

基于 YOLOv8 的红外无人机小目标检测研究

李燕, 曲毅, 胡健生

(武警工程大学信息工程学院, 西安, 710086)

摘要 红外目标检测是无人机反制技术常用的一种手段。针对复杂环境下红外小目标图像特征不明显, 常淹没在噪声中的问题, 提出了一种改进的 YOLOv8 目标检测模型。首先, 引入注意力机制, 自适应调节感受野大小; 其次, 构建小目标检测层, 更加关注网络的浅层信息, 增强细粒度特征提取能力; 最后, 使用深度可分离卷积改进检测头, 提高检测准确度的同时更加轻量化。实验结果表明, 与原 YOLOv8 模型相比, 精确率、召回率、mAP50、mAP50-95 分别提升了 5.3%、8.1%、9.1%、21.1%, 在无人机小目标检测中取得了很好的效果。

关键词 YOLOv8; 红外; 小目标检测; 注意力机制; 深度可分离卷积

DOI 10.3969/j.issn.2097-1915.2025.03.013

中图分类号 TN391.41 **文献标志码** A **文章编号** 2097-1915(2025)03-0106-06

Research on Small Target Detection of Infrared UAV Based on YOLOv8

LI Yan, QU Yi, HU Jiansheng

(School of Information Engineering, Engineering University of PAP, Xi'an 710086, China)

Abstract Infrared target detection is a commonly used means to UAV countermeasure technology. Aimed at the problems that infrared small target image is not obvious in feature and is often submerged in noise under conditions of complex environments, an improved YOLOv8 target detection algorithm is proposed. Firstly, the introduction of an attention mechanism into research is utilized for adaptively adjusting the size of the receptive field. Secondly, a small target detection layer is constructed to pay more attention to the shallow information of the network, enhancing the ability of fine-grained feature extraction. Finally, the detection head improved by depth separable convolution is used to improve the detection accuracy and to simultaneously become even more lightweight. The experimental results show that the precision rate, recall rate, mAP50 and mAP50-95 are improved by 5.3%, 8.1%, 9.1% and 21.1% respectively in comparison with the original YOLOv8 algorithm. The result comes off well in the detection of small targets by UAV.

Key words YOLOv8; infrared detection; small target detection; attention mechanism; depthwise separable convolution

近年来,随着无人机技术的迅猛发展,无人机数量激增,给区域安全带来了巨大的威胁。怎样快速、准确地对入侵无人机进行检测成为无人机反制技术的研究重点。目前,无人机检测技术主要有雷达探测、射频检测、声波探测和光学图像检测(可见光和

红外)^[1]。其中红外目标检测是基于目标的热辐射特性进行被动成像,再利用图像处理、分析等手段检测目标的技术^[2]。与其他技术相比,红外检测技术具有许多突出优点,作用距离远,抗干扰能力强,隐蔽性强,可全天时、全天候工作,因此,广泛应用于军

收稿日期: 2024-12-18

作者简介: 李燕(1983—),女,山东泰安人,副教授,研究方向为计算机视觉。E-mail: yanlichina@163.com

引用格式: 李燕,曲毅,胡健生. 基于 YOLOv8 的红外无人机小目标检测研究[J]. 空军工程大学学报, 2025, 26(3): 106-111. LI Yan, QU Yi, HU Jiansheng. Research on Small Target Detection of Infrared UAV Based on YOLOv8[J]. Journal of Air Force Engineering University, 2025, 26(3): 106-111.

