

融合先验掩膜与 YOLOv8 的 FOD 检测方法研究

费春国 陈世洪

(中国民航大学电子信息与自动化学院 天津 300300)

摘要:为了解决 YOLOv8 在检测机场跑道异物时由于某些异物体积较小、空间位置随机及异物间尺度差距大而引起的漏检误检问题,本文对 YOLOv8 进行改进,提出应用先验掩膜的针对小目标的 AMMS-YOLOv8。首先在主干网络中,引入各向同性边缘检测算子并构建 EIEStim,增强模型的小边缘感知及预处理能力。同时替换下采样,将改进后的接受场注意力应用在检测领域,构建 LDFDS,增强空间感知并保护微小语义信息;其次重构 Neck 层结构,允许多尺度特征聚合,构建出 CCFPN 以增强模型对多尺度异物的语义感知;最后向检测头中嵌入先验异物掩膜特征,并将其与深度特征级联形成了 MSN-Head,以强化模型空间感知力。使用自建的复杂场景异物数据集对模型检测能力进行验证分析,在该数据集上 AMMS-YOLOv8 的 mAP50 及 mAP50-95 分别提升了 1.8% 及 1.7%,准确率、召回率、F1 函数分别达到了 0.971、0.976 和 0.973,相比原网络有很大提升。实验结果验证了改进方法的有效性,同时应用向复杂场景异物数据集中加入 FOD-A 的混合数据集和复杂输电线路异物数据集对 AMMS-YOLOv8 做了泛化性及鲁棒性对比实验,经实验表明各项指标均有提升。

关键词: YOLOv8; 机场跑道异物(FOD); 目标检测; 深度学习

中图分类号: TP751; TN919.8 **文献标识码:** A **国家标准学科分类代码:** 520.6040

Research on FOD detection method combining a priori mask and YOLOv8

Fei Chunguo Chen Shihong

(School of Aeronautic Information and Automation, Civil Aviation University of China, Tianjin 300300, China)

Abstract: To address the missed and false detection issues of YOLOv8 in identifying foreign object debris (FOD) on airport runways—caused by small object sizes, random spatial distribution, and significant scale variations among debris—this paper proposes AMMS-YOLOv8, an enhanced model incorporating prior masks specifically for small targets. In the backbone network, an isotropic edge detection operator is introduced to construct EIEStim, strengthening the model's perception and preprocessing capabilities for subtle edges. Simultaneously, downsampling is replaced by an improved receptive field attention mechanism applied to the detection domain, forming LDFDS to enhance spatial awareness and preserve minute semantic information. Subsequently, the Neck layer is restructured to enable multi-scale feature aggregation, developing CCFPN to improve semantic perception of multi-scale debris. Finally, prior FOD mask features are embedded into the detection head and concatenated with deep features to create MSN-Head, thereby amplifying spatial perception. The model's detection capability was validated using a self-built complex-scenario FOD dataset. On this dataset, AMMS-YOLOv8 achieved improvements of 1.8% and 1.7% in mAP50 and mAP50-95 respectively, with precision, recall, and F1-score reaching 0.971, 0.976, and 0.973—marking significant enhancements over the baseline network. Experimental results confirm the efficacy of these improvements. Furthermore, robustness and generalizability were evaluated through comparative experiments using a hybrid dataset (combining complex-scenario FOD data with FOD-A) and a complex transmission line FOD dataset, demonstrating performance gains across all metrics.

Keywords: YOLOv8; foreign object debris; object detection; deep learning

0 引言

伴随着民航业的起源与发展,安全问题始终处于首位。

随着近年来航线增加,同一机场飞机起降日益频繁,机场跑道异物(foreign object debris, FOD)成了威胁航空安全的重要因素,因此对其准确识别的重要性不言而喻。FOD指

任何本不属于机场道面,却被留在机场运行区域的物质,其以金属物质、石头、塑料制品和橡胶碎片为主。它们中较大的物质容易被吸入发动机中,并造成发动机的永久性损伤,或造成飞机的侧滑,引发大量经济损失和致使人员受伤。目前,国内大部分机场依旧采用人工巡检的方式清除此类 FOD,这种方法不仅效率低下,无法做到实时检测,而且极易产生漏检,从而遗留安全隐患。所以,准确识别 FOD 对民航发展有着重要的作用。

FOD 的识别主要有 3 种方式:雷达探测^[1]的自动识别方法;基于多传感器融合探测的自动识别方法^[2]以及基于光学成像探测的自动识别方法。雷达检测是目前使用较为广泛的方法,英国的 Tarsier 系统、以色列的 FODetect 系统以及美国的 FOD Finder 系统都是应用毫米波的雷达检测系统,其优点是可以做到不受外界光照等环境因素影响,全天候地对机场跑道进行高精度检测,且检测效果较为稳定。但是因为雷达造价较高并且缺少目标语义描述,在感知异物时无法对其进行分类进而影响处理策略的制定,且毫米波雷达的发射频率也会对机场的安全产生威胁,因此没有在国内大量应用。基于红外探测的检测由于其极易发生虚警及漏检且其受目标表面材质影响较大,并且在融合多传感器红外信息时常常因为技术原因导致多源信息无法融合,所以很少现实应用。

相比之下,基于光学成像的检测方法以其较为稳定的检测效果,较强的技术更新能力,较好的检测实时性得到大量应用。在基于光学成像的检测方法中,深度学习凭借强大的计算力成为现代研究的热点,更在很多类似任务^[3-12]中得到了较好的效果。YOLO 系列作为目标检测的热门算法更是在 FOD 检测的领域得到了很多应用。郭晓静等^[13]针对异物尺寸大小不一造成的漏检,以 YOLOv5 为基础网络,提出一种基于多尺度的特征融合识别算法,优化其中的双向特征金字塔网络(bidirectional feature pyramid network, BIFPN),达到了平均精度 94.7% 的效果;张敬博等^[14]针对因光照不足所造成漏检的情况,以 YOLOv5 检测算法为框架,设计了可以自适应提取深层语义信息的自适应局部金字塔模块以及多尺度定位损失函数,最后获得了 95.83% 的平均精度和 0.943 的召回率。以上对模型特征融合方法的改进虽然有效果,但是却不能自适应地融合多层次多尺度语义信息。程擎等^[15]也在 YOLOv5 网络上进行改进,研究了注意力机制对于模型检测较小异物效果的提升效果,加入有效通道注意力机制模块并替换了基础网络的特征融合步骤,提高了提取全局特征的能力,使优化后的模型平均精度提升 1.7%,并通过修改损失函数加强模型的收敛能力;Ren 等^[16]提出双向流通 YOLO 模型(bidirectional YOLO, Bi-YOLO),在 YOLOv5 中添加一个改进的多尺度融合网络(path aggregation network, PANet),使网络能够自适应地提高对小对象的关注,并使用无锚方式及多分支头来降低模型的复杂度并提高模型检

测小目标的性能,准确率最终提升了 3.86%。以上模型针对较小异物的改进虽有效果,却依然无法避免下采样损失较小异物特征,无法从根本上解决问题。

可以看出 YOLO 系列方法可以在检测 FOD 的任务上发挥很好的效果,但也存在一定缺陷,同时,针对且近些年随着更多模型的提出,为 FOD 的检测提供了更多方法,YOLOv8^[17]网络就是其中效果较好的一种。但是在实际研究中发现检测 FOD 时,基础 YOLOv8 网络在检测机场跑道异物时由于某些异物体积较小、空间位置随机及异物间尺度差距大容易产生的漏检误检问题。针对此类问题,本文提出应用先验掩膜的针对小目标的 YOLOv8 网络模型(apply the mask to mapping optimize the small target-YOLOv8, AMMS-YOLOv8)其主要进行以下四点创新:

1) 整合主干网络第 0、1 层的两层卷积并提出边缘信息敏感处理模块(edge information examination sensitive to the isotropic edge processing module, EIEStim),首次尝试使用 Isotropic Sobel 算子提取目标边缘信息,并与空间信息对照,用其增强模型微小边缘感知能力。

2) 面对较小的异物,本文创新性地提出可以保护语义信息的轻量化自适应权重采样模块(lightweight adaptive weight downsampling module trobust feature downsampling, LDFDS)替换主干中的下采样卷积层,该模块将接受场注意力机制(receptive-field attention, RFA)^[18]应用在检测领域,并针对 RFA 检测不同尺度目标注意力不足的问题,修改其结构并加入深层鲁棒特征采样(deep robust feature downsampling, DRFD)^[19]的三支处理结构,用以保护语义信息的同时提升检测性能。

3) 针对多尺度异物,YOLOv8 网络的 Neck 层不能大量综合上下文的多尺度信息,导致对浅层网络提取到的特征利用效果不佳,本文创新性地提出上下文汇聚的图像金字塔特征融合网络(contextual convergence feature pyramid networks, CCFPN),提出使用汇集结构增强多个不同层级间语义联系,并尝试选用尺度感知选择性继承模块(dimension-aware selective integration DASD)^[20]作为各层语义汇聚节点,希望可以增强多尺度异物的检测能力。

