

基于 MHD-YOLO 的轻量化绝缘子缺陷检测算法^{*}

关燕鹏 傅芃博 要会娟

(山西大学自动化与软件学院 太原 030013)

摘 要: 针对无人机巡检输电线路时航拍绝缘子图像中存在背景复杂,目标大小不一致和待检缺陷区域占比小三个问题,提出了一种轻量化的绝缘子缺陷检测算法 MHD-YOLO。首先在 YOLOv8 的主干网络中引入特征提取网络 MAFNet,使用混合卷积来增强网络在复杂背景下的特征提取能力。其次,使用特征融合网络 HS-FPN 来实现不同尺度的特征融合,结合轻量化的动态上采样方法 DySample 来提升上采样的质量和效率。然后设计了一种轻量化检测头 CSH,通过使用共享卷积的方法大幅减少检测层的参数量和计算量。最后引入 NWD 损失函数来提高模型对小型目标的定位精确性。实验结果证明,MHD-YOLO 目标检测算法与 YOLOv8 相比,参数量降低了 43.8%,在绝缘子缺陷检测数据集上检测精度提高了 5.1%。改进后的算法在检测绝缘子缺陷方面的效果有明显提升,且模型复杂度大大降低,为部署到嵌入式设备上提供了更有效的方法。

关键词: 小型目标检测;绝缘子缺陷检测;YOLOv8;轻量化改造

中图分类号: TP391.4;TN911.73 **文献标识码:** A **国家标准学科分类代码:** 510.80

Lightweight insulator defect detection algorithm based on MHD-YOLO

Guan Yanpeng Fu Pengbo Yao Huijuan

(School of Automation and Software Engineering, Shanxi University, Taiyuan 030013, China)

Abstract: Aiming at the three problems of complex background, inconsistent target size and small proportion of defective areas to be inspected in aerial insulator images taken by UAVs during transmission line inspection, a lightweight insulator defect detection algorithm, MHD-YOLO, is proposed. Firstly, a feature extraction network MAFNet is introduced into the backbone network of YOLOv8, and hybrid convolution is used to enhance the feature extraction capability of the network under complex background. Second, a feature fusion network, HS-FPN, is used to realize feature fusion at different scales, and combined with a lightweight dynamic up-sampling method, DySample, to improve the quality and efficiency of up-sampling. Then, a lightweight detection head CSH is designed, which significantly reduces the number of parameters in the detection layer and the computation amount by using the shared convolution method. Finally, the NWD loss function is introduced to improve the localization accuracy of the model for small targets. The experimental results demonstrate that the MHD-YOLO target detection algorithm reduces the number of parameters by 43.8% compared with YOLOv8, and improves the detection accuracy by 5.1% on the insulator defect detection dataset. The improved algorithm is significantly more effective in detecting insulator defects, and the model complexity is greatly reduced, providing a more effective method for deployment on embedded devices.

Keywords: small target detection;insulator defect detection;YOLOv8;lightweight improvement

0 引 言

绝缘子作为输电线路中的重要部件,可以隔离导线与输电塔,防止电流泄漏或接地。但是由于输电线路会受到恶劣自然环境的影响,绝缘子出现故障的概率较高,若没有

及时更换,可能会对周围环境和人员造成危害,甚至导致大范围的停电事故,产生巨大的经济损失^[1]。近年来,随着计算机视觉与人工智能的兴起,基于深度学习的目标检测算法获得了长足发展^[2]。将目标检测算法与无人机巡检相结合,由无人机采集输电线路现场图像数据,通过目标识别算

法进行数据处理,能够有效解决人工巡检所面临的问题,是目前线路巡检的重要方式之一^[3]。

目标检测算法可以分为两类:双阶段算法和单阶段算法。双阶段算法先生成候选区域,然后再对这些区域进行分类和精细定位,代表算法有Faster R-CNN^[4]、Mask R-CNN^[5]等。双阶段目标检测算法由于需要生成候选区域,导致其检测速度较慢,在实时检测方面不适用。单阶段算法直接从输入图像中预测目标类别和边界框,代表算法有SSD^[6]、YOLO^[7]系列等。俞俊等^[8]在SSD算法中引入残差网络,加强了网络对绝缘子缺陷小目标检测的能力;廖丽瑛等^[9]在YOLOv8模型中引入BiFPN-SimAM模块来实现多尺度特征融合,提高了模型复杂背景下对绝缘子缺陷的识别能力;熊炜等^[10]提出了一种基于YOLOv8-GCS的检测方法,通过结合坐标注意力模块和卷积网络,提升了对复杂背景下绝缘子缺陷的检测能力;曾业战等^[11]提出一种基于MFCE-YOLOv8的检测算法,通过引入多尺度注意力机制MIAA与跨层特征融合模块CFFM,优化了模型对尺度大小不一致目标以及小型缺陷的检测能力;He等^[12]提出一种基于MFI-YOLO的检测算法,针对背景复杂与故障大小不一的问题,引入了多尺度特征融合结构ResPANet与不对称卷积,提高了模型的识别精度;Zhang等^[13]针对现有检测方法识别速度慢、准确率低的问题,在YOLOv8的基础上引入多尺度大核注意力模块MLKA与GSConv卷积块,提高了模型性能。

由上述文献可知,目前在绝缘子缺陷识别方面主要存在3个问题:绝缘子图像背景复杂、目标大小不一致以及缺陷区域占比小。而现有的研究都集中于其中的一个或两个问题进行研究,且模型的复杂度较高,部署在嵌入式设备上较为困难。本文将同时针对上述的3个问题,以YOLOv8为基准模型进行专项改进,并同时为模型进行轻量化改造,以便于改造后的模型部署在嵌入式设备中进行工作。首先在主干网络中加入MAFBlock(mixed aggregation faster block)模块提升主干网络的运算速度,并且加强了主干网络对图片中小尺度特征的提取能力;然后在颈部网络中引入HS-FPN(high-level screening-feature fusion pyramid networks)模块^[14]和DySample模块^[15]实现多层级特征融合,提升特征图上采样的质量和效率。其次设计了一种轻量化的检测头CSH(convolutional-shared head)优化检测层,使用共享卷积的方法,大大减少了检测层的计算量。最后采用NWD(normalized wasserstein distance)评价指标^[16]代替CIoU,提升了网络对小目标的检测能力,解决了CIoU对小目标检测时的尺度敏感和位置敏感问题。

