

DOI:10.19651/j.cnki.emt.2416278

# 改进 YOLOv8 的输电线路异物检测方法<sup>\*</sup>

易磊 黄哲玮 易雅雯

(华中科技大学电气与电子工程学院 武汉 430074)

**摘要:** 针对无人机对输电线路异物检测准确度有限,模型计算复杂度高、计算速度有限的问题,提出一种改进 YOLOv8 的输电线路异物检测方法 SC-YOLO。该方法引入 StarNet 以构造 C2f\_Star 模块实现 Neck 网络轻量化,有效降低模型参数量与计算量,同时通过增加特征空间维度提升 Neck 部分特征提取能力;在骨干网络输出特征图后添加卷积注意力融合模块,提升骨干网络对输入特征图的初步特征提取能力,增强模型整体检测效果;将原检测头替换为动态检测头,提升模型对不同输入的动态调整能力与对关键信息的关注程度;使用 WIoU 作为边界框损失函数,EMA-Slide Loss 作为分类损失函数,提升模型泛化能力与检测性能。实验结果表明,提出的 SC-YOLO 计算量较原始模型下降 8.02%,mAP 提升 1.4 个百分点,达到了 95.2% 的检测精度,在降低模型计算复杂度的同时实现了较高的检测准确率,具有高可行性与实用性。

**关键词:** 输电线路异物检测;YOLOv8;StarNet;卷积注意力融合模块;动态检测头;WIoU 损失函数;EMA-Slide Loss 损失函数

中图分类号: TP391.4;TN60 文献标识码: A 国家标准学科分类代码: 520.6040

## Improved YOLOv8 foreign object detection method for transmission lines

Yi Lei Huang Zhewei Yi Yawen

(School of Electrical &amp; Electronic Engineering, Huazhong University of Science and Technology, Wuhan 430074, China)

**Abstract:** Aiming at the problems of limited accuracy of UAV detection of foreign objects on power transmission lines, high model computational complexity and limited computational speed, a power transmission line foreign object detection method SC-YOLO which improves YOLOv8 is proposed. This method introduces StarNet to construct C2f\_Star module to realize the lightweight of Neck network, effectively reducing the number of model parameters and calculation amount, and at the same time improves the feature extraction ability of Neck by increasing the dimension of feature space; adds convolution attention fusion module after the backbone network outputs feature map to improve the backbone network's preliminary feature extraction ability of input feature map, and enhance the overall detection effect of the model; replaces the original detection head with dynamic detection head to improve the model's dynamic adjustment ability to different inputs and the degree of attention to key information; uses WIoU as the bounding box loss function and EMA-Slide Loss as the classification loss function to improve the model's generalization ability and detection performance. Experimental results show that the proposed SC-YOLO has 8.02% fewer computational amount than the original model, and mAP is increased by 1.4 percentage points, reaching a detection accuracy of 95.2%. RC-YOLO reduces the computational complexity while achieving a high detection accuracy, and is highly feasible and practical.

**Keywords:** transmission line foreign body detection; YOLOv8; StarNet; convolutional attention fusion module; dynamic detection head; WIoU loss function; EMA-Slide loss function

## 0 引言

电力系统作为我国供电基础设施,对电力的安全供应

对经济发展和社会的各个方面都至关重要。输电线路作为电力系统的关键部分,需要实时维护与检测以维持电力系统的安全、稳定工作。近年来,鸟巢、风筝、气球等异物出现

收稿日期:2024-06-20

\* 基金项目:教育部 2020 年第二批新工科研究与实践项目(E-NYDQHGC20202219)资助

在输电线路上的频率逐渐上升,对电力系统的安全稳定运行产生不利影响。为了及时清除输电线路异物以维持电力系统的可靠性,有关部门需对输电线路进行实时巡检。基于无人机的灵活性、实时性与高效性,相关部门目前常使用无人机搭载摄像头对输电线路异物进行实时检测。然而,无人机在采集输电线路图像以进行检测时,易受环境、天气等因素影响,进而导致误检、漏检情况。因此,如何使用无人机进行高效准确的输电线路异物检测成为电网领域的关键问题。

当前,输电线路检测领域已存在众多研究,基于无人机平台的目标检测算法为研究热点。其中,深度学习目标检测算法由于其强大的学习能力与准确的检测效果,被广泛应用于输电线路检测领域。现有深度学习目标检测算法被分为两类:二阶段算法和一阶段算法<sup>[1]</sup>。二阶段算法将目标检测任务划分为两个阶段,首先通过候选框生成器产生候选区域,然后对此区域进行分类和定位,常见算法包括 R-CNN<sup>[2-4]</sup> 系列算法等。该类算法检测准确率有限,且计算量大、响应速度慢,不适用于计算资源有限的无人机平台和对实时性要求高的电力系统运行场景。一阶段算法则直接在单个神经网络中完成目标的检测和定位,常见算法包括 YOLO<sup>[5-9]</sup> 系列算法与 SSD<sup>[10-11]</sup> 系列算法等。该类算法检测准确率高、运行速度快,有效节约无人机平台计算资源,满足电力系统运行场景的高实时性要求。

其中,YOLO 系列算法基于其高效、易于部署、检测效果好、普适性强等优点,在输电线路检测领域的应用尤其广泛。张红民等<sup>[12]</sup>提出一种改进 YOLOv3 的输电线路异物检测方法,采用 RepVGG 模块提高模型对图像的特征提取能力,增加网络的多尺度检测框提升检测精度,并采用 CIOU 损失函数提升模型泛化能力,获得了较好的检测效果。然而,模型整体性能不足,检测精度有限,并且 RepVGG 模型参数量大、训练时间长,不适用于计算资源有限的无人机平台与电力系统运行场景。Li 等<sup>[13]</sup>提出一种基于改进 YOLOv5s 的输电线路异物检测方法,将 GC 模块与 C3 模块融合构成 C3GC 注意力机制,有效提升模型的特征提取能力,引入 WZ 解耦检测头以实现解耦检测,同时将损失函数替换为 SIOU 进一步提升模型整体性能,获得了较为准确的检测结果。然而,该模型所使用的 C3GC 注意力机制检测计算复杂度仍较高,需要进一步降低计算复杂度以提升模型整体效率。张剑锐等<sup>[14]</sup>提出一种基于改进 YOLOv7 的输电线路绝缘子检测方法,引入 CBAM 注意力机制,加强模型对绝缘子的特征提取能力,并在深层添加集中金字塔,实现更加全面的特征提取,以提高绝缘子检测精度,获得了准确的单目标检测结果。然而,上述改进模型的基线模型较老,目前已有 YOLOv8 等检测效果更好的 YOLO 系列基线模型,且上述模型整体参数量过大,计算速度有限,并且非多目标检测算法,不适用于无人机平台进行多种异物检测,无法满足电力系统运行场景对实时性的要求。

