

基于改进 YOLOv8n 的无人机目标检测研究

闫泽雨, 杨 洋, 王国琿

(西安工业大学 光电工程学院, 西安 710021)

摘要: 无人机反制系统中, 无人机目标具有显著的多尺度特征, 特别是在小目标检测方面, 往往检测精度较低; 针对此问题提出一种基于动态样本注意力尺度序列的 YOLOv8n 优化改进算法 YOLOv8-C2f-RFCBAMConv; 分析了目标无人机在多尺度和小目标识别中的挑战, 提出通过改进主干网络和融合 C2f-RFCBAMConv 模块来优化特征提取能力, 提升多尺度特征处理能力; 采用 RFCBAM 机制, 并引入残差融合和上下文注意力机制, 提升了特征表达能力并减少计算复杂度; 此外, 使用 WIoU 损失函数改善了小目标低质量数据对梯度的影响, 加快了网络收敛速度; 实验结果显示, 改进模型在自采无人机数据集上的 $mAP@0.5$ 和 $mAP@0.5:0.95$ 分别提升了 3.1% 和 1.7%, GFLOPs 提升了 0.7, 表现出更高的检测精度和更低的计算复杂度。

关键词: YOLOv8; 目标检测; 注意力机制; C2f 模块; RFCBAM; 无人机

Research on UAV Target Detection Based on Improved YOLOv8n

WEN Zeyu, YANG Yang, WANG Guohui

(School of Opto-electronic Engineering, Xi'an Technological University, Xi'an 710021, China)

Abstract: In unmanned aerial vehicle (UAV) countermeasure systems, UAV targets exhibit prominent multi-scale characteristics, particularly in small target detection, which often have a low detection accuracy. To address this issue, a YOLOv8n-based optimization algorithm, YOLOv8-C2f-RFCBAMConv, is proposed, which incorporates a dynamic sample attention scale sequence. this paper makes a analysis of the challenges of multi-scale and small target recognition in UAVs, optimizes feature extraction capability by improving the backbone network and integrating the C2f-RFCBAMConv module, and improves multi-scale feature extraction capability. By adopting the RFCBAM mechanism and introducing residual fusion and context attention mechanisms, the feature representation ability is improved, and the computational complexity is reduced. Furthermore, the WIoU loss function is used to improve the impact of low-quality data from small targets on gradient propagation, accelerating the convergence speed of the network. Experimental results demonstrate that the proposed model achieves a 3.1% and 1.7% improvement in the $mAP@0.5$ and $mAP@0.5:0.95$, respectively, on a self-collected UAV dataset, with a 0.7 GFLOP increase indicating a higher detection accuracy and lower computational complexity.

Keywords: YOLOv8; object detection; attention mechanism; C2f module; RFCBAM; UAV

0 引言

近年来, 无人机等移动平台在安防、侦查和监控等领域的应用日益广泛^[1]。反无人机技术雷达系统、光学摄像头和热成像技术可在不干扰无人机飞行的情况下进行监测和识别。^[2]

目前针对大中型目标的检测已取得显著成效。然而, 小目标检测依然面临诸多挑战。当前对小目标的界

定主要存在两种标准: 其一为基于绝对尺寸的判定方法, 通常将图像中长宽均不超过 32 像素的检测目标视为小目标; 其二则依据相对尺度进行划分, 参照根据国际光电工程学会采纳的标准, 在 256×256 像素的图像中, 当目标成像区域的像素总数低于 80, 或其占图像总面积的比例小于 0.12% 时, 该目标即被判定为小目标, 同样将其归为小目标范畴^[7]。

在无人机目标检测任务中, 目标与光学摄像头的距

收稿日期: 2024-10-25; 修回日期: 2024-12-09。

基金项目: 陕西省教育厅科学研究计划项目(22JY025)。

作者简介: 闫泽雨(2000-), 男, 硕士。

通讯作者: 王国琿(1980-), 男, 博士, 教授, 博导。

引用格式: 闫泽雨, 杨 洋, 王国琿. 基于改进 YOLOv8n 的无人机目标检测研究[J]. 计算机测量与控制, 2025, 33(11): 133-141.

离变化导致目标存在多尺度特征并以小目标为主,并且需部署在边缘计算设备上,对实时性也有很高要求。以 YOLO 系列^[8]为主的一阶段目标检测方法对于复杂场景下的任务具有较高的实时性和准确性,常见的结构设计是在此基础上集成注意力模块与卷积运算,以压缩模型复杂度与参数量,从而适应实时检测场景。例如, Du-an 等人^[14]通过将 CBAM^[15]与 Swim-transformer 结合,有助于提升模型效率,但带来了额外的计算负担和参数复杂度。王春梅等人^[16]则将 SPD 模块集成至主干网络中,虽然可以扩展感受野并提升模型的多尺度特征提取能力,但也会带来计算开销增加、过拟合等问题。针对小目标检测, Cao 等^[17]在 YOLOv5 网络结构中引入了多尺度特征融合模块,以提高小目标的检测精度。然而,这种方法需要处理和融合来自不同尺度的特征图,从而增加了模型的计算复杂度和推理时间。Yi 等^[18]提出了一种基于 YOLOv7 的网络,通过融合注意力机制、增强 FPN 网络拓扑结构及结合损失函数等手段,集中提取小目标的关键特征信息。然而,复杂的结构设计也带来了潜在的过拟合问题,限制了模型的泛化效能。Zhang 等^[19]提出一种新颖的接受域注意力机制(RFA),它全面地解决了卷积核的参数共享问题,并考虑了接受域中每个特征的重要性。RFA 设计的卷积运算(RFAConv)是一种开创性的方法,可以取代目前神经网络中的标准卷积运算。Li 等^[20]于颈部集成 Bi-PAN-FPN 与 GhostblockV2,增强了整体检测效能,却难以应对多尺度复杂场景下的小目标。史涛等^[21]提出的 SPPF-LSKA 与 CF-FPN 组合旨在解决长程依赖与特征冲突,但对遮挡目标鲁棒性不足。潘玮等^[22]通过引入多种注意力机制与改进损失函数来强化小目标特征,但付出了模型复杂度增加的代价。相较之下,程期浩等^[23]设计的 SConv 与 C2f-LSK 模块则侧重于灵活捕捉多尺度特征,以提升模型对小目标的感知能力。