4) 最后针对模型检测随机空间异物能力欠缺的问题,本文借鉴并改变大核选择卷积网络思想(large selective kernel network, LSKNet)^[21],提出了一个特征先验掩膜融合检测头(fusion mask and selection network-head, MSN-Head)以增强模型对异物的空间检测能力,该检测头开创性地将掩膜信息加入检测头,由掩膜图像携带空间特征,直观突出异物目标,再通过 Transformer 模型中提出的自注意力提取掩膜特征,并将其与图像特征加权融合,以提升整体模型效果。

本文希望通过改进的 AMMS-YOLOv8 更好地对机场跑道异物进行检测,并保持较小参数量以适配搭载在

硬件设备上进行检测。

1 AMMS-YOLOv8 算法

面对传统YOLOv8算法上下文多尺度语义信息联系

不足,面对较小异物识别效果不佳,对检测目标空间感受能力不足等问题,本文改进YOLOv8基础结构,提出AMMS-YOLOv8算法,算法整体结构如图1所示,其中改进部分用红色标出。

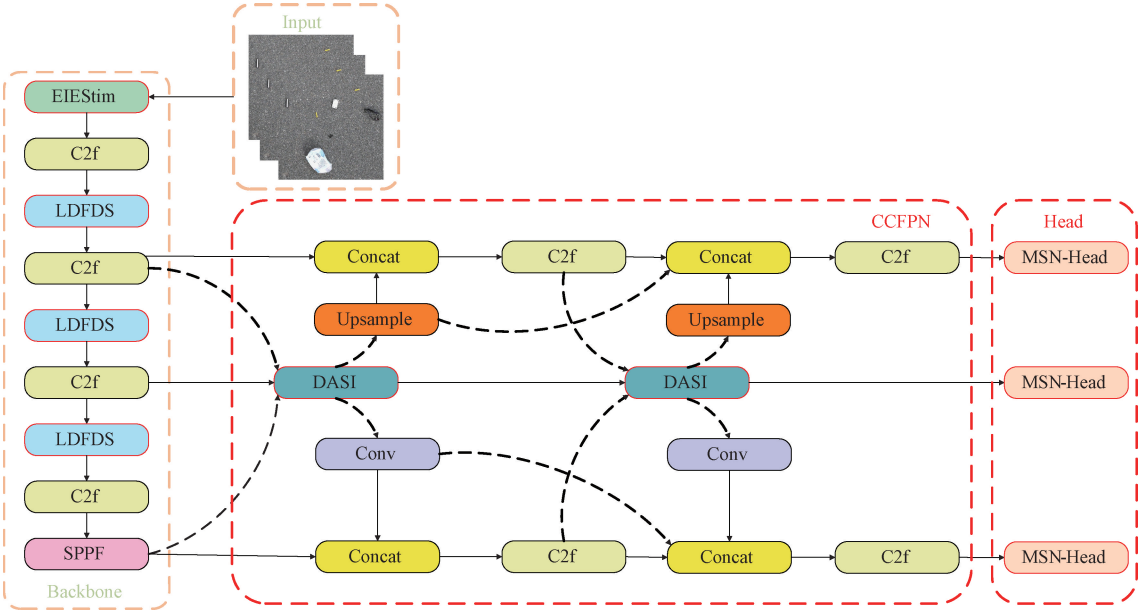


图1 改进的AMMS-YOLOv8网络结构

Fig. 1 Improved AMMS-YOLOv8 network structure

其中,在第0层,通过边缘信息各向同性敏感的处理模块,可以确保模型在目标语义未受下采样损失时更好地提取细微特征;在下采样中,通过保护语义信息的轻量化自适应权重采样模块可以在下采样过程中尽量少损失细微特征,便于接下来的信息融合;在语义融合阶段,通过聚合多尺度特征,改变原本简单的信息传播方式,增强特征融合能力;在检测头中,开创性地融合先验掩膜信息,提取异物空间信息,通过加权融合来增强模型空间位置感受能力,具体改进原理以下将分点介绍。

1.1 边缘信息各向同性敏感处理模块

在FOD的检测中,经常会遇到较小异物,其占据像素较少,且易存在目标与背景之间视觉差异较弱的情况,而检测边缘可以突出异物与背景的差异,有效解决这一问题。针对该情况,本文尝试在浅层网络中先学习目标边缘特征,Sobel算子以其较快的计算效率和较好的抗噪性适合提取异物的边缘信息,在异物较小且难以与背景区分的情况下进行识别。

然而,传统Sobel算子虽然可以提取边缘特征,但却因为其明确的各向异性导致其会忽略不规则异物的部分边缘信息,而FOD又大多不具有规则形状,所以改善算子的各向异性非常重要。所以本文首次应用各向同性Isotropic Sobel算子提取异物边缘信息,提出对目标边缘各向同性敏感的信息处理模块(EIEStim),其算子结构如式(1)~(2)所示。

$$\text{Isotropic Sobel-x} = \begin{pmatrix} -1 & 0 & 1 \\ -\sqrt{2} & 0 & \sqrt{2} \\ -1 & 0 & 1 \end{pmatrix} \quad (1)$$

$$\text{Isotropic Sobel-y} = \begin{pmatrix} 1 & \sqrt{2} & 1 \\ 0 & 0 & 0 \\ -1 & -\sqrt{2} & -1 \end{pmatrix} \quad (2)$$

其中,式(1)为横向Isotropic Sobel算子,式(2)为纵向Isotropic Sobel算子,分别关注模型 x 与 y 方向特征。

EIEStim由类残差结构组成,兼顾图像边缘及空间信息,具体结构如图2所示。由输入端将图像特征 X 输入网络主干部分,其经过一个 3×3 大小的卷积后可以将其切分为两个分支,分支一被输入SobelConv结构,其由3D组卷积实现,用以特定提取边缘特征,为保持其空间尺度不变,提前对特征图进行填充;分支二被送入最大池化层以保留图像主要特征。

在分支一中,输入特征经过水平与垂直的各向同性SobelConv提取得:

$$F_x = 3DConv_x(\text{Isotropic Sobel-x}(\text{Conv}_{3 \times 3}(X))) \quad (3)$$

$$F_y = 3DConv_y(\text{Isotropic Sobel-y}(\text{Conv}_{3 \times 3}(X))) \quad (4)$$

其中,式(3)为水平方向各向同性算子提取流程, F_x 为水平边缘特征, $3DConv_x$ 表示水平方向组卷积操作, $\text{Conv}_{3 \times 3}(X)$ 表示第一分支的输入特征,对原特征进行 3×3 卷积处理,Isotropic Sobel-x为 x 方向各向同性提取边

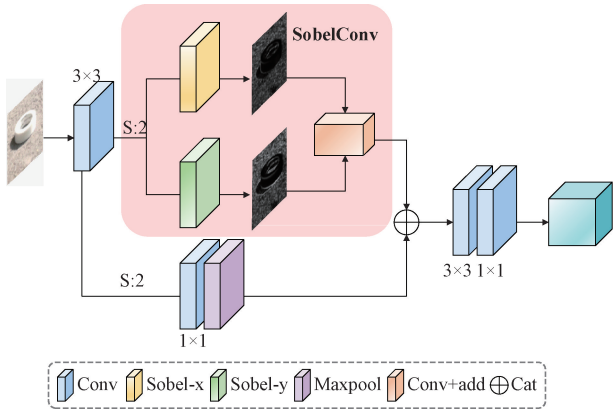


图 2 EIEStim 模块结构

Fig. 2 EIEStim module structure

缘信息,式(4)为同操作 y 方向处理。

随后,将两方向边缘信息在通道上进行拼接,使模型可以综合学习各向边缘信息,同时压缩通道,减少参数数量,提高计算效率,得到 SobelConv 的结果 E 为:

$$E = \text{Conv}_{1 \times 1}(\text{cat}(F_x + F_y)) \quad (5)$$

在第二分支中,输入特征信息被送入最大池化层进行操作,用以保留主要特征,在后面与 SobelConv 得到的边缘特征结合,其保留的特征 W 为:

$$W = \text{Max pool}(\text{Conv}_{1 \times 1}(\text{Conv}_{3 \times 3}(X))) \quad (6)$$

在双分支模式中,第一分支采取 3 D 组卷积高效地实现算子,降低了计算复杂度的同时尽量减少信息损失,快速提取边缘信息;第二分支通过最大池化保护浅层丰富特征,并最终与边缘信息结合,得到最终输出 X_{edge} ,使模型从空间整体与边缘局部多角度理解图像内容。

$$X_{edge} = \text{Conv}_{1 \times 1}(\text{Conv}_{3 \times 3}(\text{cat}(E + W))) \quad (7)$$

