

改进 YOLOv10 的架空输电线路多缺陷检测方法^{*}

李坤祥 刘大明

(上海电力大学计算机科学与技术学院 上海 201306)

摘要: 针对输电线路缺陷检测任务中目标尺度多样、背景复杂、目标遮挡,以及现有目标检测算法难以在实时检测的前提下保证检测精度而出现漏检、误检的问题,提出改进 YOLOv10 的输电线路无人机巡检缺陷检测算法 TLDDet。首先设计融合部分卷积和上下文锚点注意力的高效特征融合模块(FC2FA),在降低模型参数数量的同时提升模型的特征集成能力。然后使用基于多头自注意力机制的尺度内特征交互模块 AIFI 增强对小目标的检测效果,该模块通过加强特征图中高级语义信息的表达从而提高模型检测的准确率。最后设计遮挡感知注意网络检测头 SEAM-Head,减少由于遮挡问题导致的特征丢失的问题。实验结果表明,所提出的 TLDDet 较原始模型 YOLOv10s 参数量减少 33%,计算量减少 30%,对输电线路多种缺陷的 Precision、Recall 和 mAP50 分别提高 4.3%、2.4%和 3.7%,检测速度达到 143 FPS,且与其他实时检测算法的对比中具有更好的检测性能。

关键词: 输电线路缺陷检测;YOLOv10;实时检测;特征融合;上下文锚点注意力;尺度内特征交互;遮挡感知注意网络
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:** 520.6040

Detection method of overhead transmission line defects
based on improved YOLOv10

Li Kunxiang Liu Daming

(School of Computer Science and Technology, Shanghai University of Electric Power, Shanghai 201306, China)

Abstract: In response to the issues of missed and false detections in transmission line defect detection tasks caused by varying target scales, complex backgrounds, and occlusion, where existing object detection algorithms struggle to maintain detection accuracy while ensuring real-time performance, an improved YOLOv10-based UAV transmission line defect detection algorithm, TLDDet, is proposed. First, a faster C2F with attention module (FC2FA) incorporating context anchor attention is designed to enhance feature integration capabilities while reducing the model's parameters. Then, an attention-based intrascale feature interaction module (AIFI) based on multi-head self-attention is used to improve the model's detection performance for small objects by enhancing the representation of high-level semantic information in the feature map, thereby increasing the detection accuracy. Finally, an occlusion-aware attention detection head (SEAM-Head) is designed to reduce feature loss caused by occlusion. Experimental results show that the proposed TLDDet reduces the parameters by 33% and the computational cost by 30% compared to the original YOLOv10s model, while improving precision, recall, and mean average precision for various transmission line defects by 4.3%, 2.4%, and 3.7%, respectively. The detection speed reaches 143 FPS, and comparative experiments with other real-time detection algorithms demonstrate superior model performance, making TLDDet more suitable for transmission line defect detection tasks.

Keywords: transmission line defect detection; YOLOv10; real-time detection; feature fusion; context anchor attention; intra-scale feature interaction; occlusion-aware attention network

0 引言

电力系统的安全稳定运行是社会经济发展的重要支

柱,而架空输电线路作为电能传输的核心通道,其健康状况直接影响着电力系统的可靠性。然而,由于输电线路长期暴露在复杂的自然环境中,常会出现诸多隐患,如绝缘子破

损、防震锤损坏,以及鸟巢侵占等问题。这些缺陷可能导致电力系统运行中断,甚至引发大规模停电事故。随着深度学习技术的快速发展,基于计算机视觉的目标检测算法广泛应用于无人机采集图像的缺陷检测任务中^[1]。通过精准快速地识别输电线路缺陷并及时处理,能够有效保障电力系统的安全运行。因此,提升输电线路缺陷检测的准确性与效率具有较高的研究价值。

目前,目标检测算法中 YOLO(you only look once)系列^[2-5]和 RTMDet^[6]为代表的单阶段检测器,凭借其无需生成锚框、仅需通过非极大值抑制(non-maximum suppression, NMS)^[7]处理置信度较高的检测框,达到了实时检测的性能。而 Faster R-CNN^[8]等双阶段检测器尽管具备更高的检测精度,但其复杂的区域提议过程增加了计算成本。基于 Transformer 的端到端检测器 RT-DETR^[9]以其简洁高效的检测通道而被广泛关注,但是由于视觉 Transformer 中通常具有较高的计算复杂度,所以在模型的部署应用中仍存在困难。

在输电线路缺陷检测的研究中,多数研究者倾向于采用单阶段检测器,因为这类模型参数量较少且推理速度较快。如赵霖等^[10]针对传统的输电线路鸟巢识别方法效率低下且安全性不足等问题提出了基于 YOLOv5 的输电线路鸟巢检测算法,改进模型通过在主干网络中加入卷积注意力模块(convolutional block attention module, CBAM)注意力模块提升主干网络的特征提取能力,在颈部网络中引入自适应特征融合模块替换原始结构,加强多尺度特征融合效果,从而对输电线路鸟巢检测表现出了较好的效果。而对于绝缘子缺陷的检测,肖粲俊等^[11]使用 YOLOv5s 进行改进设计出缺陷检测算法,其主要是利用 K-means 算法对数据集进行聚类分析得到最佳锚框尺寸,加入全局注意力(global attention mechanism, GAM)来解决无效特征对识别精度的影响,最后引入空间自适应网络(adaptive structure feature fusion, ASFF)模块增强网络的特征提取能力。