1 YOLOv8 模型

本文使用YOLOv8作为基准模型,其基础网络结构主要由主干网络(backbone)、颈部网络(neck)和头部网络(head)组成^[12],如图1所示。

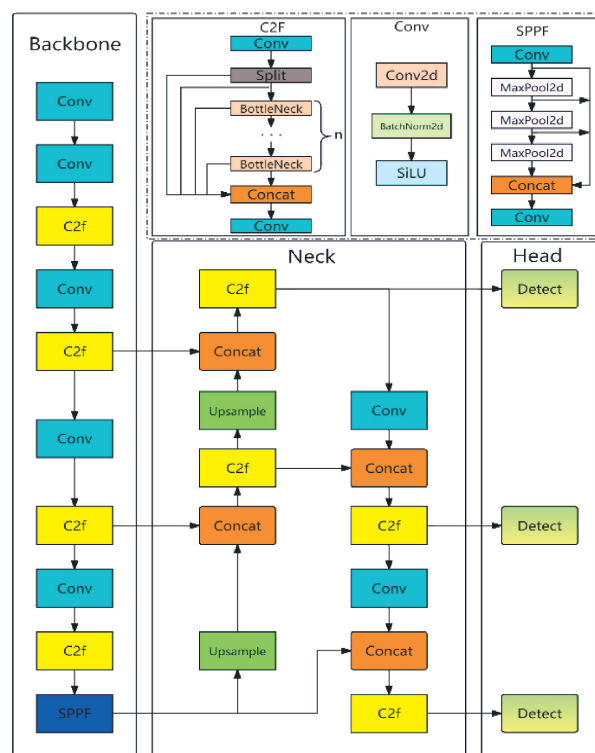


图1 YOLOv8模型结构图

Fig.1 YOLOv8 model structure diagram

主干网络主要负责对输入图片进行特征提取,颈部网络将主干网络所提取不同深度的特征进行融合,头部网络包含3个不同尺度的检测头,使用颈部网络生成的特征图来预测目标的类别、位置以及预测框的置信度。

2 改进 MHD-YOLO 模型

在使用无人机航拍巡检输电线路中的绝缘子时,需要与输电线路保持安全距离,这会导致图片中绝缘子缺陷区域占比较小且目标大小不一致;由于输电线路分布广泛,所以航拍的图像还会存在背景复杂的问题;考虑到该算法的应用场景,需要进行轻量化设计以便于嵌入式设备的应用。综合以上问题,本文在YOLOv8模型的基础上进行的以下改进,改进后的网络结构如图2所示。

2.1 主干网络

考虑到图片中背景复杂的问题,在主干网络部分引入MANet(mixed aggregation net)来增强其在复杂背景下的特征提取能力^[17]。MANet使用混合卷积结构,将3种典型的卷积变体(用于通道特征再校准的 1×1 旁路卷积,用于空间特征处理的深度可分离卷积,用于增强特征层次集成的C2f模块)融合在一起,从而在训练阶段产生更加多样化的梯度流,使图像的特征在空间特征、通道特征和层次特征上得到全面的表达。面对背景复杂的图片,具有更强的特征提取能力,其结构如图3所示。