除此之外,在 YOLOv8 模型结构改进方面,已有众多学者进行相关研究,主要包括注意力机制改进、损失函数改进、轻量化结构改进等<sup>[15-19]</sup>。Liu 等<sup>[15]</sup>提出一种针对苹果检测的改进 YOLOv8 模型(Faster-YOLO-AP),该模型调整网络缩放因子以降低模型参数,并使用部分深度卷积(partial depthwise convolution, PDWconv)与深度可分离卷积(depthwise separable convolution, DWConv)替换标准卷积实现卷积的轻量化,同时引入 EIOU 损失函数。然而,过度轻量化导致检测精度有所降低,并且 EIOU 损失函数在处理小尺度物体时对异常值检测敏感,不适用于本文小型异物检测场景。Kang 等<sup>[16]</sup>提出一种基于注意力尺度序列融合的 YOLO 模型(ASF-YOLO),采用尺度序列特征融合模块来增强模型的多尺度信息提取能力,并引入通道和位置注意力机制(channel and position attention mechanism, CPAM)以提升对重点信息的关注。该模型有效提升模型多尺度特征能力,然而模型参数量较大,不适用于无人机平台实时检测。Kang 等<sup>[17]</sup>提出了一种针对车辆检测的 YOLOv8 改进模型(YOLO-FA),引入 1 型模糊注意力机制(type-1 fuzzy attention, T1FA)以降低特征图的不确定性,提升对目标中心的关注,同时使用 MDFormer 中的混合深度卷积作为令牌混合器来捕获多尺度感知场。然而,该模型采用的 T1FA 注意力机制对特征提取的能力有限,同时未考虑损失函数的改进与模型的轻量化,难以满足无人机实时检测的准确率与计算量要求。上述基于 YOLOv8 的改进模型难以综合考虑高准确度与高实时性,难以同时满足无人机平台对计算资源的要求、电力系统安全运行对实时性与检测准确性的要求。

基于无人机对输电线路异物检测准确度有限,难以同时满足模型轻量化、高运行速度与检测准确率的问题,本文提出一种改进 YOLOv8 的输电线路异物检测方法。该方法引入轻量化网络 StarNet 以构造 C2f\_Star 模块替换 Neck 网络中的 C2f 模块,有效降低模型参数量与计算量,并通过增加特征空间维度提升模型对输电线路异物多尺度特征提取能力;引入 CAFM 注意力机制模块实现骨干网络优化,提升骨干网络对输入特征图的初步特征提取能力,增强模型整体检测精度;将原检测头替换为动态检测头,提升对不同种类异物的动态调整能力与对关键特征的注意力;引入 WIIOU 作为边界框损失函数,并提出一种基于 EMA 注意力机制改进的 Slide Loss 作为分类损失函数,综合提升模型泛化性与收敛速度,以及分类与定位效果。实验结果证明,本文方法在降低模型计算复杂度的同时实现了较高的检测准确率,有助于及时清除输电线路异物,助力电力系统安全稳定运行。

## 1 YOLOv8 模型

YOLOv8 由 Redmon 等<sup>[20]</sup>于 2023 年提出,YOLOv8 延续了 YOLO 系列的快速、高效检测特点,并在网络结构

和检测性能上进行了多方面的改进,有效提升了检测精度和速度,被广泛应用于目标检测领域。

YOLOv8网络结构主要由骨干网络、Neck部分与Head部分构成。骨干网络包括卷积模块、C2f模块和SPPF模块,实现部分特征提取。卷积模块包括卷积层、批量归一化层与SiLU激活函数,提升模型特征提取能力。C2f模块中包含Split和Bottleneck层,在有效减少计算量的同时,捕获全局特征并融合多尺度信息。SPPF模块对不同尺度进行池化与卷积操作,拼接不同尺度的特征图,以提高对不同尺寸目标的检测能力。Neck部分使用PAN与FPN实现不同层次的特征融合,提升对不同尺度目标的检测能力。Head部分采用解耦式,包括分类头与检测头,实现对目标的检测与分类。

YOLOv8采用自适应锚框机制,具有高灵活性,有助于提升检测精度。同时,YOLOv8采用改进损失函数,包括分类损失、定位损失和置信度损失,并对这些损失函数进行加权组合,保证模型检测精度与运行速度的平衡。

基于上述优点,本文采用YOLOv8作为基线模型,其网络结构如图1所示。

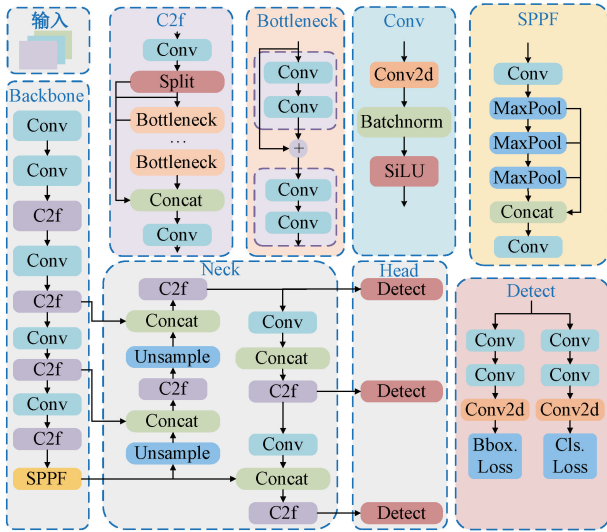


图1 YOLOv8网络结构图

Fig. 1 Architecture of YOLOv8

## 2 改进YOLOv8检测模型

### 2.1 Neck网络优化

YOLO系列模型的网络常采用基于线性投影与非线性激活函数组合,为了提升模型的特征提取能力,常常引入基于此组合的自注意力机制模块。其通过点积相乘构建注意力矩阵,计算复杂度会随着特征提取能力的提升明显增加,降低模型效率。为了解决该问题,近年来研究人员提出了星型操作,该操作在自然语言处理、计算机视觉等领域表现出良好的性能与效率,被逐渐引入轻量级网络结构中。超轻量级网络StarNet<sup>[21]</sup>采用基于星型操作的模型,保证

模型复杂度低的同时,显著提升模型将输入映射到高维特征空间的能力,进而提升模型的特征提取能力。星型操作通过逐元素乘法融合两个线性变换的特征,计算公式如下:

$$\begin{cases} \mathbf{O}_{star} = (\mathbf{W}_1^T \mathbf{X}) * (\mathbf{W}_2^T \mathbf{X}) \\ \mathbf{W} = \begin{bmatrix} \mathbf{W} \\ \mathbf{B} \end{bmatrix} \\ \mathbf{X} = \begin{bmatrix} \mathbf{X} \\ 1 \end{bmatrix} \end{cases} \quad (1)$$

式中: $\mathbf{O}_{star}$ 为进行星型操作后的结果; $\mathbf{W}$ 为线性层的权重矩阵; $\mathbf{B}$ 为线性层的偏置; $\mathbf{X}$ 为输入;\*代表星型操作。