然而,当前文献中多数研究仅聚焦于单一缺陷目标的检测,对多缺陷目标的检测研究仍存在不足。目前,针对输电线路中多缺陷检测方法研究,刘兰兰等^[12]使用超分辨率网络对巡检图像进行重建,提升清晰度后在 YOLOX 中嵌入卷积块注意力机制并新增浅层检测尺度进行特征融合,由于其所提出的方法首先需要提升巡检图像质量,因此在输电线路多缺陷的实时检测任务中仍需进一步优化。李鑫等^[13]针对输电线路中多种缺陷目标尺寸小与密集遮挡问题,使用 Cascade R-CNN 为基础模型,通过改进模型主干网络和 NMS 达到了较高的检测精度,但是此方法是基于两阶段模型,因此推理速度较慢,无法达到无人机实时巡检需求。易磊等^[14]通过改进 YOLOv8 提出了输电线路多种异物检测方法 SC-YOLO,检测对象主要为鸟巢、塑料袋、漂浮物等,该方法通过引入 StarNet 构造出特征融合模块

C2f-Star 实现网络轻量化,并通过增加特征空间维度来提升模型的特征提取能力,SC-YOLO 在降低计算量的同时实现了较高的检测准确率,但是对输电线路中如绝缘子,防震锤等重要部件未进行检测。

在输电线路多类别缺陷检测中,无人机巡检通常需要较高的检测精度以及实时检测速率,YOLO 系列算法在电力行业的实时应用场景中虽然展示了高效的目标检测能力,但在复杂场景中,特别是在低光、遮挡以及小目标检测的场景中表现不佳,电力领域中的数据集也存在正负样本差异大、样本稀缺等问题,限制了 YOLO 算法的泛化能力。而 YOLO 系列最新模型 YOLOv10 算法相较其他算法具有很好的检测实时性,但是在检测精度方面,YOLOv10 仍需不断改进升级,且现有模型在多缺陷目标检测中的表现有限,尤其是在小目标和复杂遮挡条件下的检测性能仍需进一步改进^[15]。因此,为了使模型检测速度和检测精度达到更好的平衡,本文选择 YOLOv10 s 版本进行改进,针对输电线路多类别缺陷检测的挑战,提出了 TLDDet 算法,主要研究内容包括:

1) 设计高效多尺度特征融合模块(faster C2F with attention, FC2FA):通过融合上下文锚点注意力机制(context anchor attention, CAA)和部分卷积(partial convolution, PConv)提升对复杂场景特征提取能力的同时降低模型的参数量和计算量。

2) 融合尺度内特征交互模块(attention-based intrascale feature interaction, AIFI):将原始模型中的空间金字塔池化层(spatial pyramid pooling-fast, SPPF)和部分自注意力模块(partial self-attention, PSA)替换为 AIFI 模块。在减少冗余参数的同时利用多头自注意力机制更好地捕获绝缘子、防震锤等小目标的细节特征,减少了漏检和误检。

3) 设计遮挡感知注意检测头(spatially enhanced attention module-head, SEAM-Head):在原始检测头中引入 SEAM 模块解决复杂背景下的目标遮挡问题。该模块能够动态关注有效特征,削弱背景干扰,从而减少特征丢失和遮挡问题。

1 相关工作

1.1 YOLOv10 模型概述

YOLOv10 主要由主干网络(Backbone)、路径聚合网络(path aggregation network, PAN)和双检测头(One-to-many Head、One-to-one Head)组成,其结构如图 1 所示。与之前的 YOLO 系列相比,YOLOv10 的主要特点是采用了双标签分配策略,设计了 One-to-many 检测头和 One-to-one 检测头。在训练阶段,两个检测头共享权重,而在推理阶段,仅导出 One-to-one 检测头,从而避免了非极大值抑制操作,减少冗余目标框并加快推理速度。另外模型使用整体效率-准确度驱动的设计策略来增强整体性能,通

