

DOI: 10.19650/j.cnki.cjsi.J2413551

BG-YOLO: 复杂大视场下低慢小无人机目标检测方法

王迎龙, 孙 备, 丁 冰, 卜德森, 孙晓永
(国防科技大学智能科学学院 长沙 410072)

摘 要:针对现有无人机目标检测模型在体积、计算资源需求以及小目标检测效果方面的不足,提出了一种改进的无人机目标检测算法 BG-YOLO。该算法基于 YOLOv8,通过在高分辨率特征层添加检测头,有效减少了图像下采样过程中的信息丢失,显著提升了模型对小目标的检测能力。同时,引入 Biformer 注意力机制可以精准捕捉图像的远程依赖关系,增强模型对不同尺度目标的感知能力。此外,NWD 损失函数的引入解决了传统损失函数在小目标检测中对位置偏差敏感的问题,显著提高了模型的鲁棒性。基于 GhostNetV2 的模型轻量化则通过替换传统卷积模块,在减少模型参数和计算量的同时,保持了模型的检测精度。实验结果表明,BG-YOLO 在 Det-Fly 数据集上相比 YOLOv8 的 $mAP@0.5$ 提高了 10.3%,参数量减少了 33.18%,而与 YOLOv9 相比提高了 7.9%。此外,该算法在自采集数据集上也表现出色,对天空、山地、建筑等不同场景的低慢小目标分别实现了 96.2%、88.1% 和 86.2% 的平均精度,检测速度分别为 150.36、128.21、112.53 fps,实现了高检测精度和高检测速度的要求。综上所述,BG-YOLO 通过检测头设计、注意力机制引入、损失函数改进以及模型轻量化,显著提升了对低慢小无人机目标的检测精度和实时性,具有广阔的应用前景。

关键词: YOLOv8; Biformer; GhostNetV2; 低慢小无人机目标检测; 复杂大视角场景

中图分类号: TP391.4 **文献标识码:** A **国家标准学科分类代码:** 510.4050

BG-YOLO: A low-altitude slow-moving small UAV targets detection method in complex large field of view

Wang Yinglong, Sun Bei, Ding Bing, Bu Desen, Sun Xiaoyong

(College of Intelligence Science and Technology, National University of Defense Technology, Changsha 410072, China)

Abstract: This article proposes an improved UAV target detection algorithm, BG-YOLO, to address the limitations of existing UAV detection models in terms of model size, computational resource requirements, and the detection performance of small targets. Based on YOLOv8, BG-YOLO adds detection heads to the high-resolution feature layers, effectively reducing information loss during image downsampling and significantly enhancing the model's ability to detect small targets. The introduction of the Biformer attention mechanism enables precise capture of long-range dependencies in images, thereby strengthening the model's perception of targets at different scales. Additionally, the incorporation of the NWD loss function overcomes the issue of traditional loss functions being sensitive to positional deviations in small target detection, thereby significantly improving the model's robustness. The model's lightweighting based on GhostNetV2 replaces traditional convolutional modules, reducing model parameters and computational load while maintaining detection accuracy. Experimental results show that BG-YOLO achieves a 10.3% improvement in $mAP@0.5$ on the Det-Fly dataset compared to YOLOv8, with a 33.18% reduction in model parameters, and a 7.9% improvement compared to YOLOv9. Moreover, on the self-collected dataset, BG-YOLO demonstrates excellent performance in detecting low, slow, and small targets in various scenarios, including sky, mountain, and urban backgrounds, achieving average precisions of 96.2%, 88.1% and 86.2%, respectively, with detection speeds of 150.36, 128.21 and 112.53 fps. These results meet the real-time requirements of high detection accuracy and speed. In summary, BG-YOLO significantly enhances the detection accuracy and real-time performance for low, slow, and small UAV targets through the design of detection heads, incorporation of attention mechanisms, refinement of the loss function, and model lightweighting, thereby offering broad application prospects.

Keywords: YOLOv8; Biformer; GhostNetV2; detection of low, slow, small UAV targets; complex and large-viewing scenes

0 引言

“低慢小”是“低空慢速小目标”的简称,通常将飞行高度小于1 000 m、体型较小、飞行速度低于100 km/h的旋翼和固定翼无人机称为“低慢小”无人机^[1]。近年来,以“低慢小”为特点的民用无人机技术发展迅猛,这使其被广泛应用于农业监测、环境监测、物流配送、新闻报道、影视拍摄、紧急救援等多个领域^[2]。民用无人机的广泛应用也导致了无人机的“黑飞”现象日益严重,进而引发一系列公共安全和侵犯隐私问题^[3]。因此,对于以“低慢小”为特点的民用无人机的监管探测技术的研究迫在眉睫。

目前,无人机的探测手段主要有雷达探测、射频探测、声学探测和光电探测等方式,常用的探测方式为光电探测。基于光电的传统无人机目标检测方法,首先需要提取出目标的特征信息,比如颜色特征、纹理特征以及尺度不变特征等,接着通过机器学习方法对这些特征进行检测和分类。深度学习近年来发展非常迅速,并在目标检测领域取得了很大的进展。相较于传统方法,基于深度学习的检测方法的精度和实时性均处于领先地位,且基于深度学习的无人机目标检测技术在反无人机光电探测的实际应用中取得了许多成果。但是目前大多算法只适用于简单场景的通用大目标检测,在一些复杂场景和挑战条件下难以有效。在实际应用中,小视场观察区域极为狭窄,难以有效发现目标,而采用大视场观测则意味着在一幅3 840 pixels×2 160 pixels的高清图像中探测像素尺寸仅有几十个像素的低慢小目标,这对检测算法进一步提出了很大的挑战^[4]。此外,复杂的环境可能包含大量干扰物体,动态的光照变化等因素,无人机在复杂环境中飞行时,这些干扰因素都可能影响到目标检测算法的性能,会给检测增加难度,增加了误检测的可能性。

为应对此问题,本研究以YOLOv8(you only look once version 8)为基础,首先增加了一个高分辨率特征层的检测头,减少图像下采样过程中的信息丢失;其次,在骨干网络中引入了双层路由注意力机制(vision transformer with Bi-Level routing attention, Biformer),实现动态稀疏的注意力查询,保持高计算效率并捕捉长距离依赖关系^[5];然后,在回归损失函数中加入归一化瓦瑟斯坦距离损失函数(normalized wasserstein distance loss, NWD loss),以提升小目标检测性能^[6]。经过上述改进可提升模型在特定场合的检测性能,但也导致模型参数量和复杂度上升,增加计算成本,不利于移动设备的部署。因此,又基于GhostNetV2网络,使用Ghost模块替换原有卷积模块,在保证模型精度的前提下显著减少模型参数和学习成本^[7]。改进后的模型称之为BG-YOLO(biformer-

ghostnetv2 YOLO with nested winged distance loss and small object detection head)。实验结果显示,BG-YOLO在Det-Fly数据集上相比YOLOv8的 $mAP@0.5$ (mean average precision at intersection over union equals 0.5)提高了10.3%,参数减少33.18%,而与YOLOv9相比提高了7.6%。Det-Fly数据集^[8]上的实验验证了改进方法的有效性。