用和民用领域。

对红外目标进行检测分析的技术有传统方法和基于深度学习的方法。与传统方法相比,深度学习由于能够学习到更丰富的特征表示,在复杂场景下具有更好的鲁棒性,实时性强等优势,现已成为国内外研究的热点,代表算法有 Faster R-CNN^[3], SSD^[4] 以及 YOLO^[5-11] 系列算法等。其中, YOLO 系列算法引入 Anchor 机制,将目标检测的预测和回归同时进行,减少了候选区域提取过程,是常用的一种实时目标检测算法。近些年来,很多研究者对 YOLO 算法如何应用到红外无人机检测做了研究。文献[12]提出一种多通道多尺度特征融合算法,但是对弱小目标的识别难度依然很大。文献[13]提出一种改进的 YOLOv8_SG 算法,通过构建小目标检测层、引入注意力机制使算法能够融合更深层特征,但数据集不够多样,只展现了云天背景。文献[14]通过构建多尺度融合特征金字塔机制,引入了边界框损失函数,挖掘特征空间信息,加快模型的收敛速度,但降低网络模型的参数量、浮点运算量等指标,更好平衡检测精度和处理速度是需要考虑的问题。文献[15]将红外与可见光图像融合,构建了一个通用的跨模态配准框架,提供更为广泛的特征互补信息,但是图像配准方法较为复杂。尽管研究者在提高检测精度或实时性方面做出了努力,但复杂环境下对弱小目标的检测仍具有挑战性。YOLOv8 在保持高精度的同时,实现了较高的检测速度;模型可扩展性强,易于部署和应用;在各种复杂场景下表现出较高的鲁棒性和稳定性。因此,本文以 YOLOv8 为基础模型,提出一种改进模型,在主干网络引入 SK 注意力机制(selective kernel networks)^[16],每个神经元自适应调整卷积核的大小,增强无人机表征信息;构建小目标检测层,注重小目标特征的提取能力,提高模型检测精度;在检测头使用深度可分离卷积(depthwise separable convolution, DWConv)^[17]代替标准卷积,提高精度的同时使模型更加轻量化。

1 YOLOv8 网络结构

YOLOv8 是一种基于全局图像信息进行目标预测的单阶段检测算法^[18],可支持物体检测、实例分割、图像分类。其网络结构主要包含 Backbone 主干网络、Neck 颈部网络和 Head 头部网络 3 个部分,如图 1 所示。

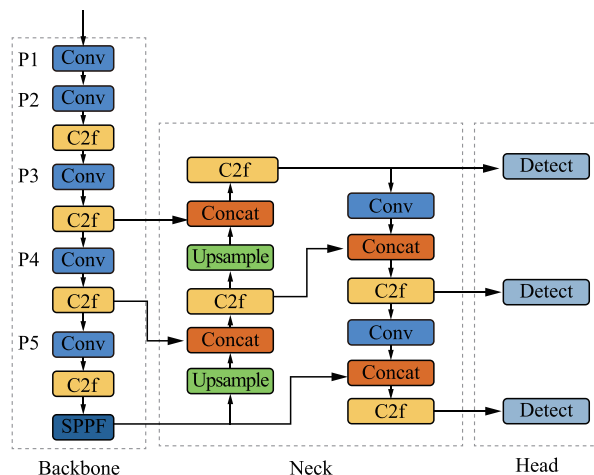


图 1 YOLOv8 网络结构

Fig. 1 YOLOv8 network structure

Backbone 主干网络的作用是从输入图像中提取特征,借鉴了 CSPDarkNet-53 网络结构^[19],通过 5 个降采样阶段提取不同的尺度特征。此外,通过采用 C2f 结构替代 YOLOv5 中的 C3 结构^[9],增强了梯度流动,提升了模型性能。Neck 颈部网络使用了特征金字塔 FPN^[20] 和路径聚合网络 PANet^[21],通过结合 FPN-PANet 结构和 C2f 模块,整合了不同尺度的特征图,融合了浅层和深度信息,增强了特征表达能力。Head 头部网络采用解耦头结构,将分类和检测分离。

2 改进的 YOLOv8 网络结构

在对红外无人机小目标检测时,能够获得的纹理信息较少,为了实现复杂背景下对红外小型无人机高效检测的目的,本文采用 YOLOv8 作为基础模型,从主干网络、颈部网络及头部网络的 Detect 模块 3 个方面进行优化,红色虚线为主要改动模块。改进后的模型结构如图 2 所示。

