

基于 YOLOv8 的输电线路绝缘子表面缺陷识别算法^{*}熊 伟^{1,2} 路 鑫¹ 邱维进¹ 王平强¹

(1. 华北电力大学计算机系 保定 071000; 2. 河北省能源电力知识计算重点实验室 保定 071000)

摘 要: 针对当前绝缘子表面缺陷识别存在的图像背景复杂、缺陷小目标识别效果差的问题,提出一种基于 YOLOv8 的输电线路绝缘子表面缺陷识别算法。首先,在主干网络引入 CAF 模块,增强模型对复杂图像场景的解析,增强全局和局部特征的提取能力;其次,在模型的颈部网络增加 GD 机制,减少特征融合过程中信息的丢失,提升小目标检测能力;最后,采用 ATFL 分类损失函数,削弱复杂背景对小目标检测的干扰,引入 PIOUS 边界框损失函数,提高识别精度,加快模型收敛速度。实验结果表明,该算法的 mAP50 达到 94.1%,精确率达到 92.5%,召回率达到 91.3%,相较于基线模型分别提高了 3.1%、0.7%、3.9%,且综合性能优于最近的 YOLOv9s、YOLOv10s 等代表性算法。

关键词: 目标检测;绝缘子表面缺陷识别;小目标;卷积和注意力融合;边界框损失函数;分类损失函数

中图分类号: TM216;TN919.8 **文献标识码:** A **国家标准学科分类代码:** 520.6040

Surface defect detection algorithm of transmission line
insulators based on YOLOv8Xiong Wei^{1,2} Lu Xin¹ Qiu Weijin¹ Wang Pingqiang¹

(1. Department of Computer, North China Electric Power University, Baoding 071000, China;

2. Hebei Key Laboratory of Knowledge Computing for Energy & Power, Baoding 071000, China)

Abstract: Aiming at the problems of complex image background and poor recognition of small defect targets in the current insulator surface defect recognition, a transmission line insulator surface defect recognition algorithm based on YOLOv8 is proposed. Firstly, the CAF module is introduced in the backbone network to enhance the model's analysis of complex image scenes and enhance the ability to extract global and local features; secondly, the GD mechanism is added to the neck network of the model to reduce the loss of information in the feature fusion process and improve the small target detection ability; finally, the ATFL classification loss function is used to weaken the interference of complex background on small target detection, and the PIOUS bounding box loss function is introduced to improve the recognition accuracy and accelerate the model convergence speed. Experimental results show that the mAP50 of the algorithm reaches 94.1%, the precision rate reaches 92.5%, and the recall rate reaches 91.3%, which are 3.1%、0.7% and 3.9% higher than the baseline model, respectively, and the comprehensive performance is better than the recent YOLOv9s, YOLOv10s and other representative algorithms.

Keywords: object detection; insulator surface defect recognition; small targets; convolution and attention fusion; bounding box loss function; classification loss function

0 引 言

绝缘子是输电线路的重要组件,用于支撑和固定输电线路,保障输电线路安全稳定。由于绝缘子长期暴露于户外环境中,受到风雨、灰尘、污染等因素的影响,容易产生自爆、破损、污闪等缺陷,从而影响输电系统的安全运行^[1]。随着深度学习在图像识别领域的快速发展,在绝缘子缺陷

识别中应用深度学习成为研究热点。

基于深度学习的目标检测算法主要包括以 Fast R-CNN^[2]、Faster R-CNN^[3]为代表的基于候选区域的双阶段检测模型和以 SSD^[4]、YOLO 系列^[5]为代表的单阶段检测模型。翟永杰等^[6]以 Faster R-CNN 模型为基础,将原有的主干网络替换为 ResNet101,并加入 SENet 注意力机制成功提高模型的表征能力。高伟等^[7]同样在 Faster R-CNN

模型上加入 SENet 注意力机制,改进候选框选取过程中的纵横比,并将 NMS 算法改进为 Aera-NMS 算法。上述算法一定程度上提高了模型的检测精度,但模型的速度较慢,且需要较多的计算资源。由于单阶段检测模型在推理速度上更具有优势,其在电力巡检中得到广泛应用。王韵琳等^[8]在 YOLOX-s 中融合注意力与多尺度特征提高对绝缘子缺陷小目标的识别效果,但未考虑到复杂背景对绝缘子缺陷识别的影响。为了降低输电线路航拍图像复杂背景的影响,实现绝缘子小目标缺陷检测,部分研究者将绝缘子缺陷检测视为两级检测问题。魏良玉等^[9]先使用改进的 YOLOv5 框架定位绝缘子位置,去除复杂背景的干扰后,再采用 DeepLabv3+ 对绝缘子缺陷进行识别。赵振兵等^[10]先使用未改进的 YOLOv8 模型提取绝缘子部件图像,再通过 ConvNeXt V2 主干网络和边缘知识融合模块改进 YOLOv8 的提取识别能力,使用改进后的模型对绝缘子部件图像进行检测。两级检测模型的缺陷检测精度一般优于单级检测模型,但模型复杂度高,检测速度慢^[11]。邬开俊等^[12]重新设计 YOLOv5 的 Neck 模块,提出 Bi FPN-F 特征融合模块,并在主干网络中设计 DFC-Faster Net 模块,增大感受野,解决自爆缺陷目标较小和背景较复杂的问题。贾晓芬等^[13]在 YOLOv5s 的基础上提出了特征提取模块 NewC3 和轻量化模块 DC-SE,用来提高绝缘子自爆缺陷检测精度并减少模型参数。窦雷雨等^[14]在 YOLOv7s 的基础上设计一个小目标检测层并引入 SimAM 注意力机制,绝缘子自爆缺陷检测精度有所提升,但仍有一定的进步空间。

上述方法在某种程度上实现了模型绝缘子缺陷识别能力的提升,但识别缺陷存在种类单一、缺陷明显、易于识别的问题,在处理绝缘子多缺陷识别和小目标检测时,常会出现漏检和误检的现象,在处理复杂背景干扰下的特征提取

时不够到位。为解决这些问题,本文提出了一种基于 YOLOv8 的输电线路绝缘子表面缺陷识别算法。本文的创新与贡献如下:1)在主干网络的 C2F 模块中引入 CAF 模块,提高网络对全局和局部特征的提取能力,从而增强模型对复杂图像场景的解析能力。2)为解决小目标特征在特征融合过程中的损失问题,在模型的颈部网络增加 GD 机制,汇集浅层信息,提高多尺度特征融合能力,提升小目标检测能力。3)使用 ATFL 分类损失函数,减轻背景相关损失,削弱复杂背景对小目标检测的干扰。4)将 CIOU 边界框损失函数替换为 PIOU,加快模型收敛速度,增强对中等质量锚框的关注能力,提高识别精度。

1 改进的 YOLOv8 模型

Ultralytics 公司于 2023 年推出了 YOLOv8 算法^[15],该算法由 Backbone、Neck 和 Head 三大部分构成。Backbone 主要包括卷积模块、C2F 模块、SPPF 模块,其中 C2F 模块具有分支处理和特征融合的特点,增强了网络的非线性能力,丰富了特征的表达能力。Neck 部分使用 PAN-FPN 结构,通过上采样和卷积操作对不同尺度的特征进行融合,提高模型的多尺度识别能力。Head 部分采用解耦头结构,将分类和检测任务分离。YOLOv8 的损失函数包括分类损失函数和回归损失函数两部分。其中使用 BCE(binary cross-entropy)作为分类损失函数,使用 DFL(distribution focal loss)和 CIOU(complete intersection over union)作为回归损失函数。