基于式(1),定义 $\omega_1, \omega_2, x \in R^{(c+1) \times 1}$ ,其中 $c$ 为输入通道数。则星型操作可以进一步写为:

$$\begin{aligned} (\omega_1^T x) * (\omega_2^T x) &= \\ \left( \sum_{m=1}^{c+1} \omega_1^m x^m \right) * \left( \sum_{n=1}^{c+1} \omega_2^n x^n \right) &= \sum_{m=1}^{c+1} \sum_{n=1}^{c+1} \omega_1^m \omega_2^n x^m x^n = \\ \underbrace{\alpha_{(1,1)} x^1 x^1 + \dots + \alpha_{(4,5)} x^4 x^5 + \dots + \alpha_{(c+1,c+1)} x^{c+1} x^{c+1}}_{\substack{(c+2)(c+1) \\ 2} \text{项}} \end{aligned} \quad (2)$$

式中: $m, n$ 为通道下标, $\alpha$ 为各子项的系数,定义如下:

$$\alpha_{(m,n)} = \begin{cases} \omega_1^m \omega_2^n, & i = j \\ \omega_1^m \omega_2^n + \omega_1^n \omega_2^m, & i \neq j \end{cases} \quad (3)$$

由式(2)可知,在 $c$ 维度空间使用星型操作,可以得到维度为 $\frac{(c+1)(c+2)}{2} \approx \left(\frac{c}{2}\right)^2$ 的隐式维度特征空间。当堆叠 $k$ 层时,能够以指数的方式递归隐式维度至 $\left(\frac{c}{2}\right)^{2^k}$ :

$$\mathbf{O}_k = \mathbf{W}_{k,1}^T \mathbf{O}_{k-1} * \mathbf{W}_{k,2}^T \mathbf{O}_{k-1} \in R^{\left(\frac{c}{2}\right)^{2^k}} \quad (4)$$

式中: $\mathbf{O}_k$ 为第 $k$ 层星型操作后的输出。

StarNet采用四级分层结构,每一级使用卷积层进行下采样,并采用Star Blocks进行特征提取。其整体结构如图2所示。Star Blocks首先采用深度卷积DWConv实现特征初步提取,进而对其进行批量归一化处理,得到两个宽度加倍、通道扩展因子为4的特征图,经过ReLU激活函数后输入至第二个DWConv,实现特征进一步提取,获得输出特征图。与传统卷积相比,DWConv将卷积操作分为对输入数据的每个通道进行卷积操作、将这些通道的结果进行逐元素相加获得输出2个步骤,能够有效减少参数量和和计算量,进一步提升模型的训练和推理过程。

本文引入StarNet结构对原YOLOv8模型Neck部分进行优化,但是,StarNet整体构造相对繁琐,直接引用会导致模型结构臃肿,不利于模型轻量化。因此,本文模型仅引入Star Blocks结构,将Neck部分的原C2f模块中的Bottleneck结构替换为StarNet中的Star Blocks,进而构成C2f\_Star模块。该改进后的C2f\_star模块的结构如图3所示。

C2f\_Star模块有效提升了输入映射特征维度,提升了模型对于特征的提取能力,进而有效提升模型的目标检测

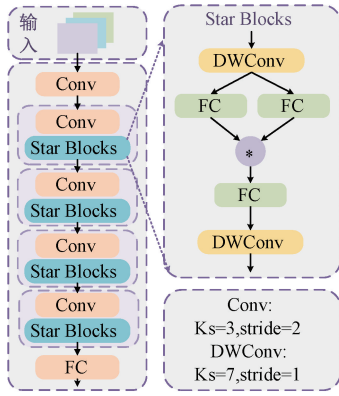


图 2 StarNet 结构  
Fig. 2 Architecture of StarNet

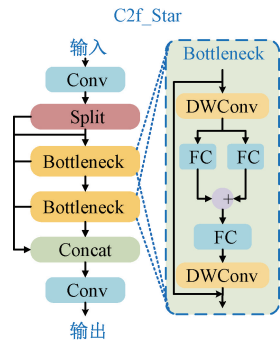


图 3 C2f\_Star 模块结构  
Fig. 3 Architecture of C2f\_Star block

准确率。同时, Star Blocks 结构参数量小, 有效降低了原 C2f 模块的参数量, 进而有效降低 YOLOv8 网络的参数量, 提升了模型的计算速度, 提高了无人机对输电线路异物检测的实时性。

### 2.2 骨干网络优化

原始 YOLOv8 模型对图像的特征提取能力有限, 为了进一步提高骨干网络的特征提取能力, 已有研究常常在骨干网络引入注意力模块以提升模型对关键信息的关注能力。当前改进 YOLOv8 通常引入的注意力模块包括 SENet、GAM、CBAM 等, 此类注意力机制特征提取能力有限, 对特定异物局部特征关注能力有限, 同时会导致模型计算量显著增加, 不利于模型轻量化。为了在有效提升特征提取能力的同时保证轻量化, 本文改进模型引入卷积注意力融合模块<sup>[22]</sup> (CAFm)。该模块将卷积与注意力融合, 有效捕捉长距离依赖关系, 以增强全局和局部特征建模能力, 进而提升骨干网络特征提取能力。CAFm 的网络结构如图 4 所示。

CAFm 由局部分支和全局分支组成。在局部分支中, 首先使用  $1 \times 1$  卷积来调整通道维度, 以增强跨通道交互能力与模型对输入输电线图像的整合能力。此后, 进行通道混洗操作以混合通道信息。通道混洗操作 (Channel shuffle) 将输入张量沿通道维度划分成组, 每组内使用深度

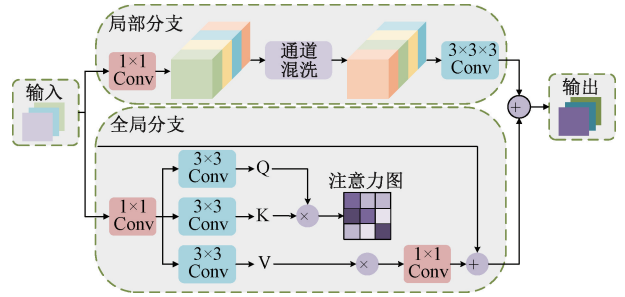


图 4 CAFM 结构  
Fig. 4 Architecture of CAFM

可分离卷积 (DWConv) 来诱导通道混洗。随后, 将每组得到的输出张量沿着通道维度进行级联, 以生成新的输出张量, 接下来对新的输出张量使用  $3 \times 3 \times 3$  卷积来提取特征。局部分支结构的计算可以用公式表示为:

$$P_{conv} = C_{3 \times 3 \times 3} (CS(C_{1 \times 1}(X))) \quad (5)$$

式中:  $P_{conv}$  为局部分支的输出特征图;  $C_{3 \times 3 \times 3}$  表示  $3 \times 3 \times 3$  卷积;  $C_{1 \times 1}$  表示  $1 \times 1$  卷积;  $CS(\cdot)$  表示通道混洗操作;  $X$  为输入特征图。