过设计空间通道解耦下采样代替传统卷积以最大限度地减少信息损失和计算成本;PSA 模块作用于骨干网络的最后一层改进全局表示学习;排序引导块根据固有冗余调整块

设计,确保最佳参数利用率。这些改进使得 YOLOv10 在提升检测精度的同时,大幅提高了推理速度和计算效率,非常适合在实时目标检测任务中应用。

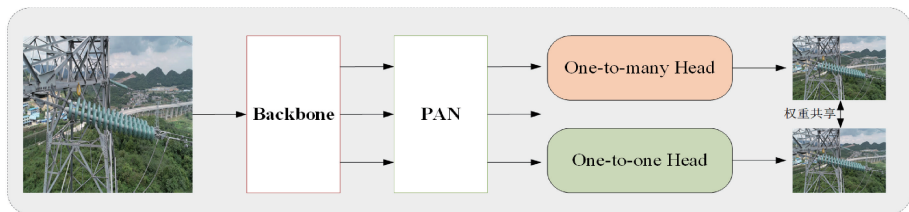


图 1 YOLOv10 网络结构

Fig. 1 YOLOv10 network structure

1.2 多尺度特征处理与尺度内特征处理

在目标检测中,特征图的特征处理分为多尺度特处理和尺度内特征处理。多尺度特征处理包括特征的提取、转换和融合,目的是确保模型能够在不同尺度特征图上同时检测大小目标。由于待检测目标在图像中的尺度往往存在显著差异,因此多尺度特征处理是提升检测精度的关键,常见的多尺度特征融合方法是特征金字塔网络(feature pyramid networks, FPN)^[16],其通过将不同尺度和层次的特征图进行融合,使模型能够更好地捕捉到不同大小的目标。而在输电线路缺陷检测任务中对模型的特征提取改进方法,徐业东等^[17]在特征融合网络部分嵌入 ConvMixer 层,ConvMixer 层的结构设计有助于在特征信息中更好的捕捉空间和通道的关系,提升模型对于小目标的检测能力;然后在特征融合网络部分引入 ODConv 模块,从而提高模型对复杂场景和小目标的检测性能。

尺度内特征处理则是指在单张特征图内进行信息交互,以增强特征表达能力,这种交互对于复杂背景中的小目标检测尤为重要。常见的尺度内特征处理方法是添加注意力机制,主要包括通道注意力^[18]、空间注意力^[19],以及 Transformer 中的自注意力机制^[20]。具体而言,通道注意力通过池化等操作捕捉通道间的关系,重新分配每个通道的权重;空间注意力则关注输入数据的空间位置信息,增强模型对局部特征的捕捉能力,文献[21]中结合了空间和通道注意力,通过逐层处理提升了特征的表达能力。此外,Transformer 中的多头自注意力机制将输入向量划分为多个子空间,分别计算多个注意力头,从而增强了模型的整体表达能力。而在使用注意力机制对 YOLOv10 的改进策略中,黄毅等^[22]为了提升模型在复杂背景中对目标的特征提取和融合能力,其在 YOLOv10 的原始 C2f 模块中引入 CBAM 注意力机制得到 C2f-CM,增强了复杂背景中对特征目标的定位能力。

1.3 遮挡感知注意网络

在无人机电力缺陷检测中,遮挡问题是常见难点之一,由于无人机巡检位置固定,绝缘子、防震锤等缺陷部件通常被高压铁塔等结构遮挡,使得模型对局部区域过

度依赖而导致对齐错误,局部混叠和特征丢失问题。传统卷积神经网络或单一特征提取的方法,通常假设检测目标清晰可见,然而,由于它们不能很好地捕捉到全局信息或重构被遮挡的部分,因此目标被部分遮挡时,这些方法往往难以有效处理。为解决这一问题,本文引入遮挡感知注意网络 SEAM 模块^[23],此模块通过动态聚焦图像中的显著区域来增强模型的检测能力。具体来说,SEAM 模块通过全局特征的有效提取,减少了由于遮挡导致的误对齐现象;同时,通过注意力机制消除局部特征重叠问题,从而提高了模型在复杂背景下的检测效果,且 SEAM 可以有效利用全局上下文信息推断被遮挡部分的特征,减少了因遮挡造成的信息丢失,显著提升了模型在遮挡场景中的表现。

2 本文方法

2.1 FC2FA 多尺度特征融合模块

在 YOLOv10 中,骨干网络输出 3 个尺度的特征图,由输电线路巡检得到原始图像中存在的目标像素较少,导致特征图中包含大量冗余信息。在进行多尺度特征融合时,这些冗余信息不仅会增加计算量,还可能导致目标语义信息的丢失。现有研究中对 YOLOv10 的改进策略通常是直接添加注意力机制,如直接在模型的 C2F 模块中添加 CBAM 注意力机制来增强特征捕捉能力,然而,这样虽然可以使模型提高对目标特征的有效关注,也会造成计算负担,因此如何在保持较少参数数量和计算量的同时实现高效的特征融合成为了目前的主要难点。在文献[24]中,不同于以往对特征图中所有输入通道都进行卷积操作,其所提出的 PConv 只对其中一部分进行卷积,这是由于不同通道的特征图高度相似且具有大量冗余信息,因此使用 PConv 既能避免多余参数的产生,又能够有效提取关键特征;文献[25]中设计的上下文锚点注意力机制 CAA 旨在增强中心特征的同时掌握远距离像素之间的上下文相互依赖性,适用于不同尺度目标的特征捕捉与区分。受以上工作启发,在输电线路多缺陷检测任务中,由于无人机拍摄图片目标像素占比非常小,使用部分卷积可以在充分提取目