YOLOv8 按照模型参数量递增分为 YOLOv8n、YOLOv8s、YOLOv8m、YOLOv8l、YOLOv8x,参数量越多,模型表达能力越强,计算资源的需求也会越大,同时也会增加过拟合发生的概率,本文选取 YOLOv8s 作为改进的基础模型。本文提出的改进 YOLOv8 算法如图 1 所示。

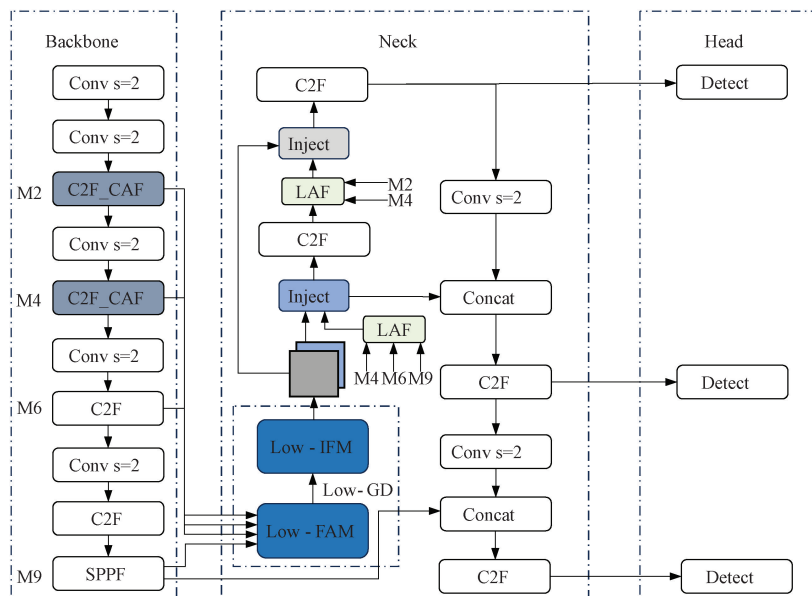


图 1 改进后的 YOLOv8 模型图

Fig. 1 Improved YOLOv8 model

1.1 卷积和注意力融合模块

输电线绝缘子图像背景复杂,包括天空、城市建筑物、植被、其他输电设备等,对于绝缘子缺陷识别的影响较大。因此,本文在主干网络的 C2F 模块中增加卷积和注意力融合(convolution and attention fusion, CAF)模块形成 C2F_CAF 模块,如图 2 所示,提高主干网络对绝缘子全局和局部特征的提取能力,增强模型对复杂图像场景的解析效果。

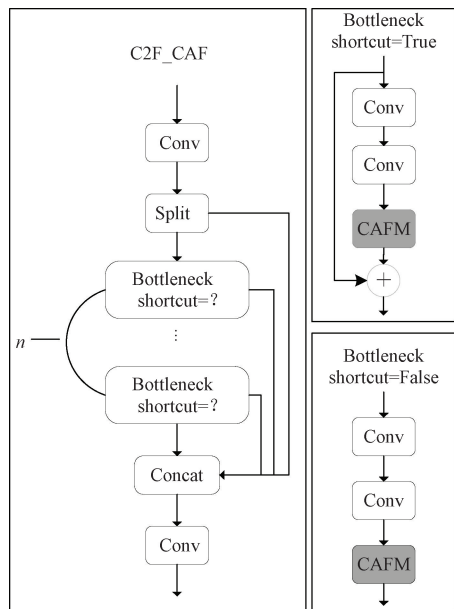


图 2 C2F_CAF 结构图

Fig. 2 C2F_CAF structure

卷积运算受局部性质和感受野的限制,在建模全局特征时存在不足,而基于注意力机制的 Transformer 擅长提取全局特征并捕捉远距离依赖关系。卷积和注意力机制在功能上相辅相成,可以同时建模全局和局部特征。因此, Hu 等^[16]提出 CAF 模块。在全局分支中采用自注意力机制来捕获更广泛的全局特征信息,在局部分支中使用卷积模块提取细致的局部特征,以实现综合去噪。CAF 模块结构如图 3 所示。

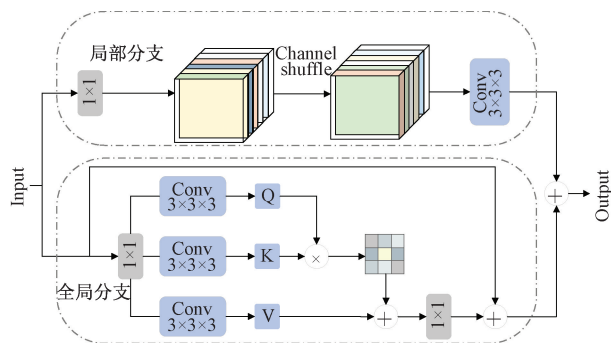


图 3 卷积和注意力融合模块结构图

Fig. 3 Structure of convolution and attention fusion module

CAF 模块由全局分支和局部分支两部分构成。在局部分支中,为了增强通道间交互和信息融合,首先使用 1×1 卷积来调整通道维度,然后进行通道重排操作,进一步融合通道信息。通道重排将输入张量沿通道维度划分为多个组,在每个组内使用深度可分离卷积来实现通道重排。随后,将每个组的输出张量在通道维度上进行拼接,以生成新的输出张量。最后,利用 $3 \times 3 \times 3$ 卷积进一步提取高级特征。局部分支可以表示为:

$$\mathbf{F}_{\text{conv}} = \mathbf{W}_{3 \times 3 \times 3} (\mathbf{CS}(\mathbf{W}_{1 \times 1}(\mathbf{Y}))) \quad (1)$$

其中, \mathbf{F}_{conv} 是局部分支的输出, $\mathbf{W}_{1 \times 1}$ 表示 1×1 卷积, $\mathbf{W}_{3 \times 3 \times 3}$ 表示 $3 \times 3 \times 3$ 卷积, \mathbf{CS} 代表通道重排操作, \mathbf{Y} 是输入特征。

在全局分支中,首先通过 1×1 卷积和 3×3 深度卷积生成查询向量 \mathbf{Q} 、键向量 \mathbf{K} 和值向量 \mathbf{V} ,从而得到形状为 $\hat{H} \times \hat{W} \times \hat{C}$ 的 3 个张量。接着,将 \mathbf{Q} 和 \mathbf{K} 分别重塑为 $\hat{\mathbf{Q}} \in \mathbb{R}^{\hat{H}\hat{W} \times \hat{C}}$ 和 $\hat{\mathbf{K}} \in \mathbb{R}^{\hat{C} \times \hat{H}\hat{W}}$,然后通过 $\hat{\mathbf{Q}}$ 和 $\hat{\mathbf{K}}$ 的交互来计算注意力图 $\mathbf{A} \in \mathbb{R}^{\hat{H}\hat{W} \times \hat{H}\hat{W}}$ 。这样,相较于计算尺寸为 $\mathbb{R}^{\hat{H}\hat{W} \times \hat{H}\hat{W}}$ 的大型常规注意力图,计算负担得以减轻。全局分支的输出定义如下:

$$\mathbf{F}_{\text{att}} = \mathbf{W}_{1 \times 1} \text{Attention}(\hat{\mathbf{Q}}, \hat{\mathbf{K}}, \hat{\mathbf{V}}) + \mathbf{Y} \quad (2)$$

$$\text{Attention}(\hat{\mathbf{Q}}, \hat{\mathbf{K}}, \hat{\mathbf{V}}) = \hat{\mathbf{V}} \text{Softmax}(\hat{\mathbf{K}} \hat{\mathbf{Q}} / \alpha) \quad (3)$$