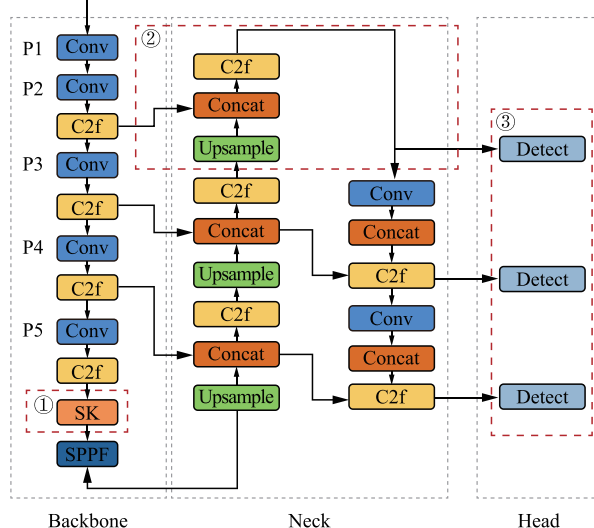


图 2 改进的 YOLOv8 网络结构

Fig. 2 Improved YOLOv8 network structure

2.1 注意力机制

对于某个特定任务,标准卷积神经网络的卷积核是固定的,感受野较小,卷积过程中只能得到局部特征信息。因此容易出现漏检、误检的情况。为此,本文引入 SK 注意力机制,将其嵌入到 Backbone 主干网络 4 个 C2f 层后,动态调整卷积核大小,加强网

络对目标的特征表达能力。

SK 注意力机制网络结构如图 3 所示,通过选择性内核单元 SK 动态调整卷积核大小,根据输入信息尺度自适应调节感受野的大小,以适应不同尺度的目标。SK 网络主要包括 Split、Fuse 和 Select 3 个操作。

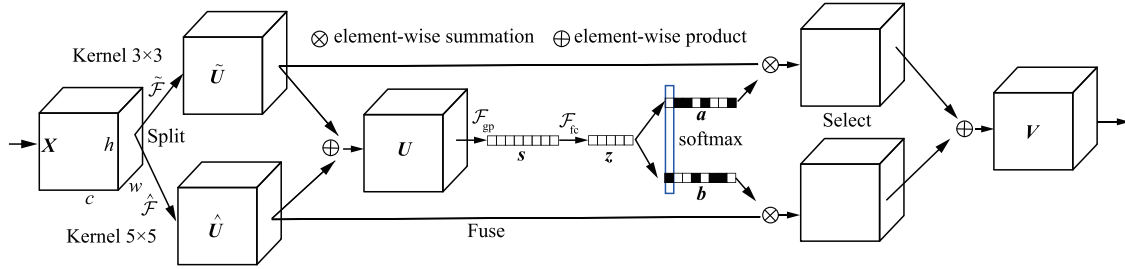


图 3 SK 注意力机制网络结构

Fig. 3 SK attention mechanism network structure

对于输入特征图 $X \in \mathbb{R}^{H' \times W' \times C'}$, 维度是 $[C', H', W']$, Split 算子产生多条不同核大小的路径, 图 3 中的模型只设计了 2 个不同大小的卷积核。经过分组/深度卷积、BN 批处理归一化、ReLU 函数进行变换 $\tilde{F}$ 、 $\hat{F}$, 分别得到相应的特征图 $\tilde{U}$ 、 $\hat{U}$, 表示为:

$$\begin{aligned} \tilde{F}: X &\rightarrow \tilde{U} \in \mathbb{R}^{H \times W \times C} \\ \hat{F}: X &\rightarrow \hat{U} \in \mathbb{R}^{H \times W \times C} \end{aligned} \quad (1)$$

Fuse 运算符聚合来自多个路径的信息, 得到如下特征图 U :

$$U = \tilde{U} + \hat{U} \quad (2)$$

为获得用于选择权重的全局和综合表示, 对 U 进行全局平均池化得到通道统计值 s 嵌入全局信息, 其中 s 的第 c 个元素:

$$s_c = \mathcal{F}_{gp}(U_c) = \frac{1}{H \times W} \sum_{i=1}^H \sum_{j=1}^W U_c(i, j) \quad (3)$$

Select 操作符根据选择权重聚合不同大小内核的特征图, 最后输出的特征图 V :

$$V_c = a_c \cdot \tilde{U}_c + b_c \cdot \hat{U}_c \quad (4)$$

式中: $V = [V_1, V_2, \dots, V_c]$; $V_c \in \mathbb{R}^{H \times W}$; a_c, b_c 为不同核的注意力权重值。

