

基于 YOLOv7 的机场跑道异物检测算法^{*}

程 擎 叶 紫 何汶键 华 翔

(中国民用航空飞行学院空中交通管理学院 广汉 618307)

摘 要: 机场跑道异物对航班安全起降构成极大威胁,准确及时地检测并清除机场跑道异物是机场安全工作的重点。针对机场跑道异物检测任务中的小目标检测精确度与实时性,提出一种基于 YOLOv7 的机场跑道异物检测算法。首先在主干网络引入 CBAM 模块,从空间注意力与和通道注意力两方面专注小目标特征信息提取;其次在加强特征提取网络结合 AFPN 思想提出 SA-PANet 结构,将相邻有效特征层进行渐进式特征融合,缓解有效特征层之间的语义差距;然后在加强特征提取网络的 PANet 结构下采样支路中引入 BiFormer 模块,聚焦小目标特征信息的进一步融合提取;最后在边界框定位损失函数计算过程中引入 MPDIoU Loss,加速模型收敛并提升机场跑道异物检测准确率与定位精度。在机场跑道异物图像数据集上实验表明,改进后算法 mAP_{50} 为 98.76%,较改进前算法提升 9.09 个百分点。与其他针对机场跑道异物检测的算法相比,改进后算法具有更高的检测精度同时将模型参数量与模型计算量增幅控制在可接受范围内,达到机场跑道异物检测任务的准确、快速需求。

关键词: 目标检测;机场跑道异物;注意力机制;CBAM;AFPN;BiFormer;MPDIoU Loss

中图分类号: TN911.73 **文献标识码:** A **国家标准学科分类代码:** 510.4050

Foreign object debris detection based on YOLOv7 algorithm

Cheng Qing Ye Zi He Wenjian Hua Xiang

(School of Air Traffic Management, Civil Aviation Flight University of China, Guanghan 618307, China)

Abstract: FOD poses a great threat to airports. Detecting and removing FOD accurately and timely is the key point of airports' safety work. A FOD detection algorithm that based on YOLOv7 is proposed to meet the requirements of accuracy and real-time. Firstly, the CBAM module is introduced into the backbone network to focus on the extraction of small target feature information from two aspects: spatial attention and channel attention. Secondly, the idea of AFPN is integrated into the feature extraction network and SA-PANet structure is proposed in combination with it. SA-PANet can asymptotically fuse adjacent effective feature layers and alleviate the semantic gap between them. Thirdly, the BiFormer module is introduced into the down-sampling branch of PANet, which can focus on further fusion extraction of small target feature information in the feature extraction network. Lastly, MPDIoU Loss is introduced into the boundary frame positioning loss calculation, which can not only accelerate the convergence of the model but also improve the detection accuracy and location accuracy. Experiments on FOD datasets show that the mAP_{50} of the improved YOLOv7 algorithm is 98.76%, which is 9.09 percentage points higher than the original YOLOv7. Comparing with other algorithms for FOD detection, the improved YOLOv7 algorithm has higher detection accuracy and the increase of Params and GELOPs is controlled within the acceptable range, which meet the accurate and fast requirements of FOD detection tasks.

Keywords: target detection; FOD; attention mechanism; CBAM; AFPN; BiFormer; MPDIoU Loss

0 引 言

机场跑道异物(foreign object debris, FOD)指不属于机场道面却出现在机场运行区域且会对机场、飞机及其安

全起降造成威胁的外来物体,例如飞机零部件、维修工具、施工碎片杂物、塑料制品及其碎片、碎石块、遗落的私人物品等。FOD 对机场安全具有极大威胁,可能造成极其严重的安全事故,导致人员伤亡和一系列经济损失^[1]。及时并

准确地对 FOD 进行检测排除对保障机场安全运行具有重要意义。

目前检测 FOD 的方法主要分为人工巡检、雷达探测以及光学成像检测三类。人工巡检在执行任务时需关闭跑道,且巡检时主要采用人员乘车完成,巡检次数和巡查时间有限,可能出现漏检误检。雷达检测可在能见度较低、光照条件较差的情况下完成探测任务,因此被广泛应用于 FOD 检测。章林等人采用圆弧合成孔径雷达提高雷达探测距离,能对雨雪等“闪烁”杂波起到良好的抑制效果^[2];俞祥威等^[3]提出分布式机场异物检测雷达设想,利用目标区别于与道面杂波的空间相关特性,有效提高微弱异物信杂比;褚昕悦等^[4]对三维激光雷达采集的点云数据进行体素特征融合,提升激光雷达在 FOD 检测的效果。雷达探测存在一定局限,如较高的天线增益在提高分辨率的同时可能造成机场未来潜在的电磁兼容问题、异物检测的实时性会受到雷达照射频率的限制、对小尺寸 FOD 检测能力有限、设备制造和维护成本高、提取的 FOD 特征信息有限并且后端信号处理过程复杂等。

传统的 FOD 光学成像检测方式多采用滑动窗口法,存在时间冗余度较高、人工设计的特征鲁棒性差等不足,且易受到光照条件、能见度条件等环境因素影响,检测精度和效率均不太理想。近年来,深度学习凭借其优秀的特征提取能力在计算机视觉的发展中体现出显著的性能优势,例如目标检测、语义分割等领域^[5]。基于深度学习的目标检测主要分为两类:单阶段目标检测与双阶段目标检测。单阶段目标检测算法直接识别生成图片目标的所属类型概率以及位置坐标,检测速度较快,其代表算法有 YOLO(you only look once)系列、单步多框检测算法(single shot multibox detector,SSD)等。双阶段目标检测算法分为两个步骤,生成目标的候选区域后,对目标进行分类并对目标位置进行精修,具有更高的识别精度但检测速度较慢,无法满足要求高实时性的检测任务需求,其代表算法有基于区域的卷积神经网络算法(region-based convolutional neural networks,R-CNN)以及以 R-CNN 为基础的衍生算法等。

考虑到 FOD 多为小物体并且检测要求精准快速,单阶段目标检测算法更能满足任务需求。YOLO 系列算法自 2015 年由 Redmon 等^[6]提出 YOLOv1 以来就受到广泛关注,不断推陈出新。2016 年 Redmon 等^[7]在 YOLOv1 的基础上进行召回率和定位精度的改进,提出 YOLO9000,即 YOLOv2;2018 年 Redmon 等^[8]改进多个独立逻辑回归分类器完成预测分类,提出 YOLOv3;2020 年 Bochkovskiy 等^[9]提出 YOLOv4;同年 Jocher 等针对解决移动端部署提出模型更精炼、检测速度更快的 YOLOv5;2022 年 Li 等^[10]针对工业应用提出 YOLOv6;同年 Wang 等^[11]提出检测速度与精度优异的 YOLOv7;2023 年 Glenn Jocher 等在先前 YOLO 系列基础上改进提出 YOLOv8;2024 年 Wang 等^[12]针对数据传输时出现的信息丢失现象提出 YOLOv9。