1.2 保护语义信息的轻量化自适应权重下采样模块

在 YOLOv8 的特征学习阶段,模型普遍使用步长较大的卷积,这可能导致一些细节特征的丢失,因此空间注意力机制被提出用来替换此类卷积,提高网络性能。但是在应用空间注意力提取信息时其产生的注意力图像往往由于权值共享不能针对每个通道单独产生独立的注意力图像,缺少丰富的信息,从而影响最终效果。

为了解决权值共享及特征损失,本文综合接受场注意力机制(RFA)和深层鲁棒特征采样(DRFD)提出可以保护语义信息的轻量化自适应权重采样模块(LDFDS),其包含了一个可以保护语义的特征采样结构替换基础大步长卷积,接下来分整体及特征采样进行介绍。

1) LDFDS 整体结构:因为卷积运算依赖共享参数提升卷积神经网络的性能,所以以往的优化方法也会具有共享参数的缺点,对位置变化带来的信息差异不敏感,无法获得更丰富的特征信息。但 RFA 则不同,其能够通过引入多感受野生成特征注意力图的方法为每个通道生成特征权重,从而生成多样化的特征注意力图,并可以通过强调

接受滑块中不同特征的重要性做到优先考虑接受空间中较为重要的特征,以此来解决这个问题,LDFDS 的模块结构如图 3 所示。

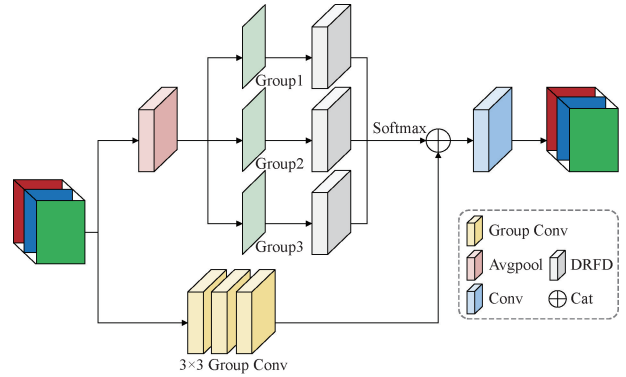


图 3 LDFDS 模块结构

Fig. 3 LDFDS module structure

在 LDFDS 中,模块在获得输入 X 之后,使用平均池化来聚合三通道每个接收域特征的全局信息,通过交互接受场特征信息来在每个通道上生成有特征性的注意图像,以此来提高网络性能,得到的特征注意力图像为 M_f 。

$$M_f = \text{Avg pool}(\text{DRFD}(X)) \quad (8)$$

式中: Avg pool 为平均值池化操作, DRFD 为可以保护丰富语义的深层鲁棒特征采样,其基础结构如图 4。最终在得到多通道独特的特征注意力图像后,采用 1×1 的组卷积整合各通道信息并使用归一化指数函数(Softmax)来强调每个特征在接受场特征中的重要性,生成特征权重。Softmax 的计算原理如式(9)所示,起到归一化的作用。

$$\text{Softmax}(x_i) = \frac{e^{x_i}}{\sum_{i=1}^n e^{x_i}} \in (0, 1) \quad (9)$$

通过该方法,可以得到的特征化权值 δ 更适合具体任务,权值 δ 的生成方法如式(10)所示,其中的 M_{fR} 、 M_{fG} 和 M_{fB} 为三通道生成的特征注意力图像。

$$\delta = \text{Softmax}(\text{Conv}_{1 \times 1}(\text{cat}(M_{fR} + M_{fG} + M_{fB}))) \quad (10)$$

在生成特征化权值 δ 的同时,LDFDS 将通过 3×3 卷积替代 Unfold 操作,通过无参数的方式提取感受野空间特征并引导到接受场空间,优先考虑感受空间特征。在 1×1 卷积整合通道后得到最终输出 X_f 。

$$X_f = \text{Conv}_{1 \times 1}(\delta \times \text{Group Conv}_{3 \times 3}(X)) \quad (11)$$

2) 保护语义的特征采样结构:在特征注意力图向的生成过程中,用到了可以保护丰富语义的深层鲁棒特征采样(DRFD),其结构如图 4 所示。

输入的特征 X 自上而下分别分别通过 L、M、C 三条通路并最终拼接结合,其中 L 执行卷积采样,由逐深度卷积(DWConv)与高斯误差线性单元(GELU)组成,DWConv 独立提取各通道的空间特征,在保留通道独立性的前提下

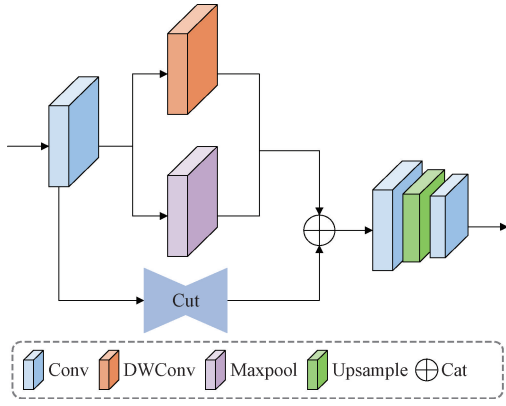


图4 DRFD模块结构

Fig. 4 DRFD module structure

提取空间特征,GELU涵盖了具有高斯分布的累积分布函数,以引入平滑的非线性,更适合深层网络训练,以融合图像局部信息,提高模型的泛化能力; M 执行最大值池化(Maxpool)进行空间采样,以确保避免关键特征的损失; C 执行切片采样,以裁剪模块为核心组成裁剪采样分支,隔行隔列裁剪 $F^{h \times w \times c}$ 的原图像得到4张 $F^{h/2 \times w \times c}$ 的子特征图并进行拼接,使通道数扩张为 $4c$,有效减少特征图尺寸的同时又保留原始特征信息。得到的最终结果 X_{DRFD} 如式(12)~(13)所示。

$$X_{DRFD} = Conv_{1 \times 1} \left(Upsample \left(Conv_{3 \times 3} \left(cat \begin{pmatrix} L \\ M \\ C \end{pmatrix} \right) \right) \right) \quad (12)$$

$$L + M + C = DWConv(Conv_{3 \times 3}(X)) + Maxpool(Conv_{3 \times 3}(X)) + cut(Conv_{3 \times 3}(X)) \quad (13)$$

LDFDS的设计极大程度增强模型在不同阶段的捕获能力,使模型生成的特征参数忽略复杂背景干扰,更关注细节特征,生成可以保护细节语义信息的注意力图,并解决基础注意力机制及传统RFA的固有缺陷。

1.3 上下文汇聚的图像金字塔特征融合网络

YOLOv8使用路径聚合网络,该网络采用先上采样再下采样的自上而下和自下而上两种路径,这样的结构固然可以融合顶层和底层的特征信息,使语义较为丰富的顶层信息可以参与交互融合,增强融合效果,但由于这样的结构各层级信息较为独立,上下文语义融合效果不佳,且FOD本身尺度及遗留位置随机并掺杂有遮挡情况,所以原本的特征融合网络显得融合能力不足。

针对此类问题,本文尝试Neck层重新设计,提出汇聚金字塔结构,与原本的路径聚合网络相比,增加横向跳连及多分支融合,改变原本线性的交互融合结构,可以同时综合3个不同尺度特征信息,切片融合,增强上下文语义联系,提高网络在复杂场景下对较小目标的检测效果。如图5为多尺度汇聚金字塔结构示意图。

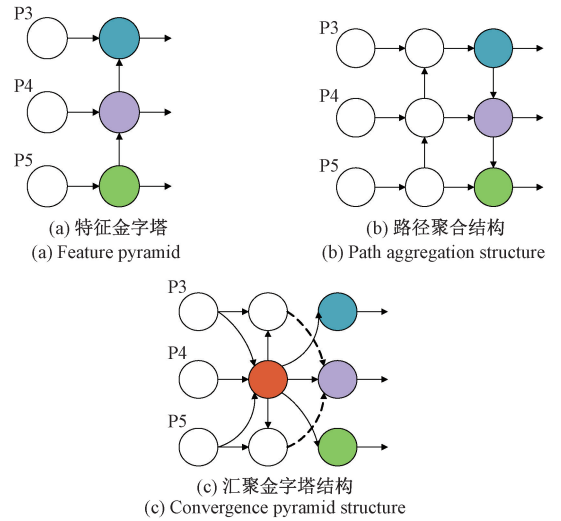


图5 颈部网络不同结构

Fig. 5 Different structures of the neck network

可以看到在原本前向传播的基础上,汇聚金字塔可以将P3、P4、P5的特征聚合到同一节点,单一节点中就可以使深层浅层与中层信息充分融合,图5中红色为该节点,应用到了尺度感知选择性继承模块(DASI),其结构如图6所示,可以根据物体的尺寸和特征自适应选择融合的特征,应用DASI用作汇集节点的选择分配器,可以同时融合多个尺度特征信息。