2.2 小目标检测层

YOLOv8 的原始结构有 3 个检测头 (P3、P4、P5), 从上到下代表对小尺度、中尺度、大尺度的检测能力。很多无人机在高速低空飞行时, 红外辐射强度弱, 信噪比低, 目标与背景之间的对比度较差, 很容易淹没在背景噪声中, 因此检测时很难获得纹理信息, 而不断地卷积操作容易导致特征被淹没。以 P3 检测头为例, 假设原有的输入图像尺寸为 $[640, 640]$, 经过 8 倍下采样后特征层大小变为 $[80, 80]$, 假设无人机的尺寸大小为 $[32, 32]$, 经过 P3 层下采样变为 $[4, 4]$ 大小的特征图, 小目标的特征信息

和位置信息逐层减少。

考虑本文的研究对象为小目标, 为了提高对目标的特征提取能力, 抑制背景的影响, 构建了小目标检测层, 如图 2 红色虚线标注②处。P2 层处于网络的较浅层, 具有更高的分辨率, 较高的分辨率能够提供更多的空间信息, 使模型能够捕捉到更多的细粒度特征, 有助于检测小目标, 同时增加 P2 检测头以提高对小目标的检测能力。但是, P2 检测头增加后, 原有的网络结构变为 (P2、P3、P4、P5) 4 个检测头, 参数量增加, 考虑到本文的研究目标无人机尺寸均小于 $[32, 32]$, P5 检测头对这类目标的特征提取能力有限, 为减少模型的复杂性, 删除 P5 检测头, 将原有的 YOLOv8 结构变为 (P2、P3、P4) 3 个检测头。

2.3 深度可分离卷积

在 Detect 模块, 预测每个锚点的位置信息和类别信息时, 使用 3×3 深度可分离卷积代替每个分支中的标准 3×3 卷积, 在保证模型准确度的同时实现轻量化, 如图 4 所示。

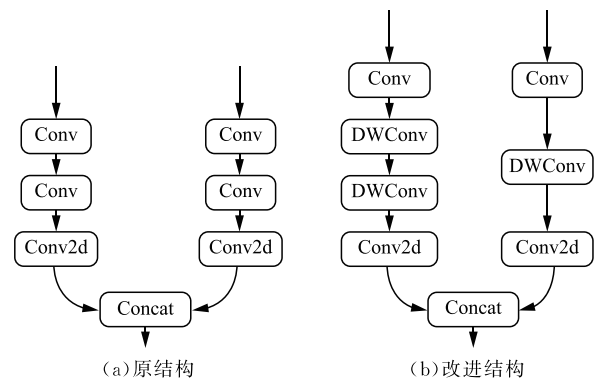


图 4 检测头结构

Fig. 4 Detection head structure

深度可分离卷积是一种常用的卷积操作。相较

于标准卷积,它由深度卷积和逐点卷积 2 部分组成。深度卷积是一种通道之间独立进行卷积操作的方式,主要用于捕捉输入数据的空间信息。在深度卷积中,每个输入通道都与 1 个单独的滤波器进行卷积操作,这意味着每个输入通道都会生成 1 个对应的输出通道。但该操作不能有效地利用同一空间位置处不同通道的特征信息。因此,需要逐点卷积来组合这些特征图以生成新的特征图。逐点卷积可以看作是在输入数据的通道维度上进行的卷积操作,不涉及空间信息,用于将深度卷积生成的各个通道的特征图进行线性组合。逐点卷积的输出通道数量可以根据需要调整,控制输出特征图的深度。

深度可分离卷积只有 1 个权重矩阵对应 1 个通道,通过逐个处理通道,保留了更多的空间信息,提高了特征的局部感知能力。同时,模型更加轻量化,参数量约为 3.1 M,计算量大大减少,FLOPs 值由原来的 8.1 GFLOPs 降至 7.7 GFLOPs。