标特征时减少大量冗余参数,而且与输电线路中单缺陷检测任务不同,多缺陷检测任务中不同目标的边界像素距离以及特征差异过大,使用 CAA 可以专注于提取多尺度目标的上下文信息,因此本文基于以上工作设计 FC2FA 模块,具体结构如图 2 所示。FC2FA 中 Efficient Fusion Block 为本文设计的轻量化特征融合块,右侧虚线内为其具体结构。将模型原始 C2F 模块中的 DarknetBottleneck 替换为本文设计的高效融合块 Efficient Fusion Block,其中使用普通卷积和 PConv 进行特征提取,多层感知机 (multilayer perceptron, MLP) 进行有效特征表示,Drop path 正则化用以提高泛化能力,最后使用 CAA 加权多种缺陷目标特征。

在 Efficient Fusion Block 中,输入一个张量 $X \in \mathbf{R}^{c \times h \times w}$ 后,首先通过卷积核为 1 的卷积层调整特征通道数,为了减少模型产生过多冗余参数和计算量,使用部分卷积 PConv 处理输入特征,此过程可以表示为:

$$\{X_1, X_2\} = \text{Split}(\text{Conv}_{1 \times 1}(X)) \quad (1)$$

$$Y = \text{Concat}(\text{PConv}_{2d_{3 \times 3}}(X_1), X_2) \quad (2)$$

其中,部分卷积块将输入特征分割成两部分 $\{X_1, X_2\}$, X_1 经过一个卷积核为 3,步长和填充为 1,输出通道为 $\text{dim}/4$ 的卷积层,之后与 X_2 连接输出 Y 并传递至 MLP 进行处理:

$$Y' = \text{Conv}_{2d_{1 \times 1}}(\text{Conv}_{1 \times 1}(Y)) \quad (3)$$

式中:MLP 由两个卷积核为 1 的卷积层组成,第 1 层的输出通道为 $\text{dim}/2$,第 2 层为标准的二维卷积层,将输出通道还原。为避免模型过拟合,增强泛化能力,模型通过 drop path 随机丢弃 10% 的连接。然后使用上下文锚点注意力机制对目标特征进行加权处理,过程如下:

$$Z = \epsilon(\text{Conv}_{1 \times 1}(\omega_1(\omega_2(\text{Conv}_{1 \times 1}(\beta(Y')))))) \quad (4)$$

式中: $\epsilon(\cdot)$ 为 Sigmoid 激活函数, $\omega_1(\cdot)$ 和 $\omega_2(\cdot)$ 分别表示在垂直方向和水平方向的 DWConv, $\beta(\cdot)$ 表示平均池化 (AvgPool)。

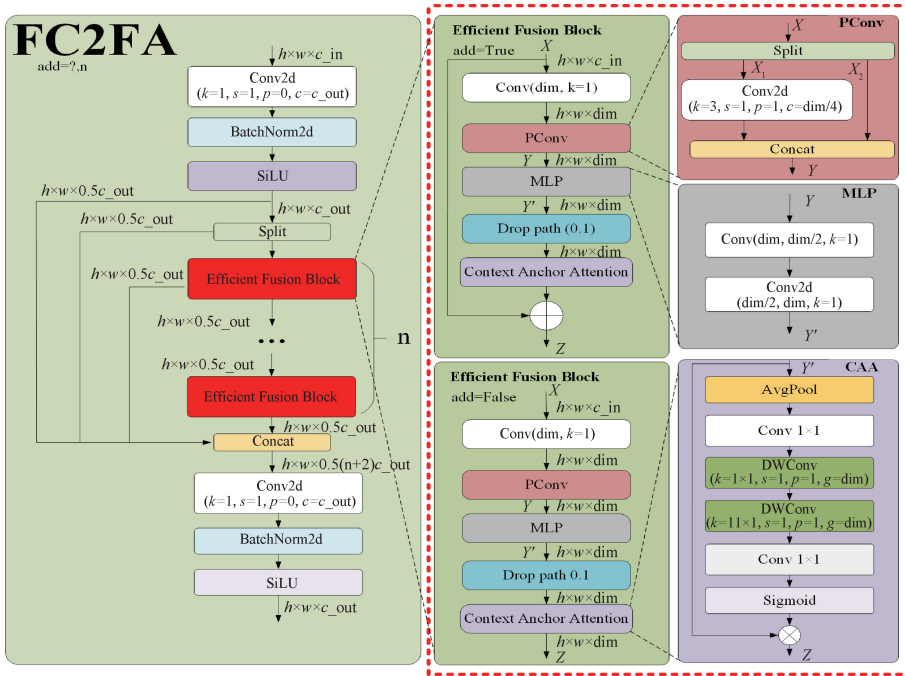


图 2 FC2FA 模块结构

Fig. 2 FC2FA module structure

在上下文锚点注意力中,为了得到局部区域特征,初始特征首先进行平均池化层和卷积核为 1 的卷积操作,随后使用两个轻量级的 DWConv 对特征进行远程上下文信息的提取,在第 1 个 DWConv 中,使用卷积核为 11,步长和填充为 1,并对每一个通道进行分组,对特征图进行水平方向的卷积操作,第 2 个 DWConv 参数相同,但是对特征图进行垂直方向的卷积操作,最后经过 $\text{Conv } 1 \times 1$ 和 Sigmoid 激活生成注意力权重,并对特征进行加权后输出加权特征 Z 。