其中, $\hat{\mathbf{K}}$ 是一个可学习的缩放参数,用于控制矩阵 $\hat{\mathbf{K}}$ 和 $\hat{\mathbf{Q}}$ 相乘时的幅度,以便在应用 Softmax 函数之前调整值的大小。

最后,CAF 模块计算输出为:

$$\mathbf{F}_{\text{out}} = \mathbf{F}_{\text{att}} + \mathbf{F}_{\text{conv}} \quad (4)$$

1.2 低阶信息聚集-分发模块

YOLOv8 在颈部结构中采用 PAN-FPN 结构来融合不同尺度的特征图,增强模型的多尺度目标检测能力。但 PAN-FPN 结构只能直接获得相邻层的特征,对于其他层的信息,只能有损失地间接获得。如果第一层想要使用第三层的信息,必须先将二、三层信息融合,然后一层信息通过结合二层信息间接获得层级信息。由于层与层之间的信息交互仅限于中间层选择的信息,其余信息会在传输过程中被丢弃,造成有用信息丢失。因此,某一层的信息只能有效地帮助相邻层,对其他全局层的支持则较弱,这种情况可能会限制整体的信息融合效果。

小目标通常在低层级特征图中呈现较为细致的结构信息,融合不同尺度的特征图时造成的信息丢失对小目标的识别精度影响更大。为此,在 YOLOv8 的颈部结构中嵌入低阶信息聚集-分发(low-stage gather-and-distribute, Low-GD)机制^[17],充分融合不同尺度的特征图信息,在不显著增加延迟的情况下提高模型的小目标检测能力。如图 4 所示,Low-GD 机制主要分为 3 个模块:特征对齐模块

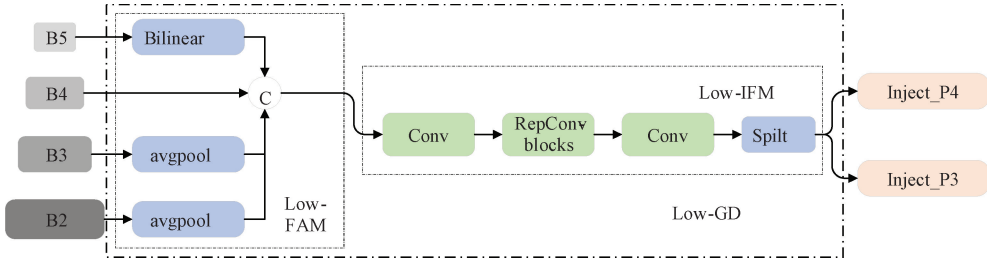


图4 低阶信息聚集-分发结构图

Fig. 4 Low-level information aggregation-distribution structure

(feature alignment module, FAM)、信息融合模块 (information fusion module, IFM)和信息注入模块 (Inject)。

在 Low-FAM 中,通过平均池化 (AvgPool) 操作对输入特征进行下采样,用来统一特征图的尺寸。特征图尺寸越大,保留的底层信息就越多,但后续的计算延迟就会增加。为了达到速度和精度之间的平衡,将统一特征图的尺寸调整为 B_4 的大小,得到 F_{align} , 公式如下:

$$F_{align} = \text{Low_FAM}([B_2, B_3, B_4, B_5]) \quad (5)$$

其中, B_2, B_3, B_4, B_5 为图片不同尺寸的特征图, F_{align} 为 Low-FAM 的输出特征图。

Low-IFM 包括 RepBlock 模块和拆分操作, F_{align} (通道数为 $C_{B_2}, C_{B_3}, C_{B_4}$ 和 C_{B_5} 的总和) 作为 RepBlock 的输入, 生成 F_{fuse} (通道数为 $C_{B_4} + C_{B_5}$)。RepBlock 生成的特征随后在通道维度上被分割为 F_{inj_P3} 和 F_{inj_P4} , 分别与不同层级的特征进行融合, 公式如下:

$$F_{fuse} = \text{RepBlock}(F_{align}) \quad (6)$$

$$F_{inj_P3}, F_{inj_P4} = \text{Split}(F_{fuse}) \quad (7)$$

其中, F_{fuse} 为 RepBlock 模块的输出特征图, F_{inj_P3} 和 F_{inj_P4} 为 Split 模块的输出特征图。

如图 5 所示,在 Inject 模块中,将局部信息 F_{local} 和全局注入信息 F_{inj} (IFM 生成特征) 分别作为输入,利用两个不同的卷积层对 F_{inj} 进行处理,并采用平均池化或双线性插值的方法调整两者的尺寸,得到 F_{inj_cs} 和 F_{inj_c} 。在 F_{local} 经卷积层处理后,对 3 个分支进行融合。在融合结束后,通过 RepBlock 进一步提取和融合信息。公式如下:

$$F_{inj_cs} = \text{resize}(\text{Sigmoid}(\text{Conv}(F_{inj_P3}))) \quad (8)$$

$$F_{inj_c} = \text{resize}(\text{Conv}(F_{inj_P4})) \quad (9)$$

$$F_{cal} = \text{Conv}(F_{local}) * F_{inj_cs} + F_{inj_c} \quad (10)$$

$$F_{out} = \text{RepBlock}(F_{cal}) \quad (11)$$

其中, F_{inj_cs} 和 F_{inj_c} 是 F_{inj_P3} 和 F_{inj_P4} 通过不同的卷积层得到的特征图, F_{cal} 是 F_{inj_cs}, F_{inj_c} 和 F_{local_c} 三者融合得到, F_{out} 是 RepBlock 模块的输出特征图。

为了进一步在信息注入的过程中融合更多的特征信息,使用轻量层级融合模块 (lightweight adjacent layer fusion, LAF) 将当前层级特征与其他两层的特征融合作为 F_{local} 的输入。由于不同层次特征尺寸不同,通过双线性插值调整小尺度特征,平均池化调整大尺度特征, 1×1 卷积