1 低慢小无人机检测相关工作

目前,无人机的探测手段主要有雷达探测、射频探测、声学探测和光电探测等方式^[9]。其中,雷达探测由于设备成本高昂、辐射较强以及对于小目标回波效应不明显等问题,不易单独应用^[10];基于射频的探测手段通常受到探测距离和探测波段的限制,不易广泛应用^[11];基于声波的探测手段通常受到探测距离和噪声干扰的限制,也不易广泛应用^[12]。几种常用探测方式的比较如表1所示。

表1 不同无人机探测方式比较

Table 1 Comparison of different drone detection methods

探测方式	优点	缺点
雷达探测	远距离全天候探测,可探测大多数无人机类型	成本高,存在近距离盲区,易受杂波干扰
光电探测	成本低,易于部署,成像直观清晰	易受环境光线影响,探测距离较近
射频探测	远距离探测,全天候工作	易受其他射频信号干扰,鲁棒性差
声波探测	成本低,易于部署	不具有鲁棒性,探测距离有限

目前比较流行的探测方式通常为雷达探测、射频探测、光电探测等方式的一种或几种结合进行协同探测^[13]。如图1所示,为常用的无人机侦测反制设备,该设备结合雷达、光电和无线电探测“低慢小”无人机目标,在无人机侦测系统发现“低慢小”目标后,目标的信息被传送到无人机反制设备进行反制。这种综合的探测方式既能利用雷达进行远距离探测,又能利用光电探测捕捉更多目标细节信息,可以显著提升目标识别的准确性。本研究主要聚焦于对低慢小无人机的光电探测方式的研究。

1.1 图像小目标检测

目前,基于深度学习的目标检测方法的精度和实时性均处于领先地位,这里主要对基于深度学习的单帧图像目标检测的研究工作进行介绍。深度学习中流行的目

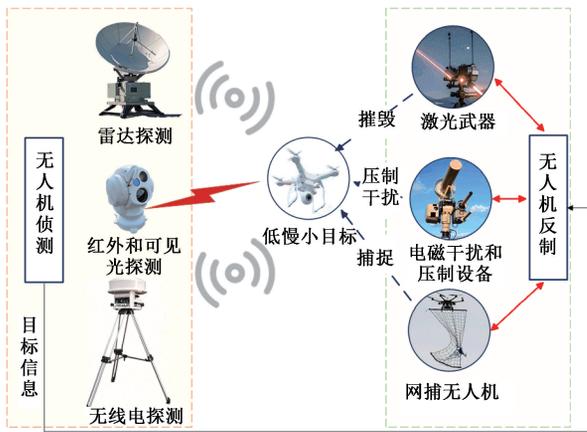


图1 典型无人机侦测反制系统结构

Fig.1 Typical drone detection and countermeasure system structure

标检测方法有两类,即以 R-CNN (region-based convolutional neural networks) 系列^[14]为代表的两阶段目标检测算法和以 YOLO (you only look once) 系列^[15]为代表的单阶段目标检测算法。

单阶段检测算法通过直接将输入图像送入卷积神经网络,省略候选区域提取步骤,实现快速检测。2015年 Redmon 等^[15]提出的 YOLO 算法将图像划分为网格,每个网格负责目标检测,并回归预测位置和大小,该工作奠定了以 YOLO 系列为代表的单阶段目标检测基础。相对而言,R-CNN 通过提取候选区域进行分类和回归,奠定了双阶段目标检测的基础。

由于单阶段检测算法实时性好,许多学者基于 YOLO 系列算法进行改进以更好地应对复杂场景下的小目标检测任务,常用的改进方法主要有添加小目标检测头、添加注意力机制、改进损失函数以及改进特征融合等。徐光达等^[16]基于 YOLOv5,在 4 倍下采样的 160 pixels×160 pixels 大小的特征图中添加检测头,实现了模型对无人机航拍图像中的小目标检测性能的显著提升;张寅等^[17]基于 YOLO 模型,通过添加增强的通道注意力机制(enhanced-efficient channel attention, E-ECA)与空间注意力模块(spatial attention mechanism, SAM)所组成的级联注意力机制(cascade attention mechanism, ESM),使模型可更精确地捕获小目标的特征信息;张宏群等^[18]基于 YOLOv5 模型,将检测头的分类任务与回归任务解耦,使用 Alpha-IoU (alpha-intersection over union) 优化损失函数定义,提升了模型对无人机航拍影像中小目标的检测效果。

此外,Qin 等^[19]基于 YOLOv7 模型,通过优化特征融合、增加小目标检测层和改进损失函数等方法,提高了微小目标检测能力;冯哲等^[20]基于 YOLOv8 模型,通过添

加小目标检测层、添加多头自注意力机制、添加多尺度特征融合模块,显著增强了模型对弱小目标的检测能力;Zhu 等^[21]基于 YOLOv5 模型,通过添加小目标检测头和 CBAM (convolutional block attention module) 注意力机制等方法,提升了对不同尺度目标的检测能力,并在 VisDrone2021 挑战赛中获第 5 名;Liu 等^[22]基于 YOLOv5l 模型,添加 ECA (efficient channel attention) 注意力机制和 Ghost 模块,实现了小目标检测性能的提升和模型的轻量化。

总的来说,单阶段检测算法因其快速处理能力而被广泛用于实时性要求高的场合。但是单目标检测算法在特定场合可能存在精度不高的问题,对此,可以通过添加检测头、添加注意力机制、改进损失函数和优化特征融合等方法,使模型更好地适应不同的任务场景。

1.2 检测模型轻量化

模型轻量化旨在减少计算复杂度和内存消耗,同时保持检测性能,主要手段包括模型压缩与剪枝、深度可分离卷积、网络结构优化、跨层特征融合和硬件加速。模型压缩与剪枝技术,如权重剪枝、通道剪枝、量化和蒸馏,能够有效减少模型大小和计算需求。引入深度可分离卷积等轻量级结构可以降低参数数量和计算复杂度。郝晋渊等^[23]基于 YOLOv5 模型,通过全维动态卷积(omni-dimensional dynamic conv, ODConv)替代普通卷积块,使用空间到深度卷积(space to depth conv, SPDConv)来实现下采样过程,在保证准确率的情况下,显著降低了模型的参数数量和计算量;Cao 等^[24-25]对 YOLOv5 模型进行改进,通过构建基于 GhostNet 的骨干网络,设计新的无人机小目标检测头以及改进损失函数,实现了在参数数量和计算量显著降低的同时,模型的平均精度显著提高。

2 改进方法

2.1 方法概述

YOLOv8 目标检测算法模型由 Backbone、Neck 和 Head 这 3 部分构成。Backbone 借鉴了 CSP (cross stage partial) 和 ELAN (efficient layer aggregation networks) 结构,设计 C2f (CSP bottleneck with 2 convolutions and faster implementation) 实现进一步的轻量化的同时进一步提高了梯度流的丰富度。同时,该模型还沿用了 YOLOv5 中的 SPPF (spatial pyramid pooling-fast),并微调不同尺度模型,提升模型的鲁棒性。该模型的 Neck 部分参考了 PAN-FPN (path aggregation network-feature pyramid network) 结构,可以更有效地融合不同尺度的特征图。在检测头部分采用目前主流的解耦头结构(decoupled-head)和无锚框(anchor-free)方法。