在本文设计的 Efficient Fusion Block 中,由于使用的卷积层多数为 DWConv 和 PConv,使得参数量和浮点计算

量显著减少。此外,使用上下文锚点注意力机制使得模型既能关注局部信息,又能获取远距离的上下文信息,从而提升了模型性能和鲁棒性。

2.2 尺度内特征交互模块

骨干网络输出的最后一层特征图通常具有较小的尺寸,但包含更丰富、更高级的语义信息^[9],将自注意力操作应用于高级特征图能够有效提升小目标检测和识别的效果。目前在改进 YOLOv10 提升小目标检测效果的工作中,多数策略为添加 P2 小目标检测头,而 P2 检测头属于较大尺度的特征图,直接添加会导致模型检测速度大幅降

低,不适合无人机实时巡检任务,为了解决这一问题,本文在多个特征图进行多尺度特征融合之前,首先对骨干网络输出的最小尺度特征图进行尺度内特征交互处理,具体做法是通过 AIFI 模块替换原始模型中 SPPF 和 PSA 模块,仅使用单层 Transformer 中的多头自注意力机制增强小目标特征的表达效果。

基于多头自注意力机制的尺度内特征交互 AIFI 模块如图 3 所示,通过该模块可以有效地对具有丰富语义的目标特征进行加权。在 AIFI 模块中,输入张量 \mathbf{X} 的形状为 $[\text{batch size}, \text{channels}, \text{height}, \text{width}]$,首先对 \mathbf{X} 进行展平 (Flatten) 和转置 (Permute) 操作后得到符合解码层 (Transformer Encoder Layer) 所需要的序列格式 X_{flat} ,其过程可以表示为:

$$X_{flat} = X, \text{Flatten}(2), \rho(0, 2, 1) \quad (5)$$

式中: $\rho(0, 2, 1)$ 表示通过转置 (Permute) 操作调整序列的维度,然后利用多头自注意力机制对其嵌入位置编码,该序列进行加权处理得到 X_{attn} ,并通过位置嵌入 (Positional Embedding) 来保持特征的空间关系,公式如下:

$$Q = K = \text{PosEmbed}(X_{flat}, \text{pos_embed}) \quad (6)$$

$$X_{attn} = \text{MultiHead SelfAttention}(Q, K, X_{flat}) \quad (7)$$

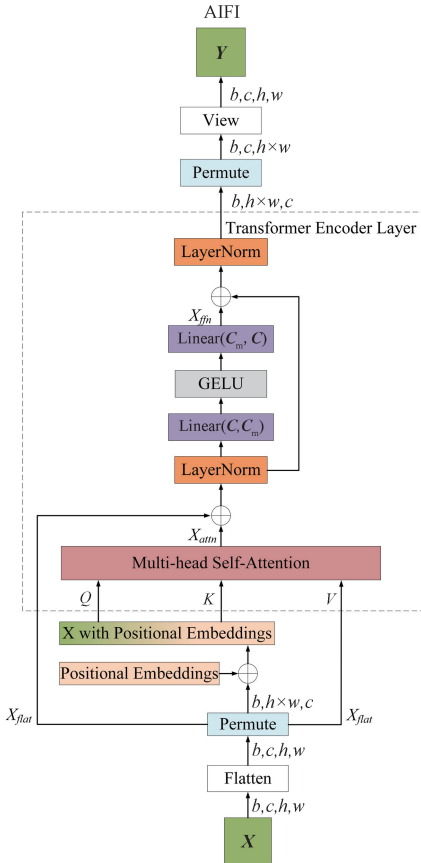


图 3 AIFI 网络结构

Fig. 3 AIFI network structure

在 Transformer Encoder Layer 中,包含两个线性层 (Linear),第 1 个线性层将输入特征的通道数 C 映射到更大的维度 C_m ,即隐藏层的维度,以学习更复杂的特征组合。第 2 个线性层将维度还原为原始通道数,确保输出与输入一致,以便与原始的输入特征进行残差连接,之后得到全连接输出 X_{ffn} ,其公式表示为:

$$X_{ffn} = \text{fc2}(\text{GeLU}(\text{fc1}(\delta(X_{flat} + X_{attn})))) \quad (8)$$