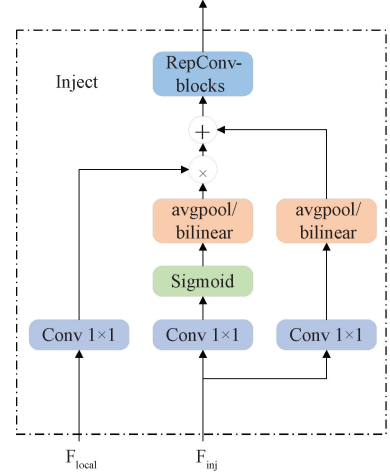


图5 信息注入模块结构图

Fig. 5 Information injection module structure

调整与目标通道不同的特征,如图 6 所示。

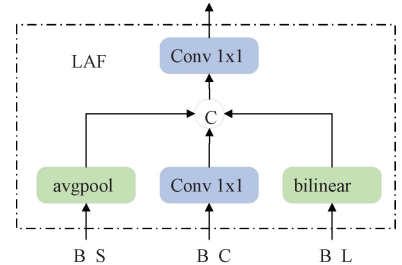


图6 轻量层级融合模块结构图

Fig. 6 Lightweight hierarchical fusion module structure

1.3 分类损失函数设计

绝缘子破损和污闪缺陷往往只占绝缘子航拍图像的极少部分,模型识别目标与背景之间存在极大的不平衡现象,这会导致模型在训练时更容易关注到背景特征,从而降低目标识别的准确性。为解决此类问题,在分类损失函数中引入自适应阈值焦点损失 (adaptive threshold focal loss, ATFL) 函数^[18],将识别目标和背景解耦,并利用自适应机制调整损失权重,使模型更加关注目标特征。

在 YOLOv8 中使用的是二元交叉熵 BCE 损失函数, 公式如下:

$$L_{BCE} = -(y \log(p) + (1 - y) \log(1 - p)) \quad (12)$$

其中, y 是真实标签, 取值为 0 或 1, p 是模型对真实类别的预测概率。

BCE 损失函数无法解决样本之间的不平衡问题。焦点损失函数(Focal Loss)^[19]引入 $(1 - p_i)^\gamma$, 通过调整聚焦参数 γ 来降低易分类样本对损失函数的影响, 公式如下:

$$FL(p_i) = -(1 - p_i)^\gamma \log(p_i) \quad (13)$$

Focal Loss 可以通过调节参数 γ 降低易分类样本的损失权重, 但同时也会对正常样本产生相同影响。为解决此类问题, 使用 ATFL 损失函数进行样本分类。

ATFL 损失函数使用阈值将样本分为易分类样本和难分类样本, 并使超参数具有自适应改进的能力, 公式如下:

$$ATFL = \begin{cases} -(\lambda - p_i)^\eta \log(p_i), & p_i \leq 0.5 \\ -(1 - p_i)^\eta \log(p_i), & p_i > 0.5 \end{cases} \quad (14)$$

$$\hat{p}_c = 0.05 \times \frac{1}{t-1} \sum_{i=0}^{t-1} \overline{p_i} + 0.95 \times p_i \quad (15)$$

其中, p_i 表示当前平均预测概率值, \hat{p}_c 表示下一个 epoch 的预测值, $\overline{p_i}$ 表示每个训练 epoch 的平均预测概率值。

1.4 边界框回归损失函数设计

在目标检测中, 边界框回归损失函数用来衡量模型预测的边界框与真实边界框之间的差异, 并通过最小化差异来优化模型。在 YOLOv8 算法中, 采用 CIOU 作为模型的边界框回归损失函数。CIOU 在 IOU 的基础上引入预测框和真实框中心点的欧氏距离惩罚项和边界框宽高比惩罚项, CIOU 边界框回归损失函数定义如下:

$$IOU = \frac{\mathbf{B}_{gt} \cap \mathbf{B}_{prd}}{\mathbf{B}_{gt} \cup \mathbf{B}_{prd}} \quad (16)$$

$$CIOU = IOU - \frac{\rho^2(\mathbf{b}_{gt}, \mathbf{b}_{prd})}{C^2} - \alpha V \quad (17)$$

$$V = \frac{4}{\pi^2} \left(\arctan \frac{w_{gt}}{h_{gt}} - \arctan \frac{w_{prd}}{h_{prd}} \right)^2 \quad (18)$$

$$\alpha = \frac{V}{1 - IOU + V} \quad (19)$$

其中, \mathbf{B}_{gt} 和 \mathbf{B}_{prd} 分别代表真实框和预测框, \mathbf{b}_{gt} 和 \mathbf{b}_{prd} 为真实框和预测框的中心坐标, ρ^2 为两个中心点的距离的平方, C 是包含预测框和真实框的最小外接框的对角线长度, α 和 V 为衡量边界框宽高比的参数, w_{gt} 和 h_{gt} 代表真实框的宽和高, w_{prd} 和 h_{prd} 代表预测框宽和高。

如式 (20) 所示, CIOU 在惩罚项中使用了 $R = \frac{\rho^2(\mathbf{b}_{gt}, \mathbf{b}_{prd})}{C^2}$ 。当 C 增大时, R 变小, 反之亦然。因此, 当真实框和预测框不重叠时, L_{IOU} 保持不变, 只需增加预测框的尺寸即可减少 L_{CIOU} , 这表明在惩罚项中使用 C 作为分母是不合理的。

针对这一问题, 引入强大交并比 (powerful intersection over union, PIOUS)^[20] 替代 CIOU, 将目标框大小作为分母

的惩罚因子, 并结合自适应锚框质量函数, 增强对中等质量锚框的关注能力, 引导锚框沿着更直接的路径高效回归, 使模型具有更快的收敛速度和更高的准确性。

PIOU 边界框回归损失函数定义如下:

$$P = \left(\frac{dw_1}{w_{gt}} + \frac{dw_2}{w_{gt}} + \frac{dh_1}{h_{gt}} + \frac{dh_2}{h_{gt}} \right) / 4 \quad (20)$$

$$f(x) = 1 - e^{-x^2} \quad (21)$$

$$PIoU = IoU - f(P), -1 \leq PIoU \leq 1 \quad (22)$$

$$L_{PIoU} = L_{IoU} + f(P), 0 \leq L_{PIoU} \leq 2 \quad (23)$$

其中, P 为适应目标尺寸的惩罚因子, dw_1 、 dw_2 、 dh_1 和 dh_2 分别为预测框对应边缘与目标框之间距离的绝对值, w_{gt} 和 h_{gt} 为目标框的宽度和高度, $f(x)$ 为自适应锚框质量函数。

惩罚因子 P 的分母只依赖于目标框的大小, 而与锚框和目标框的最小外部框无关, 锚框的大小不会影响 P 。除非锚框与目标框完全重叠, 否则 P 永远不会退化为 0。

对于非常差的锚框, 如完全没有重叠或位置非常不准确的锚框, 在优化过程中难以调整到正确的位置, 其梯度应较小, 以避免对模型产生过大的影响; 对于非常好的锚框, 如接近真实框的锚框, 误差很小, 梯度也应较小, 以避免过度调整, 导致预测框偏离正确位置。如图 7 所示, 函数 $f(x)$ 符合上述要求, 当 P 较大时, 锚框与目标框之间存在显著差异, 由图 7 中的分图(a)可知, $f(P)$ 也较大, 由图 7 中的分图(b)可知, $f'(P)$ 取较小的值, 可以抑制来自低质量锚框的有害梯度; 当 P 趋近于 0 时, 锚框接近目标框, 随着锚框质量的提高, $f(P)$ 也趋近于 0, 同时 $f'(P)$ 逐渐减小, 从而使锚框能够稳定地优化, 直至与目标框完全对齐。这样能够使模型专注于中等质量的锚框, 从而提高模型精度。