针对现有模型对微小目标无人机检测效果不佳的问题,在 P2 高分辨率特征层中新增检测头,减少下采样过程的特征丢失;引入视觉 Transformer 架构的双层路由注意力机制 Biformer,添加至主干网络末端的 SPPF 模块,提升长距离依赖捕捉能力并保持高计算效率;在回归损失中加入 NWD 损失函数,增强

小目标检测的鲁棒性;基于 GhostNetV2,改进卷积模块,降低模型参数量和关键特征学习成本,同时保持模型的精度。改进的 BG-YOLO 结构如图 2 所示。其中, Biformer 模块、Ghost 模块和 GhostNetV2 的 Bottleneck 模块将分别在 2.3 和 2.5 节中进行详细介绍。

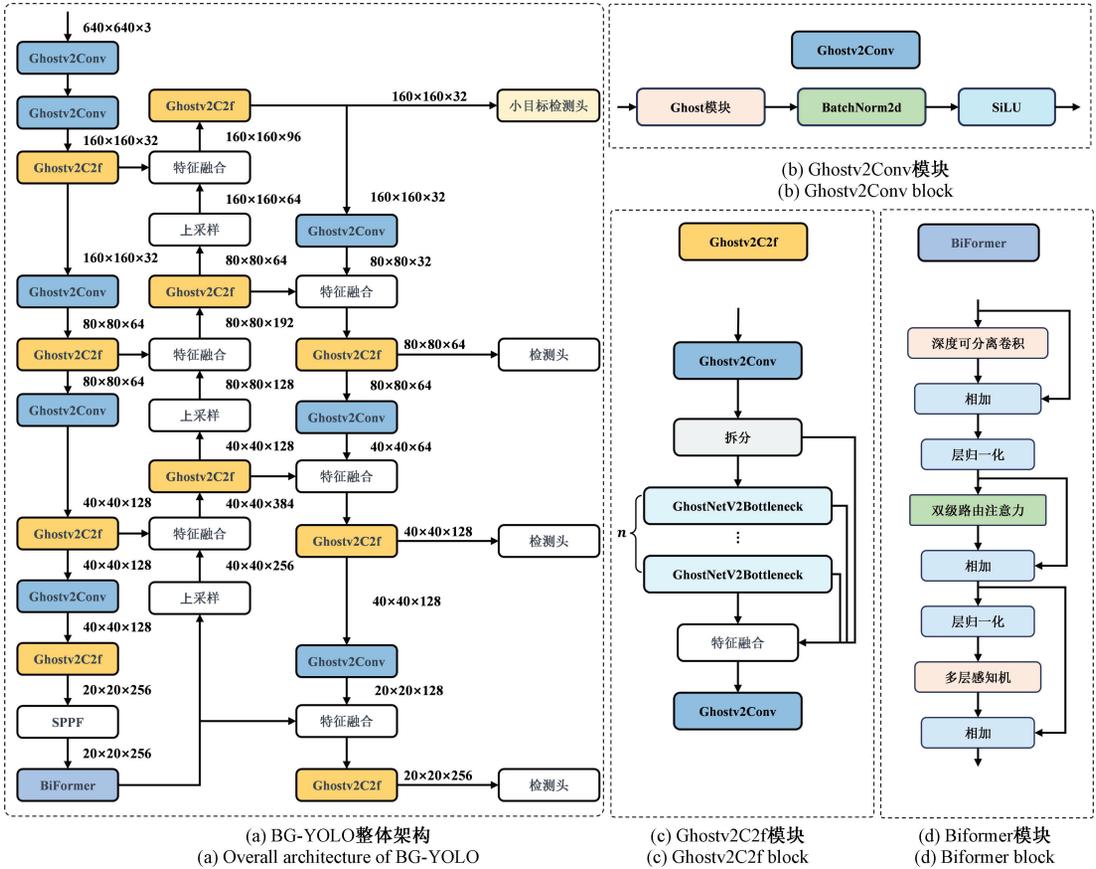


图 2 BG-YOLO 模型架构

Fig. 2 BG-YOLO model architecture

2.2 用于小目标检测的检测头

小尺寸无人机目标在一张图像中的像素占比很小,在经过多次下采样之后,其特征信息变得更加模糊,从而导致模型在处理微小的无人机目标时的定位与分类能力差,如图 3 所示。

因此,为了有效地检测这些微小目标,本研究设计了专门的检测头来增强模型对微小目标的感知能力。

YOLOv8 的 P2 特征层像素大小为 160 pixels × 160 pixels,具有高分辨率的特征信息,本研究从该特征层添加小目标检测头,充分利用小目标的特征信息,从而增强模型对微小目标检测的性能,添加的检测头如图 2(a) 所示。添加小目标检测头后,4 个检测头可以显著减少目标尺度变化造成的检测困难问题,微小目标的检测性能也大大提升,但是相对不足的是添加检测头后模型的计算和内存成本显著增加。

2.3 Biformer 注意力机制

Biformer 引入了双层路由注意力机制,将注意力分为全局和局部两个层次,以精准捕捉不同尺度的目标。全局注意力关注大目标或整体结构,而局部注意力则专



图 3 YOLOv8 中无人机目标下采样示例

Fig. 3 Downsampling example of drone targets in YOLOv8

注于小目标或细节信息。这种分层机制不仅提升了检测精度,还显著降低了计算复杂度,使得模型可以同时捕捉全局和局部特征信息,有效捕捉目标物体之间的关联性,保留特征提取过程中更细粒度的上下文信息,从而提高对小目标检测的性能。双层路由注意力机制的计算过程如算法1所示。

算法1 双层路由注意力机制的计算过程

输入:

特征图: $X \in \mathbf{R}^{H \times W \times C}$

步骤:

S1: 将特征图 X 分成 $S \times S$ 个非重叠区域, 得到: $X' \in$

$$\mathbf{R}^{\frac{H \times W \times C}{S^2} \times S^2};$$

S2: 使用全连接层对 X' 进行线性投影, 得到 Q, K 和 V , 并计算 Q, K 的均值得到 Q', K' ;

S3: 计算邻接矩阵以衡量不同区域之间的语义相似度:

$$A' = Q'(K')^T;$$

S4: 对矩阵 A' 进行过滤, 挑选出语义相似度最高的前 k 个区域, 得到: $I' = \text{topkIndex}(A')$;

S5: 对 K 和 V 进行过滤和收集, 得到 K^s 和 V^s :

$$K^s = \text{gather}(K, I'), V^s = \text{gather}(V, I')$$

S6: 对 Q, K^s 和 V^s 集中应用细粒度的 token-to-token 注意力, LCE 代表深度可分离卷积:

$$O = \text{Attention}(Q, K^s, V^s) + LCE(V)$$

整个计算过程如图4所示。双层路由注意力机制在粗糙区域过滤掉大部分不相关的键值对, 去除冗余信息, 只保留小部分路由区域, 并在这部分路由区域再应用注意力机制, 因此比直接在全局使用自注意力机制可以降低计算量。