其中 YOLOv3、YOLOv5 以及 YOLOv7 的基础框架成熟完整且可扩展性强,适用于各种实际目标检测任务中,因此目前 FOD 检测任务普遍从上述 3 个主要算法延伸展开。郭晓静等^[13]在 YOLOv3 算法基础上,替换特征提取网络并增加检测尺度以进行多尺度特征融合、在边界框定位损失引入广义交并比损失函数(generalized intersection over union loss,GIoU Loss),取得平均精度 86.5%的检测效果;李小军等^[14]再 YOLOv5 算法基础上,移除大目标检测层并引入卷积注意力模块(convolutional block attention module,CBAM)、采用重参数化模块进行特征融合,在边界框定位损失引入简化交并比损失函数(simplified intersection over union loss,SIoU Loss),取得平均精度 95.01%的检测结果;郭晓静等^[15]在 YOLOv7 算法基础上,引入双向加权特征金字塔(bi-directional feature pyramid network,BiFPN)融合多尺度特征,取得平均精度 94.7%的检测效果。

本文选择近年来在目标检测领域应用广泛的 YOLOv7 作为基础模型框架。针对 FOD 检测任务需求对 YOLOv7 进行兼顾检测精度与效率的改进提升,改进 YOLOv7 模型在以机场外来物(foreign object debris in airports,FOD-A)图像数据集为基础并添加自制图像的数据集上进行实验。

本文对 YOLOv7 算法的具体改进如下:

改进 YOLOv7 主干网络结构,在下采样模块前添加 CBAM 使主干网络提取特征时加强对小目标信息的关注度^[16],为后续网络结构提供更有利的特征信息输入。

改进 YOLOv7 加强特征提取网络,结合渐进特征金字塔(asymptotic feature pyramid network,AFPN)思想^[17],提出简化渐进式路径聚合结构(simplified asymptotic PANet,SA-PANet),将模型参数数量和模型计算量增幅控制在可接受范围内的同时对主干网络输出的相邻有效特征图进行渐进式特征融合,缓解非相邻有效特征图之间的语义差距。在路径聚合网络(path aggregation network,PANet)结构下采样支路中添加由双层路由注意力(bi-level routing attention,BRA)模块构建的 BiFormer 模块^[18],聚焦小目标特征信息的融合提取。

在边界框定位损失计算中引入新的损失函数——基于最小点距离的交并比损失函数(minimum point distance based intersection over union loss,MPDIoU Loss)^[19],加速模型收敛速度,提高针对 FOD 的检测精度。

通过上述对 YOLOv7 算法的改进,提高对 FOD 图像的检测精度的同时控制模型参数数量与计算量增幅使得模型避免过于庞大,达到 FOD 检测任务的精确、快速需求,对提升机场 FOD 检测工作的效率具有参考意义。

1 原始 YOLOv7 算法

YOLOv7 分为主干网络、加强特征提取网络以及头部

预测网络 3 大部分。主干网络对输入图像进行特征提取后得到表示图像特征集合的 3 个有效特征图。加强特征提取网络对特征图中的 3 个有效特征图进行特征融合与提取得

到特征加强后的特征图。头部预测网络对加强特征图进行先验框对应物体的种类分类与位置回归,完成图像目标检测任务。YOLOv7 原始网络结构如图 1 所示。

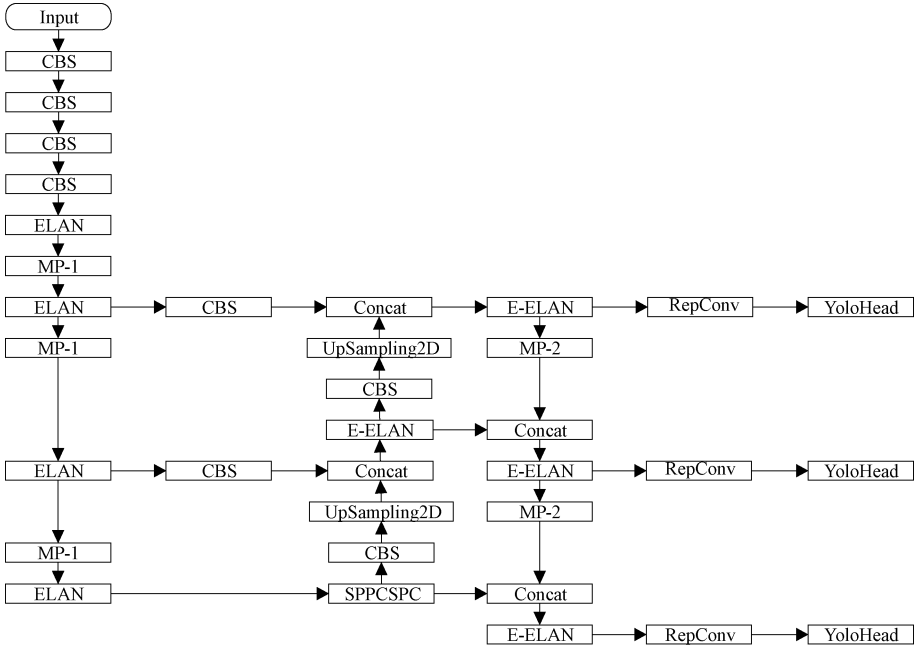


图 1 YOLOv7 的原始网络结构图
Fig. 1 Network structure of original YOLOv7

YOLOv7 在网络结构方面的创新性主要体现在高效聚合网络(efficient layer aggregation networks, ELAN)模块、最大池化降维(max-pooling reduction, MP)模块以及空间金字塔通道交叉池化(spatial pyramid pooling cross stage partial channel, SPPCSPC)模块三部分。

ELAN 模块将输入进行多分支的卷积-标准化-激活操作后堆叠,再执行卷积-标准化-激活操作。通过增加深度提高准确率,用残差块的跳跃连接来缓解深度增加引发的梯度消失。