综上所述,本文提出了优化改进的 YOLOv8 目标检测算法 Yolov8-C2f-RFCBAMConv。有效地提升了特征的表达能力,并减少了模型的计算复杂度;改善了小目标低质量数据对梯度的不利影响,在达到实时性检测要求和模型轻量化前提下提高了模型性能,在无人机防御系统和城市反无人机作业中能快速的识别定位无人机目标。

1 Yolov8n 模型

YOLOv8 引入了一种 SOTA 模型架构,类似于 YOLOv5,并提供了根据缩放系数区分的 N/S/M/L/X 不同规模的模型。本研究选择了体积小但精度高的 YOLOv8n。YOLOv8 模型的整体架构主要由 Backbone(骨干网络)、Neck(颈部网络)与 Head(检测头)三

部分构成。Backbone 基于 CSPDarkNet 结构^[24],整合了 Conv、C2f 与 SPPF 三个核心模块,共同完成多层次图像特征的高效提取。具体而言,Conv 模块执行初步特征提取;C2f 模块采用融合 CSP 分流与残差连接的设计,并嵌入 Bottleneck 单元以达成轻量化处理;而 SPPF 模块则执行空间金字塔池化操作,以实现异构特征图的尺度标准化。在 Neck 部分,该模型借鉴了 YOLOv7 中的 ELAN 设计^[25],并引入 PANet 结构以加强特征金字塔中的信息流动,有效促进底层细节特征向高层语义特征的传递。检测头(Head)采用了解耦头结构,将分类任务与检测任务分离,从而提升模型的学习效率。在损失函数设计上,模型分别为分类任务与回归任务设定了损失函数:分类任务使用二元交叉熵(BCE Loss)^[26],回归任务则采用分布式聚焦损失(DFL)与完全交并比损失(CIoU Loss)联合作用于边界框优化。YOLOv8n 的整体结构如图 1 所示。

2 改进算法 Yolov8-C2f-RFCBAMConv

尽管 YOLOv8n 在通用目标检测中表现良好,但其原始结构在应对无人机视角所特有的复杂背景、小目标及多尺度挑战时却存在明显不足。为克服这些瓶颈,本研究旨在对 YOLOv8n 进行针对性结构改进,提出一种更适应无人机场景的目标检测算法—Yolov8-C2f-RFCBAMConv。该算法对 YOLOv8n 的主干网络进行了重新设计,引入了 RFCBAMConv 模块,以替代基线模型主干网络中的普通卷积,从而增强了网络对小目标特征信息的提取能力。此外,将 RFCBAMConv 模块与 C2f 模块相融合,对骨干网络结构进行了改造,进一步提升了模型的性能。为增强模型的泛化能力,本研究采用了加权交并比损失函数(WIoU Loss)^[27]。Yolov8-C2f-RFCBAMConv 的网络结构如图 2 所示。

2.1 RFAConv 模块

感受野注意卷积(RFAConv, receptive-field attention convolution)针对感受野内的空间特征,不仅强调了不同特征的重要性,还对这些特征进行了优先排序。该方法通过根据卷积核的大小动态生成感受野空间特征,使 RFAConv 成为一个固定的卷积组合。与自注意力机制相比,RFAConv 在解决远程信息的参数共享和建模问题时所需的参数和计算资源大幅减少。

为提升模型对无人机小目标多方向特征的提取能力,同时减少因方向变化导致的特征信息流失,本研究选择在骨干网络中引入 RFCBAMConv 模块。该模块基于 RFAConv 开发,通过空间注意机制关注感受野的空间特征,从而解决了卷积参数共享的问题。感受野空间特征是专门为卷积核设计的,并根据核的大小动态生成。如图 3 所示,以卷积核为例,其中“Spatial Fea-