MANet的运算过程可以用以下的式(1)表达。式中

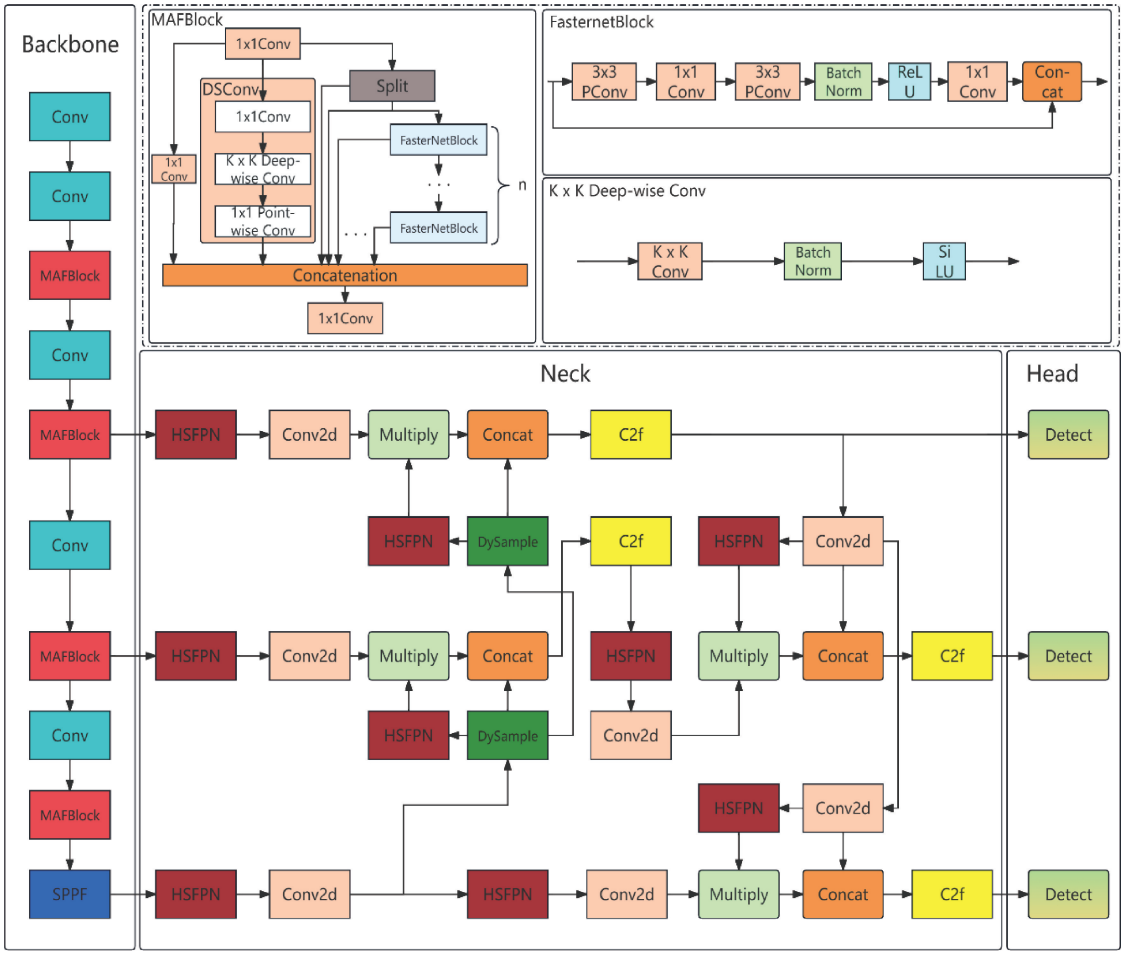


图 2 MHD-YOLO 模型结构图

Fig. 2 MHD-YOLO model structure diagram

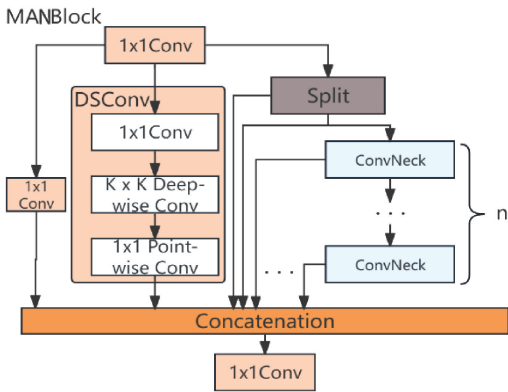


图 3 MANet 结构图

Fig. 3 MANet structure diagram

F_{mid} 的通道数为 $2c$, F_1, F_2, \dots, F_n 的通道数为 c 。输入首先经过 $Conv_1$ 变成 F_{mid} , 再分为 3 部分, 左侧经过标准卷积分支 $Conv_2$ 生成 F_1 , 中间经过深度可分离卷积 $DSConv$ (depthwise separable convolution) 分支生成 F_2 , 右侧经过 $C2f$ 模块分支生成 F_3, F_4, \dots, F_n 。生成的 F_1, \dots, F_n 的

过一个相加操作后再经过一个 1×1 的卷积来融合、压缩 3 种类型特征的语义信息, 生成通道号为 $2c$ 的 F_{out} , 表达式如式(2)所示, 其中 \parallel 表示 Concat 操作。

$$\begin{cases} F_{mid} = Conv_1(F_{in}) \\ F_1 = Conv_2(F_{mid}) \\ F_2 = DSConv(Conv_3(F_{mid})) \\ F_3, F_4 = Split(F_{mid}) \\ F_5 = ConvNeck_1(F_4) + F_4 \\ F_6 = ConvNeck_2(F_5) + F_5 \\ \dots \\ F_n = ConvNeck_n(F_{n-1}) + F_{n-1} \end{cases} \quad (1)$$

$$F_{out} = Conv_o(F_1 \parallel F_2 \parallel \dots \parallel F_n) \quad (2)$$

FasternetBlock^[18] 利用特征图中通道之间的冗余性, 使用 PConv(partial convolution), 仅对输入特征图的一部分通道进行卷积操作, 减少了冗余的计算, 从而更高效地提取图片的空间特征, 其结构如图 4 所示。

一般情况下, 输入和输出特征图具有相同数量的通道。此时, PConv 的每秒浮点运算次数仅为:

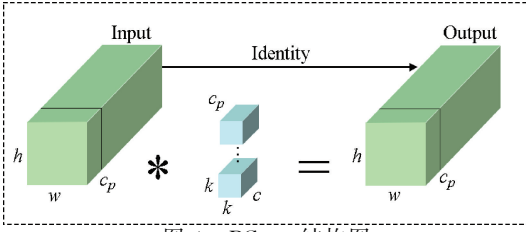


图 4 PConv 结构图

Fig. 4 PConv structure diagram

$$G = h \times w \times k^2 \times c_p^2 \quad (3)$$

其中, h 、 w 分别表示图片的宽度和高度, k 表示卷积核的尺寸, $c_p = c \times r$, c 是输入特征图的总通道数, r 是部分比例, 当取 $r = \frac{c_p}{c} = \frac{1}{4}$ 时, PConv 的运算次数仅为普通

卷积的 $\frac{1}{16}$ 。取 $r = \frac{1}{4}$ 时, 访问次数仅为普通卷积的 $\frac{1}{4}$ 。

本文中的 FasterBlock 均选取 $r = \frac{1}{4}$ 。

将 MANet 的 ConvNeck 部分替换为 FasternetBlock 得到 MAFNet, 可以在保留 MANet 的特征提取能力的基础上进行更快的进行图像识别。MAFNet 的结构如图 2 上方所示。

2.2 颈部网络

由于图片中待识别目标存在大小不一致的问题, 这会增加模型准确识别缺陷的难度。故引入 HS-FPN 模块来实现多尺度的特征融合。HS-FPN 的结构如图 5 所示。

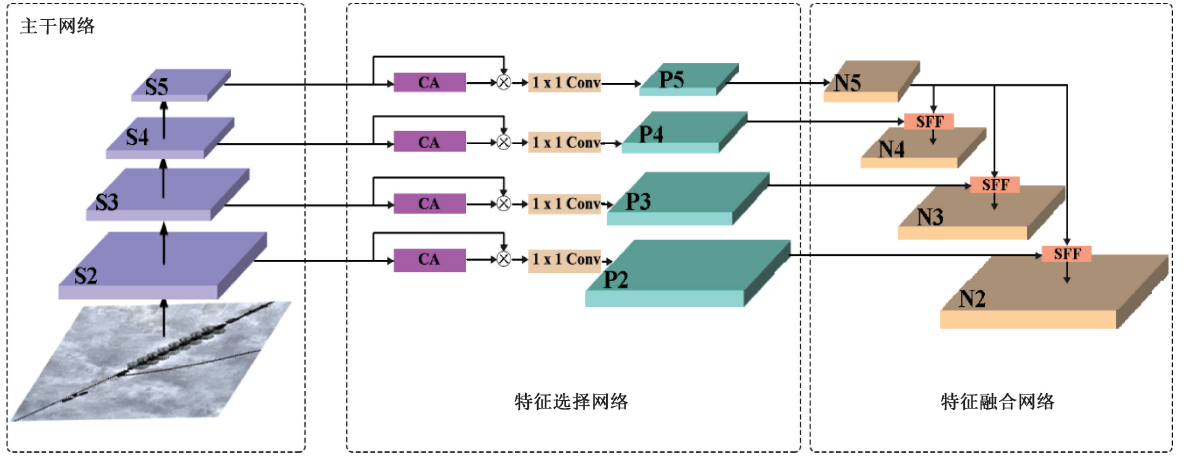


图 5 HS-FPN 结构图

Fig. 5 HS-FPN structure diagram

HS-FPN 主要由两部分组成: 特征选择模块和特征融合模块。输入特征图的大小为 $f_{in} \in R^{C \times H \times W}$, 其中 C 表示通道的数量, H 表示特征图的高度, W 表示特征图的宽度。在特征选择模块中, 输入的特征图会首先经过 CA (channel attention) 模块处理, 其结构如图 6 所示。