3 实验结果与分析

本实验使用的硬件平台包括 18vCPU AMD EPYC 9754 128-Core Processor、RTX 4090D(24 GB) GPU、60 G 内存,python 版本为 3.8,CUDA11.3。实验所用训练数据集来自 CVPR Anti-UAV 无人机挑战赛提供的无人机红外视频,经过视频分帧处理筛选得到 3 160 张图像。训练集、验证集和测试集按照 7:2:1 的比例进行划分,训练轮数设定 100 epochs。

3.1 评估参数

本文采用精确率、召回率、mAP50、mAP50-95 作为评价模型的性能指标。精确率表示在所有被模型预测为正例的样本中,实际为正例的比例。召回率表示在所有实际为正例的样本中,模型成功预测为正例的比例。精确率、召回率的数学表达式分别为:

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}} \tag{5}$$

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}} \tag{6}$$

式中:TP 为将正类预测为正类的数量;FP 为将负类预测为正类的数量;FN 为将正类预测为负类的数量。

mAP 表示在不同 IoU 值下所有类别的平均检测精度,mAP50 是指 IoU 阈值为 0.5 时模型的平均精度。在精确、召回曲线(P-R)上绘制精确性和召回率,利用式(7)计算曲线下的面积 AP,平均精

度 mAP 的数学表达式为:

$$\text{AP} = \int_0^1 P(R) \text{d}R \tag{7}$$

$$\text{mAP} = \frac{1}{N} \sum \text{AP} \tag{8}$$

式中:N 为训练数据集中的总类别数。

3.2 结果分析

3.2.1 模型对比

本文选用 YOLOv8n 为基础模型,将基础模型和改进之后的模型分别进行了实验,通过各项指标对比,验证改进模型的检测性能提升效果,实验结果如表 1 所示,改进后的模型精度较基础模型有明显提升,精确率、召回率、mAP50、mAP50-95 分别提升了 5.3%、8.1%、9.1%、21.1%。

表 1 不同模型效果对比
Tab. 1 Effect comparison of various algorithms

模型	精确率/ %	召回率/ %	mAP50/ %	mAP50-95/ %
YOLOv8n	81.0	76.0	81.8	44.9
本文模型	86.3	84.1	90.9	66.0

3.2.2 消融实验

为了评估各个模块互相组合对于基础模型的价值性,本文设计了消融实验,实验结果如表 2 所示,其中,√ 表示加入该模块,否则表示未加入该模块。

表 2 消融实验结果
Tab. 2 Results of ablation experiment

SK 注意力 机制	小目 标检 测层	深度可 分离 卷积	精确率/ %	召回率/ %	mAP50/ %	mAP50-95/ %
			81.0	76.0	81.8	44.9
√			83.6	77.3	84.1	47.3
	√		83.8	80.3	87.9	62.3
		√	83.6	79.8	83.9	48.6
√	√	√	86.3	84.1	90.9	66.0

从表 2 中可以看出,模型在精度方面有了不同程度的提升。加入 SK 注意力机制模块,精确率由原来 81% 提升至 83.6%,召回率由 76% 提升至 77.3%,mAP50 由 81.8% 提升至 84.1%,mAP50-95 由 44.9% 提升至 47.3%,其中,精确率和 mAP50 较基础模型提升幅度较大。添加 P2 检测层,更加关注小目标的特征提取能力,结果显示性能指标均得到了有效提升,尤其是 mAP50-95 提升了 17.4%。而修改检测头结构后,精确率、召回率的提升效果更加明显。组合使用上述 3 种结构,精确率、召回率、mAP50、mAP50-95 精度最高。

3.2.3 复杂背景环境下的实验结果

为了直观展示小目标的检测性能,本文选择了

5 处场景进行模型验证,如图 5 所示,分别是天空背景(772. bmp),树林背景(8. bmp),山脉、树林下的复杂背景(2923. bmp),山脉、树林、楼宇下的复杂背景(2947. bmp),田地背景(2543. bmp),图 5(a)为原始 YOLOv8 检测图,图 5(b)为本文模型检测图。