ELAN 模块的网络结构图如图 2 所示。

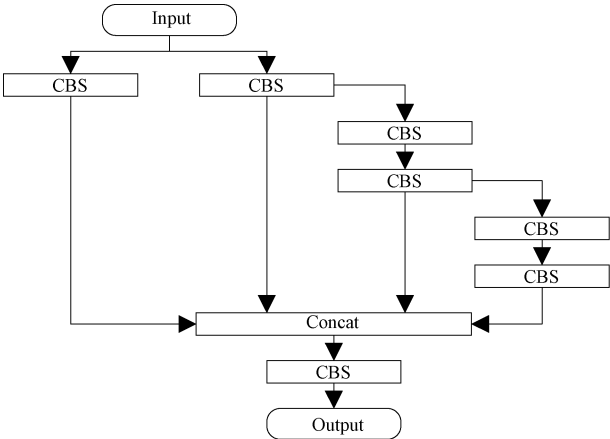


图 2 ELAN 模块的网络结构图
Fig. 2 Network structure of ELAN

MP 模块由两条支路组成,一条对输入依次进行最大池化与卷积-标准化-激活操作,另一条对输入进行两次不同的卷积-标准化-激活操作,两条支路的结果进行堆叠后输出。在图 1 中,主干网络与加强特征提取网络中分别使用的 MP-1 模块与 MP-2 模块结构上一致,仅存在通道数差异。MP 模块的网络结构图如图 3 所示。

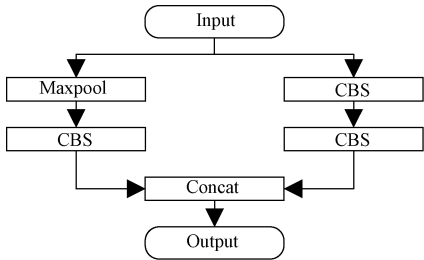


图 3 MP 模块的网络结构图
Fig. 3 Network structure of MP

SPPCSPC 模块通过不同尺度的最大池化扩大感受野,兼顾简化模型计算量与提升运算速度。SPPCSPC 模块的网络结构图如图 4 所示。

2 改进 YOLOv7 算法

本文结合 FOD 检测任务中目标尺寸偏小、检测需快速准确的任务需求改进原始 YOLOv7 算法,改进后 YOLOv7 的网络结构如图 5 所示。

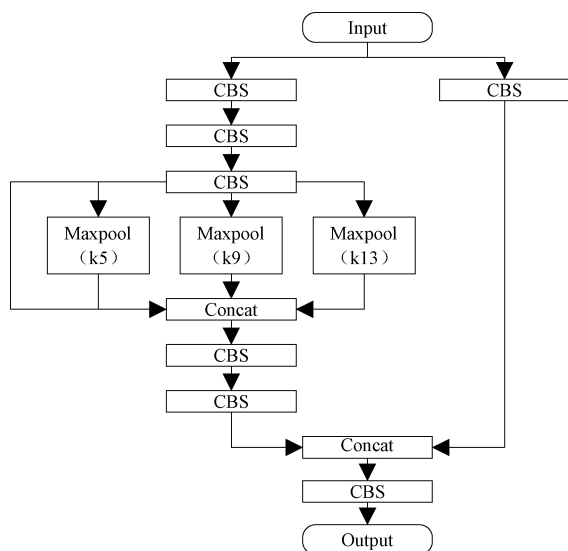


图 4 SPPCSPC 模块的网络结构图

Fig. 4 Network structure of SPPCSPC

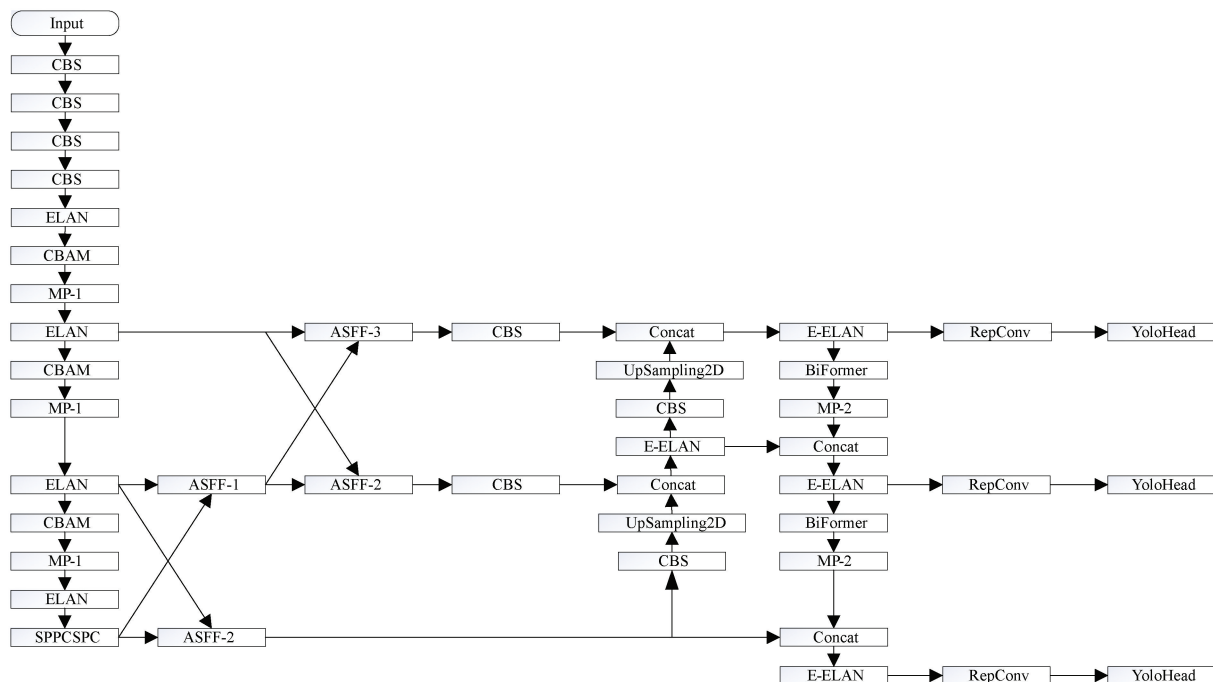


图 5 改进后 YOLOv7 的网络结构图

Fig. 5 Network structure of our improved YOLOv7

调重要特征并抑制不重要特征,使得主干网络输出的 3 个有效特征层具备更优质的特征信息。

CBAM 模块在 YOLO 算法改进中应用十分广泛,根据其添加位置主要分为两类。第一类是添加在主干网络。刘一诺等^[20]将主干网络第二和第四层卷积替换为 CBAM 模块,在处理图像的早期阶段引入 CBAM 模块可以提高模型对深层特征的提取能力,达到减轻模型后续计算量以及检测精度的效果。范先友等^[21]在主干网络 ELAN 模块后添加 CBAM 模块,达到抑制复杂背景干扰进而提升检测精