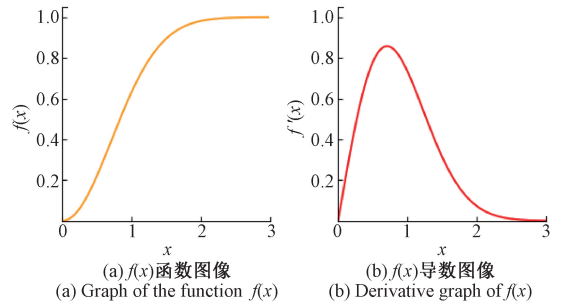


图 7 $f(x)$ 及其导数的函数图像
Fig. 7 Graph of $f(x)$ and its derivative

2 实验与结果分析

2.1 实验环境及数据集

本实验在 Ubuntu20.04 系统下采用 pytorch 深度学习框架进行训练, 具体配置如表 1 所示。

在 YOLOv8 训练过程中, 参数设置如下: 早停机制 (patience) 50, 迭代轮次 (epochs) 为 400 次, 批量大小 (batch) 为 8, 优化器 (optimizer) 为 AdamW, 学习率 (lr) 为

表 1 实验环境配置

Table 1 Experimental environment configuration

配置	版本参数
操作系统	Ubuntu20.04
处理器(CPU)	Intel(R) Core(TM) i7-12700F
运行内存	64 GB
显卡(GPU)	NVIDIA GeForce RTX 3080
显存	10 GB
Pytorch	2.1.0
torchvision	0.16.0
CUDA	12.0
Python	3.8.18

0.002,动量大小(momentum)为 0.9,权重衰减(decay)为 0.000 5。

本实验采用的数据集为常州大学和中国科学院于 2022 年发布的 Insulator-Defect Detection Dataset^[21],该数据集包含 1 600 张图像,其中标记对象个数为 5 373,包含绝缘子、绝缘子破损和绝缘子污闪 3 类标注对象。将该数据集按照 8:1:1 的比例划分为训练集、验证集和测试集对模型进行训练,各类绝缘子缺陷样本如图 8 所示。

2.2 评价指标

为了验证本文对 YOLOv8 的改进效果,实验采用平均精度(average precision, AP)评估单个类别的识别效果,使用平均类别精度(mean average precision, mAP)评估所有类别平均识别效果,采用参数量(Params)衡量模型的复杂程度。计算公式如下:

$$AP = \int_0^1 P(r)dr \tag{24}$$

表 2 消融实验

Table 2 Ablation experiment

实验	PIOU	Low-GD	ATFL	C2F_CAF	mAP50/%	mAP50:95/%	Params/M
1					91.0	61.2	11.1
2	✓				91.8	62.0	11.1
3		✓			92.7	62.1	12.1
4			✓		91.7	62.2	11.1
5				✓	92.6	62.3	11.2
6	✓	✓			93.4	62.2	12.1
7	✓	✓	✓		93.7	62.6	12.1
8	✓	✓	✓	✓	94.1	63.1	12.2

实验 1 为 YOLOv8s 基础模型,未加入任何改进模块,mAP50 值为 91%,mAP50:95 为 61.2%,参数量为 11.1 M,以此实验作为改进效果的评判标准。在实验 2 中将 CIOU 损失函数替换为 PIOU 损失函数,增强对中等质量锚框的关注能力,mAP50 提高了 0.8%,mAP50:95 提升了 0.8%。在实验 3 中引入 GD 机制,减少特征融合过程中有用信息的损失,提高多尺度特征融合能力,mAP50 提高了 1.7%,



图 8 绝缘子缺陷样例图

Fig. 8 Insulator defect sample

$$mAP = \frac{1}{n} \sum_{i=0}^n AP_{(i)} \tag{25}$$

其中, $P(r)$ 为精确率-召回率(Precision-Recall)曲线, n 为识别类别数,本实验中 $n=4$ 。mAP50 代表 IOU 大于等于 50 的模型平均精度,mAP50:95 代表 IOU 从 0.5~0.95(步长为 0.05)10 个阈值下 mAP 的均值。

2.3 消融实验

为了验证本文改进方法中每个模块的合理性和有效性,在绝缘子缺陷检测数据集上通过消融实验进行验证,其中,“✓”代表采用改进方法,实验结果如表 2 所示。

mAP50:95 提升了 0.9%,参数量小幅度上升。在实验 4 中将 BCE 替换为 ATFL 损失函数,削弱复杂背景对小目标检测的干扰,mAP50 提高了 0.7%,mAP50:95 提升了 1%。

在实验 5 中设计 C2F_CAF 模块,增强模型对复杂图像场景的解析能力,mAP50 提高了 1.6%,mAP50:95 提升了 1.1%,参数量小幅度上升。为进一步验证改进后 C2F_CAF 模块的优势,对 C2F 模块改进前后进行热力图

可视化对比,如图 9 所示。在热力图中,颜色亮度的高低与模型的关注程度成正比,由对比结果可知,输电线路绝缘子图像的复杂背景对 C2F 模块影响较大,从而容易出现

漏检、错检的现象。在 C2F 模块中加入 CAF 模块后,模型重新将注意放在绝缘子的缺陷上,降低复杂背景的干扰,从而提高模型的小目标检测能力。

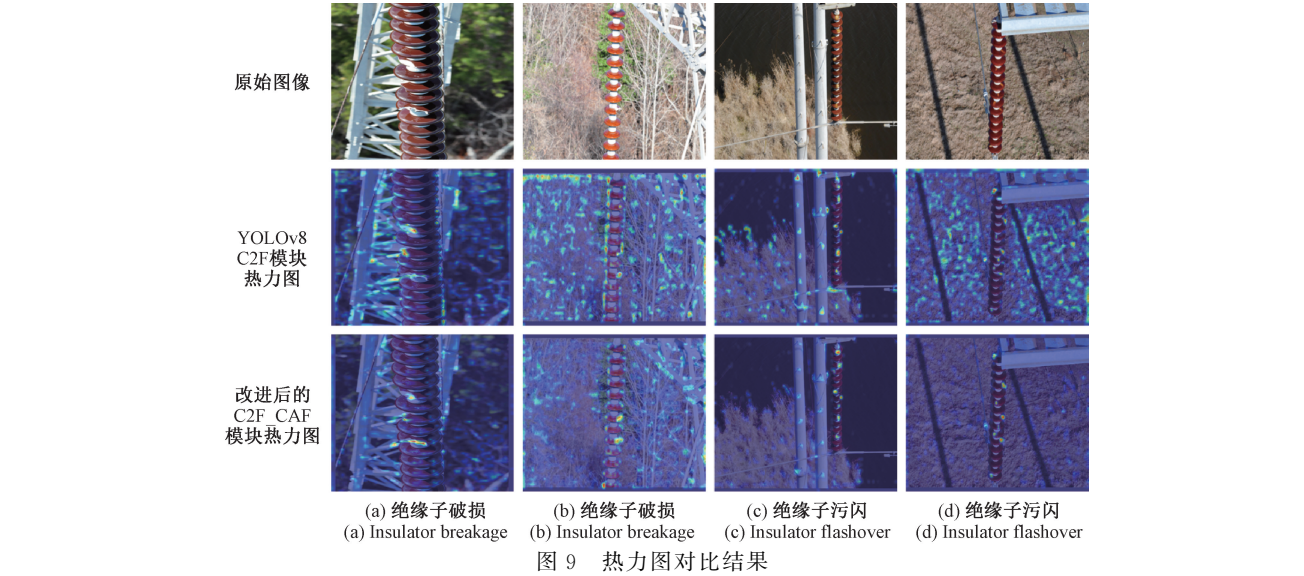


Fig. 9 Comparison of heatmap results

2.4 对比实验

C2F 模块中分别加入 ECA^[22]、CA^[23]、PSA^[24]、PPA^[25]、CoT^[26] 模块与 CAF 模块进行对比,实验结果如表 3 所示。