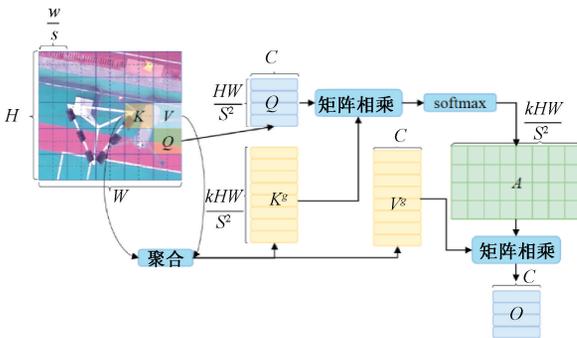


图4 双层路由注意力机制

Fig. 4 Dual-layer routing attention mechanism

BiFormer 结合双层路由注意力机制构建的 BiFormerBlock 结构如图2(d)所示。在开始时使用深度可分离卷积(depthwise convolution, DWConv)隐式编码输入的相对位置信息, 然后依次使用双层路由注意力机制

和多层感知机(multilayer perceptron, MLP)模块对输入的信息进行跨位置关系建模和逐位置嵌入。

2.4 基于NWD的损失函数改进

相较于之前的YOLO模型, YOLOv8采用解耦头和Anchor Free的策略, 其损失函数由分类损失和定位损失组成, 其中分类损失使用二元交叉熵损失 BCE Loss (binary cross entropy loss), 定位损失使用 VFL Loss (variFocal loss) 和 CIoU Loss (complete intersection over union) 作为回归损失。引入 IoU (intersection over union) 损失是为了消除训练和测试之间的性能差距, 然而当预测的边界框和真实的边界框之间没有重叠或者两个框相互包含的时候, IoU 损失无法为优化网络提供梯度。此外, IoU 对不同尺寸的物体敏感性差异很大。对于微小目标而言, 微小目标的位置偏差会导致 IoU 的显著下降。

为解决该问题, 本研究使用 NWD 损失函数。该方法一致地反映了模型检测到的物体分布之间的距离, 克服原有损失函数对不同尺寸的物体敏感性差异很大的问题。

具体来说, 对于水平边界框 $R = (cx, cy, w, h)$, 其中 (cx, cy) 、 w 和 h 分别表示中心坐标、宽度和高度。其内切椭圆的方程可以表示为:

$$\frac{(x - \mu_x)^2}{\sigma_x^2} + \frac{(y - \mu_y)^2}{\sigma_y^2} = 1 \quad (1)$$

其中, (μ_x, μ_y) 为椭圆的中心坐标, σ_x 和 σ_y 为沿 x 、 y 轴的半轴长度。 $\mu_x = cx, \mu_y = cy, \sigma_x = w/2, \sigma_y = h/2$, 二维高斯分布的概率密度函数为:

$$f(x|\mu, \Sigma) = \frac{\exp\left(-\frac{1}{2}(x - \mu)^T \Sigma^{-1}(x - \mu)\right)}{2\pi |\Sigma|^{\frac{1}{2}}} \quad (2)$$

其中, μ 和 Σ 表示高斯分布的坐标 (x, y) 、均值向量和协方差矩阵, 当:

$$(x - \mu)^T \Sigma^{-1}(x - \mu) = 1 \quad (3)$$

式(2)中的椭圆为二维高斯分布的密度轮廓, 因此, 水平边界框 $R = (cx, cy, w, h)$ 可以被建模成一个二维高斯分布:

$$\left\{ \begin{aligned} \mu &= \begin{bmatrix} c_x \\ c_y \end{bmatrix} \\ \Sigma &= \begin{bmatrix} \frac{w^2}{4} & 0 \\ 0 & \frac{h^2}{4} \end{bmatrix} \end{aligned} \right. \quad (4)$$

边界框 A 和 B 之间的相似度可以转换为两个高斯分布之间的分布距离。NWD 损失函数采用 Wasserstein 距离来进行衡量。

对于两个二维高斯分布 $\mu_1 = N(\mathbf{m}_1, \Sigma_1)$ 和 $\mu_2 = N(\mathbf{m}_2, \Sigma_2)$, μ_1 和 μ_2 之间的二阶 Wasserstein 距离定义为:

$$W_2^2(\mu_1, \mu_2) = \mathbf{m}_1 - \mathbf{m}_2^2 + \text{Tr}(\Sigma_1 + \Sigma_2 - 2(\Sigma_1^{1/2} \Sigma_1 \Sigma_2^{1/2})^{1/2}) \quad (5)$$

其中, $\|\cdot\|_F$ 为 Frobenius 范数, 代入两个目标边界框 $A = (cx_a, cy_a, w_a, h_a)$ 和 $B = (cx_b, cy_b, w_b, h_b)$, 式(5)可以进一步简化为:

$$W_2^2(N_a, N_b) = \left(\left[cx_a, cy_a, \frac{w_a}{2}, \frac{h_a}{2} \right]^T, \left[cx_b, cy_b, \frac{w_b}{2}, \frac{h_b}{2} \right]^T \right)_2^2 \quad (6)$$

最后对其进行归一化得到:

$$NWD(N_a, N_b) = \exp\left(-\frac{\sqrt{W_2^2(N_a, N_b)}}{C}\right) \quad (7)$$

其中, C 是与数据集密切相关的常数。

2.5 基于 GhostNetV2 的模型轻量化

通过模型轻量化,可以节省计算资源、提高模型检测的实时性,以及降低存储需求。本研究基于 GhostNetV2 轻量化网络结构,对 YOLOv8 的卷积模块进行改进,在牺牲模型较少的检测精度的情况下,大大减少了模型的参数和复杂程度。

首先来对 GhostNetV1 网络中的 Ghost 模块进行介绍。给定输入数据 $\mathbf{X} \in \mathbf{R}^{c \times h \times w}$, 其中 c 是输入通道的数量, h 和 w 是输入数据的高度和宽度,用于产生 n 个特征图的任意卷积层的操作可以公式化为:

$$\mathbf{Y} = \mathbf{X} * f + b \quad (8)$$

其中, $*$ 表示卷积操作, b 是偏差项, $\mathbf{Y} \in \mathbf{R}^{h' \times w' \times n}$ 是具有 n 个通道的输出特征图, $f \in \mathbf{R}^{c \times k \times k \times n}$ 是该层的卷积核。用 h' 和 w' 分别表示输出数据的高度和宽度, $k \times k$ 是卷积核 f 的核大小。在该卷积过程中,所需的计算量为:

$$\text{FLOPs} = n \times h' \times w' \times c \times k \times k \quad (9)$$

由此得到的计算量通常达到数十万。为了解决传统卷积操作的计算量大的问题, Ghost 模块应运而生。具体来说,使用初级卷积生成 m 个固有特征映射:

$$\mathbf{Y}' = \mathbf{X} * f' \quad (10)$$

其中, $\mathbf{Y}' \in \mathbf{R}^{h' \times w' \times m}$, $f' \in \mathbf{R}^{c \times k \times k \times m}$ 是使用的卷积核, $m \leq n$, 为简单起见省去偏置项。为了进一步获得所需的 n 个特征映射,对 \mathbf{Y}' 中的每个固有特征应用一系列简易的线性运算,以根据以下函数生成 s 个虚影特征:

$$\mathbf{y}_{ij} = \Phi_{i,j}(\mathbf{y}'_i), \quad i = 1, \dots, m, \quad j = 1, \dots, s \quad (11)$$