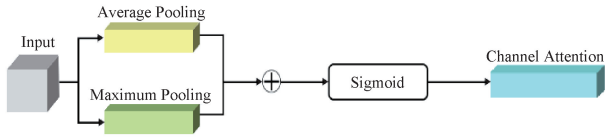


图 6 CA 模块结构图

Fig. 6 CA module structure diagram

在 CA 模块中, 特征图会分别进行全局平均池化和全局最大池化, 将得到的两个结果组合得到输出特征, 随后使用 Sigmoid 激活函数来确定每个通道的权重值, 得到 $f_{CA} \in R^{C \times 1 \times 1}$ 。其中最大池化从每个通道中提取最相关的数据, 而平均池化从特征图中均匀地获取所有数据。因此 CA 模块同时使用这两种池化方法可以从每个通道提取最具代表性的信息, 同时最小化信息损失。随后, 通过将权

重信息与对应尺度的特征图相乘来生成过滤后的特征图。

在特征融合模块中, 引入 SFF (select feature fusion module) 模块来进行不同维度的特征融合, 其结构如图 7 所示。

SFF 模块通过使用高维度特征作为权重来选择性地融合低维度特征, 过滤掉嵌入在低维度特征中的基本信息。如图中所示, 对于输入的一对高维度特征 $f_{high} \in R^{C \times H \times W}$ 和低维度特征 $f_{low} \in R^{C \times H_1 \times W_1}$, 首先使用转置卷积 (transpose convolution) 来扩展高维度特征, 得到转置后的特征大小为 $f_{high} \in R^{C \times 2H \times 2W}$; 然后, 为了统一高维度特征与低维度特征的维数, 使用双线性插值 (bilinear interpolation) 来对高维度特征进行上采样或者下采样得到 $f_{att} \in R^{C \times H_1 \times W_1}$; 获得两个维度一致的特征后, 使用 CA 模块将高维度特征转换成相应的注意力权重用来过滤低维度特征; 最后将过滤后的低维度特征与高维度特征融合, 增强模型的特征表达能力, 得到 $f_{out} \in R^{C \times H_1 \times W_1}$, 以下展示了特征选择的融合过程。

$$f_{att} = BI(TConv(f_{high})) \quad (4)$$

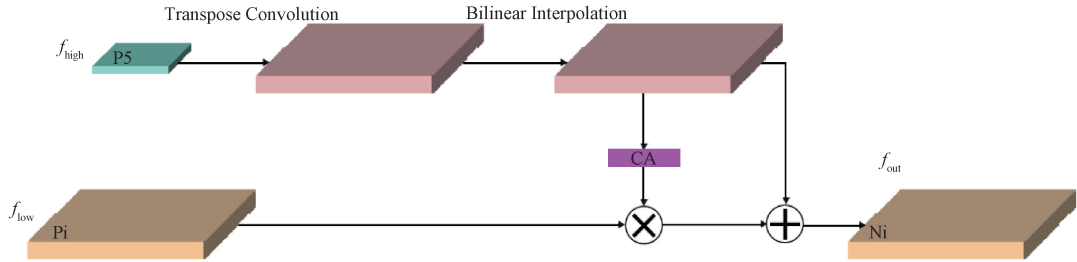


图 7 SFF 模块结构图

Fig. 7 SFF model structure diagram

$$f_{out} = f_{low} \times CA(f_{att}) + f_{att} \quad (5)$$

式中: BI 表示双线性插值, $TConv$ 表示转置卷积。双线性插值可以简单快速的实现了对像素的直接操作, 实现了图像的缩放; 转置卷积可以通过可学习的参数来适应数据, 使输出放大特征图的同时以卷积的形式重构了输入。在图像采样过程中, 利用转置卷积和双线性插值相结合的方法可以更好的恢复高层特征的尺度。

目前主流的高效上采样器(如 CARAFE^[19], FADE^[20]和 SAPA^[21])通常采用基于内核的动态上采样模式。实验证明, 它们可以使上采样过程获得显著的性能提升, 但是由于使用了计算较慢的动态卷积和用于生成动态内核的附加子网络, 大大增加了模型的复杂程度, 延长了检测时间, 在性能受限的场景下无法使用。而 DySample 绕过动态卷积, 从点采样的角度进行上采样, 这不仅节省了资源, 而且可以直接使用 Pytorch 中的标准内置函数实现, 加快了计算的速度。

2.3 轻量化检测头

在 YOLOv8 模型中, 包含 3 个结构相同的检测头, 其内部结构如图 8 所示。

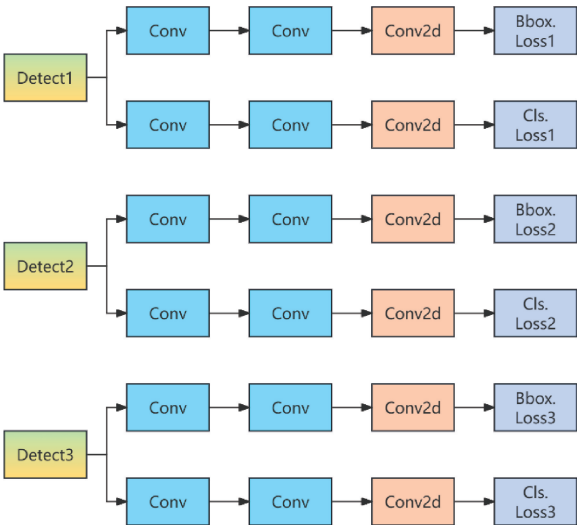


图 8 检测头结构图

Fig. 8 Detection head structure diagram

可以看到, 每个检测头都分为两部分, 分别计算 Bbox.