1)注意力机制对比实验
为了进一步验证 C2F_CAF 模块设计的合理性,在

表 3 注意力机制对比结果
Table 3 Comparison of attention mechanism results

模块	AP50/%	AP50/%	AP50/%	mAP50/%	mAP50:95/%
	绝缘子	破损	污闪		
baseline	99.2	93.3	80.4	91.0	61.2
ECA	99.3	92.7	83.1	91.7	61.7
CA	99.3	94.7	81.1	91.7	62.1
PSA	99.4	93.7	81.6	91.6	61.4
PPA	99.4	93.5	80.4	91.1	61.3
CoT	99.2	96.0	81.3	92.2	62.1
CAF	99.4	95.8	82.7	92.6	62.3

从对比结果可以看出,注意力机制的加入对网络的识别能力有一定的提升。其中,PSA、PPA、CAF 对绝缘子 AP50 的提升效果最大,提高了 0.2%。CoT 对破损缺陷 AP50 的提升效果最大,提高了 2.7%。ECA 对污闪缺陷 AP50 的提升效果最大,提高了 2.7%。CAF 对 mAP50 和 mAP50:95 的提升效果最大,分别提高了 1.6%和 1.1%,且在破损和污闪识别方面,与最高水准仅相差 0.2%和 0.4%。综合来看,CAF 模块对模型整体的提升效果最好,进一步证明其提高模型识别能力的有效性。

2)边界框损失函数对比实验
为了体现 PIOUS 损失函数针对绝缘子缺陷识别的优越性,将其与 CIOUS、SIOUS^[27]、DIOUS^[28]、GIOUS^[29]、Focal CIOUS、

Focal DIOUS 进行对比实验,实验结果如图 10、图 11 所示。
以 YOLOv8 使用的 CIOUS 作为评价指标,对各类损失函数进行分析。由图 10 可知,Focal CIOUS 和 Focal DIOUS 相较于未改进前的 CIOUS 和 DIOUS,mAP50:95 得到一定提升,但是 mAP50 有着不同程度的下降。SIOUS 与 CIOUS 相比,mAP50 没有明显的提升,mAP50:95 没有提升。DIOUS 与 CIOUS 相比,mAP50 提升了 0.4%,mAP50:95 下降了 0.2%。GIOUS 与 CIOUS 相比,mAP50 提升了 0.8%,mAP50:95 没有提升。PIOUS 与 CIOUS 相比,mAP50 提升了 0.8%,mAP50:95 提升了 0.8%,且在比较的损失函数中,mAP50 和 mAP50:95 最高。且由图 11 可知,PIOUS 相较于其他模型,有着最快的收敛速度,且最终收敛值最小,

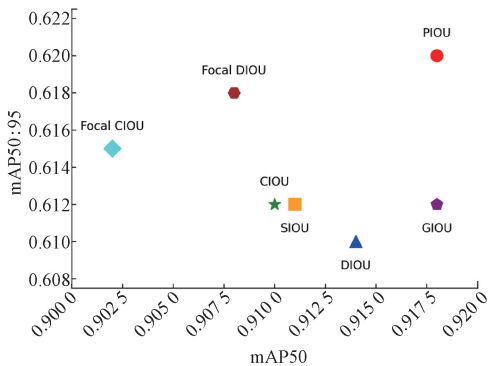


图 10 损失函数对比图
Fig. 10 Comparison of loss functions

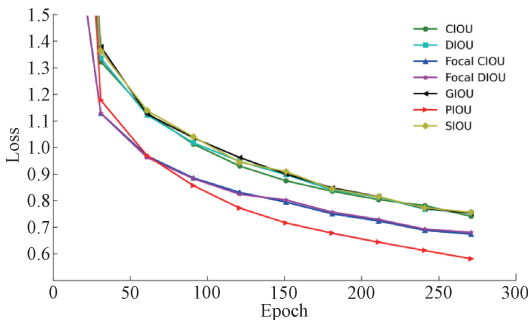


图 11 损失函数收敛性能对比图
Fig. 11 Comparison of loss function convergence performance

验证了 PIOU 损失函数针对绝缘子缺陷识别的优越性。

2.5 不同算法对比实验

为了验证本文算法的有效性,在相同实验环境下对其和现阶段常用目标检测算法进行对比实验,对比方法包括 RetinaNet、YOLOv5s、YOLOv6s、YOLOv8s、Gold-YOLO、YOLOv9s、YOLOv10s、文献[30]和文献[31]实验结果如表 4 所示。

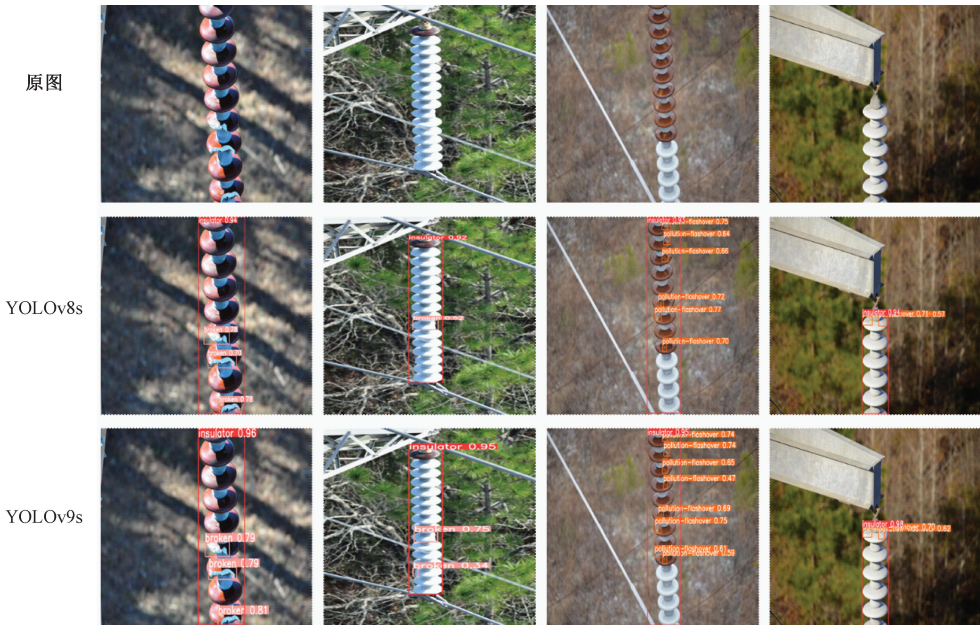
表 4 不同算法对比结果

Table 4 Comparison of different algorithms results

算法	参数量/ M	Precision	Recall	mAP50/ %	FPS
RetinaNet	36.4	97.2	59.3	72.7	—
YOLOv5s	9.11	90.9	87.4	90.9	190
YOLOv6s	16.3	91.9	88.0	91.7	174
YOLOv8s	11.1	91.8	87.4	91	173
Gold-YOLO	21.5	89.8	78.0	87.6	149
YOLOv9s	9.6	93.9	88.7	92.7	156
YOLOv10s	8.0	92.7	86.4	91.2	178
YOLOv10m	16.5	93.5	85.1	91.4	121
文献[30]	22.8	92.3	91.0	93.2	66
文献[31]	—	92.1	91.5	92.0	53
本文	12.2	92.5	91.3	94.1	132