在全局分支中, 首先通过  $1 \times 1$  卷积和  $3 \times 3$  深度卷积获得查询 ( $Q$ )、键 ( $K$ ) 和值 ( $V$ ), 得到 3 个形状为  $(\hat{H} \times \hat{W} \times \hat{C})$  的张量。然后, 将  $Q$ 、 $K$ 、 $V$  重塑至  $R^{(\hat{n} \times \hat{W} \times \hat{C})}$ 。进而根据重塑后  $Q$  与  $K$  的关系计算注意力矩阵  $A$ , 得到全局分支的输出特征图为:

$$\begin{cases} P_{att} = C_{1 \times 1} Attention(\hat{Q}, \hat{K}, \hat{V}) + X \\ Attention(\hat{Q}, \hat{K}, \hat{V}) = \hat{V} \text{softmax}\left(\frac{\hat{K}\hat{Q}}{\alpha}\right) \end{cases} \quad (6)$$

式中:  $P_{att}$  为全局分支输出特征图;  $Attention(\cdot)$  为注意力分数计算操作;  $\hat{Q}$  为由  $Q$  重塑至  $R^{(\hat{n} \times \hat{W} \times \hat{C})}$  的查询矩阵;  $\hat{K}$  为由  $K$  重塑至  $R^{(\hat{c} \times \hat{H} \times \hat{W})}$  的查询矩阵;  $\hat{V}$  为由  $V$  重塑后的查询矩阵;  $\alpha$  为一个可学习的缩放参数, 用于在应用 softmax 函数之前控制  $\hat{Q}$  和  $\hat{K}$  的矩阵乘法大小。

最终, 通过相加全局分支与局部分支输出可得 CAFm 的输出为:

$$P_{out} = P_{att} + P_{conv} \quad (7)$$

式中:  $P_{out}$  为 CAFm 模块的输出特征矩阵。

本文在原 YOLOv8s 的骨干网络中的末端的 SPPF 模块后添加 CAFm 注意力机制模块, 进一步提取 SPPF 模块输出特征图中关键局部与全局特征, 从而增强骨干网络对输入输电线路图像的初步特征提取能力, 有助于提升模型对输电线路异物的检测准确率。

### 2.3 动态检测头

原 YOLOv8 模型中的目标检测头对不同尺度与类型目标特征提取能力有限, 进而导致模型对不同种类异物的

检测效果有限。为了提升模型在检测不同种类与尺度的异物时的自适应性能与特征提取能力,本文采用动态检测头(dynamic head, DyHead)<sup>[23]</sup> 替换原 YOLOv8 网络的普通检测头。DyHead 根据输入输电线路图像不同特征与任务需求动态调整模型输出层结构,有效提升模型泛化能力;同时,其融入自注意力机制,在没有任何计算开销的情况下显著提升了目标检测头的表示能力,有助于精准检测不同输电线路异物。

动态检测头包含 3 种不同的注意力机制,分别针对尺度感知注意力、空间感知注意力和任务感知注意力。动态检测头将注意力机制分为 3 个维度,对应一个特征层  $\mathbf{M} \in \mathbf{R}^{L \times S \times C}$  的注意力权重计算公式如下:

$$\mathbf{W}(\mathbf{M}) = \pi_c(\pi_s(\pi_L(\mathbf{M}) \cdot \mathbf{M}) \cdot \mathbf{M}) \cdot \mathbf{M} \quad (8)$$

式中:  $\mathbf{W}(\mathbf{M})$  为特征层  $\mathbf{M}$  的注意力权重矩阵;  $\mathbf{M}$  为特征层;  $\pi_c(\cdot)$  为尺度感知注意力操作;  $\pi_s(\cdot)$  为空间感知注意力操作;  $\pi_L(\cdot)$  为任务感知注意力操作。

尺度感知注意力对不同层次的特征曾给予不同权重,有利于模型自适应地根据层次特征重要程度进行融合,计算公式如下:

$$\pi_L(\mathbf{M}) \cdot \mathbf{M} = \sigma\left(f\left(\frac{1}{SC} \sum_{S,C} \mathbf{M}\right)\right) \cdot \mathbf{M} \quad (9)$$

式中:  $\sigma$  为 Hard-sigmoid 激活函数;  $f$  为全连接层操作;  $\frac{1}{SC} \sum_{S,C} \mathbf{M}$  为  $HW$  维度下的 Avg pool 操作;  $S$  为特征层的  $S$  维度;  $C$  为特征层的  $C$  维度。

空间感知注意力考虑根据可变形卷积进行叙述采样与在相同空间位置上聚合各个层次的特征。计算公式如下:

$$\pi_s(\mathbf{M}) \cdot \mathbf{M} = \frac{1}{L} \sum_{l=1}^L \sum_{a=1}^A \mathbf{w}_{l,q} \cdot \mathbf{M}(l; o_q + \Delta o_q; c) \cdot \Delta n_q \quad (10)$$

式中:  $A$  为稀疏采样位置的数目;  $o_q + \Delta o_q$  为通过自学习的空间偏移量;  $\Delta n_q$  为位置  $o_q$  的自学习重要性标量;  $l$  为特征层的层次;  $L$  为特征层的总层次数目;  $\mathbf{w}$  为权重矩阵。

任务感知注意力将通道自适应激活,实现 dyrelu 操作,该操作为特征层每一个通道学习 4 个值,控制通道激活强度,计算公式如下:

$$\pi_c(\mathbf{M}) \cdot \mathbf{M} = \max(\alpha^1(\mathbf{F}) \cdot \mathbf{F}_c + \beta^1(\mathbf{F}), \alpha^2(\mathbf{F}) \cdot \mathbf{F}_c + \beta^2(\mathbf{F})) \quad (11)$$

式中:  $[\alpha^1, \beta^1, \alpha^2, \beta^2]^T = \theta(\cdot)$  为控制激活阈值的超函数;  $\mathbf{F}_c$  为第  $c$  个通道的特征切片。

动态检测头的详细结构图如图 5 所示,分别对输入进行了 3 个维度的处理。

本文将基线 YOLOv8 模型的普通检测头替换为动态检测头,可以根据输入特征自适应调整检测策略,加强了模型检测头对不同尺度与尺寸的输电线路异物的自适应调节能力,同时动态检测头包含的注意力机制有效提升了模型检测头对输入特征图的局部与全局特征提取能力,进而提