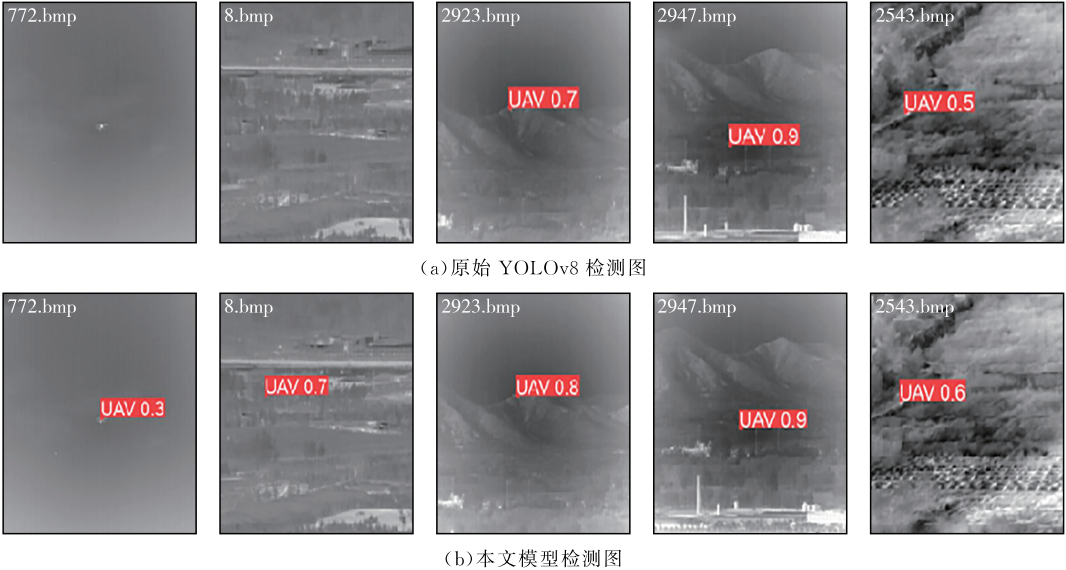


图 5 复杂背景环境下检测结果
Fig. 5 Detection results in complex background environments

4 结语

针对红外无人机小目标,目标与背景之间的对比度较差,检测时很难获得纹理信息等问题,提出基于 YOLOv8 的改进方法,该方法使用注意力机制获得更为丰富的局部特征信息,同时构建小目标检测层关注网络较浅层提取细粒度特征,改进检测头提高检测精度的同时更加轻量化。实验结果表明,改进后的模型提高了检测精度,减少了漏检率,能很好地适应复杂环境下的无人机检测。但论文中的模型是在硬件配置较高的实验环境下运行的,考虑到实际应用及经济成本,未来的研究重点为轻量化模型的部署应用。

参考文献

[1] 陈小龙,陈唯实,饶云华,等. 飞鸟与无人机目标雷达探测与识别技术进展与展望[J]. 雷达学报,2020,9(5):803-827.
CHEN X L, CHEN W S, RAO Y H, et al. Progress and Prospects of Radar Target Detection and Recognition Technology for Flying Birds and Unmanned Aerial Vehicles[J]. Journal of Radars, 2020, 9(5): 803-827. (in Chinese)

[2] 孙晓亮,郭良超,张文龙,等. 复杂背景下红外弱小运动目标检测半仿真数据集[J]. 中国科学数据,2024,9

从图中可以直观看出,图 5(a)中有 2 幅图像未检测出无人机位置,而在图 5(b)中无人机被检出,有效减少了漏检率。并且相较于原始模型,本文模型在复杂的背景环境下能有效减少环境的噪声干扰,表现出更好的检测置信度。

(3):315-331.
SUN X L, GUO L C, ZHANG W L, et al. A Dataset of Semi-Synthetic Detection for Small Infrared Moving Targets under Complex Backgrounds[J]. China Scientific Data, 2024, 9(3): 315-331. (in Chinese)

[3] 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.

[4] LIU W, ANGUELOV D, ERHAN D, et al. SSD: Single Shot MultiBox Detector[C]//Computer Vision- ECCV 2016. Cham: Springer International Publishing, 2016: 21-37.