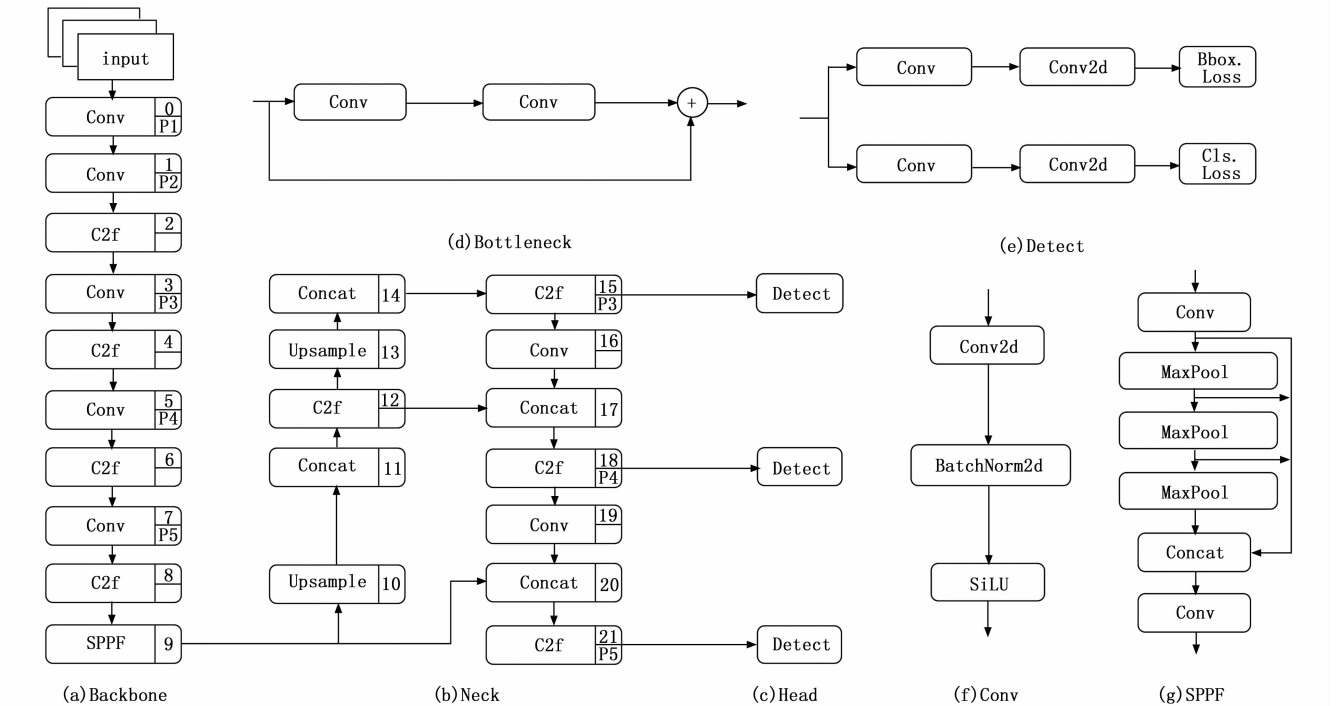


图 1 YOLOv8n 模型结构图

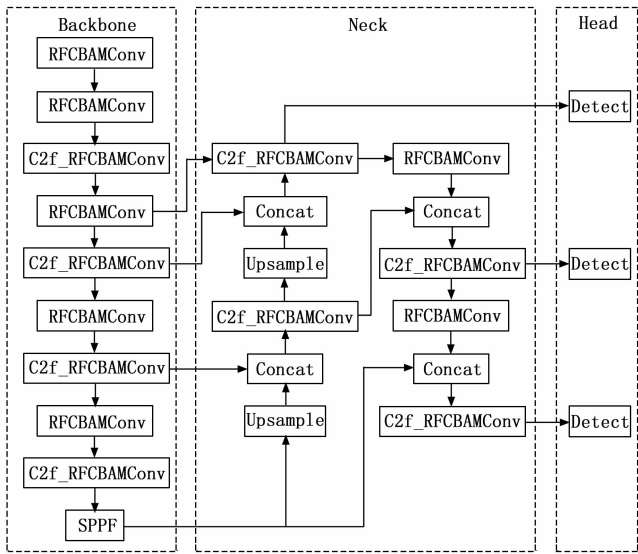


图 2 Yolov8-C2f-RFCBAMConv 模型结构图

ture”指的是原始的 Feature map。“感受野空间特征”是由空间特征变换后的特征映射，由不重叠的滑动窗口组成。感受野空间特征中的每个大小窗口表示一个感受野滑块。为减少计算负担及模型参数数量，采用了平均池化 (AvgPool) 操作以汇聚每个感受野的全局特征信息。同时，利用卷积操作实现特征之间的信息交互。最终，通过 Softmax 函数对每个特征在感受野中的重要性进行加权强调。

一般情况下，RFA 的计算过程可以表示为：

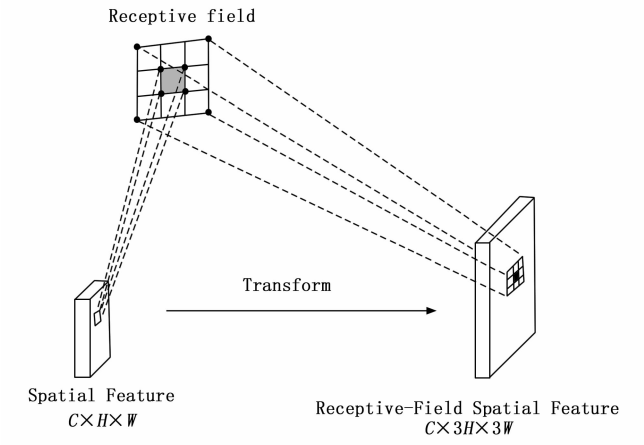


图 3 空间特征特征变换

$$F = \text{Softmax}(g^{1 \times 1}[\text{AvgPool}(X)]) \times \text{ReLU}(\text{Norm}[g^{k \times k}(X)]) = A_{rf} \times F_{rf} \quad (1)$$

RFAConv 的整体结构如图 4 所示。

2.2 C2f_RFCBAMConv 模块

C2f 模块基于传统的 Bottleneck^[28]设计，主要由一个卷积层和两个卷积层构成。该模块通过残差连接将输入特征图与经过卷积操作后的输出特征图相加，从而有效降低计算复杂度，并提高网络的学习能力。C2f 模块进一步通过增加跨层连接的分支，并省略分支中的卷积操作，实现了不同层级特征的融合。此外，引入的额外 split^[29]操作帮助网络更有效地捕捉目标的多尺度和语义信息。这种设计不仅丰富了提取的特征信息和模型的

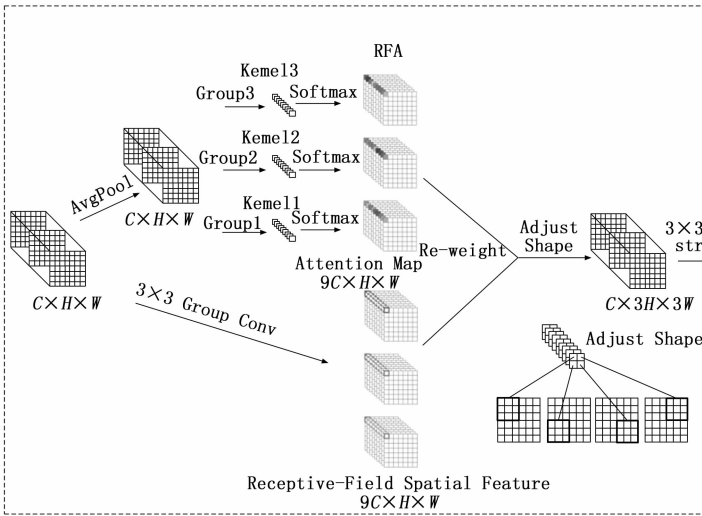


图 4 RFACConv 的详细结构

梯度流，还在性能与效率之间实现了优化平衡。

本文改进用到的 RFCBAMConv (Recalibration and Fusion of Contextual Block Attention Module Convolution) 模块通过集成 RF-CBAM 注意力机制显著增强了卷积神经网络的特征表达能力。RFCBAMConv 专注于感受野的空间特征。与 RFA 类似，它首先通过一个卷积层提取初步特征，然后在第一个卷积层后插入 RF-CBAM 机制，这一机制通过通道注意力和空间注意力模块来优化特征图的上下文信息。具体而言，通道注意力模块通过全局平均池化生成通道统计信息并进行权重调整，而空间注意力模块通过卷积生成空间统计信息并加权处理。最终，这些优化后的特征图通过卷积层进一步处理，残差连接将输入特征图与处理后的输出特征图相加，确保了信息的完整性和网络的稳定性。此模块在复杂视觉任务中如目标检测和图像分类中，能够显著提升模型的性能和精度。具体结构如图 5 所示。