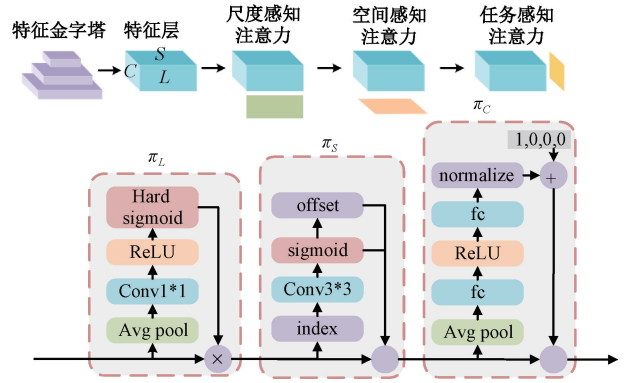


图 5 动态检测头结构

Fig. 5 Architecture of dynamic head

升模型对输电线路异物的检测准确率。

## 2.4 损失函数优化

### 1) 边界框损失函数优化

边界框损失函数在目标检测中有重要意义,精确的边界框损失函数有利于提升目标检测模型性能。原 YOLOv8 模型所使用的 CIoU 损失函数重点考虑距离与长宽比等几何因素,会加剧对低质量样本的惩罚,从而降低模型的泛化性。当前,边界框损失函数的优化策略研究颇多,EIOU、SIOU、Focal loss 等被广泛应用于 YOLOv8 模型的改进中。然而,EIOU 计算开销较大,影响模型推理速度;SIOU 增加了超参数调节的难度与计算复杂性;Focal loss 则主要用于解决不平衡问题,可能会导致过度关注难样本,而忽视易样本的学习。基于上述问题,本文使用基于动态非单调的聚焦机制的 Wise-IOU(WIOU)<sup>[24]</sup> 作为边界框损失函数。WIOU 使用“离群度”替代 IoU 对锚框进行质量评估,并提供了明智的梯度增益分配策略,实现对普通质量锚框的动态聚焦,同时削弱几何因素的惩罚,有助于提高模型整体泛化效果与检测效果。WIOU 的计算公式如下:

$$L_{WIOU} = R_{WIOU} L_{IoU} \quad (12)$$

$$R_{WIOU} = \exp\left(\frac{(a - a_{gt})^2 + (b - b_{gt})^2}{(W_{min}^2 + H_{min}^2)}\right)$$

式中:  $L_{IoU}$  为交并比,用于衡量模型预测与实际目标间重叠程度;  $R_{WIOU}$  为普通质量锚框  $L_{IoU}$  的放大倍数;  $W_{min}$ 、 $H_{min}$  为最小封闭框的宽与高;  $a$ 、 $b$  分别为图像像素点的纵横坐标;  $a_{gt}$ 、 $b_{gt}$  分别为标记框的中心纵横坐标。

本文模型将原 YOLOv8 模型中的 CIoU 边界框损失函数替换为 WIOU 边界框损失函数,有效提升对低质量框识别的能力,进而提升模型整体的泛化能力与对输电线路异物的检测准确率。

### 2) 分类损失函数优化

分类损失函数用于评估模型在进行分类任务时预测结果与真实标签之间的差异,良好的分类损失函数可帮助模

型对正确的分类进行学习,从而提高模型的检测率。原 YOLOv8 模型分类损失函数为 CIOU,对于难样本的识别程度有限。目前,已有的基于 YOLOv8 的改进模型在对损失函数改进时,一般仅改进边界损失框函数,忽略对分类损失函数的优化,导致模型训练过程的收敛能力无法进一步提升。基于上述问题,本文提出一种基于 EMA 注意力机制改进的 EMA Slide loss 作为分类损失函数,加强模型对难样本的关注,以提升模型对较小异物目标与遮挡物的整体检测性能。EMA Slide loss 基于 Slide loss<sup>[25]</sup>进行改进,Slide loss 基于各样本难度完成对难样本的重点关注,样本难度权重计算公式如下:

$$m(z) = \begin{cases} 1, & z \leq \lambda - 0.1 \\ e^{-z}, & \lambda \leq z \leq \lambda + 0.1 \\ e^{-z}, & z \geq \lambda \end{cases} \quad (13)$$

式中: $m$  为 slide 加权函数; $z$  为样本比例; $\lambda$  ( $0 \leq \lambda \leq 1$ ) 为比例参数,用于将样本分为正样本和负样本。

本文模型为了进一步提升模型对难样本的关注能力,以提升模型对输电线路异物检测性能,引入 EMA 注意力机制优化 Slide Loss。EMA 注意力机制通过对历史值的加权平均来平滑数据,使得训练过程中模型参数的更新更加稳定,有助于提高收敛速度;其能够根据输入信息动态调整注意力权重,提升分类损失函数的自适应性能;并且计算开销小,有助于模型轻量化,适用于本文无人机实时检测平台。本文将 YOLOv8 模型的 CIOU 分类损失函数替换为 EMA Slide loss,有效提升了模型对难样本的分类能力,进而提升了模型整体对输电线路异物的检测性能。

总体而言,与目前已有的改进模型相比,本文模型对损失函数优化时同时考虑边界框损失函数与分类损失函数,综合提升了各损失函数的性能,有效提升了模型泛化能力与检测效率,进一步提升了模型对输电线路异物的检测精度,有助于维持电力系统安全稳定运行。

### 2.5 SC-YOLO 模型

本文提出基于 YOLOv8s 的改进模型——SC-YOLO 模型,该模型结构如图 6 所示。

本文首先在 YOLOv8s 的 Neck 网络,采用 Star Block 替换 C2f 模块的 Bottleneck 构成 C2f\_Star 模块,在保证高准确率的同时有效实现 Neck 网络轻量化。其次,在 YOLOv8s 的 Backbone 末端 SPPF 模块后引入 CAFM 注意力机制层加强特征感知,增强 Backbone 网络对输入特征图的局部与全局特征提取能力。将 YOLOv8s 的检测头替换为 DyHead 动态检测头进行预测输出,增强网络对不同尺度与尺寸的输电线路异物的自适应检测能力。最后,将边界框损失函数替换为 WIOU,将分类损失函数替换为 EMA Slide loss,提升对较小异物目标与被遮挡异物的识别能力,以及模型收敛速度与泛化性能。具体而言,Star Block 作为 Bottleneck 降低了模型参数量以及计算量,同时保持原有 Bottleneck 几乎相同的精度。CAFM 通过卷积

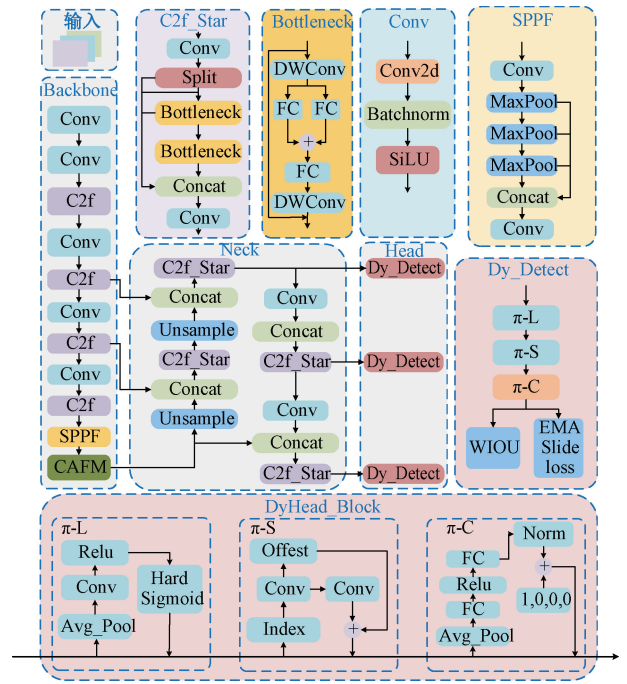


图 6 SC-YOLO 网络结构  
Fig. 6 Architecture of SC-YOLO

与注意力融合,拥有更长的上下文理解能力,提升了全局特征提取效果。DyHead 检测头则通过结合多种注意力机制,在多尺度、多空间以及多目标的场景中能够拥有更广泛的注意力,使模型在小目标、复杂背景的条件下的性能得到进一步提升。

