performance table after adding noise

Model	噪声	Precision	Recall	mAP@0.5	mAP@0.5:0.95
YOLOv8-n	有	84.4	91.6	91.9	61.7
YOLO-BioFusion	有	86.5	90.6	93.5	63.8
YOLO-BioFusion	无	87.5	90.9	94.0	65.2

在表 10 中,展示了本文模型在不同光照和对比度条件下的性能表现。实验结果表明,模型在不同光照和对比

度设置下表现稳定,能够在大多数光照条件下保持良好的性能,确保了模型在复杂环境下的适应性。

表 10 模型在不同光照和对比度下性能表现

Table 10 Performance table of the model under different illumination and contrast

光照	对比度	Precision	Recall	mAP@0.5	mAP@0.5:0.95
0.2	0.2	86.4	90.7	93.4	65.0
0.3	0.3	85.8	90.8	93.1	64.9
0.4	0.4	85.5	89.8	92.9	64.2

9) 泛化实验

为了验证本文算法的通用性,在 TXL-PBC 数据集上与 YOLOv8-n 进行了对比实验。TXL-PBC 数据集是一个公开可获取、经过专家标注的外周血细胞数据集,包含多种类型的血细胞图像,如白血球、红血球和血小板等。数据集中每张图像都被标注了细胞类型及其状态,适用于血液细胞的分类与检测任务。该数据集总计

包含 1 440 张图像,实验中使用的输入图像分辨率为 640 pixel×640 pixel。实验结果如表 11 所示。从表 11 可以看出,改进后的算法在 Precision 和 Recall 上分别达到了 95.1% 和 95.7%,均高于基线模型。mAP@0.5 提升至 97.9%,mAP@0.5:0.95 为 83.6%,各项指标均实现了有效提升。这表明本文模型在不同任务中的泛化性能优异。

表 11 泛化实验结果

Table 11 Results of the generalization experiments

Model	Precision	Recall	mAP@0.5	mAP@0.5:0.95	GFLOPs
YOLOv8-n	93.8	94.8	97.4	82.2	8.1
YOLO-BioFusion	95.1	95.7	97.9	83.6	6.8

3 结论

本研究提出的 YOLO-BioFusion 模型通过引入 ACFN 模块、C2f-DPE 模块、SPPF-LSK 模块和 Inner-CIoU 损失函数,显著提升了血细胞检测的精度和计算效率。与现有文献相比,YOLO-BioFusion 在小目标检测、多尺度特征融合及复杂背景处理等方面实现了关键突破。ACFN 模块采用动态调整卷积操作,有效增强了对细小目标和重叠目标的识别能力;C2f-DPE 模块利用扩展感受野机制,提高了多尺度特征的提取与融合效果;SPPF-LSK 模块结合大卷积核与注意力机制,增强了模型在复杂背景下的鲁棒性;Inner-CIoU 损失函数则优化了目标定位精度。实验结果表明,YOLO-BioFusion 在 BCCD 和 TXL-PBC 数据集上均表现优异,与主流检测模型相比,mAP 和 GFLOPs 等关键指标上取得了显著提升。然而,模型在处理极为密集的重叠红细胞时仍有漏检问题,推理速度在资源极为有限的设备上也有优化空间。整体而言,YOLO-BioFusion 为血细胞检测和小目标检测任务提供了有效的解决方案,并展现出较强的泛化能力。未来的研究应继续优化模型的计算效率,以更好地支持移动设备或嵌入式环境下的应用。

参考文献

[1] KHODASHENAS M, EBRAHIMPOUR K H,

- NICKFARJAM A M. White blood cell detection and counting based on genetic algorithm [C]. 2019 Advances in Science and Engineering Technology International Conferences(ASET). IEEE, 2019: 1-4.
- [2] ACHARYA V, PRAKASHA K. Computer aided technique to separate the red blood cells, categorize them and diagnose sickle cell anemia[J]. Journal of Engineering Science & Technology Review, 2019, 12(2), DOI: 10.25103/jestr.122.10.
- [3] BISWAS S, GHOSHAL D. Blood cell detection using thresholding estimation based watershed transformation with Sobel filter in frequency domain [J]. Procedia Computer Science, 2016, 89: 651-657.
- [4] REDMON J, DIVVALA S, GIRSHICK R, et al. You only look once: Unified, real-time object detection[C]. IEEE Conference on Computer Vision and Pattern Recognition, 2016: 779-788.
- [5] LIU W, ANGUELOV D, ERHAN D, et al. SSD: Single shot MultiBox detector[C]. Computer Vision-ECCV, 2016: 21-37.
- [6] REN SH Q, HE K M, GIRSHICK R, et al. Faster R-CNN: Towards real-time object detection with regionproposal networks[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2017,