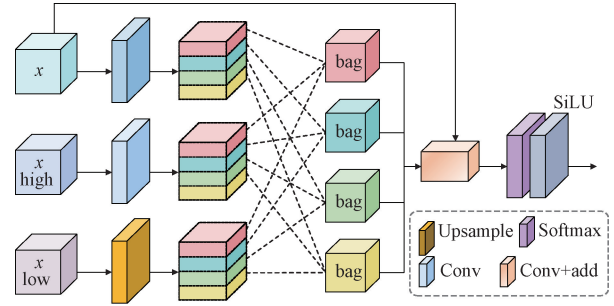


图6 DASI模块结构

Fig. 6 DASI module structure

在DASI中,面对输入的3个不同尺度特征,DASI将其视为高维、中维、低维3种特征并对低维输入 $X_l \in R^{l_h} \times W^l \times C^l$ 进行上采样,对高维输入 $X_h \in R^{h_h} \times W^h \times C^h$ 进行 3×3 的卷积降采样,对中维输入 $X_u \in R^u \times W \times C$ 进行 1×1 的整合操作,进行维度初步对齐,该过程如式(14)所示。

$$\begin{aligned} X_l &= Upsample(X_l) \\ X_u &= Conv_{1 \times 1}(X_u) \\ X_h &= Conv_{3 \times 3}(X_h) \end{aligned} \quad (14)$$

对齐后再通过切分的操作将各维度分别划分为4个相等的部分,设计专门的打包(bag)操作,其可以通过中维的特征得到权重分配从而将3个维度中同一层级的切分片段加权融合在一起,其融合公式为:

$$F = \text{sigmoid}(F_u) \times F_l + (1 - \text{sigmoid}(F_u)) \times F_h \quad (15)$$

其中, F 为最终生成的融合输出特征图, 这种结构增强了不同维度之间的特征融合效果, 有效地整合来自不同尺度的特征信息, 增强特征表示能力。通过这种方式就可以做到将多个尺度特征融合。使用 DASI, 可以提出上下文汇聚的图像金字塔特征融合网络 (CCFPN) 具体结构如图 7 所示。

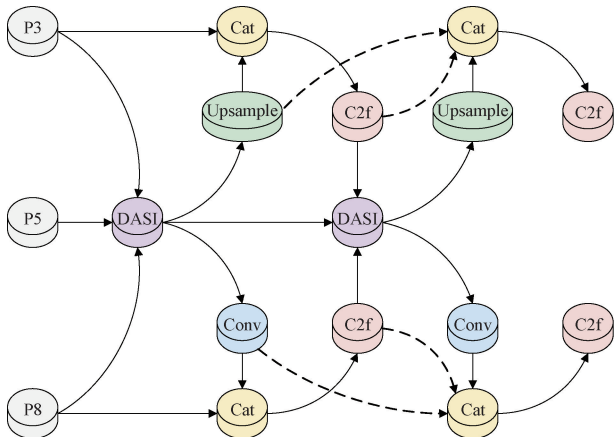


图 7 CCFPN 整体结构图

Fig. 7 Diagram of the overall structure of CCFPN

在原本的 YOLOv8 网络中, 模型选用第 3、5、8 三层的 80×80 、 40×40 、 20×20 特征输入 Neck 层进行特征融合, 在改进的特征融合金字塔中, 同样选用该三层特征, 将该三层不同尺度特征输入 DASI, 切片融合后, 又分别对融合后的特征进行上下采样, 恢复其原本维度并与输入的对应维度特征进行拼接。以低维特征为例, 在第一次拼接时其结果如式 (16) 所示。

$$X_{cat_1} = \text{cat}(X_l + \text{Upsample}(DASI_{X_{l-1}})) \quad (16)$$

式中: X_{cat_1} 为第一次拼接后的结果, $DASI_{X_{l-1}}$ 为经第一次多维特征融合输出的结果。

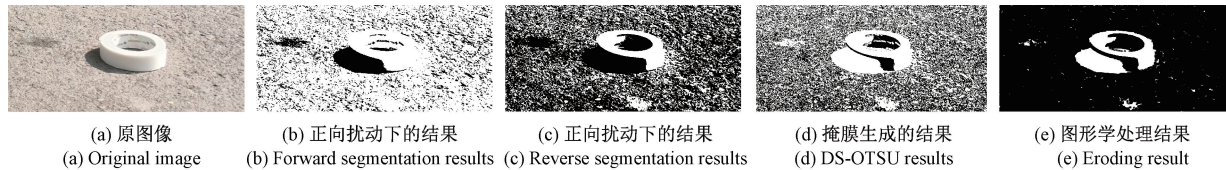


图 8 掩膜生成步骤及结果

Fig. 8 Mask generation steps and results

得到异物图像的掩膜之后, 本文受轻量化大核选择卷积网络 (LSKNet) 中选择融合思想的启发, 将原本的多个大核卷积特征融合改变为掩膜与输入特征进行特征融合, 建立特征先验掩膜融合检测头 (MSN-Head) 结构。改进后的检测头结构如图 9 所示, 可以看出 MSN-Head 检测头与传统检测头相比, 是在传统的检测头前添加了 MSN 模块,

在第一次多维特征融合后, 可以进行第二次特征融合, 将已经恢复维度特征的信息再次汇聚进入 DASI, 切片融合后, 得到第二次融合后的结果, 将该结果分别与第一次融合后的结果和第一次拼接后的结果进行拼接。以低维特征为例, 第二次拼接后的结果如式 (17) 所示。

$$X_{cat_2} = \text{cat}\left(\begin{matrix} X_{cat_1} + \text{Upsample}(DASI_{X_{l-1}}) \\ \text{Upsample}(DASI_{X_{l-2}}) \end{matrix}\right) \quad (17)$$

式中: $DASI_{X_{l-2}}$ 为经第二次多维特征融合输出的结果。这种汇集的特征融合方法改变了原本只能自上而下或自下而上的特征融合方式, 实现了层级之间信息的跳连融合。删除单一的输入输出节点, 增加同一层级下接受及输出特征的通道数, 使模型可以兼顾更多信息, 减少信息丢失。

随着训练深入, 这样的模型结构可以在尽量不增加模型大小的情况下更加准确地挖掘和整合来自不同层级的丰富信息, 对比各类金字塔结构可以看出本文的 Neck 层结构更紧凑, 可以更高效地聚合特征。

1.4 特征先验掩膜融合检测头

YOLOv8 网络有 3 个尺度的检测头, 用于检测不同尺度的物体, 考虑到 FOD 在实际检测过程中位置较为随机且可能存在与背景较为相似的异物, 对于此类异物原网络的检测头空间感知能力不足, 检测效果可能受限, 本文考虑应用改进的 OTSU 法先得到异物的掩膜, 并在 Head 层将此掩膜作为先验信息与特征进行融合, 增强模型的空间感知能力。

通过双通道麻雀该改进 OTSU 的方法^[22]可以生成一个异物的掩膜, 该方法在双通道内并行运行, 用麻雀搜索算法优化 OTSU 的计算速度, 并向类间方差的计算方法中分别引入正反向扰动得到用于对比输出的正反向结果图, 对比两结果图可以得到异物的分割图像, 该方法应对与背景相似的异物时效果较好。如图 8 为浅色道路上白色胶带的分割效果, 可以看出分割后的掩膜结果可以明显地突出异物的空间位置, 使提取掩膜信息增强检测效果成为可能。

其由两个全连接层和一个 GELU 子块、一个 MSNBlok 子块串联组成。全连接层处理输入特征后, 交由 GELU 激活函数添加非线性模式, 之后特征进入 MSN-Blok 子块, 该子块负责应用先验掩膜提取异物空间位置, 其分为双通道特征输入、特征选择融合两部分, 其结构如图 10 所示。

MSN-Blok 子块由 LSKNet 架构改进而来, 在原本架

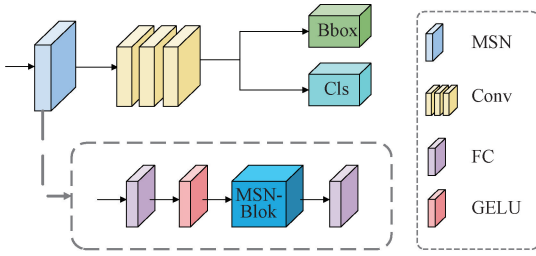


图9 MSN-Head检测头结构

Fig. 9 MSN-Head inspects the head structure

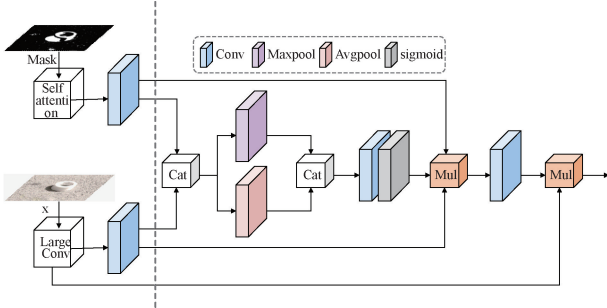


图10 MSN-Blok模块结构

Fig. 10 MSN-Blok module structure

构中,模型应用独立的不同大小的大核卷积对输入特征分别进行处理,通过筛选可以自适应地得到最适合不同数据图片的卷积核大小,从而规避单一卷积核大小不能应对所有检测目标的情况。在MSN-Blok中,该思想则被用来动态选择融合掩膜信息与输入信息。