尽管 C2f (CSP Bottleneck) 模块在许多任务中表现出色，但在处理特定复杂场景或小目标检测时，仍然存在进一步优化的潜力。为此，本研究引入了 C2f_RFCBAMConv 模块，在 C2f 模块的基础上结合了 C2f_RFCBAMConv 模块以提升模型的整体性能。

C2f_RFCBAMConv 模块由一个卷积层和两个卷积层构成。首先，卷积层用于初步提取特征。随后，在第一个卷积层后插入 RF-CBAM 机制，以增强特征图的上下文信息处理能力。接下来，使用两个卷积层对特征进行进一步的压缩和扩展。这种模块的替换确保了新的 C2f_RFCBAMConv 模块在输入和输出通道数上与原有 C2f 模块保持一致。最终，整体网络需要进行微调，以适应引入的 C2f_RFCBAMConv 模块。

RFCBAM 机制包括通道注意力和空间注意力两个

部分，分别针对特征通道和空间位置进行优化。该注意力模块以双分支结构协同工作：通道分支通过全局平均池化建模通道间依赖关系，并利用缩放因子进行精细化调整；空间分支则依赖群范数提炼空间统计特征以生成注意力图。最终，通过整合双支输出并引入“Channel Shuffle”算子，促进了跨维度特征交互，从而全面增强了网络的表征能力与精度。

通道注意力子模块通过使用全局平均池化来嵌入全局信息，生成通道统计数据 $s \in R^{C/2G \times 1 \times 1}$ ， $F_{gp}(X_{k1})$ 是对通道 $k1$ 的特征图 X_{k1} 进行全局平均池化操作，得到该通道的一个标量值，代表该通道的全局信息。可以通过空

间维度 $H \times W$ 缩小 X_{k1} 来计算：

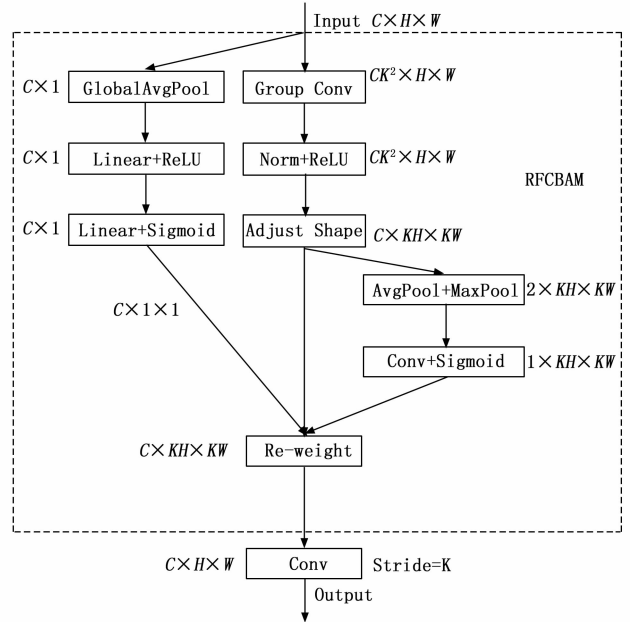


图 5 RFCBAMConv 的详细结构

$$S = F_{gp}(X_{k1}) = \frac{1}{K \times W} \sum_{i=1}^H \sum_{j=1}^W X_{k1}(i, j) \quad (2)$$

通道注意力的输出函数为：

$$X'_{k1} = \sigma(F_c(s)) \cdot X_{k1} = (W_{1s} + b_1) \cdot X_{k1} \quad (3)$$

$W_1 \in R^{C/2G \times 1 \times 1}$ ， $b_1 \in R^{C/2G \times 1 \times 1}$ ，二者用于缩放和移位 s 。其中， W_1 是一个用于缩放和变换的权重矩阵，其大小为 $R^{C/2G \times 1 \times 1}$ ，它主要用于对通道注意力进行调整； b_1 是一个与 W_1 同形状的偏置项，用于在计算中进行加法偏移； s 是通道注意力的输出，是通道的权重，表示每个通道的关注度； W_{1s} 是将通道权重 s 通过 W_1 变换后的结果，用于最终的通道加权。

空间注意力子模块在 X_{k2} 上使用群范数 (GN) 来获得空间统计, $F_c(\cdot)$ 是指一个由全连接层 (Fully Connected Layer) 组成的操作, 它的作用是通过权重矩阵对输入特征进行线性变换来增强 $\hat{X}_{k2}$ 的表示。空间注意力的最终输出为:

$$X'_{k2} = \sigma(W_2 \cdot \text{GN}(X_{k2}) + b_2) \cdot X_{k2} \quad (4)$$

其中: W_2 和 b_2 是 $R^{C/2 \times 1 \times 1}$ 的权重矩阵。对两个支路的输出进行融合, 并使最终通道数与输入特征图对齐, 以维持网络的信息流通能力, 即 $X'_k = [X'_{k1}, X'_{k2}] \in R^{C/G \times W \times H}$ 。

2.3 WIoU Loss

在 YOLOv8 模型的回归分支中, 采用了 Deep Feature Loss (DFL) 和 Complete Intersection over Union (CIoU) 作为损失函数。相较于多数边界框回归损失函数, 这些方法在优化具有相同纵横比但宽高比例不同的预测边界框时展现出更高的效率。具体而言, CIoU 的计算公式如下:

$$L_{\text{CIoU}} = 1 - \text{IoU} + \frac{\rho^2(b, b^g)}{c^2} + \omega \quad (5)$$

在 CIoU 损失函数中, 预测框和真实框的中心点分别表示为 (b, b^g) , 即预测框和真实框的中心点坐标, 式中 IoU 是预测框与真实框之间的重叠区域相对于它们联合区域的比值; ρ 代表二者之间的欧几里得距离; c 代表包围预测框和真实框的最小矩形框的对角线; 在 CIoU 损失函数中, 参数 α 用于平衡中心点距离与长宽比一致性之间的权重, 变量 v 用于表征预测框与真实框之间长宽比的一致性。然而, CIoU 损失仅建模了中心点距离, 未能有效评估边界对齐情况, 且对目标形变适应性不足。相比之下, WIoU 通过引入权重因子动态调整样本损失贡献, 从而更灵活地处理复杂场景, 实现精准回归。鉴于上述局限性, 本研究基于 Focal-EIoU^[30] 思想, 将回归损失函数由 CIoU 替换为 WIoU。