## 3 实验结果与分析

### 3.1 实验环境

本文实验基于 Linux20.04 操作系统,使用基于 Python3.10.8 的 Pytorch2.0.1 搭建深度学习模型。在模型训练方面,使用 NVIDIA GeForce RTX 4090D GPU 作为模型训练平台,并采用 CUDA12.4 对 GPU 进行加速。本文深度学习模型参数设置如表 1 所示。

表 1 模型参数

Table 1 Model parameters

参数	设置	参数	设置
类别数量	4	学习阈值	0.01
输入宽	320	WARMUP	True
输入高	320	批大小	320
学习率	0.001 25	训练轮数	300

### 3.2 实验数据集

本文实验采用铁路输电线路异物检测公开数据集 RailFOD23<sup>[26]</sup>,该数据集包含 14 615 张图像,40 541 个标注物体,标注物体包含鸟巢、气球、塑料袋、漂浮物共 4 种常

见异物。

由于原始数据集较大,训练量过大。为了降低模型大小与计算复杂度,本文对原始数据集进行数据蒸馏,得到2 435张图像,按照7:3的比例将2 435张图像随机划分为训练集与测试集,其中训练集包括1 705张图像,测试集包括730张图像。

### 3.3 评价指标

为了衡量本文模型对输电线路异物检测的准确率,采用准确率(precision, P)、召回率(recall, R)与IoU为0.5时的平均精度均值(mean average precision, mAP)对检测准确率进行评价。其中, mAP为本文实验准确率关键评价指标。P、R、mAP越接近于1表明模型识别准确性越高,相应计算公式如下:

$$P = \frac{TP}{TP + FP} \quad (14)$$

$$R = \frac{TP}{TP + FN} \quad (15)$$

$$mAP = \frac{1}{n} \sum_{i=1}^N AP_i = \frac{1}{n} \sum_{i=1}^N \int P dR \quad (16)$$

式中: TP为图像中检测正确的目标数量; FP为图像中检测错误的目标数量; FN为图像中漏检的目标数量; AP为平均精度,即PR曲线下的面积; N为目标检测类别的数量。

同时,为了衡量本文模型对输电线路异物检测的效率,采用参数量(Params)与计算量(giga floating point operations per second, GFLOPs)作为模型效率评价指标,参数量与计算量越小说明模型越高效。

### 3.4 消融实验

为了证明本文提出的模型所进行的每一部分改进的有效性,对各个改进模块进行消融实验。实验过程严格控制模型参数一致。消融实验结果如表2所示。

表2 消融实验结果

Table 2 Results of ablation experiment

C2f_Star	CAFm	DyHead	Params/ ×10 <sup>6</sup>	GFLOPs	mAP/ %
—	—	—	11.13	28.7	0.938
✓	—	—	9.53	25.2	0.937
—	✓	—	12.35	29.6	0.946
—	—	✓	10.85	28.3	0.950
✓	✓	—	11.49	26.9	0.946
✓	—	✓	9.27	25.1	0.948
—	✓	✓	12.07	29.3	0.952
✓	✓	✓	11.21	26.4	0.952

由表2可知,每种改进均一定程度提高了模型的精度与性能。与原模型相比,引入C2f\_Star模块后模型参数量下降了14.1%,计算量下降了12.2%,而其对精度影响几

乎可以忽略不计;在检测端引入动态检测头DyHead后,动态卷积通过对不同尺度的目标采用不同大小的卷积核,降低了模型的参数量以及计算量,模型在进行检测的流程中能够利用注意力机制关注到不同检测层之间的关联,提高了模型的特征提取能力, mAP提升了1.2%,表明改进后的检测头能够有效提升检测性能;模型骨干网络引入CAFm后,能够在特征提取过程中对图像不同区域语义进行连接,提升模型对长距离依赖关系的提取能力, mAP提升了0.8%,表明CAFm卷积注意力融合模块能够有效改进模型性能。值得一提的是,在C2f\_Star与DyHead检测头组合后,精度相较单DyHead实验降低了0.2%。原因在于C2f\_Star引入的特征损失被DyHead检测头的注意力机制放大,导致了明显的精度下降。在引入CAFm注意力机制后,通过在骨干网络捕获更多特征,弥补了特征损失,实现了更高的检测精度。

改进的网络在参数量略低于基准网络的条件下,计算量降低了8.02%, mAP提升了1.4%,达到95.2%的准确率。实验结果证明SC-YOLOv8模型能够获取更丰富的特征细节,在实现轻量化的前提下有效提升模型性能。

### 3.5 模型对比实验

#### 1) 整体模型对比

为了证明本文所提出的SC-YOLO模型整体效果,将本文模型与二阶段模型与单阶段模型进行对比分析,其中单阶段模型包括各类基线模型与改进模型,严格控制实验参数,实验结果如表3所示。

表3 不同模型实验结果对比

Table 3 Comparison of different models

模型	Params/ ×10 <sup>6</sup>	GFLOPs	mAP/ %
Faster-RCNN	136.8	369.8	88.2
SSD	24.1	61.2	82.8
RetinaNet	18.7	45.2	84.5
ATSS	20.3	49.2	85.7
YOLOv7s	36.5	105.2	55.5
YOLOv4-tiny	9.0	26.8	91.6
YOLOv7-tiny	6.0	13.2	39.4
YOLOv5s	7.0	16.5	91.1
YOLOv8s	11.1	28.7	93.9
Faster-YOLO-AP <sup>[15]</sup>	0.66	0.57	68.3
YOLO-FA <sup>[16]</sup>	11.0	26.7	94.5
AFS-YOLO <sup>[17]</sup>	11.1	27.6	94.1
本文模型	10.9	26.4	95.2

由表3可知,本模型相较二阶段架构的Faster-RCNN模型 mAP提升7%,并且模型参数以及计算量显著降低。相较其他一阶段模型,包括传统SSD模型相比,本文模型

在参数量与计算量均有所降低的情况下, mAP 提升 13%。与最新一阶段算法 RetinaNet 和 ATSS 相比, 本文模型参数量分别降低 41.7% 与 46.3%, 同时 mAP 分别提升 10.7% 与 9.5%, 在显著降低模型参数的同时有效提升了输电线路异物检测精度。与 YOLO 系列经典模型相比, 本文模型检测精度显著高于 YOLOv7 系列模型, 较参数量较小的 YOLOv7-tiny 与参数量较大的 YOLOv7s, mAP 分别提升 55.8% 与 39.7%。与最新 YOLO 系列改进模型相比, 本文模型在保证参数量较低的同时, 较轻量级网络 Faster-YOLO-AP, mAP 显著提升 26.9%; 相较 YOLO-FA 与 AFS-YOLO, 本文模型在参数量与计算量均有所降低的同时具有更高的 mAP, 分别提升 0.7% 与 1.1%。总体而言, 与其他单阶段与二阶段目标检测模型相比, 本文模型在保证计算量较低的同时保证了较高检测准确率。