其中, \mathbf{y}'_i 是 \mathbf{Y}' 中的第 i 个本征特征图, 上述函数中的 $\Phi_{i,j}$ 是用于生成第 j 个重影特征图 \mathbf{y}_{ij} 的第 j 个 ($j \neq s$) 线性运算, 即 \mathbf{y}'_i 可以有一个或多个重影特征图。通过式(10),可以获得 $n = m \times s$ 个特征图 $\mathbf{Y} = [\mathbf{y}_{11}, \mathbf{y}_{12}, \dots, \mathbf{y}_{ms}]$, 并将其作为 Ghost 模块的输出特征。如图 5 所示, 其中的线性运算 Φ 表示简易操作, Φ 在每个通道上操作, 其计算成本远低于普通卷积。

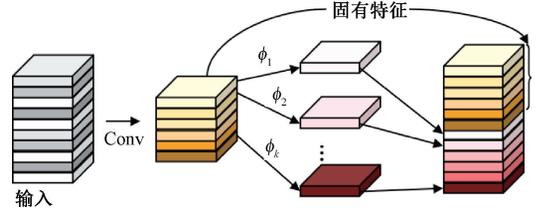


图 5 输出相同数量特征图的 Ghost 模块

Fig. 5 Schematic diagram of the Ghost module outputting the same number of feature maps

通过以上操作, Ghost 模块可以显著减少特征通道的冗余, 以降低计算成本。 Ghost 模块组成的 Ghostv2Conv 模块如图 2(b) 所示。

相较于 GhostNetV1, GhostNetV2 在原有 Ghost 模块的基础上添加了一种长距离注意力机制模块 (decoupled fully connected, DFC), 其不仅可以在常见硬件上快速执行, 也可以捕获长距离像素之间的依赖关系, 有效解决了原始 GhostNetV1 无法有效提取全局感受野的问题。 DFC 注意力机制的结构如图 6 所示。



图 6 DFC 注意力机制

Fig. 6 DFC attention mechanism

具体来说, DFC 注意力机制通过将全连接层分解为水平和垂直两个方向的全连接层, 来分别聚合 2D 特征图中沿着这两个方向的长距离信息。给定一个特征 $\mathbf{Z} \in \mathbf{R}^{H \times W \times C}$, 它可以被视为 $H \times W$ 个 token $\mathbf{z}_i \in \mathbf{R}^C$, 即 $\mathbf{Z} = \{\mathbf{z}_{11}, \mathbf{z}_{12}, \dots, \mathbf{z}_{HW}\}$ 。

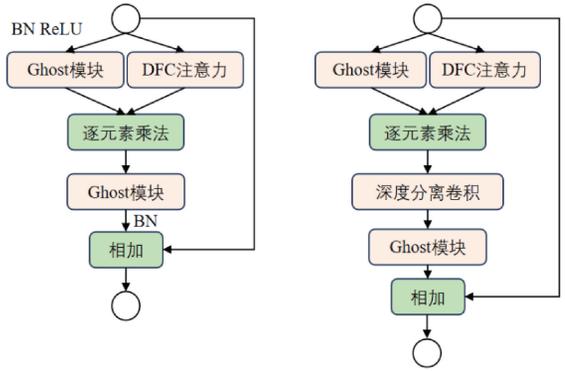
$$\mathbf{a}'_{hw} = \sum_{h'=1}^H F_{h, h'w} \odot \mathbf{z}_{h'w} \quad (12)$$

$$\mathbf{a}_{hw} = \sum_{w'=1}^W F_{w, hw'} \odot \mathbf{a}'_{hw'} \quad (13)$$

式(12)为水平方向的全连接过程, $F_{h, h'w}$ 是可学习的权重, \odot 表示逐元素乘法, $\mathbf{a}' = \{\mathbf{a}'_{11}, \mathbf{a}'_{12}, \dots, \mathbf{a}'_{HW}\}$ 是水平方向聚合后的特征。

式(13)是垂直方向的全连接过程, $F_{w, hw'}$ 是另一组可学习的权重, $\mathbf{a} = \{\mathbf{a}_{11}, \mathbf{a}_{12}, \dots, \mathbf{a}_{HW}\}$ 是最终的注意力图。

通过这两个步骤, DFC 注意力机制能够以较低的计算复杂度捕获特征图中的长距离依赖关系。整个 GhostNetV2 Bottleneck 的结构如图 7 所示。由 Ghostv2Conv 和 GhostNetV2 Bottleneck 组成的 Ghostv2C2f 结构如图 2(c) 所示。



(a) 步长=1的GhostNetV2 Bottleneck (b) 步长=2的GhostNetV2 Bottleneck
 (a) GhostNetV2 Bottleneck with a stride of 1 (b) GhostNetV2 Bottleneck with a stride of 2

图 7 GhostNetV2 Bottleneck 结构

Fig. 7 GhostNetV2 Bottleneck structure

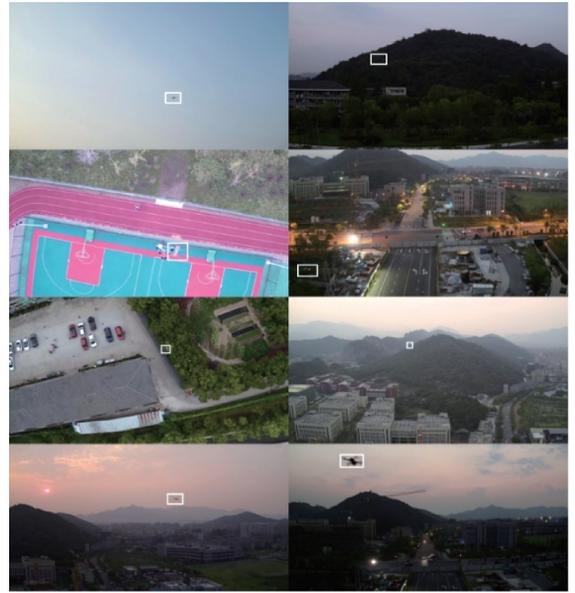


图 8 Det-Fly 数据集示例

Fig. 8 Det-Fly dataset example

3 实 验

3.1 数据集和评估指标

本研究的对比实验和消融实验使用的数据集是 Det-Fly 数据集,并最终在自采集数据集上进行了验证。Det-Fly 数据集和自采集数据集示例分别如图 8、9 所示。Det-Fly 数据集为空对空的无人机检测场景,其中目标的平均像素占比为 0.117%,低于小目标 0.12% 的像素占比,且包括不同场景、视角和相对距离。自采集数据集包含了天空、山地、建筑等场景,且目标的平均像素占比为 0.045%,远低于小目标 0.12% 的像素大小占比。图像中低慢小无人机的位置均用白色框表示。

在评估指标方面,本研究采取平均精度均值(mean average precision, mAP)、模型参数数量和帧率(frames per second, FPS)作为模型的主要检测指标。