式中: $\text{fc1}(\cdot)$ 和 $\text{fc2}(\cdot)$ 分别表示第 1 个和第 2 个线性层, $\delta(\cdot)$ 表示层归一化 (LayerNorm) 操作,最终特征经过归一化后进行转置和还原操作,保证输出特征 \mathbf{Y} 的形状与原始特征 \mathbf{X} 相同,过程如下:

$$\mathbf{Y} = \varphi(\rho(\delta(X_{ffn}) + \delta(X_{flat} + X_{attn}))) \quad (9)$$

式中: $\varphi(\cdot)$ 表示调整张量形状的操作 (View),确保数据形状不变。

通过在特征图上添加 AIFI 模块,能够充分利用包含丰富高级语义信息的最小尺寸特征图,此外,由于多头自注意力机制的计算复杂度与特征图尺寸成 $O(n^2)$ 的关系,仅对最小尺度特征图进行尺度内特征交互,能够避免过多的计算开销。这一改进使得模型在复杂场景中能够更准确地定位和识别输电线路缺陷中的小目标,减少漏检和误检现象。

2.3 遮挡感知注意检测头

为了解决输电线路检测任务中的遮挡问题,本文设计 SEAM-Head,其网络结构如图 4 所示。SEAM-Head 通过获取特征图中目标的全局和局部信息,能够动态感知被遮挡的部分,并将注意力集中在有效特征上,降低因遮挡带来的检测误差。

在 YOLOv10 中,One-to-one 检测头和 One-to-many 检测头具有相同的结构,区别在于前者只保留置信度最高的检测框,而后者保留多个置信度较高的检测框。本文在原始 YOLOv10 模型的基础上,增加了遮挡感知注意网络 SEAM 模块。具体操作如下,首先,将原始检测头中 CV2 操作序列中的最后一个 CBS(Conv2d, BatchNorm, SiLU) 操作序列替换为 SEAM 模块;而在 CV3 操作序列中,减少一个 CBS 的同时将最后一个 CBS 替换为 SEAM 模块,保证模型特征捕获能力的同时降低冗余计算负担。最终, CV2 和 CV3 的输出通过 Concat 连接进行融合。

当特征 $\mathbf{X} \in (\mathbb{R}^{C \times h \times w})$ 输入 SEAM 模块中时,首先经过由 3 个通道空间混合模块 (channel and spatial mixing module, CSMM) 组成的残差网络部分,即融合通道空间信息模块,3 个模块的 patch 分别为 6、7、8。CSMM 模块为具有残差连接的深度可分离卷积。使用深度可分离卷积按深度逐层操作,在学习不同通道特征的同时减少计算量,但是它忽略了通道之间的信息联系,而为了弥补这一损失,使用逐点卷积将不同深度的卷积的输出结合起来。然后通过两层全连接网络来融合各个通道之间的信息得到权重 \mathbf{Y} ,通过这种方式,模型能够学习到遮挡目标与非遮挡

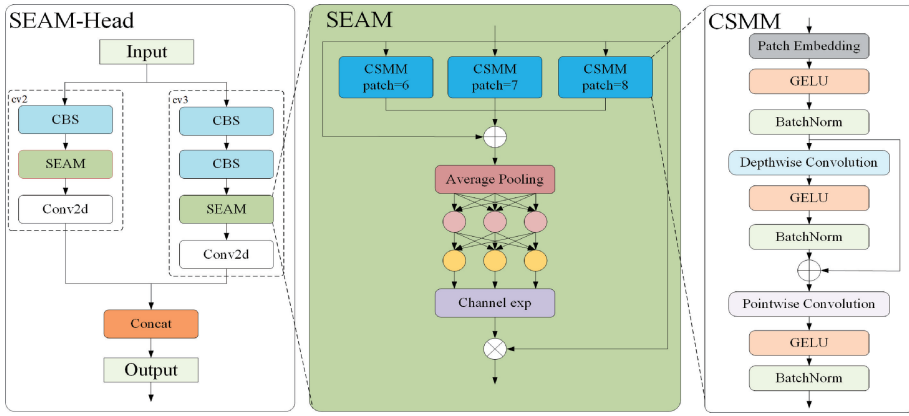


图 4 SEAM-Head 网络结构

Fig. 4 SEAM-Head network structure

目标之间的关系,在遮挡场景下弥补上述损失。此过程可以使用公式表示为:

$$Y = \exp(\epsilon(\text{Linear}(\text{ReLU}(\text{Linear}(X)))))) \quad (10)$$

式中: $\epsilon(\cdot)$ 为 Sigmoid 激活函数。通过全连接层学习到的输出经过指数函数处理(exp),将值范围从 $[0, 1]$ 扩展到 $[1, e]$,此操作提供了单调的映射关系,使模型对遮挡引起的位置信息误差更具容忍性。最后,SEAM 模块的输出被扩展到与输入特征 X 相同的形状,并与原始特征相乘,得到注意权重 W ,公式表示为:

$$W = X \cdot Y.\text{expand_as}(X) \quad (11)$$

式中:expand_as 操作表示将 Y 扩展到与 X 相同的形状。SEAM 模块通过突出重要特征并抑制不相关的信息,从而更好地应对遮挡场景。

2.4 TLDDet 整体架构

本文基于 YOLOv10 s 模型进行改进而设计出输电线路多缺陷检测模型 TLDDet,其结构如图 5 所示。TLDDet 接收一个输入尺寸 640×640 的 RGB 原始图像,模型的骨干网络承担初步特征提取的功能,其中 FC2FA 模块中的 Efficient Fusion Block 将 add 参数设置为 True,表示进行残差连接,从而在初步特征提取的时候充分保留原始特征信息。而颈部网络(Neck)负责对骨干网络输出的特征图进行多尺度融合,以便在检测不同大小的目标时具有更好的鲁棒性。Neck 部分的多尺度特征融合同样基于 FC2FA 模块,但是此时 add 参数为 False,在减少计算量时更专注于提取新的特征,特别是远距离拍摄的小目标。而为了更好地利用骨干网络输出的高级语义信息,本文引入 AIFI 模块来重点关注最小尺寸特征图中语义丰富的目标特征,有效地提升了小目标的检测性能。最后针对遮挡问题,SEAM-Head 检测头采用了遮挡感知注意网络 SEAM 模块,在检测过程中动态聚焦于图像中的显著区域,消除因遮挡引发的错位、局部特征混叠等问题,确保对目标的精确检测。此外,SEAM-Head 仍然保留了双标签分配策略,使得模型无需 NMS 即可完成端到端的目标检测,显著提

高了模型推理效率。

3 实验与结果分析

3.1 数据集及实验平台

本文实验数据集为输电线路无人机巡检航拍真实图像数据集(transmission line defect dataset, TLDD),数据集包含 2 000 张真实缺陷图像,涵盖 4 种缺陷类别,部分缺陷样本如图 6 所示,分别为绝缘子缺陷、防震锤缺陷、鸟巢缺陷和蜂巢缺陷。数据集按照 8:2 的比例分为训练集和验证集,具体样本标签及解释如表 1 所示。

训练输电线路缺陷数据集时,首先对数据集进行了数据增强处理,包括几何变换、马赛克增强、混合增强、随机扰动以及颜色变换、以提升训练数据的多样性,缓解样本不均衡的问题,并增强模型对不同尺度变化的适应性。所有训练图像统一调整为 640×640 pixels, batch size 设置为 8,训练总轮次为 300 epoch。本文所有实验均在相同的实验环境中完成,具体环境为 Windows 11 操作系统,处理器为 i7-12700 h,内存 16 GB,显卡为 RTX 3070 laptop GPU,深度学习框架为 PyTorch 2.0.1, CUDA 版本为 11.8。

3.2 实验评价指标

本文实验主要通过以下指标来评估模型的缺陷目标检测性能:检测精度(Precision)、平均精度均值(mAP50:95, mAP50)、召回率(Recall)、模型检测速度(FPS),以及模型的参数量(Params)和浮点计算量(GFLOPs)。检测精度 Precision 和召回率 Recall 的计算公式如下:

$$\text{Precision} = \frac{TP}{TP + FP} \quad (12)$$

$$\text{Recall} = \frac{TP}{TP + FN} \quad (13)$$

式中: TP 表示模型正确预测的正样本数量, FP 表示将负样本错误预测为正样本的数量, FN 表示将正样本错误预测为负样本的数量。精度衡量的是模型在预测为正样本的所有实例中,实际为正样本的比例;召回率则表示模型