Loss 与 Cls. Loss, 两部分都分别包含了两个 Conv 卷积模块(结构与图 2 中相同)与一个 Conv2d 模块。由于位于输出层, 图像的通道数量通常较大, 所以在进行卷积操作时所需要的参数数量和消耗的计算量都会很大。

通过编写程序输出模型每层的参数量, 发现 YOLOv8 模型中 Detect 层的参数量占总参数量的 29.7%, 计算量占比更是高达 41.6%。如果可以减少 Detect 层的卷积数量, 那么就可以大大降低模型的复杂度。基于这个思想, 本文提出了一种共享卷积的轻量化检测头 CSH。其结构如图 9 所示。

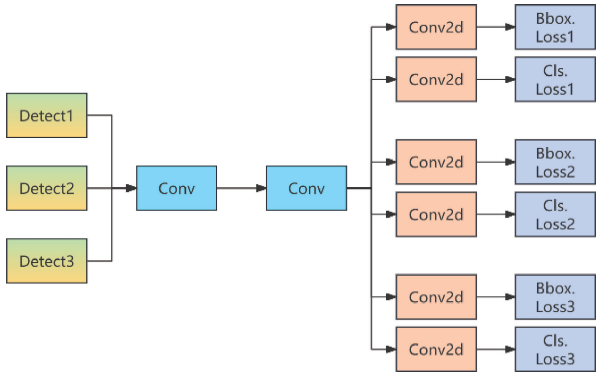


图 9 CSH 结构图

Fig. 9 Convolutional-shared head structure diagram

由于每个 Conv 模块中都包含一个 3×3 的卷积, 多次使用会带来巨大的参数量和计算量, 故尝试让 3 个检测头共用两个 Conv 模块, 使原本的 6 个 Conv 模块减少为 2 个, 之后再连接各自的 Conv2d 模块实现不同的输出。实验表明, 改进后模型参数量与计算量均显著下降, 同时其平均准度均值几乎不受影响。

2.4 NWD 损失函数

YOLOv8 作为当前先进的目标检测框架, 广泛应用于各类检测任务。然而, 现有的 IoU 度量方法在微小目标检测中对目标尺度的变化敏感, 对位置偏移的敏感, 这会导致微小目标的 IoU 计算值波动剧烈, 从而导致误检或漏检。由于本文检测目标绝缘子属于微小目标检测的范畴, 为了解决以上问题, 提出了在 YOLOv8 框架中使用 NWD 作为 IoU 的替代方案。

NWD 通过将目标边界框建模为二维高斯分布,计算其 Wasserstein 距离,并进行指数归一化,使其值保持在 0~1 之间。相比 IoU, NWD 具有更强的尺度适应性和更平滑的梯度变化,可有效缓解 IoU 在微小目标检测中的局限性。其详细运算过程如下:

在 NWD 方法中,目标检测框 $B = (C_x, C_y, w, h)$ 被建模为二维高斯分布 $N(\mu, \Sigma)$ 。其中, $\mu = \begin{bmatrix} C_x \\ C_y \end{bmatrix}$ 为高斯分布的均值向量,表示边界框的中心坐标,其中 C_x 和 C_y 分别表示边界框的中心横坐标和纵坐标;协方差矩阵 $\Sigma = \begin{bmatrix} \frac{w^2}{4} & 0 \\ 0 & \frac{h^2}{4} \end{bmatrix}$, 表示边界框的尺度信息,其中 w 和 h 分别表示边界框的宽度和高度。

Wasserstein 距离是一种衡量两个分布相似度的方法,对于两个二维高斯分布 $N_1(\mu_1, \Sigma_1)$ 和 $N_2(\mu_2, \Sigma_2)$, 其二阶 Wasserstein 距离计算公式为:

$$W_2^2 = \|\mu_1 - \mu_2\|_2^2 + \text{Tr}(\Sigma_1 + \Sigma_2 - 2(\Sigma_1^{\frac{1}{2}} \Sigma_2 \Sigma_1^{\frac{1}{2}})^{\frac{1}{2}}) \quad (6)$$

为了使 Wasserstein 距离与 IoU 兼容,并归一化到 [0,1] 区间, NWD 通过指数归一化定义为:

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

其中, C 是一个与数据集相关的归一化常数,通常设为数据集中目标的平均尺度。该归一化方式确保 NWD 值在 0~1 之间,使其能直接替换 IoU。

3 实验设置

3.1 数据集

本文使用的数据集来自于中国输电线路绝缘子数据集(chinese power line insulator dataset, CPLID)^[22], 该数据集共包含绝缘子与缺陷绝缘子图片共 848 张。将这些图片按照 8:1:1 的比例将图片划分为训练集,验证集和测试集。

由于图片数量较少,训练过程中可能出现过拟合的情况,故使用图像增强的方式来扩充图片数量。包括添加高斯噪声、旋转、裁剪、拼接和调整对比度,来模拟实际航拍过程中可能遇到的气候问题和由于拍摄角度导致的图片质量下降问题。先进行数据集划分再进行图像增强,可以有效的避免训练集、测试集和验证集出现图像交叉的情况。增强后的数据集包含 5 088 张图片。

3.2 评价指标

本文使用模型参数量(parameters)与模型计算量(GFLOPs)来评估模型的复杂性;使用精确率(precision, P)、召回率(recall, R)与平均精度均值(mAP, 本文取 IoU 阈值为 0.5 计算)来评估模型检测精度;使用检测帧数

(FPS)来评估模型的检测速度,相关参数的计算公式如下:

$$P = \frac{TP}{TP + FP} \quad (8)$$

$$R = \frac{TP}{TP + FN} \quad (9)$$

$$FPS = \frac{Numbers}{Time} \quad (10)$$

$$AP = \int_0^1 P(R) d(R) \quad (11)$$

$$mAP = \frac{\sum_{i=1}^N AP_i}{N} \quad (12)$$

式中: TP (true positive) 表示模型预测为正类且目标为正类的数量, FP (false positive) 表示模型预测为正类但目标为负类的数量, FN (false negative) 表示模型预测为负类但目标为正类的数量。 $Numbers$ 表示图片总数, $Time$ 表示检测一轮照片所需要的时间。 AP 表示目标的 PR 曲线面积, AP_i 表示第 i 类目标的 PR 曲线面积, N 表示检测目标的类别总数。

3.3 实验环境

实验设备详细参数及软件配置如表 1 所示。实验中的超参数设置如表 2 所示。

表 1 实验设备详细参数及软件配置

Table 1 Experimental equipment detailed parameters and software configuration

名称	参数
GPU	NVIDIA GeForce RTX 4090 24G
CPU	i9-13900k
内存	124 G
操作系统	Ubuntu 22.04
编程语言	Python3.10.15
深度学习框架	Torch2.3.0+CUDA11.8

表 2 超参数设置

Table 2 Hyper-parameter setting

参数名	参数值
训练轮数	300
批次大小	64
学习率	0.01
动量	0.937
学习率衰减系数	0.000 5
优化器	SGD
图像尺寸	640×640

4 实验结果与分析

4.1 主干网络改进实验

为了验证 MAFNet 对主干网的优化效果,进行实验使

用了多种目前主流的特征提取模型来优化 YOLOv8 的主干网络,表中的模型 1、2、3、4、5、6,分别对应 YOLOv8 原模型、Efficientnet^[23]、HGNet^[24]、DyConv-HGNet、MANet 与 MAFNet。对比实验结果如表 3 所示。