本文针对原始 YOLOv7 算法做出如下改进:首先在主干网络融入 CBAM 模块,其次在加强特征提取网络结合 AFPN 思想提出 SA-PANet 结构并且在其 PANet 结构下采样支路中引入 BiFormer 模块,最后在将边界框定位损失函数更换为 MPDIoU Loss。

2.1 CBAM 注意力模块

2018年Woo等^[16]提出CBAM模块,包含两个独立的子模块:通道注意力模块(channel attention module, CAM)以及空间注意力模块(spatial attention module, SAM)。CBAM模块的网络结构图如图6所示。

CAM 对输入特征图分别进行全局平均池化和全局最大池化,再由共享的全连接层处理,结果相加后再使用 sigmoid 函数归一化得到特征权重。

SAM 对输入特征图的各特征点取最大值与平均值并进行通道堆叠,通过卷积操作调整通道数后使用 sigmoid 函数归一化得到空间权重。CBAM 模块利用 SAM 聚焦图像中 FOD 相关区域,利用 CAM 显重要特征通道,实现强

度的效果。第二类是添加在加强特征提取网络。张敬博等^[22]分别在加强特征提取网络的上采样和下采样支路中添加 CBAM 模块,提升模型对待测目标的特征提取效果。王宸等^[23]在加强特征提取网络的 C3 特征层支路添加 CBAM 模块,提升模型对小目标特征的提取能力。

对本文检测算法而言,在主干网络 MP-1 模块进行下采样操作前添加 CBAM 模块,总计添加模块数为 3。随着主干网络深度的增加,各 CBAM 模块处理的特征图尺寸依次为 160×160 、 80×80 以及 40×40 ,特征图通道数依次为

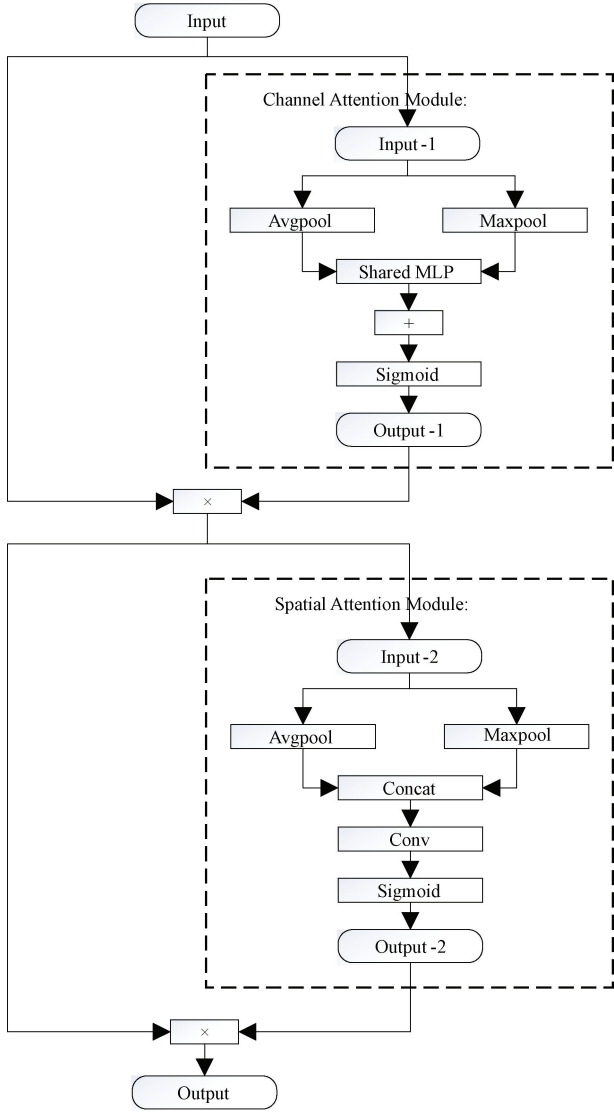


图 6 CBAM 模块的网络结构图

Fig. 6 Network structure of CBAM

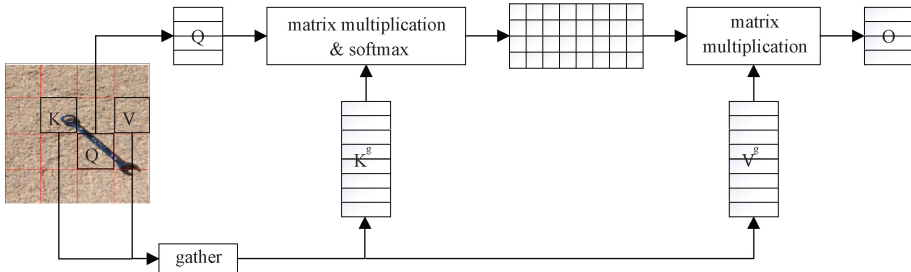


图 7 BiFormer 模块的网络结构图

Fig. 7 Network structure of BiFormer

256, 512 以及 1 024。

2.2 BiFormer 注意力模块

2023 年 Vaswani 等^[24]针对 Transformer 中自注意力机制的内存占用大及计算代价高问题,提出一种新的动态稀疏注意力机制——BiFormer 模块,达到提高计算效率及降低计算复杂度的效果。BiFormer 模块在实际运用中对小目标检测展现出优秀的性能,例如无人机搜寻海上遇险人员^[25]、无人机探测森林火灾烟雾^[26]以及检测识别农作物病害^[27]等。

对本文检测算法而言,在 PANet 结构中引入 BiFormer 模块,在对 3 个有效特征层进行加强特征提取与融合过程中提升对小目标关注度。在 MP-2 模块进行下采样操作前添加 BiFormer 模块,总计添加模块数为 2,随着下采样操作方向,各 BiFormer 模块处理的特征图尺寸依次为 80×80 以及 40×40 ,特征图通道数依次为 128 以及 256。BiFormer 模块的网络结构图如图 7 所示。

BiFormer 模块的处理过程为先将输入特征图 $X \in \mathbb{R}^{H \times W \times C}$ 均分为 $S \times S$ 个区域,各区域包含 $\frac{HW}{S^2}$ 个特征向