为了直观展现本文提出的 SC-YOLO 相较 YOLOv8s 模型检测准确率的提升, 将 YOLOv8 与 SC-YOLO 的 PR 曲线进行对比, 如图 7 所示。

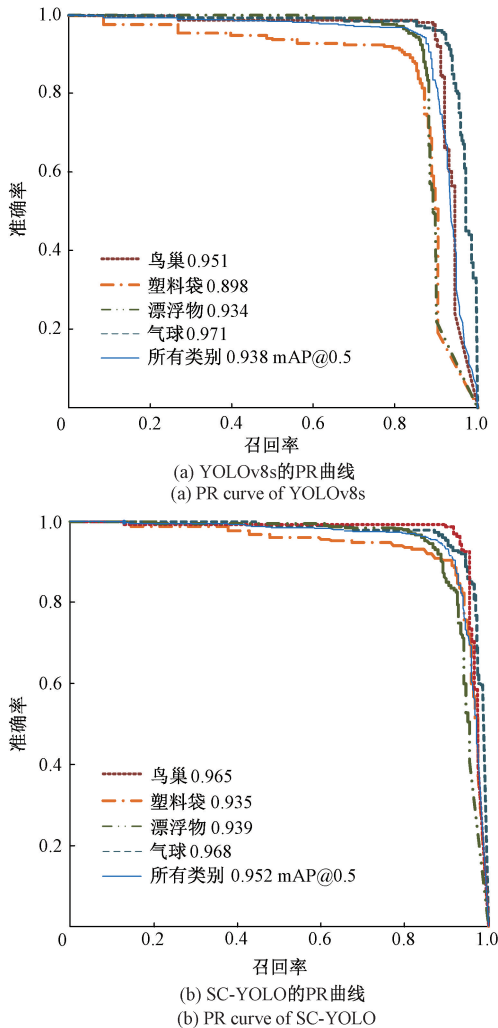


图 7 YOLOv8s 与 SC-YOLO 的 PR 曲线对比

Fig. 7 Comparison of PR curves between YOLOv8s and SC-YOLO

由图 6 可知, 本文模型相较于原始 YOLOv8s 模型整体 PR 曲线前段斜率更低, 精度坠落点更加靠后, 对各类识别对象识别精度都有所提高, 尤其在塑料袋识别方面, 提升幅度达到了近 4 个百分点。总而言之, 改进后的 SC-YOLO 模型整体检测精度水平更高。

### 2) 注意力机制对比

为了证明本文采用的 CAFM 注意力模块的优越性, 将原始模型结合 CAFM 与结合多个其他常用于最新 YOLO 改进模型中的注意力模块的组合模型进行对比分析, 严格控制实验参数保持一致, 实验结果对比如表 4 所示。

表 4 不同注意力机制对比

注意力机制	参数/ $\times 10^6$	GFLOPs	mAP/%
CA	10.0	26.6	94.8
GAM	11.7	28.0	93.7
CBAM	11.3	27.7	94.1
PSA	11.0	27.2	94.4
CAFM	11.1	26.4	95.2

由表 4 可知, 本文所使用的 CAFM 在参数量与计算量都适中的条件下实现了最高的准确度, 相较相对轻量的 PSA, mAP 有 0.8% 的提升, 相较最轻量的 CA, mAP 有 0.4% 的提升; GAM 与 CBAM 参数量较大, 而效果却相对较差, 经过分析, 将原因归结为 GAM 与 CBAM 缺乏细粒度的注意力, 导致对细节特征的感知能力较弱, 使得对不同种类输电线路异物检测精度较低。

### 3) 损失函数对比

为了证明本文所采用的 WIOU 与 EMA-Slide loss 损失函数的优越性, 将原始模型结合 WIOU+EMASlide loss 与结合多个其他常用于当前改进 YOLO 模型的损失函数的组合模型进行对比分析, 严格控制实验参数, 实验结果对比如表 5 所示。

表 5 不同损失函数对比

损失函数	mAP/%
CIOU	93.8
EIOU	93.7
WIOU	94.1
CIOU + EMA-Slide loss	94.3
WIOU + EMA-Slide loss	94.4

由表 5 可知, 将 WIOU 在本文的检测场景中表现最佳, 相比 YOLOv8 原始 CIOU 损失函数提升 0.3%; 将分类损失函数替换为 EMA-Slide loss 模型检测精度能够得到进一步提升, 相较 CIOU 本文模型 mAP 提升 0.4%, 相较



WIOU 本文模型 mAP 提升 0.3%。综上所述,本文所使用的 WIOU+EMA-Slide loss 损失函数组合提升了模型泛化能力,并提升了模型整体检测精度。

### 3.6 实验结果可视化分析

为了直观对比本文提出 SC-YOLO 模型与 YOLOv8s 的检测效果,从测试集中随机选取部分图像进行验证,验证结果如图 8 所示。



图 8 YOLOv8s 与 SC-YOLO 检测结果对比

Fig. 8 Comparison of detection results between YOLOv8s and SC-YOLO

由图 7 可知,本文分别在 3 幅不同光照、不同清晰度的图像中测试,左侧为黑夜场景,中间为晴天高清晰度场景,右侧为低清晰度场景。对比结果显示本模型在各类检测场景中均有精确度提升。从黑夜场景图像中可知,SC-YOLO 能够检测出难以被检测出的目标,捕捉更低细节的物体;从白天场景图像可知,相较 YOLOv8s,SC-YOLO 区分常被混淆的塑料袋与气球两类的精度更高,有效降低混淆程度;从低清晰度图像可知,本文模型能够提升被正确检测目标的置信度,降低其他类别对正确检测目标的干扰。

## 4 结 论

针对无人机对输电线路异物检测准确率有限,无法综合考虑模型检测效率与准确率,难以满足电力系统安全稳定运行对高实时性与准确性要求的问题,提出了一种改进 YOLOv8 的输电线路异物检测方法,主要结论为:本文引入 StarNet 改进 Neck 部分,降低了模型的参数量与计算量,有效实现模型的轻量化的同时保证了模型检测精度;本文引入 CAFM 注意力模块优化原模型骨干网络部分,有效提升了骨干网络对输入图像特征的初步提取能力,进而提升了模型整体检测精度;本文将原检测头替换为动态检测

头,有效提升了检测头对不同尺度物体的自适应识别能力,同时通过融合注意力机制实现了模型的对输出特征图的特征提取能力,进而提升模型整体检测性能;本文使用 WIOU 作为边界框损失函数,EMA-Slide loss 作为分类损失函数,有效提升了模型的泛化能力,与模型整体检测性能。综上所述,本文提出的改进 YOLOv8 的输电线路异物检测方法相较原始 YOLOv8s 模型具有更高检测效率与检测准确率,模型整体性能良好、可行性高,有助于无人机协助输电线路巡检,进而保证电力系统安全稳定运行。