MSN-Blok同时接受异物掩膜及模型特征信息两个输入,对于输入的特征信息,对其应用 7×7 大核卷积扩张感受野,强化模型对较小目标的检测能力;对于掩膜图像,则应用较为敏感的Selfattention提取异物的位置及空间信息,之后对两输入信息进行与通道数等同次数的切分及连接,得到处理后的特征 X_d 便于之后筛选操作的进行。

$$X_d = \text{cat} \left(\begin{array}{l} \text{Conv}_{1 \times 1}(\text{Large Conv}(X)) + \\ \text{Conv}_{1 \times 1}(\text{Selfattention}(X_{Mask})) \end{array} \right) \quad (18)$$

式中: X 为输入的特征信息, X_{Mask} 为掩膜图像。

在对输入的图像特征与掩膜特征进行提取拼接之后,再将 X_d 传入筛选网络进行处理,将拼接后特征并行传入并在空间维度上进行最大值池化及平均值池化,将双通道得到的特征进行拼接,经过卷积层生成注意力图像,再通过sigmoid函数归一化空间维度上结合掩膜的注意力图像得到特征权重 φ_1 。

$$\varphi_1 = \text{sigmoid} \left(\text{Conv} \left(\text{cat} \left(\begin{array}{l} \text{Max pool}(X_d) + \\ \text{Avg pool}(X_d) \end{array} \right) \right) \right) \quad (19)$$

通过得到结合掩膜的注意力特征权重 φ_1 并用其处理原始输入 X 与 X_{Mask} 可以得到输入图像中空间特性重要性区域,并通过将该特性与原输入相乘,对掩膜与输入特征进行分配融合最终达到应用先验掩膜信息增强模型空间

感知能力的效果,通过卷积乘法计算得到最终输出 X_p 。

$$X_p = \text{Conv}(\varphi * X * X_{Mask}) * X \quad (20)$$

2 实验设置与结果

2.1 数据集介绍

目前学术范围内虽然针对机场跑道异物检测任务提出了一些数据集,如FOD-A^[23]数据集,其包含了31种不同类型的异物,异物种类及数量如图11所示,但由于机场环境特殊及特殊管制原因,公开的数据集背景过于简单且分辨率较低,其每张图片只有单一目标,缺乏复杂性,无法适应复杂检测任务,所以目前并没有一个统一有效的应对复杂机场跑道的异物数据集。

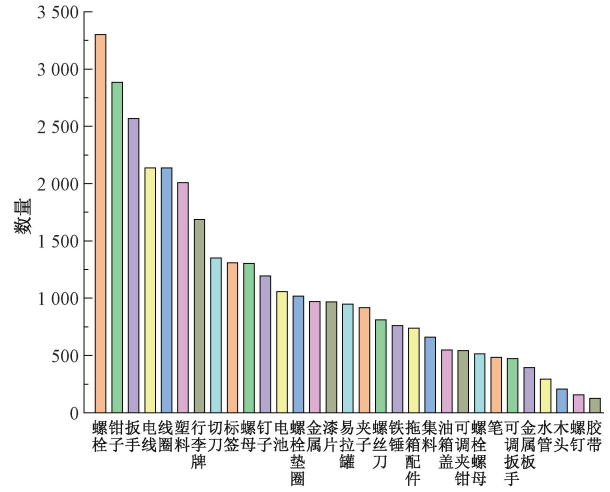


图11 FOD-A数据集中各类异物分配情况

Fig. 11 Allocation of various types of foreign objects in the FOD-A data set

基于此背景,本文针对混合复杂场景,依照美国联邦航空局(FAA)通告手册和各航司对于FOD类型的统计报告,构建了一个复杂场景异物数据集(foreign object datasets for complex scenarios, FCS-FOD),以提高数据集的鲁棒性。该数据集包含螺杆、坚果、备件、铁棒、硬币、石头、指甲、塑料袋、塑料钉、金属管、连接线、钳子、钣金、金属垫片、金属工具这15类常见的FOD类型,共计5000张图片。采集图片数据时分别在中午及傍晚进行采集以提升数据集应对不同光照情况下的泛化能力,并注重涵盖各种尺寸的FOD,以便检测模型对较小目标的检测能力,并将其分为低危、中危、高危3类,在背景中的分布情况如图12所示,可以看出异物为较为均匀的分布。在图13中展示了在各种光照条件下的数据集图片,在图14中展示了针对数据集生成的特征掩膜数据。采用LabelImg标注工具统一标注为PASCAL VOC格式并按照1:1:8的比例划分成为验证集、测试集和训练集。

由于用于验证模型性能与泛化性的数据集背景大多没有干扰,无法体现模型鲁棒性,所以本文增添了鲁棒性验证

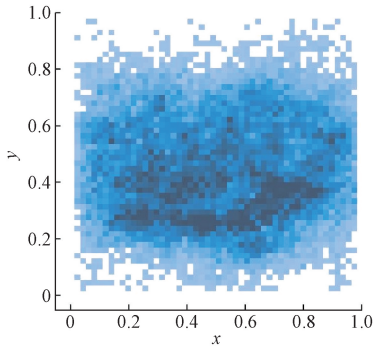


图 12 FCS-FOD 数据集中异物分布情况

Fig. 12 Distribution of foreign bodies in FCS-FOD datasets

具体信息如表 2 所示。



图 15 输电线路异物数据集示例

Fig. 15 Example of a transmission line foreign body dataset

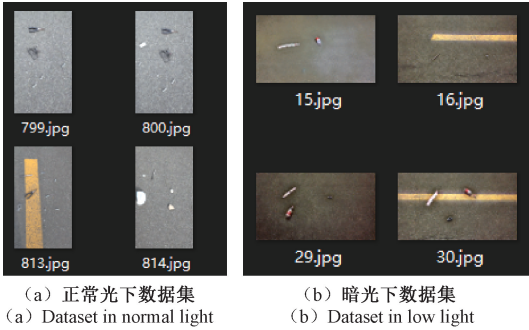


图 13 数据集图片展示

Fig. 13 Dataset Image Exhibition

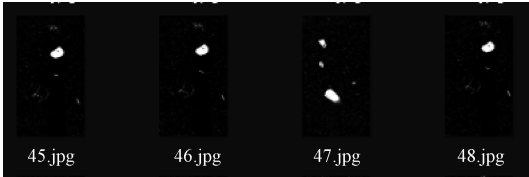


图 14 特征掩膜图片展示

Fig. 14 Feature mask visualization exhibition

实验,但是由于机场 FOD 数据集较少且较多处于无干扰的机场背景,所以选取输电线路异物检测公开数据集^[24]做鲁棒性验证。数据集共 4 516 张图片,光照情况及背景较为复杂,包含 4 种输电线路异物常见的类别,分别是气球、鸟巢、风筝以及飘浮物垃圾,其中鸟巢与飘浮物垃圾为较大异物,气球与风筝为较小异物,数据集示例如图 15 所示。

本文选取构建的复杂场景异物数据集(FCS-FOD)验证模型有效性。在泛化性验证试验中,由于 FOD-A 数据集缺乏复杂性且场景过于简单,所以本文采用向复杂场景异物数据集中增加 FOD-A 数据集的方法构建混合数据集,保证数据集的复杂性并以此验证模型的泛化性。同时本文采用输电线路异物公开数据集验证鲁棒性。

2.2 实现细节

本文实验所使用的实验设备为个人移动式电脑,型号为联想拯救者 y7000p,实验平台选用 Pycharm 2023a,硬件配置如表 1 所示,模型训练调参时所使用的一些超参数的

表 1 实验所处硬件环境配置

Table 1 The configuration of the hardware environment where the experiment is located

配置	参数
操作系统	Windows1164 位
CPU	Intel(R)Core(TM)i7-14700HX
GPU	NVIDIARTX4070
显存	16 GB
GPU 加速环境	CUDA11.8
网络架构	PyTorch
开发语言	Python

表 2 实验运行的超参数设置

Table 2 Hyperparameter settings for the experiment run

类别	参数
训练周期(epochs)	100
线程数(workers)	4
批量大小(batch)	16
优化器(optimizer)	SGD
初始学习率(lr0)	0.01
权重衰减系数(weight)	0.000 5

2.3 评价指标

FOD 检测属于目标检测领域,目前目标检测较为权威的评价指标有按 0.5 的交并比(IoU)计算的平均精度(mAP50)、按 0.5-0.95 的交并比(IoU)计算的平均精度(mAP50-95)、准确率(P)、召回率(R)以及 F1 函数(F1)。其中 mAP50 和 mAP50-95 分别表示的是模型在面对不同类型异物时所展现出来的检测精度,而 P、R、F1 三个指标