Focal-EIoU 的计算过程详见公式 (6)。这一方法通过优化损失函数的设计, 旨在提高边界框回归的精度和鲁棒性:

$$L_{\text{EIoU}} = 1 - \text{IoU} + \frac{\rho^2(b, b^g)}{c^2} + \frac{\rho^2(\omega, \omega^g)}{c_\omega^2} + \frac{\rho^2(h, h^g)}{c_h^2} \quad (6)$$

$$L_{\text{Focal-EIoU}} = \text{IoU}^\gamma L_{\text{EIoU}} \quad (7)$$

其中: L_{EIoU} 是基于 Empirical IoU (EIoU) 损失的计算, γ 作为控制曲线弯度的超参数, 聚焦机制为静态。为了充分发挥非单调聚焦机制的潜力, Tong 等人提出了 WIoU (Weighted Intersection over Union)^[32], 该方法构建了一种动态非单调聚焦机制以评估锚框质量, 并结合梯度增益分配策略来抑制低质量样本的梯度

干扰, 从而优化高质量锚框的筛选。在 WIoU 的系列版本中, 本研究采用 WIoUv3 实现边界框回归, 其计算公式如下:

$$L_{\text{WIoUv1}} = \exp\left(\frac{(x - x^g)^2 + (y - y^g)^2}{[(w^g)^2 + (h^g)^2]^*}\right) \cdot L_{\text{IoU}} \quad (8)$$

$$L_{\text{WIoUv3}} = \frac{\beta}{\delta \alpha^{\beta-\alpha}} L_{\text{WIoUv1}} \quad (9)$$

其中: w^g 、 h^g 是最小的包围框的大小。 β 、 δ 为调整因子, 用于控制梯度增益的大小。它们的作用是调节锚框的离群度, 从而在训练过程中为不同质量的锚框分配适当的梯度增益。当锚框离群度满足 $B=C$ (C 为某常值) 时, 梯度增益最大, 这个常数可以控制损失函数中各类锚框的优先级分配。WIoUv3 通过引入动态的 LIoU 机制, 使锚框质量评估标准能够自适应调整, 从而实现对梯度增益分配的实时优化。

在模型改进过程中引入 WIoU 损失函数, 通过引入权重因子, 为目标的大小、形状及其他显著特征差异的样本分配更合理的权重。这种权重调整机制使得模型在训练过程中能够更加有效地捕捉和学习这些关键特征, 从而增强了模型在无人机视角下的无人机目标检测任务中的灵活性和鲁棒性。具体而言, WIoU 损失函数通过对不同样本的重要性进行加权, 有助于模型更好地适应复杂的检测环境。这一方法的应用显著提高了整体检测性能, 使得模型在处理各种复杂场景时表现出更高的准确性和稳定性。

3 实验设计与结果分析

3.1 数据集构建

实验所用的数据集来源于自采无人机数据, 由 DJI Mini3 Pro 设备拍摄, 通过从无人机对目标录制的视频中抽取图像帧, 以构建目标检测数据集。该数据集集中的无人机目标以小目标为主, 涵盖了多尺度目标和复杂场景, 以增强目标检测模型的鲁棒性和泛化能力。由于无人机目标在不同的高度和距离上呈现不同大小使得它们具有多尺度的特征, 本数据集涵盖从小于 32×32 像素的小目标到具有明显的几何特征和在图像中有显著占比的大目标等不同尺度的无人机目标。且包含云层、飞行中的其他物体 (如鸟类、飞机)、以及天气因素 (如雾霾或阳光反射) 等复杂场景下的数据。

数据集在筛选过程中剔除了成像质量差的照片后包含 10 585 张图像, 按用途划分为训练集、验证集和测试集。其中训练集包含 5 785 张图像, 用于模型训练, 验证集包含 2 600 张图像, 用于模型训练过程中的中期评估, 帮助调整超参数和进行模型选择。测试集包含 2 200 张图像, 用于模型的最终性能评估, 提供对模型在未见数据上的泛化能力的客观评估。数据标注工作采用

LabelImg 工具对每张图像逐一完成标注。生成的标注数据包含了图像中目标的边界框位置与对应类别,为模型的监督训练与性能评估提供了坚实的基础。部分数据集示例如图 6 所示。

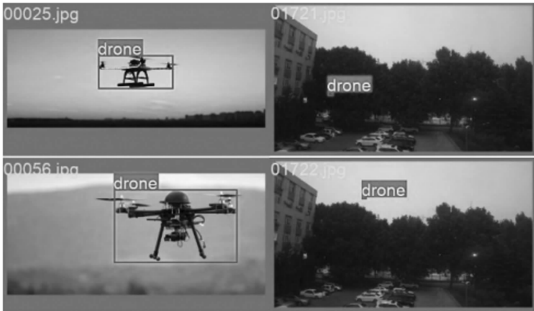


图 6 不同尺寸的无人机目标数据集示例

3.2 实验环境及参数设置

本实验在 Ubuntu18.04 系统下进行, GPU 选择 NVIDIA GeForce RTX 4090, 主机内存为 72 GB, 编程语言为 Python3.9.19, PyTorch 版本为 1.13.1, C-uda 版本为 11.8, 本文模型训练参数根据 YOLOv8n 模型训练参数进行设置, 实验所使用的有关硬件、软件及实验环境详细信息如表 1 所示, 训练参数设置如表 2 所示。

表 1 实验环境配置

配置	版本/参数
操作系统	Ubuntu18.04
CPU	Intel(R) Core(TM) i7-10610U
GPU	GeForce RTX 4090
编译环境	Python 3.9.19
CUDA	11.8
深度学习框架	Pytorch 1.13.1
编译器	Pycharm 2023.2.7

表 2 训练参数

配置	版本/参数
optimizer	SGD
epochs	300
batchesizes	16
workers	4
imgs	640
lr0	0.01
lrf	0.01
weight-decay	0.000 5
momentum	0.8
mosaic	1.0

3.3 模型评价指标

本研究对模型改进效果进行量化分析,核心评估维度涵盖检测性能与模型复杂度。所选定量指标从检测精

度维度,涵盖了精确率(P)、召回率(R)和平均精度均值(mAP),以及参数量($Params$)和十亿次浮点运算量($GFLOPs$)两项效率指标。具体计算公式如下所示:

$$Precions = \frac{TP}{TP + FP}$$
 (10)

$$Recall = \frac{TP}{TP + FN}$$
 (11)