### 参考文献

- [1] 张利丰,田莹. 改进 YOLOv8 的多尺度轻量级车辆目标检测算法[J]. 计算机工程与应用, 2024, 60(3): 129-137.  
ZHANG L F, TIAN Y. Improved YOLOv8 multi-scale and lightweight vehicle object detection algorithm [J]. Computer Engineering and Applications, 2024, 60(3): 129-137.
- [2] YANG L, ZHONG J H, ZHANG Y, et al. An improving faster-RCNN with multi-attention ResNet for small target detection in Intelligent autonomous transport with 6G [J]. IEEE Transactions on Intelligent Transportation Systems, 2023, 24(7): 7717-7725.
- [3] YASIR M, LIU SH W, XU M M, et al. ShipGeoNet: SAR image-based geometric feature extraction of ships using convolutional neural networks [J]. IEEE Transactions on Geoscience and Remote Sensing, 2024, 62(1):1-13.
- [4] CAI J H, ZHANG L, DONG J, et al. Automatic identification of active landslides over wide areas from time-series InSAR measurements using Faster RCNN[J]. International Journal of Applied Earth Observation and Geoinformation, 2023, 124(1):103516.
- [5] SHAO Y H, HUANG Q M, MEI Y Y, et al. MOD-YOLO: Multispectral object detection based on transformer dual-stream YOLO [J]. Pattern Recognition Letters, 2024, 183(1):26-34.
- [6] WANG J X, LIU M, DU Y R, et al. PG-YOLO: An efficient detection algorithm for pomegranate before fruit thinning [J]. Engineering Applications of Artificial Intelligence, 2024, 134(1):108700.
- [7] 崔勇强,李嘉轩,侯林果,等. 改进 YOLOv7 的城市小型无人机目标检测方法[J]. 计算机工程与应用, 2024, 60(10):237-245.  
CUI Y Q, LI J X, HOU L G, et al. Improved YOLOv7 target detection method for small urban UAVs [J]. Computer Engineering and Applications, 2024, 60(10): 237-245.

- [8] SUN SH Z, MO B, XU J W, et al. Multi-YOLOv8: An infrared moving small object detection model based on YOLOv8 for air vehicle [J]. *Neurocomputing*, 2024, 588(1):127685.
- [9] QIU SH, CAI B X, WANG W D, et al. Automated detection of railway defective fasteners based on YOLOv8-FAM and synthetic data using style transfer [J]. *Automation in Construction*, 2024, 162(1):105363.
- [10] LI X SH, WANG CH L, ZENG ZH. WS-SSD: Achieving faster 3D object detection for autonomous driving via weighted point cloud sampling[J]. *Expert Systems with Applications*, 2024, 249(1):123805.
- [11] LIU SH, WANG D, WANG Q, et al. NIV-SSD: Neighbor IoU-voting single-stage object detector from point cloud[J]. *Neurocomputing*, 2024, 597(1):127987.
- [12] 张红民,周豪,李顺远,等.改进 YOLOv3 的输电线路异物检测方法[J]. *激光杂志*, 2022, 43(5):82-87.
- ZHANG H M, ZHOU H, LI SH Y, et al. Improved YOLOv3 foreign body detection method in transmission line [J]. *Laser Journal*, 2022, 43 (5): 82-87.
- [13] LI SH Y, WANG ZH W, LYU Y CH, et al. Improved YOLOv5s-based algorithm for foreign object intrusion detection on overhead transmission lines[J]. *Energy Reports*, 2024, 11(1):6083-6093.
- [14] 张剑锐,魏霞,张林鎰,等.改进 YOLOv7 的绝缘子检测与定位 [J]. *计算机工程与应用*, 2024, 60(4): 183-191.
- ZHANG J R, WEI X, ZHANG L X, et al. Improving detection and positioning of insulators in YOLOv7[J]. *Computer Engineering and Applications*, 2024, 60(4): 183-191.
- [15] LIU Z F, ABEYRATHNA R M R D, SAMPURNO R M, et al. Faster-YOLO-AP: A lightweight apple detection algorithm based on improved YOLOv8 with a new efficient PDWConv in orchard[J]. *Computers and Electronics in Agriculture*, 2024, 223(1):109118.
- [16] KANG M, TING C M, TING F F, et al. ASF-YOLO: A novel YOLO model with attentional scale sequence fusion for cell instance segmentation [J]. *Image and Vision Computing*, 2024, 147(1):105057.
- [17] KANG L, LU ZH W, MENG L Y, et al. YOLO-FA: Type-1 fuzzy attention based YOLO detector for vehicle detection [J]. *Expert Systems with Applications*, 2024, 237(11):121209.
- [18] LIN Q CH, TAKAMASU K, CHEN M Y. An optical lens defect detection method for micro vision based on WGSO-YOLO[J]. *IEEE Transactions on Instrumentation and Measurement*, 2024, 73(1):5018508.
- [19] 李忠科,刘小芳.基于轻量级 YOLOv8n 网络的 PCB 缺陷检测算法[J]. *电子测量技术*, 2024, 47(4):120-126.
- LI ZH K, LIU X F. PCB defect detection algorithm based on lightweight YOLOv8n network [J]. *Electronic Measurement Technology*, 2024, 47 (4): 120-126.
- [20] REDMON J, DIVVALA S, GIRSHICK R, et al. You only look once: Unified, real-time object detection[C]. *Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition*, 2016:779-788.
- [21] MA X, DAI X Y, BAI Y, et al. Rewrite the stars[C]. *Proceedings of the 2024 IEEE Conference on Computer Vision and Pattern Recognition*, 2024:5694-5703.
- [22] HU SH, GAO F, ZHOU X W, et al. Hybrid convolutional and attention network for hyperspectral image denoising [J]. *ArXiv preprint arXiv: 2403. 10067*, 2024.
- [23] DAI X, et al. Dynamic head: Unifying object detection heads with attentions [C]. *Proceedings of the 2024 IEEE Conference on Computer Vision and Pattern Recognition*, 2021:7369-7378.
- [24] TONG Z J, CHEN Y H, XU Z W, et al. Wise-IoU: bounding box regression loss with dynamic focusing mechanism[J]. *ArXiv preprint arXiv:2301.10051*, 2023.
- [25] YU Z P, HUANG H B, CHEN W J, et al. YOLO-FaceV2: A scale and occlusion aware face detector[J]. *Pattern Recognition*, 2024, 155:110714.
- [26] CHEN ZH CH, YANG J, FENG ZH CH, et al. RailFOD23: A dataset for foreign object detection on railroad transmission lines[J]. *Scientific Data*, 2024, 11(1):72.

### 作者简介

易磊(通信作者),正高级工程师,主要研究方向为自动控制原理、电力电子学、电机学教学与科研。

E-mail: yilei\_hust@hust. edu. cn

黄哲玮,本科生,主要研究方向为深度学习、机器视觉。

E-mail: whztever@hust. edu. cn

易雅雯,本科生,主要研究方向为深度学习、机器视觉。

E-mail: yawenyi@hust. edu. cn