量。将 X 变形为 $X^r \in \mathbb{R}^{S^2 \times \frac{HW}{S^2} \times C}$,根据式(1)~(3)计算得到 $Q, K, V \in \mathbb{R}^{S^2 \times \frac{HW}{S^2} \times C}$ 。

$$Q = X^r W^q \quad (1)$$

$$K = X^r W^k \quad (2)$$

$$V = X^r W^v \quad (3)$$

对 Q, K 分别计算区域平均值可以得到 $Q^r, K^r \in \mathbb{R}^{S^2 \times C}$,根据式(4)计算得到相关性邻接矩阵 A^r 。

$$A^r = Q^r (K^r)^T \quad (4)$$

根据式(5)计算得到各行路由索引矩阵 I^r 。

$$I^r = \text{topIndex}(A^r) \quad (5)$$

根据式(6)、(7)聚集 K 和 V ,再根据式(8)对聚集后的 K 和 V 使用注意力机制并引入局部上下文增强项 $LCE(V)$,

获得 BiFormer 模块的输出 O 。

$$K^g = \text{gather}(K, I^r) \quad (6)$$

$$V^g = \text{gather}(V, I^r) \quad (7)$$

$$O = \text{Attention}(Q, K^g, V^g) + LCE(V) \quad (8)$$

2.3 SA-PANet 结构

2023 年 Yang 等^[28]针对非相邻特征图之间语义差距较大的问题提出 AFPN 结构,渐进融合各有效特征层的特征。AFPN 结构如图 8 所示,由主干网络提取四层特征

层,特征图之间进行渐进融合来缓解非相邻特征图之间的语义差距。

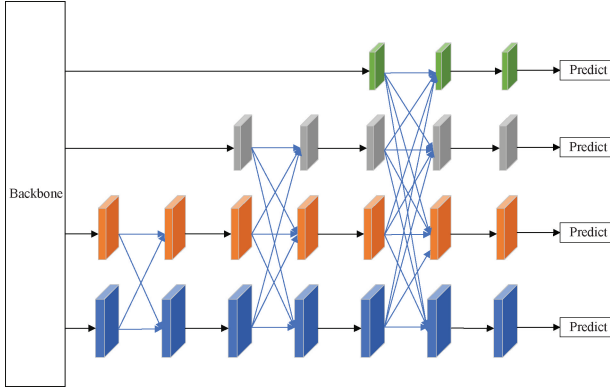


图 8 AFPN 的结构示意图

Fig. 8 Network structure of AFPN

对本文检测算法而言,在 YOLOv7 加强特征提取网络结构中,结合 AFPN 思想对 PANet 结构进行改进,提出 SA-PANet 结构。

SA-PANet 结构将主干网络提取到的有效特征图由浅层到深层依次命名为 feat1、feat2 和 feat3。为避免模型运算量过大,将 feat1、feat2 和 feat3 根据由深层至浅层的顺序,相邻特征层两两进行渐进式融合。SA-PANet 结构通过自适应空间融合 (adaptively spatial feature fusion, ASFF) 策略完成不同尺度特征图之间的渐进式融合,在特征融合与提取过程中尽可能保留各有效特征层的特征信息。每条渐进式融合支路的具体处理策略根据该条支路参与融合的特征层尺寸及通道数而定,即内部结构存在细微差别的 ASFF-1 模块、ASFF-2 模块以及 ASFF-3 模块。SA-PANet 的网络结构图如图 9 所示。

2.4 MPDIoU Loss

2023 年 Ma 等针对常用边界回归损失函数在预测框与实际标注框宽高不同但宽高比一致情况下无法进行优化的问题,提出一种基于最小点距离的新型损失函数——MPDIoU Loss,提升损失函数收敛速度与回归精确度。

YOLOv7 原始结构中损失函数包含 3 部分:边界框定位损失、分类损失以及置信度损失,使用完全交并比损失函数 (complete intersection over union loss, CIoU Loss) 计算边界框定位损失,二元交叉熵损失函数计算分类损失和置信度损失,3 种损失加权求和得到总损失函数。CIoU Loss 在目前最常用的目标检测损失函数中应用广泛且效果较好,但是仍存在计算复杂度较高、参数敏感度高、真实边界框和预测边界框相交时收敛速度变慢以及两框包含时会退化为交并比损失函数 (intersection over union loss, IoU Loss) 等不足。综合考虑 FOD 检测任务的小目标特点和 CIoU Loss 的不足,将边界框定位损失函数替换为 MPDIoU Loss,加快模型收敛速度并达到提升小目标检测精度。

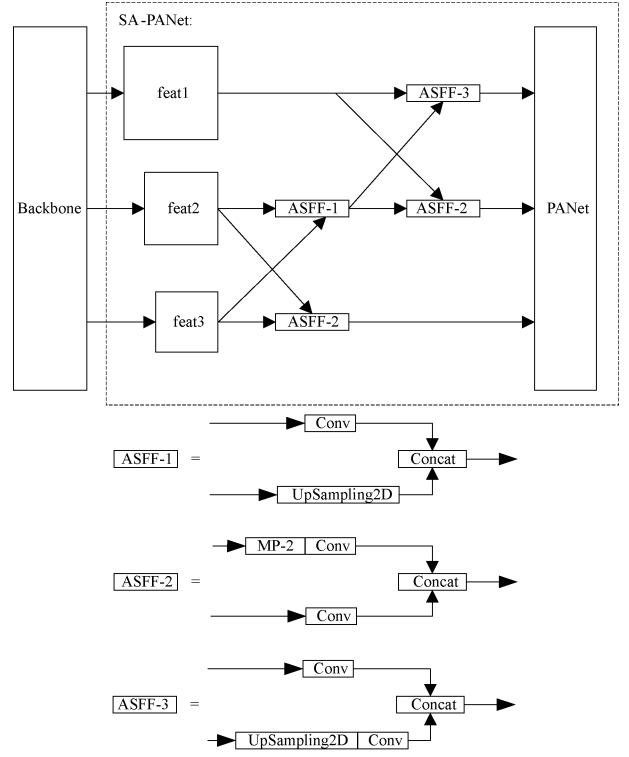


图 9 SA-PANet 的网络结构图

Fig. 9 Network structure of SA-PANet

MPDIoU Loss 的参数示意图如图 10 所示,图中黄色框表示真实边界框,红色框表示预测边界框,坐标分别表示对应边界框的左上角与右下角坐标, d_1, d_2 分别表示坐标点之间的距离, w, h 分别表示图像的宽度与高度。