其中: TP 、 FP 和 FN 分别表示被正确识别的正样本、被错误识别的负样本以及被错误忽略的正样本的数量:

$$AP = \int_0^1 p(r) dr$$
 (12)

$$mAP = \frac{\sum_{i=1}^K AP_i}{K}$$
 (13)

平均精度(AP)是衡量单一类别检测性能的指标,定义为交并比(IoU)阈值从 0.5 到 0.95 区间内,其精确率—召回率(P - R)曲线下的面积。平均精度均值(mAP)则为数据集中所有类别 AP 值的平均值, mAP 值越高,表明模型的整体检测性能越优。

3.4 实验结果分析

在自制无人机目标数据集上对本文模型进行 300 轮训练,整理实验所得数据并绘制平滑曲线如图 7 所示,分别为 $Loss$ 曲线、训练模型精度曲线、召回率曲线以及 mAP 曲线。实验结果说明模型在目标检测任务中达到了较为优异的结果,具体而言,在模型训练参数一致的情况下,随着训练轮数缓慢增加,模型 $Loss$ 值平缓减小,表

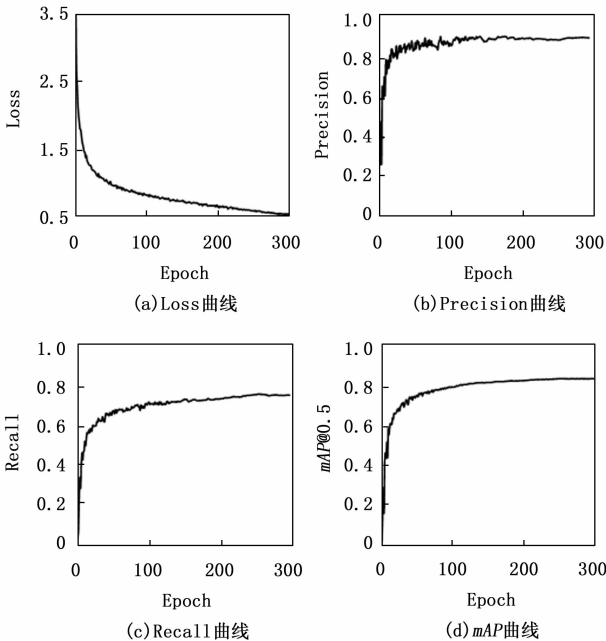


图 7 自制数据集实验结果

明模型未出现过拟合现象。在 $mAP@0.5$ 指标上,模型的检测精度趋近于 85.9%,表明在较高 IoU 阈值下,模型能够较准确地定位和识别目标。此外,模型的 $precision$ 值接近 92.1%,表明其在检测结果中的正样本精度较高,误检率较低。 $recall$ 值接近 76.4%,显示出模型在全面检测多尺度目标(尤其是较小目标)方面的良好能力,能够较全面地识别出真实目标。

3.5 消融实验

为验证改进算法 YOLOv8-C2f-RFCBAMConv 的有效性,本研究对基线模型 YOLOv8n 进行了系统的消融实验。实验中,通过分别引入不同模块及其组合,对模型进行训练,并在验证集(val 集)上进行评估, val 集因其较少的遮挡情况和较多的目标数量,适合作为最终验证结果的分析依据。表 3 所示实验结果中,字母 A、B、C 依次指代 RFCBAMConv 模块、C2f_RFCBAMConv 模块与 $WIoU$ 损失函数的引入。结果表明,引入 RFCBAMConv 模块的模型在 val 集上 $mAP@0.5$ 提升了 0.4%,同时仅增加了 0.2 M 的参数量,验证了其在提升小目标特征关注能力和检测效率方面的有效性。采用 C2f_RFCBAMConv 模块相比于传统卷积,验证集上的 mAP 分别提高了 0.8%和 0.9%,尽管 $GFLOPs$ 增加了 0.2,但显著提高了对不同形状和大小目标的适应性。将 $CIoU$ 损失函数替换为 $WIoU$ 后,各项参数均有小幅度提升。当 RFCBAMConv 模块与 C2f_RFCBAMConv 模块结合使用时,模型在提取不同尺寸目标特征方面表现出色,两种 mAP 指标分别提升了 2.8%和 0.7%,达到 85.6%和 53.2%。最终,将 3 种改进方法综合应用于 YOLOv8-C2f-RFCBAMConv 算法,模型在 val 集上的精度达到 92.1%, $mAP@0.5$ 提升至 85.9%,提高幅度为 3.1%; $mAP@0.5:0.95$ 达到 54.2%,提高了 1.7%。总体而言,3 种改进的有效融合显著提升了无人机目标识别与检测能力,消融实验结果证明了该算法的卓越性能。

表 3 总体消融实验

A	B	C	<i>Precision</i> /%	<i>Recall</i> /%	<i>mAP</i> @0.5 /%	<i>mAP</i> @ 0.5 : 0.95/%	<i>GFLOPs</i>
			92.1	74.0	82.8	52.5	8.1
✓			92.1	74.3	83.2	53.3	8.3
	✓		91.6	75.5	83.6	53.4	8.3
		✓	92.2	75.2	84.9	53.5	8.5
✓	✓		92.1	76.2	85.6	53.2	8.8
✓	✓	✓	92.1	76.4	85.9	54.2	8.8

3.6 对比实验

本研究选取了 SSD、Faster R-CNN 与部分 YOLO 系列模型,在自建的无人机数据集上进行了对比实验,以验证所提模型的性能。结果如表 4 所示。

表 4 轻量化网络性能对比

Model	<i>Precision</i> /%	<i>mAP</i> @ 0.5 /%	<i>mAP</i> @ 0.5 : 0.95 /%	Para/ M	Model size /MB	<i>GFLOPs</i>
Faster R-CNN ^[33]	86.4	74.4	46.4	40.12	47.8	167.8
SSD ^[34]	91.7	71.5	47.2	27.4	17.8	8.9
Yolov5n ^[35]	86.8	79.9	45.5	2.4	3.4	2.3
Yolov6n ^[36]	89.3	81.9	50.2	4.1	5.8	2.7
Yolov8n	92.1	82.8	52.5	3.0	6.3	8.1
Ours	92.1	85.9	54.2	3.1	6.7	8.8