[5] REDMON J, DIVVALA S, GIRSHICK R, et al. You Only Look Once: Unified, Real-Time Object Detection [C]//2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Las Vegas, NV: IEEE, 2016: 779-788.

[6] REDMON J, FARHADI A. YOLO9000: Better, Faster, Stronger[C]//2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Honolulu, HI: IEEE, 2017: 6517-6525.

[7] REDMON J, FARHADI A. YOLOv3: An Incremental Improvement[EB/OL]. (2018-04-08) [2024-11-18]. <https://arxiv.org/abs/1804.02767v1>.

[8] BOCHKOVSKIY A, WANG C Y, LIAO H M. YOLOv4: Optimal Speed and Accuracy of Object De-

- tection [EB/OL]. (2020-04-23) [2024-11-18]. <https://arxiv.org/abs/2004.10934v1>.
- [9] JOCHER G. Ultralytics YOLOv5: Version 7.0 [EB/OL]. (2020-05-01) [2024-09-23]. <https://github.com/ultralytics/yolov5>.
- [10] LI C, LI L, JIANG H, et al. YOLOv6: A Single Stage Object Detection Framework for Industrial Applications [EB/OL]. (2022-06-07) [2024-09-23]. <https://arxiv.org/abs/2209.02976>.
- [11] WANG C Y, BOCHKOVSKIY A, LIAO H M. YOLOv7: Trainable Bag of Freebies Sets New State of the Art for Real Time Object Detectors [C]//2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Vancouver, BC: IEEE, 2023: 7464-7475.
- [12] 王国刚, 孙召进, 刘云鹏. J-MSF: 一种新的多通道多尺度红外弱小目标检测算法 [J]. 红外与激光工程, 2022, 51(5): 136-145.
- WANG G G, SUN Z J, LIU Y P. J-MSF: A New Infrared Dim and Small Target Detection Algorithm Based on Multi-Channel and Multiscale [J]. Infrared and Laser Engineering, 2022, 51(5): 136-145. (in Chinese)
- [13] 岳明凯, 权康男, 张骢, 等. 基于改进YOLOv8的红外小目标检测算法研究 [J]. 红外技术, 2024, 46(11): 1286-1292.
- YUE M K, QUAN K N, ZHANG C, et al. Research on Infrared Small Target Detection Algorithm Based on Improved YOLOv8 [J]. Infrared Technology, 2024, 46(11): 1286-1292. (in Chinese)
- [14] 陈海永, 刘登斌, 晏行伟. 基于IDOU-YOLO的红外图像无人机目标检测算法 [J]. 应用光学, 2024, 45(4): 723-731.
- CHEN H Y, LIU D B, YAN X W. Infrared Image UAV Target Detection Algorithm Based on IDOU-YOLO [J]. Journal of Applied Optics, 2024, 45(4): 723-731. (in Chinese)
- [15] LUO Y, CHA H, ZUO L, et al. General Cross Modality Registration Framework for Visible and Infrared UAV Target Image Registration [J]. Scientific Reports, 2023, 13: 12941.
- [16] LI X, WANG W H, HU X L, et al. Selective Kernel Networks [C]//2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Long Beach, CA: IEEE, 2019: 510-519.
- [17] CHOLLET F. Xception: Deep Learning with Depth-wise Separable Convolutions [C]//2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Honolulu, HI: IEEE, 2017: 1800-1807.
- [18] VARGHESE R, SAMBATH M. YOLOv8: a Novel Object Detection Algorithm with Enhanced Performance and Robustness [C]//2024 International Conference on Advances in Data Engineering and Intelligent Computing Systems (ADICS). Chennai, India: IEEE, 2024: 1.
- [19] WANG C, LIAO H, WU Y, et al. CSPNet: A New Backbone that Can Enhance Learning Capability of CNN [C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops. Piscataway, NJ: IEEE, 2020: 390-391.
- [20] 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). Honolulu, HI: IEEE, 2017: 936-944.
- [21] LIU S, QI L, QIN H F, et al. Path Aggregation Network for Instance Segmentation [C]//2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Salt Lake City, UT: IEEE, 2018: 8759-8768.

(编辑: 杜娟)