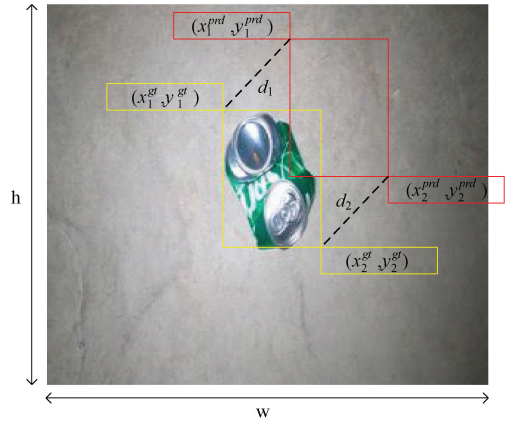


图 10 MPDIoU Loss 的参数示意图

Fig. 10 Factors of MPDIoU Loss

根据式 (9) 定义 $MPDIoU$, A 表示真实边界框覆盖的范围, B 表示预测边界框覆盖的范围, d_1^2 表示两边界框左上角之间欧氏距离的平方, d_2^2 表示两边界框右下角之间欧氏距离的平方。

$$MPDIoU = \frac{A \cap B}{A \cup B} - \frac{d_1^2}{w^2 + h^2} - \frac{d_2^2}{w^2 + h^2} \quad (9)$$

模型训练时,根据式 (10)计算损失函数 $\mathcal{L}$,使各个预测边界框 $\mathcal{B}_{prd} = [x_{prd}, y_{prd}, w_{prd}, h_{prd}]$ 趋近于真实边界框 $\mathcal{B}_{gt} = [x_{gt}, y_{gt}, w_{gt}, h_{gt}]$ 。式中 $\mathcal{B}_{gt}$ 表示真实边界框的集合, Θ 表示用于回归的深度模型参数。

$$\mathcal{L} = \min_{\Theta} \sum_{\mathcal{B}_{gt} \in \mathcal{B}_{gt}} \mathcal{L}(\mathcal{B}_{gt}, \mathcal{B}_{prd} \mid \Theta) \tag{10}$$

根据式(11)定义 MPDIoU Loss。

$$\mathcal{L}_{MPDIoU} = 1 - MPDIoU \tag{11}$$

根据图 9 坐标信息由式 (12)计算 $\mathcal{B}_{prd}$ 与 $\mathcal{B}_{gt}$ 的最小外接矩阵面积 $|C|$ 。

$$|C| = (\max(x_2^{gt}, x_2^{prd}) - \min(x_1^{gt}, x_1^{prd})) \times (\max(y_2^{gt}, y_2^{prd}) - \min(y_1^{gt}, y_1^{prd})) \tag{12}$$

由上述各式联立可知,在边界框定位损失计算中真实边界框与预测边界框的对比因素如两框的中心点坐标、宽与高、非重叠区域面积、中心点距离、宽度偏差与高度偏差,均可通过已知坐标点计算,计算过程得到简化。

MPDIoU Loss 定义,当预测真实边界框与预测边界框的宽高比相等时,预测边界框位于真实边界框外部情况下的 $\mathcal{L}_{MPDIoU}$ 取值高于预测边界框位于真实边界框内部情况下的 $\mathcal{L}_{MPDIoU}$ 取值,有助于筛选出冗余更小的预测边界框。

3 实验验证

3.1 实验环境与参数配置

实验操作系统为 Windows 10,实验环境为 python3.7, pytorch1.8.0 和 cuda11.1。GPU 为 NVIDIA GeForce GTX 1080 Ti,使用随机梯度下降 (stochastic gradient descent,SGD)优化器以及余弦退火学习率衰减策略,结合 mosaic 数据增强与 mixup 数据增强对输入图像进行处理,其他详细实验配置参数如表 1 所示。

表 1 详细实验配置参数

Table 1 Detailed experimental parameters

序号	参数名	参数值
1	冻结阶段训练轮次	50
2	冻结阶段批次大小	8
3	解冻阶段训练批次	300
4	解冻阶段批次大小	4
5	图像尺寸	640×640
6	标签平滑	0.005
7	初始学习率	0.01
8	最小学习率	0.000 1
9	权重衰减系数	0.000 5

3.2 数据集

选用 2022 年内布拉斯加大学发布的 FOD-A^[29] 图像数据集作为基础。该图像数据集采用便携式相机与无人机摄像进行图像采集,包含 33 792 个注释实例,涵盖 31 个

对象类别。在 FOD-A 数据集中,根据环境条件分为两类:干燥条件以及湿润条件。根据光照条件分为 3 类:明亮条件、昏暗条件以及黑暗条件。

由于 FOD-A 数据集中部分图像相似度较高,为避免训练过拟合导致模型泛化能力差,抽取 FOD-A 图像数据集中约 20% 的图像,与自制的 1 021 张图像建立新的图像数据集,按照比例 8 : 1 : 1 划分训练集、测试集及验证集。

3.3 评价指标

选择精确度 (precision, P)、召回率 (recall, R)、平均精度均值 (mean average precision, mAP)、模型计算量 (giga floating-point operations per second, GFLOPs) 以及模型参数量 (parameters, Params) 作为评估指标。

根据式 (13)、(14) 分别对 P, R 进行计算,式中 TP 表示真正例个数, FP 表示假正例个数, FN 表示假反例个数。

$$P = \frac{TP}{TP + FP} \tag{13}$$

$$R = \frac{TP}{TP + FN} \tag{14}$$

mAP 由目标检测所有类别的平均精度值取平均得到,目标检测中某个类别的平均精度值是该类别 R-P 曲线与坐标轴围成封闭图形的面积, mAP 值越高表示模型检测精度越高。本文使用 mAP₅₀, 表示 IoU 设为 0.5 时 mAP 的取值。

GFLOPs 用于评估模型的时间复杂度。Params 用于评估模型的空间复杂度。

3.4 测试结果可视化

本文使用抽取 FOD-A 数据集中 20% 的图像以及根据常见 FOD 种类自制的 1 021 张图像所构成的数据集进行测试。原始 YOLOv7 算法与改进后 YOLOv7 算法的部分检测结果对比如图 11 所示。

针对同一检测目标,改进后 YOLOv7 算法具有更高的检测置信度,部分原始 YOLOv7 算法未能准确识别物体种类信息的 FOD 图像在改进后 YOLOv7 算法中被正确检测出。

对改进后模型检测结果进行热力图可视化处理,热力图显示模型聚焦关注于小目标物体,满足 FOD 检测任务需求。改进模型在输入图像中对检测目标的关注程度高且定位准确。部分检测结果及对应可视化热力图如图 12 所示。

3.5 消融实验分析

为验证本文对 YOLOv7 各项改进的有效性,在相同实验环境下,依次对 4 个改进部分进行消融实验。消融实验表明本文 YOLOv7 的各项改进虽然在模型计算量和模型参数量上有所牺牲,但是对 FOD 目标检测精度具有可观的提升效果。