展现的是面对一类异物时,模型的查精能力、查全能力和均衡查精与查全的综合能力,为比较模型计算速度,同时选取GFLOPs以及处理一张图片所用时间做评价指标。

2.4 实验结果以及分析

为进一步检验AMMS-YOLOv8的性能优势,本文选取了具有代表性的检测算法与本文提出算法针对检测FOD得到的效果进行对比。包括Fast-R-CNN^[25]、SSD^[26],以及YOLO系列算法:YOLOv5^[27]、YOLOv6^[28]、YOLOv8以及最新出现提出的YOLOv9^[29]、YOLOv10^[30]、YOLOv11^[31]模型,以及近年邹松林在

文献[13]中提出的改进的YOLOv5算法^[13]、张敬博在文献[14]提出的改进YOLOv5算法^[14]在相同的实验环境下应用相同的FCS-FOD数据集进行测试,为比较算法的效果,选取map50、mAP50-95、准确率(P)、召回率(R)、F1函数、GFLOPs以及处理一张图片所用时间做评价指标,并将基础网络与AMMS-YOLOv8结果用粗体展示,在FCS-FOD数据集的实验结果如表3所示。图16为记录在该数据集下各个检测模型在训练的每一个轮次下获得Map50值绘制成的折线图,可以用以比较模型收敛速度及最终效果,同时在后文增添泛化性及鲁棒性验证环节。

表3 FCS-FOD数据集下主流检测算法结果对比

Table 3 Comparison of the results of mainstream detection algorithms under the FCS-FOD dataset

模型	mAP50/%	mAP50-95/%	准确率(P)	召回率(R)	F1函数	GFLOPs	Speed/ms
Faster R-CNN	86.7	45.3	0.853	0.869	0.860	126	30.2
SSD	80.5	44.9	0.810	0.801	0.805	136	18.8
YOLOv5	95.4	53.3	0.954	0.960	0.956	7.1	6.2
YOLOv6	94.4	53.0	0.953	0.953	0.953	11.8	4.9
YOLOv8	96.0	53.8	0.956	0.961	0.958	8.1	5.8
YOLOv9	95.5	54.3	0.951	0.960	0.955	26.8	7.3
YOLOv10	94.4	52.3	0.951	0.949	0.949	6.5	5.2
YOLOv11	95.1	54.1	0.935	0.946	0.941	6.3	6.7
文献[13]算法	95.9	53.0	0.961	0.963	0.962	7.7	6.8
文献[14]算法	96.1	52.6	0.952	0.97	0.961	8.1	6.4
Ours	97.8	55.5	0.971	0.976	0.973	27.4	6.3

从图表中可以看出,在用Faster R-CNN和SSD检测FOD时,其检测效果不佳且整体落后于YOLO系列算法。面对YOLO系列目标检测算法,由于v5、v6等版本提出时间过早,模型优化不足,所以在进行FOD检测时得到的检测精度不足,在mAP50和mAP50-95的指标下大都落后于v8版本,而v9、v10及v11两种模型虽然提出时间比较靠后,但是其只对特定数据集效果较好,在面对其他数据集时效果不及技术完善得较为成熟的v8版本,两个近年提出的针对FOD检测任务的优化算法由于其基础网络是YOLOv5,所以精度不高,本文是第一个将YOLOv8用在FOD检测上的尝试。

而本文提出的AMMS-YOLOv8模型在检测精度的指标上,mAP50相较于传统YOLOv8模型提升了1.8%,mAP50-95相较于传统模型提升了1.7%,这表明AMMS-YOLOv8对于不同种类大小的异物都能得到较好的效果,具有很好的鲁棒性和泛化性。对于准确率、召回率和F1函数指标,本文算法相较传统网络模型分别提升了0.015、0.015和0.015。图17为AMMS-YOLOv8检测每一个FCS-FOD数据集中类别异物时相比原始网络获得的mAP50效果提升,可以看出由于添加了提升针对较小FOD目标检测效果的模块以及位置空间感知模块,原本检

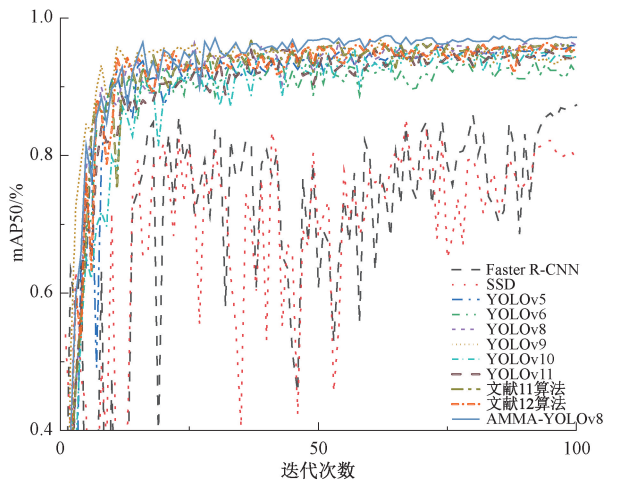


图16 各检测模型迭代曲线对比

Fig. 16 Comparison of the iterative curves of each detection model

测效果较差的较小异物获得了更好的检测效果。

为了直观对比本文改进算法相较于原始YOLOv8网络检测性能的提高,本文选取数据集内真实图片测试集样本对两种算法做定性分析,并通过可视化展示检测效果对比,最终如图18所示。

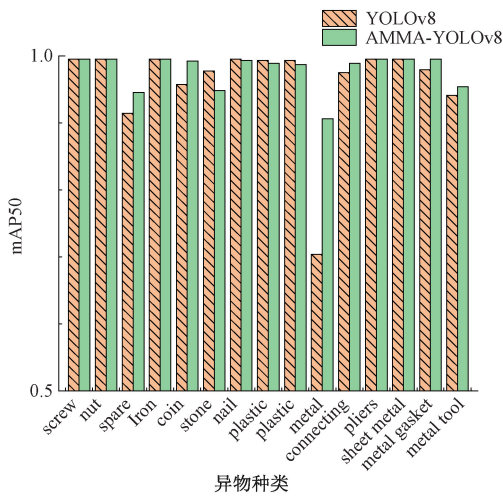


图 17 各类别异物检测 mAP50 对比

Fig. 17 Comparison of mAP50 for foreign body detection in each category

从图 18 中可以看出,对于其中图像 1 中的数据线,由

于该数据线打开放置,所以其目标较细,基础网络不能将其检测出,而 AMMS-YOLOv8 因为添加了提升较小目标检测效果的模块,所以可以将其识别。在图像 2 和图像 3 中一些较小异物,在基础网络下因为模型感知能力不佳,导致检测效果不好,AMMS-YOLOv8 对其进行优化,最终得到更精确的识别效果。在图像 4 和图像 5 中数据线呈团状放置,基础网络虽可以将其感知,但是却由于模型上下文语义描述较差,导致基础网络将其错检,在 AMMS-YOLOv8 的结果中,该错误被消除。在图像 6 中,出现了微小的螺丝被遮挡的情况,面对此类情况基础网络没有检测出螺丝,但 AMMS-YOLOv8 网络将其准确检测。在图像 6 中,这些都表明本文提出的 AMMS-YOLOv8 在检测 FOD 的效果上有很大提升。

2.5 泛化性验证

为了测试 AMMS-YOLOv8 模型的泛化性,本文向复杂场景异物数据集中增添 FOD-A 数据集,将 AMMS-YOLOv8 与其他主流模型在混合数据集中进行验证,最终得到的实验结果如表 4 所示。