实验结果分析可知,本文算法与原算法相比参数量增加了 1.1 M,mAP 提高了 3.1%,模型识别能力得到提升。与之前流行的 RetinaNet 算法、YOLOv5s 算法、YOLOv6s 算法相比,mAP 分别提高了 21.4%、3.2%、2.4%。与最近提出的 Gold-YOLO、YOLOv9s、YOLOv10s 算法、YOLOv10m 算法相比,mAP 分别提高了 6.5%、1.4%、2.9%、2.7%。与文献[30]、文献[31]相比,mAP 分别提高了 0.8%、2.1%。本文算法的精确率和召回率与原算法相比分别提高了 0.7%、3.9%,与其他算法相比虽然没有达到最好的精确率和召回率,但二者较为均衡。本文算法的参数量与其他算法相比,参数量适中,适合在有限运算资源的环境下运行。

为了进一步验证本文算法相较于其他算法的优越性,选取 4 张代表性绝缘子图像进行对比检测,包括 2 张绝缘子破损图像、2 张绝缘子污闪图像,对比结果如图 12 所示。



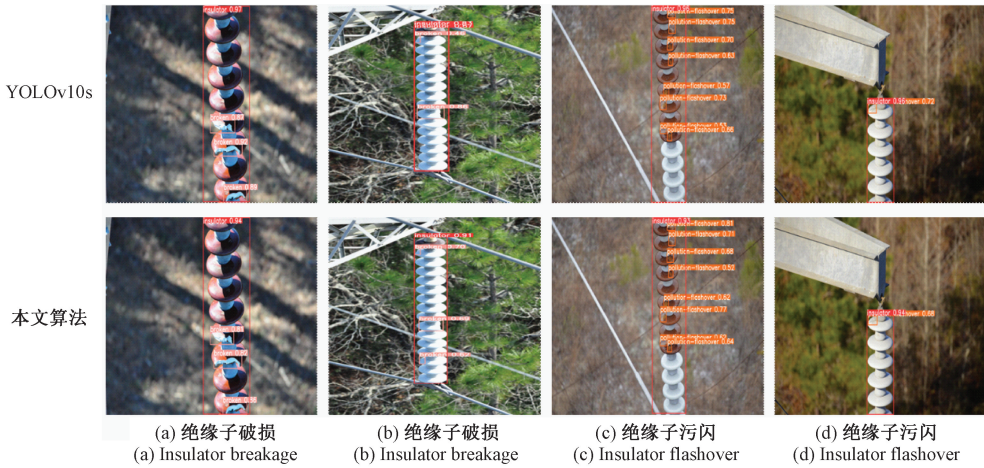


图 12 绝缘子图像可视化对比结果

Fig. 12 Comparison of insulator image visualization results

对于第 1 列图片来说,4 类算法都有着不错的识别效果,识别出了全部的绝缘子破损缺陷。对于第 2 列图片,由于复杂背景干扰,以及破损区域小的因素,YOLOv8s 和 YOLOv10s 算法漏检 1 处绝缘子破损,而 YOLOv9s 和本文算法全部认出。对于第 3 列图片,YOLOv8 算法漏检 2 处绝缘子污闪,其他算法认出全部缺陷,未出现漏检、错检现象。对于第 4 列图片,YOLOv8s 和 YOLOv9s 算法误将绝缘子上铭刻的信息误认为污闪,而 YOLOv10s 和本文算法全部正确认出。因此,在输电线路绝缘子表面缺陷识别任务中,改进后的算法针对不同复杂背景下的小目标缺陷,具有不错的识别效果。

3 结 论

针对绝缘子图像背景复杂、缺陷小目标识别效果差的问题,本文提出了一种基于 YOLOv8 的输电线路绝缘子表面缺陷识别算法。首先,设计 C2F_CAF 提高主干网络的特征提取能力。其次,在颈部网络增加 GD 机制,提高网络的特征融合能力,减少有用信息丢失。然后,采用 ATFL 损失函数降低复杂背景对模型识别的干扰。最后,使用 PIOU 边界框损失函数增强模型对中等质量锚框的关注能力,提高模型收敛速度和识别准确率。与其他模型相比,综合性能最优,全类平均精度达到 94.1%,能够基本满足绝缘子表面缺陷识别的需要。下一阶段研究重点将聚焦在模型轻量化,以满足现场各类计算资源的需求。

参考文献

[1] LI Y CH, ZHANG W B, LI P, et al. A method for autonomous navigation and positioning of UAV based on electric field array detection[J]. Sensors, 2021, 21(4): 1146.

[2] 彭豪,李晓明. 基于改进 Faster R-CNN 的小目标检测模型[J]. 电子测量技术, 2021, 44(24): 122-127.

PENG H, LI X M. Small target detection model based on improved Faster R-CNN [J]. Electronic Measurement Technology, 2021, 44(24): 122-127.

[3] MAITY M, BANERJEE S, CHAUDHURI S S. Faster R-CNN and YOLO based vehicle detection: A survey [C]. 2021 5th International Conference on Computing Methodologies and Communication (ICCMC), 2021: 1442-1447.

[4] LU X C, JI J, XING Z Q, et al. Attention and feature fusion SSD for remote sensing object detection [J]. IEEE Transactions on Instrumentation and Measurement, 2021, 70: 1-9.

[5] JIANG P Y, ERGU D J, LIU F Y, et al. A review of YOLO algorithm developments [J]. Procedia Computer Science, 2022, 199: 1066-1073.

[6] 翟永杰,王璐瑶,郭聪彬. 基于 Faster R-CNN 的复杂背景下绝缘子目标检测[J]. 电子测量技术, 2023, 46(16): 187-194.

ZHAI Y J, WANG L Y, GUO C B, et al. Insulator object detection in complex background based on Faster R-CNN [J]. Electronic Measurement Technology, 2023, 46(16): 187-194.

[7] 高伟,董云云,刘军,等. 融合注意力机制的输电部件及缺陷检测模型[J]. 计算机工程与设计, 2023, 44(3): 929-936.

GAO W, DONG Y Y, LIU J, et al. Transmission component and defect detection model that incorporating attention mechanism [J]. Computer Engineering and Design, 2023, 44(3): 929-936.

[8] 王韵琳,冯天波,孙宁,等. 融合注意力与多尺度特征的电力绝缘子缺陷检测方法[J]. 高电压技术, 2024, 50(5): 1933-1942.