精确率(Precision)表示在所有被模型预测为正类的样本中,实际为正类的比例:

$$Precision = \frac{TP}{TP + FP} \tag{14}$$

其中,TP 是真正例(true positives)的缩写,FP 是假正例(false positives)的缩写。

召回率(Recall)表示在所有实际为正类的样本中,模型预测为正类的比例:

$$Recall = \frac{TP}{TP + FN} \tag{15}$$

其中, FN 是假负例(false negatives)的缩写。

平均精度(average precision, AP)是一种综合性能评估指标,能够有效地反映模型在不同阈值下的表现。以精确率为横坐标,召回率为纵坐标绘制 P-R 曲线,计算各类别 P-R 曲线下的面积为各类别的 AP,利用各类别 AP

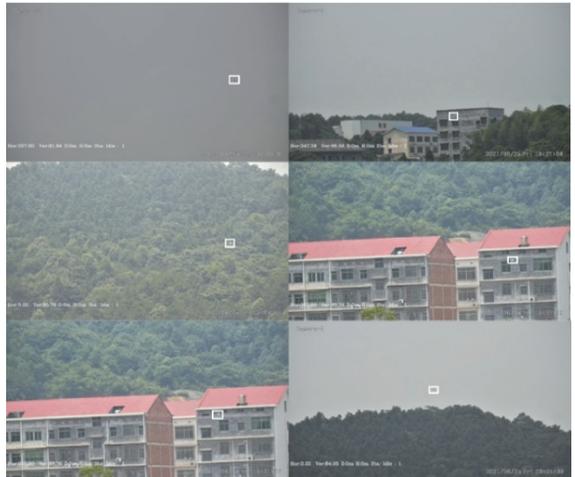


图 9 自采集数据集

Fig. 9 Self-collected dataset

的均值作为 mAP,其值越大越好。

参数量是指网络模型中需要训练的参数总数,它是衡量模型复杂度的重要指标。参数量越大,模型的复杂度也越高,从而需要更多内存来存储这些模型参数。

帧率指的是模型每秒处理的图像帧数,对检测模型性能至关重要。较高的帧率意味着模型具有更好的实时能力。

3.2 实验环境

本研究首先在 Det-Fly 数据集上进行了对比实验,证明了与目前常用的目标检测模型相比,改进的 BG-YOLO 在模型精度和推理速度均处于领先水平。接着,又在

Det-Fly 数据集上进行了消融实验,验证了每个改进部分的有效性。最后,在自采集的数据集上测试了所提算法。如图 10 所示为自采集数据集用到的实验设备。如表 2 所示为本研究算法的运行环境。

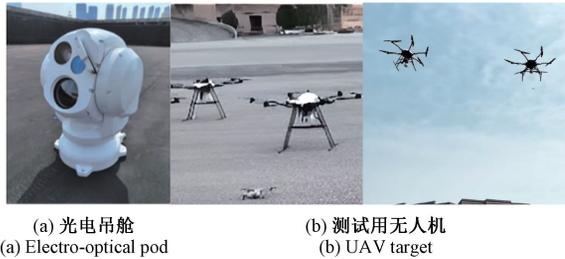


图 10 实验设备

Fig. 10 Experimental equipment

表 2 算法运行环境

Table 2 Algorithm operating environments

实验环境	版本
编程语言	Python3. 11
深度学习架构	Pytorch2. 0. 0
CUDA	11. 8
操作系统	Win10
GPU	NVIDIA RTX4060(8 GB)

3.3 对比实验

为验证所提算法的有效性,本研究在 Det-Fly 测试集上与代表性的目标检测模型进行对比,结果如表 3、4 所示。由表 3、4 可知,使用 Det-Fly 数据集,所提算法 BG-YOLO 的 $mAP@0.5$ 为 91.5%,显著优于其他算法,且其帧率并没有随着模型复杂度的增加而显著下降。与基线模型 YOLOv8 相比,改进的 BG-YOLO 算法 $mAP@0.5$ 提高了 10.3%,在小目标检测上表现尤为突出;与 YOLOv9 相比,BG-YOLO 算法的 $mAP@0.5$ 提高了 7.9%。以上数据均为模型训练了 80 个 epoch 得到的结果。召回率、精确率和平均精度曲线如图 11 所示。

表 3 Det-Fly 数据集上不同模型精度对比

Table 3 Precision comparison of different models on

Det-Fly dataset		(%)
数据集	模型	$mAP@0.5$
Det-Fly	YOLOv8	81.2
	YOLOv9e	83.6
	Faster R-CNN	70.5
	SSD512	78.7
	FPN	78.7
	本研究	91.5

表 4 Det-Fly 数据集上不同模型 FPS 对比

Table 4 FPS comparison of different models on Det-Fly dataset

数据集	模型	FPS
Det-Fly	YOLOv8	125.26
	YOLOv9e	156.21
	Faster R-CNN	17.51
	SSD512	42.33
	FPN	15.89
	本研究	127.65

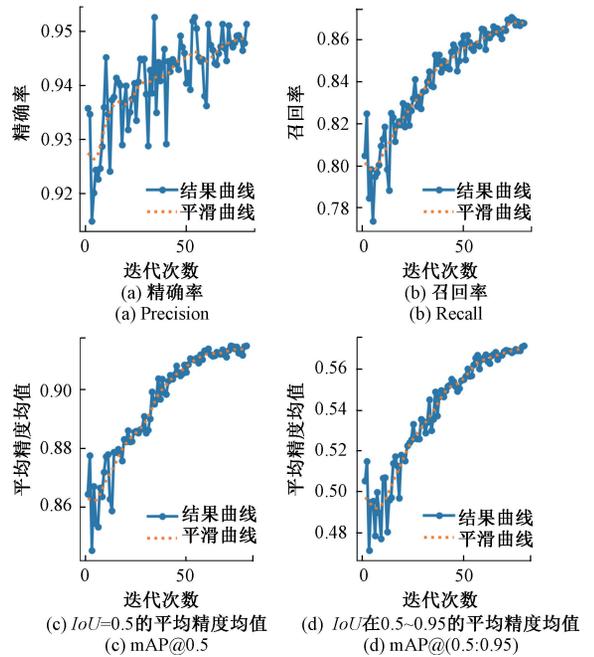


图 11 BG-YOLO 在 Det-Fly 数据集上的检测指标

Fig. 11 Detection metrics of BG-YOLO on the Det-Fly dataset

3.4 消融实验

为了评估不同改进对模型的影响,在 Det-Fly 数据集上进行了测试,结果如表 5 所示。YOLOv8 基线模型在 $mAP@0.5$ 指标上达到了 81.2% 的准确率,参数数量为 3.011 M,处理速度为 125.26 fps。

消融实验显示,结合 GhostNetV2 对 YOLOv8 基线模型进行轻量化后,模型的 mAP 略降至 80.6%,但参数数量减少至 1.855 M, FPS 提升至 153.85,虽然精度有所下降,但在参数和推理速度上有了改善。在此基础上,通过添加 Biformer 注意力、NWD 损失函数和 P2 小目标检测头,验证了这些方法对提升模型精度的有效性。特别是添加 P2 检测头后, mAP 显著提升至 89.7%,参数量降至 1.746 M, FPS 略降至 131.58。最终,在轻量化 YOLOv8 和添加 P2 小目标检测头的基础上,结合 Biformer 注意