由表 4 可知,本文提出的模型相比于 Faster R-CNN、SSD、YOLOv5n、YOLOv6n、YOLOv8n 等模型, $mAP@0.5/%$ 分别提升了 11.5%、14.4%、6.0%、4.0%、3.1%。 $mAP@0.5:0.95$ 分别提升了 7.8%、7.0%、8.7%、4.0%、1.7%。对比 YOLOV8n 基础检测算法,本文算法在 $mAP0.5/%$ 和 $mAP@0.5:0.95/%$ 两个重要精度指标上有 3.1%和 1.7%的提升,且在参数量和模型大小均未明显增加的情况下, $G-FLOPs$ 提升了 0.7。总的来说,本文所提出的模型在各个指标上相比于之前的模型均有显著提升,尤其是在 $mAP@0.5$ 和 $mAP@0.5:0.95$ 上的提升最为显著。相比于 Faster R-CNN 和 SSD,它的提升尤为明显,显示出其在无人机小目标检测精度上的优势。

3.7 实验结果可视化分析

为了评估本文改进算法 Yolov8-C2f-RFCBAMConv 在无人机视角下的小目标检测性能,我们在自制数据集中选取了一些实际图像,并与基准模型 Y-olov8n 进行了可视化对比。图 8 展示了改进前后的算法对自制无人机数据集的部分检测结果。

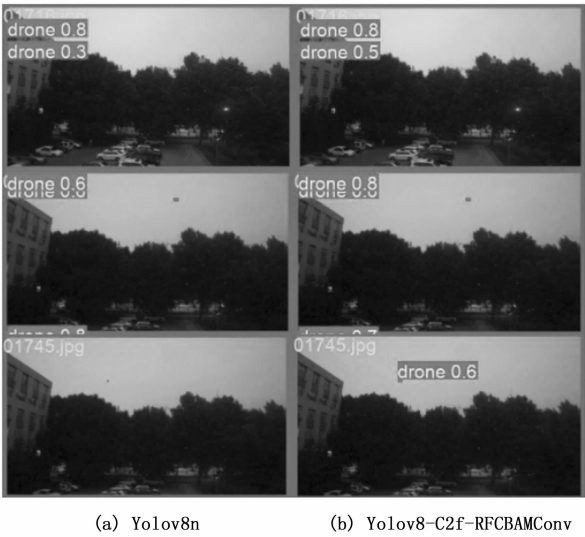


图 8 检测结果对比图

具体而言,第一行图像中的目标在改进模型中

drone 类别的检测精度从 0.3 提升至 0.5；第二行图像中的目标在改进模型中 drone 类别的检测精度从 0.6 提升至 0.8。尤其值得注意的是，第三行图像中的目标在基准模型中出现了漏检现象，而改进模型成功纠正了这一错误，显著改善了检测结果。

从这些对比结果可以明显看出，改进后的 YOLOv8-C2f-RFCBAMConv 算法在处理小目标时展现出了显著的性能提升。改进后的算法不仅提高了检测精度，还在小目标特征的恢复方面表现出更强的能力。通过与基准模型的对比，验证了 YOLOv8-C2f-RFCBAMConv 模型在复杂场景中的有效性及其在小目标检测任务中的优势。这一改进显著增强了算法对小目标的敏感性和检测准确性，展示了其在无人机视角下进行目标检测的潜力和应用价值。

4 结束语

本文提出的改进算法旨在解决无人机反制系统中小目标检测精度不足的问题。针对无人机目标具有显著多尺度特征及小目标识别难度较大的挑战，通过对 YOLOv8n 网络架构的优化与改进，提升了多尺度特征的提取与处理能力，为小目标检测提供了有效的解决方案。

通过在自采无人机数据集上的实验验证，YOLOv8-C2f-RFCBAMConv 算法相较于原始 YOLOv8n，在 $mAP@0.5$ 和 $mAP@0.5 : 0.95$ 指标上分别提升了 3.1% 和 1.7%，GFLOPs 也提高了 0.7，表明在提高检测精度的同时，算法在计算效率上的优化效果明显。这一结果验证了改进算法在小目标检测任务中的有效性，尤其适用于无人机反制系统中的实时目标检测需求。

总之，本文的研究成果为小目标检测尤其是在无人机目标检测任务中的应用提供了新的思路与方法。未来的研究可以在此基础上进一步探索如何结合更先进的特征融合网络、头部网络设计以及多模态数据融合等技术，进一步提升无人机视角下小目标的检测精度和计算效率，以满足实际应用中对实时性和高精度的双重要求。同时，也可以将该方法应用到其他具有类似小目标识别难题的领域，为计算机视觉领域的发展做出贡献。

参考文献：

- [1] 宋晓茹, 刘 康, 高 嵩. 复杂战场环境下改进 YOLOv5 军事目标识别算法研究 [J]. 兵工学报, 2024, 45 (3): 934 - 947.
- [2] 曹亚明, 肖 奇, 杨 震. 仿真图像作为模板的遥感影像小目标检测方法 [J]. 计算机工程与应用, 2022, 58 (17): 111 - 119.
- [3] MITTAL P, SINGH R, SHARMA A. Deep learning-based object detection in low-altitude UAV datasets [J]. Image and Vision computing, 2020, 104: 104046.
- [4] CHENG G, HAN J. A survey on object detection in optical remote sensing images [J]. ISPRS journal of photo-grammetry and remote sensing, 2016, 117: 11 - 28.
- [5] 董 刚, 谢维成, 黄小龙. 深度学习小目标检测算法综述 [J]. 计算机工程与应用, 2023, 59 (11): 16 - 27.
- [6] 李东臣, 任俊杰, 张志文. 基于高分辨率无人机影像的地震地表破裂半自动提取方法——以 2021 年 M_s7.4 青海玛多地震为例 [J]. 地震地质, 2022, 44 (6): 1484 - 1502.
- [7] 李红光, 于若男, 丁文锐. 基于深度学习的小目标检测研究进展 [J]. 航空学报, 2021, 42 (7): 107 - 125.
- [8] REDMON J, DIVVALA S, GIRSHICK R. You only look once: Unified, real-time object detection [C] //Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, IEEE, 2016: 779 - 788.
- [9] REDMON J, FARHADI A. YOLO9000: better, faster, stronger [C] //Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, IEEE, 2017: 7263 - 7271.
- [10] REDMON J, FARHADI A. YOLOv3: an incremental improvement [J]. ArXiv Preprint ArXiv: 1804.02767, 2018.
- [11] BOCHKOVSKIY A, WANG C Y, LIAO H. YOLOv4: optimal speed and accuracy of object detection [J]. ArXiv Preprint ArXiv: 2004.10934, 2020.
- [12] LI C Y, LI L, JIANG H L. YOLOv6: a singlestage object detection framework for industrial applications [J]. ArXiv: 2209.02976, 2022.
- [13] WANG C Y, BOCHKOVSKIY A, LIAO H Y M. YOLOv7: trainable bag-of-freebies sets new state-of-the-art for realtime object detectors [C] //Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2023: 7464 - 7475.
- [14] 段必冲, 马明涛. 基于改进 YOLOv5 算法的口罩检测研究 [J]. 计算机工程与应用, 2023, 59 (16): 223 - 231.
- [15] 郝 帅, 张 旭, 马 旭. 基于 CBAM-YOLOv5 的煤矿输送带异物检测 [J]. 煤炭学报, 2022, 47 (11): 4147 - 4156.
- [16] 王春梅, 刘 欢. YOLOv8-VSC: 一种轻量级的带钢表面缺陷检测算法 [J]. 计算机科学与探索, 2024, 18 (1): 151 - 160.
- [17] CAO S, WANG T, LI T. UAV small target detection algorithm based on an improved YOLOv5s model [J]. Journal of Visual Communication and Image Representation, 2023, 97: 103936.
- [18] YI W, WANG B. Research on underwater small target detection algorithm based on improved YOLOv7 [J].