表 3 主干网络对比
Table 3 Comparison of backbone networks

模型	Parameters/M	GFLOPs/G	P/%	R/%	FPS/fps	mAP@0.5/%
1	3.012	8.2	92.4	89.6	118.19	90.5
2	4.013	9.5	92.6	89.2	108.99	90.6
3	2.359	7.0	91.4	89.5	92.60	89.8
4	2.565	6.9	91.2	90.1	113.11	92.1
5	3.894	10.3	92.5	92.6	96.62	94.0
6	3.189	8.5	92.7	92.0	102.91	93.8

由实验数据可知,使用 Efficientnet 优化主干网络会使模型参数俩增加 33%,计算量增加 16%,降低检测速度,但是对平均检测精度的提升微乎其微,只有 0.1%;使用 HGNet 优化主干网络可以减少模型复杂度,但是会同时降低检测速度与平均检测精度;使用 DynamicConvolution^[25]改进后的 HGNet 来优化主干网络,可以在保持模型复杂度较低的前提下,较明显的提升平均检测精度;使用 MANet 优化主干网络会使平均检测精度有较大提升,但同时也会使模型变得复杂,并且明显拖慢检测速度;在此基础上加入 FasternetBlock 组成 MAFNet,网络中的 PConv 模块可以在小幅度降低平均检测精度的条件下大幅度减少模型的复杂度,并且加快检测速度。

最后将 MAFNet 与 YOLOv8 模型进行对比,可以看到模型的参数数量和计算量有小幅度增加,检测速度也有所下降,但是对平均检测精度的提升要明显高于除 MANet 以外的网络。权衡模型复杂度与平均检测精度,最终选择 MAFNet 来优化主干网络,并且在后续的改进中,通过优化其他模块来解决模型复杂度与检测速度的问题。

4.2 损失函数改进实验

实验选取 NWD 与 4 种常用的损失函数进行对比,实

验数据如表 4 所示。

表 4 损失函数对比
Table 4 Comparison of loss function

损失函数	P/%	R/%	FPS/fps	mAP@0.5/%
CIoU	94.1	90.8	118.40	95.2
DIoU	92.7	89.3	114.70	93.7
EIoU	91.8	92.2	120.40	94.2
Focal Loss	94.7	90.1	110.43	95.3
NWD	94.8	91.7	121.12	95.6

由表 4 实验数据可知,在检测精度方面,NWD 损失函数的平均检测精度较 CIoU、DIoU、EIoU、Focal Loss 分别提升 0.4、1.9、1.4、0.7;在检测速度方面,FPS 较 CIoU、DIoU、EIoU、Focal Loss 分别提升 2.7、6.4、0.7、10.5。综上所述,使用 NWD 损失函数替换原有的 CIoU 损失函数可以同时提高检测精度和检测速度。

4.3 消融实验

为了验证本文所提出的各个模块对网络改进的有效性,进行了 5 组消融实验,实验结果如表 5 所示。

表 5 消融实验结果
Table 5 Results of ablation experiments

模型	MAFNet	HS-DySample	CSH	NWD	Parameters/ M	GFLOPs/ G	P/%	R/%	FPS/fps	mAP@0.5/ %
1					3.012	8.2	92.4	89.6	118.19	90.5
2	√				3.189	8.5	92.7	92.0	102.91	93.8
3		√			1.958	7.3	92.8	91.3	120.32	91.7
4			√		2.372	6.7	91.9	88.3	135.49	90.2
5				√	3.012	8.2	93.0	90.5	124.61	91.6
6	√	√			2.061	6.7	94.5	90.8	102.79	95.4
7	√	√	√		1.694	5.5	94.1	90.7	118.40	95.2
8	√	√	√	√	1.694	5.5	94.8	91.7	121.12	95.6

其中模型 1 为 YOLOv8 模型,模型 2、3、4、5 分别添加了表中标有“√”的模块。由表中数据可知,第二章中提到的 4 种改进均对模型有不同程度的提升。在主干网络中使用 MAFNet 可以在复杂背景下更好的提取图片不同尺度的特征,检测精度为 93.8%,与 YOLOv8 模型相比提高了 3.3%。使用 HSFPN 与 DySample 改进颈部网络,可以增强模型对不同尺度特征融合的能力;由于 DySample 的点采样模式,模型的总参数量由 3.189 M 降低至 2.061 M,降低了 35%,总计算量由 8.5 G 减少至 6.7 G,降低了 21%。轻量化检测头 CSH 的引入会使检测精度减少 0.2%,但是加快模型的检测速度并进一步降低模型的复杂度,在之前改进的基础上,模型的总参数量降低了 18%,总计算量降低了 18%,FPS 增加了 15.61,综合实验数据分析,使用共享卷积的方式来优化检测头是一次成功的改进引入用于小型目标检测的 NWD 损失函数,在绝缘子缺陷检测方面要明显优于 CIoU 损失函数,FPS 由 118.40 提升至 121.12,检测精度由 95.2 提升至 95.6。

4.4 对比实验

为了验证 MHD-YOLO 网络的性能,在相同的实验环境和实验超参数下,使用常见的目标检测算法 Fast R-CNN、SSD、YOLO 系列与 MHD-YOLO 对比,实验结果如表 6 所示。

表 6 对比实验结果 1

Table 6 Results of comparison experiments 1				
模型	Parameters/ GFLOPs/		FPS/ mAP@0.5/	
	M	G	fps	%
Faster R-CNN	138.4	389.4	32.52	46.3
SSD	27.193	64.8	37.26	64.1
YOLOv3	103.698 6	283.0	47.31	91.9
YOLOv5	2.510	7.2	113.33	87.5
YOLOv7	38.742	107.3	68.67	88.3
YOLOv8	3.012	8.2	118.19	90.5
YOLOv10	2.710	8.4	304.76	85.3
MHD-YOLO	1.694	5.5	121.12	95.6

分析表 6 中数据可知,改进后的算法 MHD-YOLO 与其他常见的目标检测算法相比,主要的性能指标都有所提高。与 Fast R-CNN、SSD、YOLOv3、YOLOv5、YOLOv7、YOLOv8、YOLOv10 模型相比,在模型复杂度方面,MHD-YOLO 的参数量分别减少了 136.7、25.5、102.1、0.8、37、1.3 M 和 1.0 M,计算量分别减少了 383.9、59.3、277.5、1.7、101.8、2.7 G 和 2.9 G;在检测精度方面,平均精度均值分别提高 49.3%、31.5%、3.8%、8.2%、7.3%、5.1%和 10.4%;在检测速度方面,除了专注于提高检测速度的 YOLOv10 模型,MHD-YOLO 的检测速度要快于其他的常见模型。