消融实验具体各项数值对比如表 2 所示。在主干网络引入 CBAM 模块,加深模型深度,增加主干网络输出的



(a) 原始YOLOv7算法检测结果
(a) Test results of the original YOLOv7



(b) 改进YOLOv7算法检测结果
(b) Test results of our improved YOLOv7

图 11 部分检测结果展示

Fig. 11 Part of the detection results

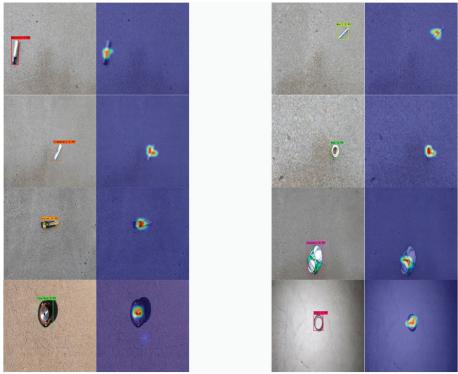


图 12 部分可视化热力图
Fig. 12 Part of the visualized heat map

有效特征层信息量, mAP_{50} 值略有提升。在此改进基础上引入 SA-PANet 结构,有效缓解非相邻特征层间的语义差距。考虑到模型轻量化需求,仅使相邻有效特征层依次进行两两渐进式特征融合而非标准 AFPN 结构操作,将模型计算量增幅控制在可接受范围内。在加强特征提取网络引入 BiFormer 模块,在对有效特征层进行加强特征提取与融合过程中提升对 FOD 小目标的关注度,提升 mAP_{50} 值。最后在边界框定位损失计算中引入 MDPIoU Loss,在加速模型收敛的同时提升对 FOD 小目标的检测精度。

3.6 对比实验分析

将双阶段目标检测算法典型算法 Faster R-CNN 以及常见单阶段目标检测算法 SSD、YOLOv3、YOLOv4、YOLOv5、YOLOv7 在本文数据集上进行对比实验,从 mAP_{50} 、GELOPs 以及 Params 三个方面进行算法性能评估,对比实验详细数据如表 3 所示。

表 2 消融实验测试结果对比

Table 2 Comparison of ablation test results

模型	CBAM	SA-PANet	BiFormer	MDPIoU Loss	$mAP_{50}/\%$	GELOPs/G	Params/M
YOLOv7					89.67	106.472	37.620
YOLOv7-1	✓				89.80	106.491	38.231
YOLOv7-2	✓	✓			95.12	138.366	46.552
YOLOv7-3	✓	✓	✓		96.31	141.241	53.748
本文	✓	✓	✓	✓	98.76	142.566	55.424

表 3 不同算法对比结果

Table 3 Comparison of different algorithm models

算法	$mAP_{50}/\%$	GELOPs/G	Params/M
Faster R-CNN	95.59	370.210	137.099
SSD	86.84	62.747	26.285
YOLOv3	83.51	66.171	61.949
YOLOv4	84.87	60.527	64.363
YOLOv5	87.06	115.918	47.057
YOLOv7	89.67	106.472	37.620
本文	98.76	142.566	55.424

实验数据表明本文算法在常见单阶段目标检测算法对比中 mAP_{50} 效果最佳,GELOPs 偏大,Params 适中。本文算法与 Faster R-CNN 相比, mAP_{50} 略高,优秀的是 GELOPs 与 Params 都更明显更小,更加轻量化。

与现有以 YOLO 系列算法为基础进行 FOD 检测研究的创新工作相比,本文算法的独创性内容主要集中在结合 AFPN 思想在 PANet 结构上提出兼顾模型计算量与检测精度的 SA-PANet 结构,在主干网络输出的有效特征层进入加强特征提取网络前进行相邻特征层间的渐进式融合。SA-PANet 结构作为本文算法改进过程中检测精度提升最

明显的部分,使模型 mAP_{50} 值提升 5.32 个百分点,在与现有 FOD 检测研究工作的对比中检测精度提升显著,可为后续检测工作提供研究思路。

4 结 论

本文针对 FOD 检测任务需求提出基于改进 YOLOv7 的 FOD 检测算法,旨在提高 FOD 检测精度。改进算法通过在主干网络引入 CBAM 模块,提升模型对小目标特征信息的敏感度;在加强特征提取网络结合 APFN 思想,提出 SA-PANet 结构,缓解有效特征层间的语义差距;在 PANet 结构下采样支路中引入 BiFormer 模块,聚焦小目标特征信息的融合提取;在边界框定位损失计算中引入 MPDIoU Loss,在提升模型对 FOD 检测精度的同时加速模型收敛。

改进后的检测模型在 FOD 图像数据集上 mAP_{50} 达到 98.76%,较改进前模型以及多种常用目标检测算法表现出更高的检测精度,可以为机场进行 FOD 检测任务提供算法参考。虽然改进后的检测模型在目标检测精度上效果较好,但是在模型参数量与计算量上均有提升,后续工作可专注于模型轻量化改进,在保障目标检测精度的同时进一步提升检测实时性。