表 5 Det-Fly 数据集上的消融实验结果

Table 5 Ablation results on the Det-Fly dataset

模型	$mAP@0.5/\%$	参数/M	FPS
YOLOv8 (baseline)	81.2	3.011	125.26
YOLOv8-Ghostv2	80.6	1.855	153.85
YOLOv8-Ghostv2-Biformer	81.7	2.121	147.05
YOLOv8-Ghostv2-NWD	83.8	1.855	153.26
YOLOv8-Ghostv2-P2	89.7	1.746	131.58
YOLOv8-Ghostv2-P2-NWD	90.2	1.746	132.14
YOLOv8-Ghostv2-P2-Biformer	90.3	2.012	128.87
YOLOv8-Ghostv2-P2-Biformer-NWD	91.5	2.012	127.65

力和 NWD 损失函数,模型精度均有所提高。当同时应用所有优化策略时,模型的 mAP 达到了 91.5%,参数量为 2.012 M, FPS 为 127.65,显著提升了模型精度并降低了复杂度。

3.5 不同场景的目标检测结果可视化分析

进一步对 BG-YOLO 在 Det-Fly 数据集和自采集数据集中不同背景下的低慢小目标检测结果进行分析,结果分别如表 6 和 7 所示。可以看出,模型在不同背景下整体上取得了较好的检测结果。

表 6 Det-Fly 数据集中不同场景的模型性能指标

Table 6 Model performance indexes of different scenarios in Det-Fly dataset

场景	精确率	召回率	$mAP@0.5/\%$	FPS
天空	0.958	0.969	97.3	140.85
山地	0.935	0.934	93.3	125.46
城市	0.921	0.856	88.2	108.70

表 7 自采集数据集中不同场景的模型性能指标

Table 7 Model performance indicators of different scenarios in the self-collecting dataset

场景	精确率	召回率	$mAP@0.5/\%$	FPS
天空	0.981	0.979	96.2	150.36
山地	0.907	0.896	88.1	128.21
建筑	0.891	0.856	86.2	112.53

具体来说,模型在较为单一的天空背景下检测性能最好,在目标平均像素占比为 0.117% 的 Det-Fly 数据集和目标平均像素占比为 0.045% 的自采集数据集上均取得了很好的检测结果,天空背景下 BG-YOLO 的 $mAP@0.5$ 分别为 97.3% 和 96.2%,帧率分别为 140.85 和 150.36。当背景变得复杂时,比如在山地背景中,模型的精度和检测帧率呈现了一定的下降趋势, $mAP@0.5$ 分别

为 93.3% 和 88.1%。尤其是在更为复杂的城市和建筑背景中,模型在两个数据集上的平均精度分别下降到了 88.2% 和 86.2%。BG-YOLO 在 Det-Fly 数据集和自采集数据集上不同场景的可视化结果分别如图 12~15 所示。



图 12 天空场景检测结果

Fig. 12 Sky scene detection results



图 13 山地场景检测结果

Fig. 13 Mountain scene detection results



图 14 城市场景检测结果

Fig. 14 Urban scene detection results

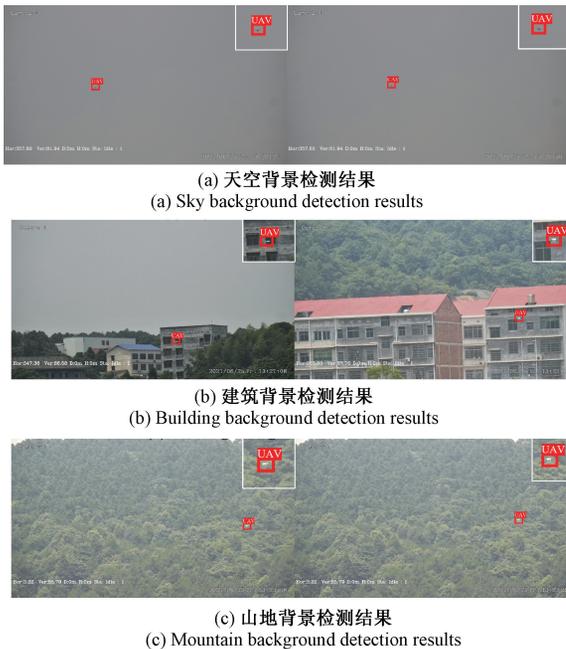


图 15 自采集数据集不同场景检测结果

Fig. 15 Detection results of different scenes from self-collected dataset

总的来说,改进后的 BG-YOLO 可以更好的对低慢小无人机进行检测,但是在一些复杂场景,比如目标和背景的纹理信息相似、环境光线暗淡、干扰物较多的情况下,检测算法仍面对一些挑战。

4 结 论

低慢小无人机目标的探测对于维护低空安全和实施有效防控至关重要。然而,由于低慢小目标体积微小、特征不明显,以及周围环境的复杂多变,探测这些低慢小目标仍然是一个充满挑战性的工作。现有的模型算法在这些挑战条件下对低慢小无人机探测时均存在一些不足,如模型的检测精度和实时性差等。为此,本研究提出了一种改进的无人机目标检测算法 BG-YOLO。BG-YOLO 基于 YOLOv8 进行改进,通过添加小目标检测头、添加 Biformer 注意力机制、改进损失函数以及基于 GhostNetV2 进行模型的轻量化,在 Det-Fly 数据集上的对比实验和消融实验证明了改进模型的有效性。BG-YOLO 的 $mAP@0.5$ 相比 YOLOv8 提升了 10.3%,参数减少了 33.18%,与 YOLOv9 相比也提高了 7.9%,实现了在保持精度的同时,显著降低了计算复杂度。

此外,本研究还在自采集数据集上对 BG-YOLO 进行了验证,自采集数据集中低慢小目标平均像素占比为 0.045%,且包含了天空、山地、建筑等不同场景。实验结果表明,BG-YOLO 对这 3 个场景分别实现了 96.2%、

88.1% 和 86.2% 的平均精度,可以较好地检测低慢小目标。

尽管本研究取得了一定的进展,但仍存在一些不足之处,例如在一些复杂背景、目标被遮挡、光线暗淡的情境下,模型的检测精度容易受到影响。未来将考虑为模型引入上下文信息模块,以增强模型在复杂背景场合的目标探测能力;综合利用时空信息来应对目标被遮挡的情况;利用多模态信息来进行全面的低慢小目标检测。

参考文献

- [1] 张鹏,徐丹. 基于光电侦察的“低慢小”目标检测综述[J]. 中国电子科学研究院学报, 2022, 17(11): 1042-1048.
ZHANG P, XU D. A survey of LSS-target detection based on photoelectric reconnaissance [J]. Journal of China Academy of Electronics and Information Technology, 2022, 17(11): 1042-1048.
- [2] 郭庆华,胡天宇,刘瑾,等. 轻小型无人机遥感及其行业应用进展[J]. 地理科学进展, 2021, 40(9): 1550-1569.
GUO Q H, HU T Y, LIU J, et al. Advances in light weight unmanned aerial vehicle remote sensing and major industrial applications[J]. Progress in Geography, 2021, 40(9): 1550-1569.
- [3] 张皓,吴虎胜,彭强. “低慢小”无人机反制装备及关键技术发展需求综述[J]. 航空兵器, 2022, 29(5): 43-52.
ZHANG H, WU H SH, PENG Q. Summary of development requirements of “low, slow and small” UAV countermeasure equipment and key technologies [J]. Aeronautical Weaponry, 2022, 29(5): 43-52.
- [4] 孙备,孙晓永,钱翰翔,等. 动态大视角场景融合帧间信息与模板匹配的低慢小无人机目标检测[J]. 仪器仪表学报, 2024, 45(7): 64-74.
SUN B, SUN X Y, QIAN H X. Low slow small UAV targets detection by fused using inter-frame information and template matching in dynamic large-view scene [J]. Chinese Journal of Scientific Instrument, 2024, 45(7): 64-74.
- [5] ZHU L, WANG X J, KE ZH H, et al. BiFormer: Vision transformer with Bi-level routing attention [C]. 2023 IEEE/CVF Conference on Computer Vision and Pattern