- WANG Y L, FENG T B, SUN N, et al. Defect detection method for power insulators based on attention and multi-scale context information [J]. High Voltage Engineering, 2024, 50(5): 1933-1942.
- [9] 魏良玉, 邹国锋, 赵新宇, 等. 面向航拍图像的绝缘子弱特征缺陷两阶段检测方法[J]. 国外电子测量技术, 2023, 42(10): 25-34.
- WEI L Y, ZOU G F, ZHAO X Y, et al. Two-stage detection method for weak feature defects of insulators in aerial images[J]. Foreign Electronic Measurement Technology, 2023, 42(10): 25-34.
- [10] 赵振兵, 韩钰, 唐辰康. 基于改进 YOLOv8 的配电线路绝缘子缺陷级联检测方法[J/OL]. 图学学报, 1-13 [2024-11-29]. <http://kns.cnki.net/kcms/detail/10.1034.t.20241112.0948.004.html>.
- ZHAO ZH B, HAN Y, TANG CH K. Cascade detection method for insulator defects in distribution lines based on improved YOLOv8[J/OL]. Journal of Graphic Engineering, 1-13 [2024-11-29]. <http://kns.cnki.net/kcms/detail/10.1034.t.20241112.0948.004.html>.
- [11] 刘传洋, 吴一全, 刘景景. 无人机航拍图像中绝缘子缺陷检测的深度学习研究方法研究进展[J/OL]. 电工技术学报, 1-21 [2024-11-29]. <https://doi.org/10.19595/j.cnki.1000-6753.tces.240610>.
- LIU CH Y, WU Y Q, LIU J J. Research progress of deep learning methods for insulator defect detection in uav based aerial images[J/OL]. Transactions of China Electrotechnical Society, 1-21 [2024-11-29]. <https://doi.org/10.19595/j.cnki.1000-6753.tces.240610>.
- [12] 邬开俊, 徐泽浩, 单宏全. 基于 FasterNet 和 YOLOv5 改进的玻璃绝缘子自爆缺陷快速检测方法[J]. 高电压技术, 2024, 50(5): 1865-1876.
- WU K J, XU Z H, SHAN H Q, et al. Rapid detection method for self-exploding defects in glass insulators based on improved FasterNet and YOLOv5[J]. High Voltage Engineering, 2024, 50(5): 1865-1876.
- [13] 贾晓芬, 吴雪茹, 赵佰亭. 绝缘子自爆缺陷的轻量化检测网络 DE-YOLO[J]. 电子测量与仪器学报, 2023, 37(5): 28-35.
- JIA X F, WU X R, ZHAO B T. Lightweight detection network for insulator self-detonation defect DE-YOLO[J]. Journal of Electronic Measurement and Instrumentation, 2023, 37(5): 28-35.
- [14] 窦雷雨, 尹向雷, 艾格帆, 等. 基于改进 YOLO 的绝缘子缺陷检测算法[J]. 电工技术, 2023(21): 30-35.
- DOU L Y, YIN X L, AI K F, et al. Improved YOLO-based insulator defect detection method[J]. Electrical Engineering, 2023(21): 30-35.
- [15] SOHAN M, SAIRAM T, RAMIREDDY V, et al. A review on yolov8 and its advancements [C]. International Conference on Data Intelligence and Cognitive Informatics, 2024: 529-545.
- [16] HU SH, GAO F, ZHOU X W, et al. Hybrid convolutional and attention network for hyperspectral image denoising[J]. IEEE Geoscience and Remote Sensing Letters, 2024, 21: 1-5.
- [17] WANG CH CH, HE W, NIE Y, et al. Gold-YOLO: Efficient object detector via gather-and-distribute mechanism [J]. Advances in Neural Information Processing Systems, 2024, 36. DOI:10.48550/arXiv.2309.11331.
- [18] YANG B, ZHANG X Y, ZHANG J, et al. EFLNet: Enhancing feature learning network for infrared small target detection[J]. IEEE Transactions on Geoscience and Remote Sensing, 2024, 62: 1-11.
- [19] MUKHOTI J, KULHARIA V, SANYAL A, et al. Calibrating deep neural networks using focal loss[J]. Advances in Neural Information Processing Systems, 2020, 33: 15288-15299.
- [20] LIU C, WANG K G, LI Q, et al. Powerful-IoU: More straightforward and faster bounding box regression loss with a nonmonotonic focusing mechanism [J]. Neural Networks, 2024, 170: 276-284.
- [21] ZHENG J F, WU H, ZHANG H, et al. Insulator-defect detection algorithm based on improved YOLOv7 [J]. Sensors, 2022, 22(22): 8801.
- [22] WANG Q L, WU B G, ZHU P F, et al. ECA-Net: Efficient channel attention for deep convolutional neural networks [C]. IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2020: 11534-11542.
- [23] HOU Q B, ZHOU D Q, FENG J SH. Coordinate attention for efficient mobile network design [C]. IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2021: 13713-13722.
- [24] LIU H J, LIU F Q, FAN X Y, et al. Polarized self-attention: Towards high-quality pixel-wise regression[J]. ArXiv preprint arXiv:2107.00782, 2021.
- [25] XU SH B, ZHENG SH CH, XU W H, et al. HCF-Net: Hierarchical context fusion network for infrared small object detection [J]. ArXiv preprint arXiv: 2403.10778, 2024.
- [26] LI Y H, YAO T, PAN Y W, et al. Contextual transformer networks for visual recognition[J]. IEEE

Transactions on Pattern Analysis and Machine Intelligence, 2022, 45(2): 1489-1500.

[27] GEVORGYAN Z. SiIoU loss: More powerful learning for bounding box regression [J]. ArXiv preprint arXiv:2205. 12740,2022.

[28] ZHENG ZH H, WANG P, LIU W, et al. Distance-IoU loss: Faster and better learning for bounding box regression [J]. AAAI Conference on Artificial Intelligence, 2020, 34(7): 12993-13000.

[29] REZATOFIGHI H, TSOI N, GWAK J Y, et al. Generalized intersection over union: A metric and a loss for bounding box regression [C]. IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2019: 658-666.

[30] 朱泓宇,程焕新,骆晓玲. 改进 YOLOv8 网络在绝缘子缺陷检测中的应用[J]. 水电能源科学, 2024, 42(5): 183-187.

ZHU H Y, CHENG H X, LUO X L. Application of improved YOLOv8 network in insulator defect detection [J]. Water Resources and Power, 2024, 42(5): 183-187.

[31] 高黎明. 基于轻量化目标检测的绝缘子缺陷识别[J]. 高压电器, 2023, 59(12): 237-244.

GAO L M. Insulator defect identification based on lightweight object detection [J]. High Voltage Apparatus, 2023, 59(12): 237-244.

作者简介

熊伟, 博士, 高级工程师, 主要研究方向为深度学习、故障诊断。

E-mail: xiongwei@ncepu.edu.cn

路鑫(通信作者), 硕士研究生, 主要研究方向为计算机视觉。

E-mail: 1763773457@qq.com

邱维进, 硕士研究生, 主要研究方向为计算机视觉。

E-mail: 1665813376@qq.com

王平强, 硕士研究生, 主要研究方向为计算机视觉。

E-mail: wangpg2001@qq.com