FFCA-YOLO^[26]和 HCF-Net^[27]在小型目标检测方面性能表现优越的,FFCA-YOLO 融合了 Transformer 编码器与自注意力机制,增强对空间细节的建模能力,提高了密集区域和复杂背景下对小目标的检测能力;HCF-Net 将多尺度特征提取和扩张卷积与通道注意力机制相结合,同时引入 CEM 模块增强语义和上下文信息,提升了小型目标检测的鲁棒性和准确性。

文献[28-30]同为基于 YOLO 改进模型,用于绝缘子缺陷检测的研究。文献[28]基于 YOLOv3 模型进行改进,通过使用双向特征金字塔网络结构进行不同尺度特征间的双向融合,解决了连续卷积下小目标信息丢失的问题。文献[29]基于 YOLOv11 模型做出优化,设计了空间注意力增强模块来补偿被遮挡区域的响应损失,来抑制背景多样化对特征图的干扰。文献[30]在 YOLOv8 网络增加了小目标检测层,并添加注意力与卷积混合模块,提出了加权双向路径聚合网络替代路径聚合网络,防止特征融合过程中原始特征的丢失,提高多尺度目标特征的融合度。

上述 5 篇文献均为近期性能较好的目标检测模型,除了 HCF-Net 以外均以 YOLO 为基准模型,将 MHD-YOLO 与其进行对比,实验结果如表 7 所示。

表 7 对比实验结果 2

Table 7 Results of comparison experiments 2				
模型	Parameters/ GFLOPs/		FPS/ mAP@0.5/	
	M	G	fps	%
FFCA-YOLO	7.120	51.20	162.23	92.4
HCF-Net	15.292	93.16	72.41	89.7
文献[28]	5.160	20.00	85.67	94.2
文献[29]	13.700	39.60	52.40	92.5
文献[30]	22.840	—	65.70	93.2
MHD-YOLO	1.694	5.50	121.12	95.6

由表中数据可知,MHD-YOLO 仅在检测速度上慢于 FFCA-YOLO,在模型复杂度与精度上均优于其他模型。

对比实验的结果说明,MHD-YOLO 可以在使用较低参数量和计算量的同时保持较高的检测精度,检测效果超过了 SSD、Fast R-CNN、传统的 YOLO 算法与近年来的新型算法,可以更好的完成绝缘子表面缺陷检测的任务。

4.5 实验结果可视化

图 10 为 YOLOv8 算法、YOLOv10 算法与 MHD-YOLO 算法对图片的检测结果。

经过对比可以看出,在闪络痕迹的检测中,YOLOv8 算法出现了误检,YOLOv10 算法出现了漏检;在破损的检测中,YOLOv8 与 YOLOv10 均出现了漏检;在缺失的检测中,两种算法都检测出了绝缘子的缺失。而改进后的 MHD-YOLO 算法对 3 种缺陷的检测都表现良好。此外,

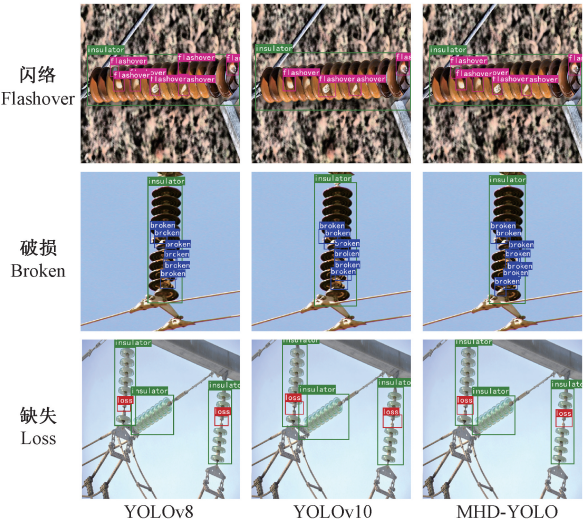


图 10 实验结果可视化

Fig. 10 Experimental results visualization

注意到 YOLOv10 算法的边界框通常比较大,可能是由于过度追求检测速度导致检测精度有所下降。

5 结 论

针对无人机航拍巡检输电线路中绝缘子的图像中存在的背景复杂、目标大小不一致和待检缺陷区域占比小 3 个问题,本文提出了一种基于 YOLOv8 的输电线路绝缘子缺陷检测算法 MHD-YOLO。针对背景复杂的问题,在主干网络中使用 MAFNet 来加强网络的特征提取能力;针对目标大小不一致的问题,在颈部网络中引入 HS-FPN 与 DySample 模块来加强网络对不同尺度特征的融合能力;针对待检缺陷区域占比小的问题,引入在小型目标检测中表现优秀的 NWD 损失函数;在此基础上还对冗余的检测头进行了轻量化改造,大大减少了模型的参数量与计算量,为模型在嵌入式设备上的部署提供了便捷。

消融实验表明,本文所提出的 4 项改进并无冲突,均对模型的改进有着积极作用;对比实验结果表明,无论是模型的复杂度还是检测的精准度,MHD-YOLO 都要优于 SSD、Faster R-CNN、YOLO 系列算法、FFCA-YOLO 与 HCF-Net。由对比实验的数据可知 MHD-YOLO 的运算速度要低于 YOLOv10,因此下一阶段的目标是继续优化算法,加快算法的检测速度。

参考文献

[1] ZHENG J, WU H, ZHANG H, et al. Insulator-defect detection algorithm based on improved YOLOv7 [J]. Sensors, 2022, 22(22): 8801.
[2] 赵朗月, 吴一全. 基于机器视觉的表面缺陷检测方法研究进展[J]. 仪器仪表学报, 2022, 43(1): 198-219.
ZHAO L Y, WU Y Q. Research progress of surface defect detection methods based on machine vision[J].