- 39(6): 1137-1149.
- [7] DAI J F, LI Y, HE K M, et al. R-FCN: Object detection via region-based fully convolutional networks [C]. 30th International Conference on Neural Information Processing Systems, 2016: 379-387.
- [8] TIAN ZH, SHEN CH H, CHEN H, et al. Fcos: Fully convolutional one-stage object detection [C]. IEEE/CVF International Conference on Computer Vision, 2019: 9627-9636.
- [9] YANG Z, LIU SH H, HU H, et al. Repoints: Point set representation for object detection [C]. IEEE/CVF International Conference on Computer Vision, 2019: 9657-9666.
- [10] ZHAO Y, LYU W Y, XU SH L, et al. Detrs beat yolos on real-time object detection [C]. IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2024: 16965-16974.
- [11] 周煜庭,余华平,肖粮钧,等.基于改进YOLOv7模型的血细胞检测算法研究[J].国外电子测量技术,2024,43(4): 1-9.
- ZHOU Y T, YU H P, XIAO L J, et al. Research on blood cell detection algorithm based on improved YOLOv7 model[J]. Foreign Electronic Measurement Technology, 2024, 43(4): 1-9.
- [12] GUO Y C, ZHANG M Y. Blood cell detection method based on improved YOLOv5 [J]. IEEE Access, 2023, 11:67987-67995.
- [13] GU W C, SUN K X. AYOLOv5: Improved YOLOv5 based on attention mechanism for blood cell detection[J]. Biomedical Signal Processing and Control, 2024, 88: 105034.
- [14] ROHAZIAT N, TOMARI M R M, ZAKARIA W N W. White blood cells type detection using YOLOv5[C]. 2022 IEEE 5th International Symposium in Robotics and Manufacturing Automation(ROMA). IEEE, 2022: 1-6.
- [15] GAO J, CHEN Y, WEI Y, et al. Detection of specific building in remote sensing images using a novel YOLO-S-CIOU model. Case: Gas station identification[J]. Sensors, 2021, 21(4): 1375.
- [16] 易见兵,黄苏泉,曹锋,等.基于改进的YOLOX血细胞检测算法研究[J].电子测量技术,2022,45(22):177-184.
- YI J B, HUANG S Q, CAO F, et al. Research on blood cell detection algorithm based on improved YOLOX [J]. Electronic Measurement Technology, 2022, 45(22): 177-184.
- [17] LU D, WANG Y X. MAR-YOLOv9: A multi-dataset object detection method for agricultural fields based on YOLOv9[J]. Plos One, 2024, 19(10): e0307643.
- [18] MEI J H, ZHU W Q. BGF-YOLOv10: Small object detection algorithm from unmanned aerial vehicle perspective based on improved YOLOv10 [J]. Sensors, 2024, 24(21): 6911.
- [19] LI Y X, CHEN Z X. Performance evaluation of machine learning methods for breast cancer prediction[J]. Applied Computational Mathematics, 2018, 7(4): 212-216.
- [20] SINGH B K. Determining relevant biomarkers for prediction of breast cancer using anthropometric and clinical features: A comparative investigation in machine learning paradigm [J]. Biocybernetics and Biomedical Engineering, 2019, 39(2): 393-409.
- [21] ONTIVEROS-ROBLES E, MELIN P. Toward a development of general type-2 fuzzy classifiers applied in diagnosis problems through embedded type-1 fuzzy classifiers[J]. Soft Computing, 2020, 24(1): 83-99.
- [22] LIU C, LI D W, HUANG P. ISE-YOLO: Improved squeeze-and-excitation attention module based YOLO for blood cells detection[C]. 2021 IEEE International Conference on Big Data, Orlando, FL, USA, 2021: 3911-3916.
- [23] RAHAMAN M A, ALI M M, HOSSEN M N, et al. DCBC_DeepL: Detection and counting of blood cells employing deep learning and YOLOv5 model [C]. Artificial Intelligence Data Science, vol. 1673, Cham, Switzerland: Springer, 2022: 203-214.
- [24] XU F, LI X K, YANG H, et al. TE-YOLOF: Tiny and efficient YOLOF for blood cell detection [J]. Biomedical Signal Processing and Control, 2022, 73: 103416.
- [25] HUANG M G, WANG B J, WAN J CH, et al. Improved blood cell detection method based on YOLOv5 algorithm [C]. IEEE 6th Information Technology, Networking, Electronic and Automation Control Conference, Chongqing, China: IEEE, 2023: 992-996.
- [26] LIU Z, LIN Y T, CAO Y, et al. Swin transformer: Hierarchical vision transformer using shifted windows[C]. IEEE/CVF International Conference on Computer Vision, 2021: 10012-10022.
- [27] LI Y Y, YUAN G, WEN Y, et al. Efficientformer: Vision transformers at mobilenet speed[J]. Advances in Neural Information Processing Systems, 2022, 35: 12934-12949.
- [28] LIU X Y, PENG H W, ZHENG N X, et al. Efficientvit: Memory efficient vision transformer with

- cascaded group attention[C]. IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2023: 14420-14430.
- [29] LIU J, CAI Z L, HE K F, et al. An efficient printing defect detection based on YOLOv5-DCN-LSK [J]. Sensors, 2024, 24(23): 7429.
- [30] QIN C X, ZHOU Z Y. YOLO-FGD: A fast lightweight PCB defect method based on fastnet and the gather-and-distribute mechanism [J]. Journal of Real-Time Image Processing, 2024, 21(4): 122.
- [31] HU J, SHEN L, SUN G. Squeeze-and-excitation networks[C]. IEEE Conference on Computer Vision and Pattern Recognition, 2018: 7132-7141.
- [32] HUANG ZH H, SU L M, WU J J, et al. Rock image classification based on EfficientNet and triplet attention mechanism [J]. Applied Sciences, 2023, 13(5): 3180.
- [33] OUYANG D L, HE S, ZHANG G ZH, et al. Efficient multi-scale attention module with cross-spatial learning [C]. ICASSP 2023-2023 IEEE International Conference on Acoustics, Speech and Signal Processing. IEEE, 2023: 1-5.
- [34] WOO S, PARK J, LEE J Y, et al. CBAM: Convolutional block attention module[C]. European Conference on Computer Vision(ECCV), 2018: 3-19.

作者简介

张傲,硕士研究生,主要研究方向为数据安全与数据处理。

E-mail:1797718602@qq.com

刘微(通信作者),博士,教授,主要研究方向为智能优化算法及其应用。

E-mail:LiuWei19781020@126.com