参考文献

- [1] Federal Aviation Administration. AC 150/5220-24: Foreign object debris detection equipment[R]. 2009: 1-13.
- [2] 章林,龙超,童建文. 基于圆成孔径体制的跑道异物监测雷达成像研究[J]. 航空工程进展, 2021, 12(4): 155-162.
ZHANG L, LONG CH, TONG J W. FOD monitoring radar imaging research using arc synthetic aperture system[J]. Advances in Aeronautical Science and Engineering, 2021, 12(4): 155-162.
- [3] 俞祥威,王杰,施赛楠,等. 分布式机场异物检测雷达处理算法研究[J]. 信号处理, 2023, 39(5): 886-898.
YU X W, WANG J, SHI S N, et al. Research on radar processing algorithm for distributed foreign object debris detection [J]. Journal of Signal Processing, 2023, 39(5): 886-898.
- [4] 褚昕悦,赵旭,李连鹏,等. 基于体素特征融合的 FOD 目标智能认知[J]. 应用激光, 2023, 43(8): 151-158.
CHU X Y, ZHAO X, LI L P, et al. Intelligent cognition of FOD based on voxel feature fusion[J]. Applied Laser, 2023, 43(8): 151-158.
- [5] 唐霖峰,张浩,徐涵,等. 基于深度学习的图像融合方法综述[J]. 中国图象图形学报, 2023, 28(1): 3-36.
TANG L F, ZHANG H, XU H, et al. Deep learning-based image fusion: A survey[J]. Journal of Image and Graphics, 2023, 28(1): 3-36.
- [6] 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.
- [7] REDMON J, FARHADI A. YOLO9000: Better, faster, stronger[C]. IEEE Conference on Computer Vision & Pattern Recognition, 2017: 6517-6525.
- [8] REDMON J, FARHADI A. YOLOv3: An incremental improvement[J]. ArXiv preprint arXiv: 1804.02767, 2018.
- [9] BOCHKOVSKIY A, WANG C Y, LIAO H Y M. YOLOV4: Optimal speed and accuracy of object detection[J]. ArXiv preprint arXiv:2004.10934, 2020.
- [10] LI CH Y, LI L L, JIANG H L, et al. YOLOv6: A single-stage object detection framework for industrial applications [J]. ArXiv preprint arXiv: 2209.02976, 2022.
- [11] WANG C Y, BOCHKOVSKIY A, LIAO H Y M. YOLOv7: Trainable bag-of-freebies sets new state-of-the-art for real-time object detectors[C]. IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2023: 7464-7475.
- [12] WANG C Y, YE H I H, LIAO H Y M. YOLOv9: Learning what you want to learn using programmable gradient information [J]. ArXiv preprint arXiv: 2402.13616, 2024.
- [13] 郭晓静,隋昊达. 改进 YOLOv3 在机场跑道异物目标检测中的应用[J]. 计算机工程与应用, 2021, 57(8): 249-255.
GUO X J, SUI H D. Application of improved YOLOv3 in foreign object debris target detection on airfield pavement [J]. Computer Engineering and Applications, 2021, 57(8): 249-255.
- [14] 李小军,邓月明,陈正浩,等. 改进 YOLOv5 的机场跑道异物目标检测算法[J]. 计算机工程与应用, 2023, 59(2): 202-211.
LI X J, DENG Y M, CHEN ZH H, et al. Improved YOLOv5's foreign object debris detection algorithm for airport runways[J]. Computer Engineering and Applications, 2023, 59(2): 202-211.
- [15] 郭晓静,邹松林. 多尺度特征融合的机场跑道异物检测与识别算法[J]. 计算机工程与应用, 2024, 60(16): 319-324.
GUO X J, ZOU S L. Multi-scale feature fusion detection and recognition algorithm for airport runway foreign object [J]. Computer Engineering and Applications, 2024, 60(16): 319-324.

- [16] WOO S, PARK J, LEE J Y, et al. CBAM: convolutional block attention module[C]. European Conference on Computer Vision(ECCV), 2018: 3-19.
- [17] YANG G Y, LEI J, ZHU ZH K, et al. AFPN: Asymptotic feature pyramid network for object detection[C]. 2023 IEEE International Conference on Systems, Man, and Cybernetics(SMC), 2023: 2184-2189.
- [18] ZHU L, WANG X J, KE ZH H, et al. BiFormer: Vision transformer with bi-level routing attention[C]. IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2023: 10323-10333.
- [19] MA S L, XU Y. MPDIoU: A loss for efficient and accurate bounding box regression[J]. ArXiv preprint arXiv:2307.07662, 2023.
- [20] 刘一诺, 张琪, 王蓉, 等. 针对航拍小目标检测的 YOLOv7 改进方法[J]. 北京航空航天大学学报, 2024: 1-10. DOI: 10.13700/j. bh. 1001-5965. 2023. 0411.
- LIU Y N, ZHANG Q, WANG R, et al. Improved YOLOv7 method for aerial small target detection[J]. Journal of Beijing University of Aeronautics and Astronautics, 2024: 1-10. DOI: 10.13700/j. bh. 1001-5965. 2023. 0411.
- [21] 范先友, 过峰, 俞建峰, 等. 基于改进 YOLOv7 的液晶面板电极缺陷视觉检测技术研究[J]. 电子测量与仪器学报, 2023, 37(9): 225-233.
- FAN X Y, GUO F, YU J F, et al. Research on visual detection technology for liquid crystal panel electrode defect by improved YOLOv7 [J]. Journal of Electronic Measurement and Instrumentation, 2023, 37(9): 225-233.
- [22] 张敬博, 任杰, 王美琪. 基于混合注意力的全天候机场跑道异物检测[J]. 北京航空航天大学学报, 2024: 1-14. DOI: 10.13700/j. bh. 1001-5965. 2023. 0500.
- ZHANG J B, REN J, WANG M Q. All-weather airport runway foreign object debris detection based on mixed attention[J]. Journal of Beijing University of Aeronautics and Astronautics, 2024: 1-14. DOI: 10.13700/j. bh. 1001-5965. 2023. 0500.
- [23] 王宸, 杨帅, 周林, 等. 基于自适应多尺度特征融合网络的金属齿轮端面缺陷检测方法研究[J]. 电子测量与仪器学报, 2023, 37(10): 153-163.
- WANG CH, YANG SH, ZHOU L, et al. Research on metal gear end-face defect detection method based on adaptive multi-scale feature fusion network[J]. Journal of Electronic Measurement and Instrumentation, 2023, 37(10): 153-163.
- [24] VASWANI A, SHAZEER N, PAEMAR N, et al. Attention is all you need[J]. ArXiv preprint arXiv: 1706.03762, 2017.
- [25] ZHANG Y J, YIN Y, SHAO Z Y. An enhanced target detection algorithm for maritime search and rescue based on aerial images[J]. Remote Sensing, 2023, 15(19): 4818.
- [26] SAYDIRASULOVICH S N, MUKHIDDINOV M, DJURAEV O, et al. An improved wildfire smoke detection based on YOLOv8 and UAV images[J]. Sensors, 2023, 23(20): 8374.
- [27] GAO X, TANG ZH W, DENG Y B, et al. HSSNet: An end-to-end network for detecting tiny targets of apple leaf diseases in complex backgrounds [J]. Plants, 2023, 12(15): 2806.
- [28] YANG G Y, LEI J, ZHU ZH K, et al. AFPN: Asymptotic feature pyramid network for object detection[C]. 2023 IEEE International Conference on Systems, Man, and Cybernetics (SMC). IEEE, 2023: 2184-2189.
- [29] MUNYER T, HUANG P CH, HUANG CH Y, et al. FOD-A: A dataset for foreign object debris in airports[J]. ArXiv preprint arXiv:2110.03072, 2021.

作者简介

程擎(通信作者), 硕士研究生, 教授, 研究生导师, 主要研究方向为空管通信、导航、监视系统、空管自动化等。

E-mail: degleqh@163.com

叶紫, 硕士研究生, 主要研究方向为目标检测。

E-mail: zy00531@163.com

何汶键, 硕士研究生, 主要研究方向为无人机视觉识别。

E-mail: z874396s@gmail.com

华翔, 硕士研究生, 主要研究方向为目标检测。

E-mail: 2532492797@qq.com