- IEEE Access, 2023; 66818–66827.
- [19] ZHANG X, LIU C, YANG D, et al. RFAConv: innovating spatial attention and standard convolutional operation [J]. ArXiv Preprint ArXiv: 2304.03198, 2023.
- [20] LI Y, FAN Q, HUANG H, et al. A modified YOLOv8 detection network for UAV aerial image recognition [J]. Drones, 2023, 7 (5): 304.
- [21] 史 涛, 崔 杰, 李 松. 优化改进 YOLOv8 实现实时无人机车辆检测的算法 [J]. 计算机工程与应用, 2024, 60 (9): 79–89.
- [22] 潘 玮, 韦 超, 钱春雨. 面向无人机视角下小目标检测的 YOLOv8s 改进模型 [J]. 计算机工程与应用, 2024, 60 (9): 142–150.
- [23] 程期浩, 陈东方, 王晓峰. 基于 NDM-YOLOv8 的无人机图像小目标检测 [J]. 计算机技术与发展, 2024: 1–9.
- [24] LI C W, et al. A video sequence face expression recognition method based on squeeze-and-excitation and 3DPCA network [J]. SENSORS, 2023, 23 (2).
- [25] LIN T Y, GOYAL P, GIRSHICK R, et al. Focal loss for dense object detection [J]. IEEE Trans Pattern Anal Mach Intell, 2020, 42 (2): 318–327.
- [26] CHENG, QIANQING, LI, et al. Drone detection method based on MobileViT and CA-PANet [J]. ELECTRONICS, 2023, 12 (1).
- [27] TONG Z, CHEN Y, XU Z, et al. Wise-IoU: bounding box regression loss with dynamic focusing mechanism [J]. ArXiv: 2301.10051, 2023.
- [28] PARK J, WOO S, LEE J, et al. Bam: bottleneck attention module [J]. ArXiv Preprint ArXiv: 1807.06514, 2018.
- [29] ZHANG H, WU C, ZHANG Z, et al. Resnest: split-attention networks [J]. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2022: 2736–2746.
- [30] QL Z, YB Y. Sa-net: Shuffle attention for deep convolutional neural networks [C] //Proceedings of International Conference on Acoustics, Speech and Signal Processing (ICASSP), IEEE, 2021: 2235–2239.
- [31] ZHANG Y F, REN W, ZHANG Z, et al. Focal and efficient IOU loss for accurate bounding box regression [J]. Neurocomputing, 2022, 506: 146–157.
- [32] TONG Z, CHEN Y, XU Z, et al. Wise-IoU: bounding box regression loss with dynamic focusing mechanism [J]. ArXiv: 2301.10051, 2023.
- [33] REN S, HEK R G, et al. Faster R-CNN: towards real-time object detection with region proposal networks [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2016, 39 (6): 1137–1149.
- [34] LIU W, ANGUELOV D, ERHAN D, et al. Ssd: single shot multibox detector [C] //Proceedings of Computer Vision-ECCV 2016: 14th European Conference, Amsterdam, IEEE, 2016: 21–37.
- [35] WEI C, TAN Z, QING Q, et al. Fast helmet and license plate detection based on lightweight YOLOv5 [J]. Sensors, 2023, 23 (9): 4335.
- [36] LI C Y, LI L, JIANG H L, et al. YOLOv6: a single-stage object detection framework for industrial applications [J]. ArXiv: 2209.02976, 2022.
- [12] 刘福顺. 无损检测基础 [M]. 北京: 北京航空航天大学出版社, 2002.
- [13] MIGLIORINO A, ZONTA D. Acoustic emission monitoring of prestressed concrete bridges: a practical case study [J]. Nondestructive Testing and Evaluation, 2020, 35 (3): 280–295.
- [14] KAPPATOS V, PAPAELIAS M. Recent advances in acoustic emission technology for industrial applications [J]. Sensors, 2023, 23 (5): 2245.
- [15] DU J B, et al. Using acoustic emission technique for structural health monitoring of laminate composite: a novel CNN-LSTM framework [J]. EngineeringFracture-Mechanics, 2024, 110447.
- [16] LUO Z, FAN X, CHEN Q. The AE signal characteristic parameters extraction and optimization of initial crack during deep drawing of metal [J]. International Conference on Electrical and Control Engineering, Yichang, China, 2011, 1798–1801.
- [17] LECUN Y, BOTTOU L, BENGIO Y, et al. Gradient-based learning applied to document recognition [J]. Proceedings of the IEEE, 1998, 86 (11): 2278–2324.
- [18] HOCHHEITER S, SCHMIDHUBER J. Long short-term memory [J]. Neural Computation, 1997, 9 (8): 1735–1780.
- [19] ALZUBAIDI L, ZHANG J, HUMAIDI A J. Review of deep learning: concepts, CNN architectures, challenges, applications, future directions [J]. Big Data 8, 53, 2021.
- [20] RUMELHART D E, HINTON G E, WILLIAMS R J. Learning representations by back-propagating errors [J]. Nature, 1986, 323 (6088): 533–536.
- [21] HE K, ZHANG X, REN S, et al. Deep residual learning for image recognition [C] // Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2016: 770–778.