Chinese Journal of Scientific Instrument, 2022, 43(1): 198-219.
[3] 熊伟, 路鑫, 邱维进, 等. 基于 YOLOv8 的输电线路绝缘子表面缺陷识别算法[J]. 电子测量技术, 2025, 48(2): 178-188.
XIONG W, LU X, QIU W J, et al. Surface defect detection algorithm of transmission line insulators based on YOLOv8 [J]. Electronic Measurement Technology, 2025, 48(2): 178-188.
[4] REN S, HE K, GIRSHICK R, et al. Faster R-CNN: Towards real-time object detection with region proposal networks[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2016, 39 (6): 1137-1149.
[5] HE K, GKIOXARI G, DOLLÁR P, et al. Mask R-CNN[C]. IEEE International Conference on Computer Vision, 2017: 2961-2969.
[6] LIU W, ANGUELOV D, ERHAN D, et al. SSD: Single shot multibox detector[C]. Computer Vision-ECCV 2016: 14th European Conference, 2016: 21-37.
[7] REDMON J. You only look once: Unified, real-time object detection[C]. IEEE Conference on Computer Vision and Pattern Recognition, 2016.
[8] 俞俊, 武丽, 付相为, 等. 基于改进 SSD 的高空绝缘子缺陷检测算法[J]. 信息技术, 2024(12): 72-79.
YU J, WU L, FU X W, et al. High-altitude insulator defect detection algorithm based on improved SSD[J]. Information Technology, 2024(12): 72-79.
[9] 廖丽瑛, 刘洪. 基于改进 YOLOv8 的绝缘子自爆缺陷检测[J]. 电子测量技术, 2024, 47(18): 138-144.
LIAO L Y, LIU H. Self-explosion defect detection of insulator based on improved YOLOv8[J]. Electronic Measurement Technology, 2024, 47(18): 138-144.
[10] 熊伟, 黄玉谦, 孟圣哲. 基于改进 YOLOv8 算法的绝缘子缺陷检测模型[J]. 电子测量技术, 2024, 47(12): 132-139.
XIONG W, HUANG Y Q, MENG SH ZH. Insulator defect detection model based on improved YOLOv8 algorithm[J]. Electronic Measurement Technology, 2024, 47(12): 132-139.
[11] 曾业战, 陈天航, 邓倩, 等. 基于多尺度融合与上下文增强的绝缘子缺陷检测[J]. 计算机工程与应用, 2025, 61(20): 368-378.
ZENG Y ZH, CHEN T H, DENG Q, et al. Insulator defect detection based on multi-scale fusion and contextual enhancement [J]. Computer Engineering and Applications, 2025, 61(20): 368-378.
[12] HE M, QIN L, DENG X, et al. MFI-YOLO: Multi-fault insulator detection based on an improved YOLOv8[J]. IEEE Transactions on Power Delivery, 2023, 39(1): 168-179.

- [13] ZHANG L, LI B, CUI Y, et al. Research on improved YOLOv8 algorithm for insulator defect detection[J]. Journal of Real-Time Image Processing, 2024, 21(1): 22.
- [14] CHEN Y, ZHANG C, CHEN B, et al. Accurate leukocyte detection based on deformable-DETR and multi-level feature fusion for aiding diagnosis of blood diseases[J]. Computers in Biology and Medicine, 2024, 170: 107917.
- [15] LIU W, LU H, FU H, et al. Learning to upsample by learning to sample[C]. IEEE/CVF International Conference on Computer Vision, 2023: 6027-6037.
- [16] WANG J, XU C, YANG W, et al. A normalized gaussian wasserstein distance for tiny object detection[J]. ArXiv preprint arXiv: 2110.13389, 2021.
- [17] FENG Y, HUANG J, DU S, et al. Hyper-YOLO: When visual object detection meets hypergraph computation [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2024, 47(4):2388-2401.
- [18] CHEN J, KAO S, HE H, et al. Run, don't walk: Chasing higher FLOPS for faster neural networks[C]. IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2023: 12021-12031.
- [19] WANG J, CHEN K, XU R, et al. Carafe: Content-aware reassembly of features [C]. IEEE/CVF International Conference on Computer Vision, 2019: 3007-3016.
- [20] LU H, LIU W, FU H, et al. FADE: Fusing the assets of decoder and encoder for task-agnostic upsampling[C]. European Conference on Computer Vision. Cham: Springer Nature Switzerland, 2022: 231-247.
- [21] LU H, LIU W, YE Z, et al. SAPA: Similarity-aware point affiliation for feature upsampling[J]. Advances in Neural Information Processing Systems, 2022, 35: 20889-20901.
- [22] TAO X, ZHANG D, WANG Z, et al. Detection of power line insulator defects using aerial images analyzed with convolutional neural networks[J]. IEEE Transactions on Systems, Man, and Cybernetics: Systems, 2018, 50(4):1486-1498.
- [23] CAI H, LI J, HU M, et al. Efficientvit: Lightweight multi-scale attention for high-resolution dense prediction[C]. IEEE/CVF International Conference on Computer Vision, 2023: 17302-17313.
- [24] YAO T, LI Y, PAN Y, et al. Hgnet: Learning hierarchical geometry from points, edges, and surfaces[C]. IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2023: 21846-21855.
- [25] HAN K, WANG Y, GUO J, et al. Parameters are all you need for large-scale visual pretraining of mobile networks[C]. IEEE/CVF Conference on Computer Vision and Pattern Recognition, Seattle, WA, USA, 2024: 17-21.
- [26] ZHANG Y, YE M, ZHU G, et al. FFCA-YOLO for small object detection in remote sensing images[J]. IEEE Transactions on Geoscience and Remote Sensing, 2024, 62: 1-15.
- [27] 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.
- [28] 杨露露, 马萍, 王聪, 等. 结合特征重用与重建的YOLO绝缘子检测方法[J]. 计算机工程, 2024, 50(7): 303-313.
- YANG L L, MA P, WANG C, et al. Insulator detection method using YOLO combining feature reuse and reconstruction[J]. Computer Engineering, 2024, 50(7):303-313.
- [29] 季星宇, 黄陈蓉, 姚军财, 等. 结合分支特征和排斥损失的绝缘子检测研究[J]. 计算机工程与应用, 2025, 61(7):142-152.
- Ji X Y, HUANG CH R, YAO J C, et al. Research on insulator detection based on diverse branches and repulsion loss [J]. Computer Engineering and Applications, 2025, 61(7):142-152.
- [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.

作者简介

关燕鹏(通信作者), 博士, 教授, 主要研究方向为计算机视觉、互联电力系统、网络化控制系统。

E-mail: y. guan@sxu. edu. cn

傅芃博, 硕士研究生, 主要研究方向为目标检测。

要会娟, 博士, 讲师, 主要研究方向为计算机视觉、模式识别、故障诊断。

E-mail: yaohuijuan@sxu. edu. cn