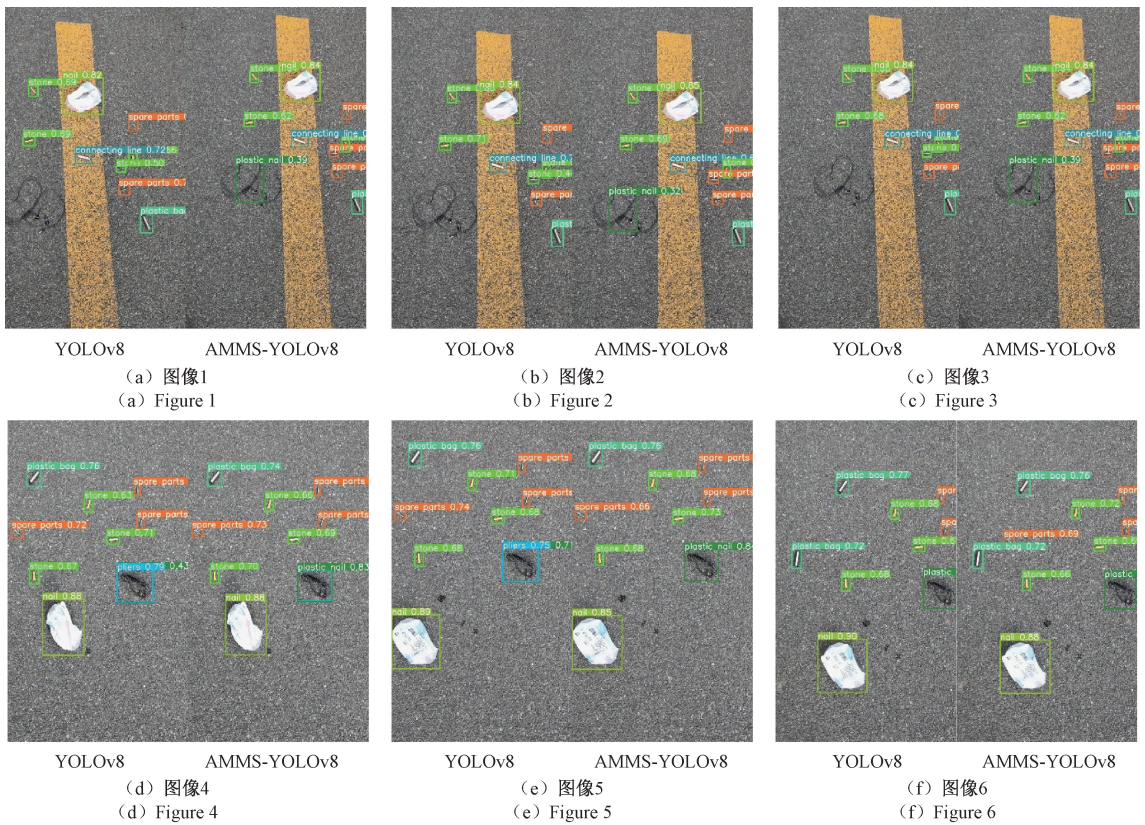


图 18 检测结果对比图

Fig. 18 Comparison chart of test results

表4 加入FOD-A的混合数据集下主流检测算法结果对比

Table 4 Comparison of the results of mainstream detection algorithms in the mixed dataset with FOD-A

模型	mAP50/%	mAP50-95/%	准确率(P)	召回率(R)	F1 函数	GFLOPs	Speed/ms
Faster R-CNN	78.4	42.8	0.785	0.780	0.782	126	30.2
SSD	82.3	58.3	0.819	0.825	0.822	136	18.8
YOLOv5	87.1	60.2	0.872	0.871	0.872	7.1	6.2
YOLOv6	87.3	59.8	0.884	0.870	0.877	11.8	4.9
YOLOv8	89.7	60.8	0.890	0.901	0.895	8.1	5.8
YOLOv9	88.6	61.1	0.882	0.890	0.886	26.8	7.3
YOLOv10	87.7	58.8	0.870	0.881	0.875	6.5	5.2
YOLOv11	89.5	63.1	0.901	0.889	0.895	6.3	6.7
文献[13]算法	91.4	65.0	0.905	0.909	0.907	7.7	6.8
文献[14]算法	93.7	63.8	0.950	0.914	0.932	8.1	6.4
Ours	95.4	69.7	0.951	0.960	0.955	27.4	6.3

从表4中可以看出,AMMS-YOLOv8网络在增添FOD-A数据集后也可以获得很好的效果,mAP50与mAP50-95两指标分别相对于基础网络提升了5.7%与8.9%,准确率、召回率及F1函数等指标均有所提升,由此可见本文模型也具有很好的泛化性。

2.6 鲁棒性验证

为了验证AMMS-YOLOv8模型的鲁棒性,本文使用输电线路异物数据集将AMMS-YOLOv8与其他主流模型在该数据集上进行验证,最终得到的实验结果如表5所示。

表5 输电线路异物数据集下主流检测算法结果对比

Table 5 Comparison of the results of mainstream detection algorithms under the transmission line foreign body dataset

模型	mAP50/%	mAP50-95/%	准确率(P)	召回率(R)	F1 函数	GFLOPs	Speed/ms
YOLOv5	96.4	67.1	0.954	0.942	0.949	7.1	6.2
YOLOv6	74.8	73.6	0.748	0.687	0.716	11.8	4.9
YOLOv8	96.2	69.7	0.936	0.947	0.941	8.1	5.8
YOLOv9	95.9	69.5	0.945	0.941	0.943	26.8	7.3
YOLOv10	95.2	66.9	0.939	0.930	0.935	6.5	5.2
文献[13]算法	96.7	68.6	0.964	0.966	0.965	7.7	6.8
文献[14]算法	96.5	70.1	0.963	0.953	0.958	8.1	6.4
Ours	97.6	71.4	0.971	0.928	0.950	27.4	6.3

从表5中可以看出,AMMS-YOLOv8网络在输电线路异物数据集中也可以获得很好地效果。由于改进了特征融合逻辑,输电线上不同尺度的异物不会影响模型检测,mAP50与mAP50-95两指标分别相对于基础网络提升了1.4%与0.7%,由此可见本模型具有很好地鲁棒性,可以应对复杂背景及多种光照情况。

2.7 消融实验

为了测试本文提出的模块是否对于提升YOLOv8网络检测FOD都有效果,本文应对了针对FOD的两个数据集采取消融实验的方法,在基础网络上逐渐加入LDFDS、EIEStim、CCFPN及MSN-Head四个模块,在双数据集中,以mAP50、mAP50-95、GFLOPs及运算每张图片所用时间为指标来测试各模块可以提升模型检测效果的能力,结果

如表6~7所示。

从表6~7中可以看出,本文以YOLOv8为基础网络在双FOD数据集中分别做了5组消融实验,在逐渐加入各个模块的过程中,由于LDFDS与EIEStim是针对小目标的优化,且FOD中小目标所占比例较多,所以在FCS-FOD数据集及混合数据集下LDFDS与EIEStim分别将mAP50提升了0.4%、0.5%及1.2%、2.3%。在添加了CCFPN后,由于其增强了模型多尺度特征融合能力,所以在双数据集中mAP分别提升0.5%及0.8%。最后在重构检测头,形成MSN-Head架构后,模型的空间特征表征能力被显著强化,具备了更准确定位异物的特性,在双数据集中mAP分别提升0.4%及1.4%。

表 6 FCS-FOD 数据集下消融实验结果对比

Table 6 Comparison of ablation experimental results under FCS-FOD dataset

基础模型	LDFDS	EIEstim	CCFPN	MSN-Head	mAP50/%	mAP50-95/%	GFLOPs	Speed/ms
YOLOv8	×	×	×	×	96.0	53.8	8.1	5.8
YOLOv8	√	×	×	×	96.4	52.8	7.9	6.8
YOLOv8	√	√	×	×	96.9	53.0	9.6	7.6
YOLOv8	√	√	√	×	97.4	54.1	10.4	6.9
YOLOv8	√	√	√	√	97.8	55.5	27.4	6.3

表 7 加入 FOD-A 的混合数据集下消融实验结果对比

Table 7 Comparison of ablation experimental results under the mixed dataset with FOD-A

基础模型	LDFDS	EIEstim	CCFPN	MSN-Head	mAP50/%	mAP50-95/%	GFLOPs	Speed/ms
YOLOv8	×	×	×	×	89.7	60.8	8.1	5.8
YOLOv8	√	×	×	×	90.9	60.5	7.9	6.8
YOLOv8	√	√	×	×	93.2	63.2	9.6	7.6
YOLOv8	√	√	√	×	94.0	65.9	10.4	6.9
YOLOv8	√	√	√	√	95.4	69.7	27.4	6.3

3 结 论

本文针对 FOD 的检测任务,选择 YOLOv8 为基础网络进行优化,提出 AMMS-YOLOv8 网络,并在双数据集中对比多种检测算法以测试其效果。本文首先针对原网络感知较小异物效果不佳的问题,本文综合各向同性的 Sobel 算子边缘检测的能力提出 EIEstim 模块通过增强感知异物边缘的方法提升检测效果。针对原模型对某些微小 FOD 检测出现漏检的问题,本文将接受场注意力机制用在 FOD 检测领域,并用可以保护小目标语义信息的 DRFD 模块替换普通卷积,组成 LDFDS 模块从而解决此类问题。针对模型检测多尺度异物效果差的问题,考虑原因是原模型各层信息比较独立,检测缺少上下文语义信息,本文重构 Neck 层的结构,形成汇集结构,增强信息的共享能力,并应用可以兼顾多尺度目标的 DASI 作为汇集节点的模块,增强模型检测能力。最后为了在空间尺度内有针对性地增强模型对异物的空间检测能力,本文用特定方法提前生成异物的掩膜,并结合 LSKNet 的思想用 Selfattention 模块提取掩膜信息以强化模型空间位置感知能力,将该信息与输入特征相结合组成融合先验掩膜特征的检测头(MSN-Head)。实验证明,本文提出的 AMMS-YOLOv8 模型较于原本 YOLOv8 网络在复杂数据集中, mAP50 检测指标提升 1.8%,同时 mAP50-95 提升 1.7%,且能保持处理时间基本不变,并能适配其他数据集,具有良好的泛化性及鲁棒性,综合效果表现优于主流各种检测网络,同时在今后的研究中可以考虑优化模型的处理速度,使其更加适配移动硬件设备。

参考文献

[1] 汤双霞. 一种毫米波雷达机场跑道异物检测算法[J].