- Recognition, 2023;10323-10333.
- [6] WANG J W, XU CH, YANG W, et al. A normalized Gaussian Wasserstein distance for tiny object detection[J]. ArXiv preprint arXiv:2110.13389, 2021.
- [7] TANG Y H, HAN K, GUO J Y, et al. GhostNetV2: Enhance cheap operation with long-range attention[C]. 36th Conference on Neural Information Processing Systems, 2022, 35: 9969-9982.
- [8] ZHENG Y, CHEN ZH, LYU D L, et al. Air-to-air visual detection of micro-UAVs: An experimental evaluation of deep learning [J]. IEEE Robotics and Automation Letters, 2021, 6(2): 1020-1027.
- [9] 屈旭涛, 庄东晔, 谢海斌. “低慢小”无人机探测方法[J]. 指挥控制与仿真, 2020, 42(2): 128-135.
- QU X T, ZHUANG D Y, XIE H B. Detection methods for low-slow-small (LSS) UAV [J]. Command Control and Simulation, 2020, 42(2): 128-135.
- [10] 周涵楚, 朱永忠, 张玲玲, 等. 旋翼无人机目标雷达探测技术研究进展[J]. 电讯技术, 2024, 64(8): 1335-1345.
- ZHOU H CH, ZHU Y ZH, ZHANG L L, et al. Advances on radar detection technology for rotorcraft unmanned aerial vehicles[J]. Telecommunication Technology, 2024, 64(8): 1335-1345.
- [11] 张巍巍, 时晨光, 周建江. 基于射频隐身的无人机集群搜索任务分配算法[J]. 战术导弹技术, 2020(6): 60-66, 119.
- ZHANG W W, SHI CH G, ZHOU J J. An algorithm for task assignment in UAV swarm search based on RF stealth[J]. Tactical Missile Technology, 2020(6): 60-66, 119.
- [12] 陈鹏, 陈洋, 王威. 无人机声学定位技术综述[J]. 华南理工大学学报(自然科学版), 2022, 50(12): 109-123.
- CHEN P, CHEN Y, WANG W. Review for UAV acoustic positioning[J]. Journal of South China University of Technology (Natural Science Edition), 2022, 50(12): 109-123.
- [13] HAIFAWI H, FIORANELLI F, YAROVY A, et al. Drone detection & classification with surveillance radar on-the-move and YOLO[C]. 2023 IEEE Radar Conference, 2023: 1-6.
- [14] REN SH 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.
- [15] 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, 2016: 779-788.
- [16] 徐光达, 毛国君. 多层次特征融合的无人机航拍图像目标检测[J]. 计算机科学与探索, 2023, 17(3): 635-645.
- XU G D, MAO G J. Aerial image object detection of uav based on multi-level feature fusion[J]. Journal of Frontiers of Computer Science and Technology, 2023, 17(3): 635-645.
- [17] 张寅, 朱桂熠, 施天俊, 等. 基于特征融合与注意力的遥感图像小目标检测[J]. 光学学报, 2022, 42(24): 140-150.
- ZHANG Y, ZHU G Y, SHI T J, et al. Remote sensing image small target detection based on feature fusion and attention[J]. Acta Optica Sinica, 2022, 42(24): 140-150.
- [18] 张宏群, 班勇苗, 郭玲玲, 等. 基于YOLOv5的遥感图像舰船的检测方法[J]. 电子测量技术, 2021, 44(8): 87-92.
- ZHANG H Q, BAN Y M, GUO L L, et al. A remote sensing image ship detection method based on YOLOv5[J]. Electronic Measurement Technology, 2021, 44(8): 87-92.
- [19] QIN ZH Y, CHEN D K, WANG H Y. MCA-YOLOv7: An improved UAV target detection algorithm based on YOLOv7[J]. IEEE Access, 2024, 12: 42642-42650.
- [20] 冯哲, 王彬, 黄鹏程, 等. 小天体表面着陆区岩石目标检测算法[J]. 仪器仪表学报, 2024, 45(4): 194-205.
- FENG ZH, WANG B, HUANG P CH, et al. Algorithm of detection rock object in landing zone of small celestial body surface[J]. Chinese Journal of Scientific Instrument, 2024, 45(4): 194-205.
- [21] ZHU X K, LYU SH CH, WANG X, et al. TPH-YOLOv5: Improved YOLOv5 based on transformer prediction head for object detection on drone-captured

scenarios[C]. 2021 IEEE/CVF International Conference on Computer Vision Workshops, 2021: 2778-2788.

- [22] LIU S Y, LIANG P, DUAN Y Y, et al. Small target detection for unmanned aerial vehicle images based on YOLOv51[C]. 2022 10th International Conference on Information Systems and Computing Technology, 2022: 210-214.

- [23] 郝晋渊, 张家明, 张少康, 等. 基于改进 YOLO 的无人机入侵检测方法[J]. 电子测量与仪器学报, 2024, 38(7): 143-151.

HAO J Y, ZHANG J M, ZHANG SH K, et al. UAV intrusion detection method based on improved YOLO[J]. Journal of Electronic Measurement and Instrument, 2024, 38(7): 143-151.

- [24] CAO J SH, BAO W SH, SHANG H X, et al. GCL-YOLO: A GhostConv-based lightweight YOLO network for UAV small object detection[J]. Remote Sensing, 2023, 15(20): 4932.

- [25] HAN K, WANG Y H, TIAN Q, et al. GhostNet: More features from cheap operations[C]. 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2020: 1577-1586.

作者简介



王迎龙, 2023 年于南京农业大学获得学士学位, 目前于国防科技大学攻读硕士学位, 主要研究方向为无人平台智能侦测与对抗。

E-mail: wyl23@ nudt. edu. cn

Wang Yinglong received his B. Sc. degree from Nanjing Agricultural University in 2023. He is currently pursuing his M. Sc. degree at the National University of Defense Technology. His main research interests include intelligent detection and countermeasures of unmanned platforms.



孙备(通信作者), 分别于 2012 年、2014 年和 2018 年于国防科技大学获得学士、硕士和博士学位, 现为国防科技大学副研究员, 主要研究方向为无人平台智能侦测与对抗。

E-mail: sunbei08@ nudt. edu. cn

Sun Bei (Corresponding author) received his B. Sc., M. Sc. and Ph. D. degrees all from National University of Defense Technology in 2012, 2014 and 2018, respectively. He is currently an associate researcher at National University of Defense Technology. His main research interests include intelligent detection and countermeasure of unmanned platforms.