激光与红外, 2022, 52(6): 820-826.

TANG SH X. An algorithm for detecting foreign objects on airport runway using millimeter wave radar[J]. Laser & Infrared, 2022, 52(6): 820-826.

[2] 王宝帅, 刘江洪, 郑小亮, 等. 基于特征谱特征的机场跑道异物分层检测算法[J]. 电子与信息学报, 2017, 39(11): 2690-2696.

WANG B SH, LIU J H, ZHENG X L, et al. Hierarchical detection algorithm of foreign bodies on airport runway based on characteristic spectrum[J]. Journal of Electronics & Information Technology, 2017, 39(11): 2690-2696.

[3] 徐尽达, 陈慈发, 张上. 基于轻量级算法的水上垃圾小目标检测研究[J]. 电子测量技术, 2024, 47(18): 145-154.

XU J D, CHEN C F, ZHANG SH. Small target detection of water surface garbage based on lightweight algorithm [J]. Electronic Measurement Technology, 2024, 47(18): 145-154.

[4] 孙铁强, 魏光辉, 宋超, 等. 基于 YOLO 的多模态钢轨表面缺陷检测方法[J]. 电子测量技术, 2024, 47(21): 72-81.

SUN T Q, WEI G H, SONG CH, et al. Rail surface defect detection method based on multimodal YOLO[J]. Electronic Measurement Technology, 2024, 47(21): 72-81.

[5] 顾杨海, 李富, 陈德基, 等. 基于多尺度特征融合与交互的路侧目标检测算法[J]. 电子测量技术, 2024, 47(23): 152-161.

GU Y H, LI F, CHEN D J, et al. Roadside object

- detection algorithm based on multi-scale feature fusion and interaction [J]. *Electronic Measurement Technology*, 2024, 47(23): 152-161.
- [6] 张宇,王延吉,马辉,等.面向无人机监控的动态多尺度目标检测模型的研究与实现[J].*电子测量技术*,2024,47(10):141-150.
- ZHANG Y, WANG Y J, MA H, et al. Research and implementation of dynamic multi-scale object detection model for drone surveillance [J]. *Electronic Measurement Technology*, 2024, 47(10): 141-150.
- [7] 曾祥津,刘耿焕,陈建明,等.基于多尺度分层残差网络的光学遥感图像微小目标检测[J].*光子学报*,2024,53(8):254-266.
- ZENG X J, LIU G H, CHEN J M, et al. Micro-target detection in optical remote sensing images based on multi-scale hierarchical residual networks[J]. *Acta Photonica Sinica*, 2024, 53(8): 254-266.
- [8] 王坤,丁麒麟.利用自适应融合和混合锚检测器的遥感图像小目标检测算法[J].*电子与信息学报*,2024,46(7):2942-2951.
- WANG K, DING Q L. Small target detection algorithm in remote sensing images using adaptive fusion and hybrid anchor detector [J]. *Journal of Electronics & Information Technology*, 2024, 46(7): 2942-2951.
- [9] 许程翔,赵明岩,梁喜凤,等.基于改进 YOLOv8 的轻量化甘薯品质分级实验研究[J].*实验技术与管理*,2024,41(6):47-56.
- XU CH X, ZHAO M Y, LIANG X F, et al. Experimental study on quality classification of lightweight sweet potato based on improved YOLOv8 [J]. *Experimental Technology and Management*, 2024, 41(6): 47-56.
- [10] 高民,陈高华,古佳欣,等. Flm-YOLOv8:一种轻量级的口罩佩戴检测算法[J].*计算机工程与应用*,2024,60(17):203-215.
- GAO M, CHEN G H, GU J X, et al. Flm-YOLOv8: A lightweight mask wear detection algorithm [J]. *Computer Engineering and Applications*, 2024, 60(17): 203-215.
- [11] 苏佳,贾泽,秦一畅,等.面向工业表面缺陷检测的改进 YOLOv8 算法[J].*计算机工程与应用*,2024,60(14):187-196.
- SU J, JIA Z, QIN Y CH, et al. Improved YOLOv8 algorithm for industrial surface defect detection[J]. *Computer Engineering and Applications*, 2024, 60(14): 187-196.
- [12] 窦智,高浩然,刘国奇,等.轻量化 YOLOv8 的小样
- 本钢板缺陷检测算法[J].*计算机工程与应用*,2024,60(9):90-100.
- DOU ZH, GAO H R, LIU G Q, et al. Lightweight YOLOv8 small sample steel plate defect detection algorithm[J]. *Computer Engineering and Applications*, 2024, 60(9): 90-100.
- [13] 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 bodies [J]. *Computer Engineering and Applications*, 2024, 60(16): 319-324.
- [14] 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*,2025,51(9):3222-3232.
- ZHANG J B, REN J, WANG M Q. All-weather airport runway foreign object detection based on mixed attention [J]. *Journal of Beijing University of Aeronautics and Astronautics*, 2025, 51(9): 3222-3232.
- [15] 程擎,王元济,李彦冬.基于注意力机制的 YOLOv5 算法在跑道异物检测中的应用[J].*现代计算机*,2023,29(3):55-59.
- CHENG Q, WANG Y J, LI Y D. Application of YOLOv5 algorithm based on attention mechanism in runway foreign body detection[J]. *Modern Computer*, 2023, 29(3): 55-59.
- [16] REN M, WAM W, YU Z, et al. Bidirectional YOLO: Improved YOLO for foreign object debris detection on airport runways[J]. *Journal of Electronic Imaging*, 2022, 31(6): 063047.
- [17] VARGHESE R, M S. 2024. 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),2024:1-6.
- [18] ZHANG X, LIU C, YANG D, et al. RFACnv: Innovating spatial attention and standard convolutional operation [J]. *ArXiv preprint arXiv: 2304.03198*, 2023.
- [19] LU W, CHEN S B, TANG J, et al. A robust feature downsampling module for remote-sensing visual tasks[J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2023, 61: 1-12.
- [20] XU S, ZHENG S, XU W, et al. Hcf-net: Hierarchical context fusion network for infrared small object detection [C]. 2024 IEEE International

- Conference on Multimedia and Expo(ICME). IEEE, 2024: 1-6.
- [21] LI Y, HOU Q, ZHENG Z, et al. Large selective kernel network for remote sensing object detection [C]. 2023 IEEE/CVF International Conference on Computer Vision(ICCV). Paris, France,2023: 16748-16759.
- [22] 费春国,陈世洪. 基于双通道麻雀改进 OTSU 的 FOD 分割方法[J]. 计算机科学,2025,52(S1):488-494.
FEI CH G, CHEN SH H. FOD segmentation method using dual-channel improved sparrow search OTSU algorithm[J]. Computer Science, 2025, 52(S1): 488-494.
- [23] MUNYER T, HUANG P C, HUANG C, et al. Fod-a: A dataset for foreign object debris in airports[J]. ArXiv preprint arXiv:2110.03072, 2021.
- [24] 李毅,徐慧英,朱信忠,等. TFD-YOLO: 基于 YOLOv10 改进的输电线异物检测算法[J]. 浙江师范大学学报(自然科学版),2025,48(4):410-419.
LI Y, XU H Y, ZHU X ZH, et al. TFD-YOLO: Transmission line foreign object detection algorithm improved based on YOLOv10[J]. Journal of Zhejiang Normal University(Natural Sciences), 2025, 48(4): 410-419.
- [25] GIRSHICK R. Fast R-CNN[C]. IEEE International Conference on Computer Vision, 2015: 1440-1448.
- [26] LIU W, ANGUELOV D, ERHAN D, et al. SSD: Single shot multibox detector[C]. Computer Vision- ECCV 2016: 14th European Conference, Amsterdam, The Netherlands, 2016: 21-37.
- [27] ZHU L, GENG X, LI Z, et al. Improving YOLOv5 with attention mechanism for detecting boulders from planetary images[J]. Remote Sensing, 2021, 13(18): 3776.
- [28] LI C, LI L, JIANG H, et al. YOLOv6: A single-stage object detection framework for industrial applications[J]. ArXiv preprint arXiv:2209.02976, 2022.
- [29] WANG C Y, YE H I, MARK L H Y. YOLOv9: Learning what you want to learn using programmable gradient information [C]. European conference on computer vision. Cham: Springer Nature Switzerland, 2024: 1-21.
- [30] WANG AO, CHEN H, LIU L H, et al. YOLOv10: Real-time end-to-end object detection[J]. Advances in Neural Information Processing Systems, 2024, 37: 107984-108011.
- [31] KHANAM R, HUSSAIN M. YOLOv11: An overview of the key architectural enhancements [J]. ArXiv preprint arXiv:2410.17725, 2024.

作者简介

费春国,博士,副教授,主要研究方向为智能优化系统、深度学习等。

陈世洪(通信作者),硕士研究生,主要研究方向为深度学习、计算机视觉等。

E-mail:2028279442@qq.com