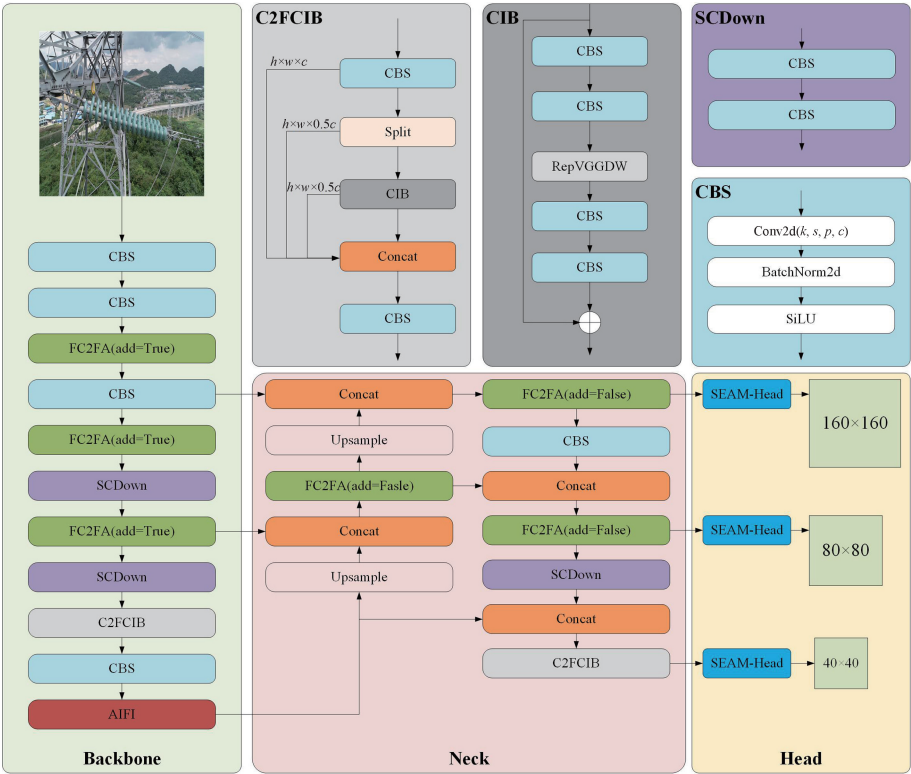


图 5 TLDDet 网络结构
Fig. 5 TLDDet network structure



图 6 TLDD 数据集部分样本
Fig. 6 TLDD dataset partial samples

表 1 TLDD 输电线路缺陷标签数据集说明

Table 1 Description of the TLDD transmission line defect label dataset

数据集	图像数量	训练集/验证集	标签名称	标签含义	缺陷标签个数
TLDD	2 000	1 600/400	Shockproof hammer	防震锤缺陷	771
			Honeycomb	蜂巢缺陷	231
			Insulator	绝缘子缺陷	904
			Nest	鸟巢缺陷	702

在所有实际为正样本的实例中,能够正确预测为正样本的比例。平均精度均值的计算公式如下:

$$mAP = \frac{\sum_{i=1}^N AP_i}{N} \quad (14)$$

式中: N 表示数据集中类别的数量,本文研究中类别数量为 4, AP_i 表示第 i 类的平均精度(AP)。AP 是评估模型

检测精度的重要指标,平均精度均值(mAP)则是所有类别 AP 的平均值。mAP50 表示交并比(IoU)阈值为 0.50, mAP50:95 是计算从 IoU 阈值 0.50~0.95(步长为 0.05)下的平均精度。

FPS(每秒帧数)是衡量模型运行速度的关键指标,通常与模型参数量(Params)和浮点计算量(FLOPs)相关。FPS 的计算公式如下:

$$FPS = \frac{G_{all}}{T} \tag{15}$$

式中: G_{all} 表示图片总数量, T 表示计算总耗时(单位为 s)。

3.3 与其他模型的检测对比实验

为了验证本文改进模型 TLDDet 的先进性,将目前主流实时检测模型以及相关改进模型与 TLDDet 进行对比,其中 SC-YOLO 为输电线路异物检测模型,LED-YOLO 为基于 YOLOv10 n 改进的无人机复杂背景下多尺度检测模型,所有模型使用 TLDD 数据集并在相同实验环境下进行实验,实验结果如表 2 所示。通过分析数据可知,TLDDet 的参数量仅为 4.61 M,浮点计算量为 15.2 GFLOPs,在检测速度上达到 143 FPS,相较于原始模型 YOLOv10s,参数量减少了 2.59 M,计算量降低了 6.4 GFLOPs,保持了较

低计算负担和内存占用的同时,TLDDet 的检测精度达到 88.9%,相较于 YOLOv10s,精度提高了 4.3%,召回率提升了 2.4%,mAP50 和 mAP50:95 分别提高了 3.7% 和 2.2%。与参数量为 TLDDet 四倍的 RT-DETR 相比,TLDDet 在 mAP50 高出 1.3%。与 YOLO 其他系列模型相比 TLDDet 各项指标均表现良好,在与输电线路异物检测算法 SC-YOLO 的对比中,TLDDet 拥有更少的计算量以及更精确的检测效果,而与 LED-YOLOv10 对比,虽然其参数量仅有 2.7 M,检测速度达到了 161 FPS,但 mAP50 仅有 69.1%,难以保证精确检测。综合实验结果表明,TLDDet 更加适合输电线路多缺陷检测任务。

图 7 展示了 TLDDet 与原始 YOLOv10s 模型在 Precision、Recall、mAP50 及 mAP50:95 的对比。

表 2 不同模型对比实验结果

Table 2 Comparison of experimental results across different models

模型	Params/M	GFLOPs	FPS	Precision/%	Recall/%	mAP50/%	mAP50:95/%
RTMDET-s ^[6]	8.99	14.8	136	0.838	0.721	0.759	0.413
RT-DETR-R18 ^[9]	20.10	60.0	87	0.876	0.788	0.819	0.435
YOLOX-s ^[2]	9.00	26.8	119	0.833	0.749	0.802	0.409
YOLOv8-s ^[3]	11.20	28.6	112	0.835	0.751	0.789	0.411
YOLOv9-s ^[4]	7.10	26.4	129	0.843	0.744	0.787	0.410
SC-YOLO ^[14]	10.90	26.4	117	0.877	0.780	0.822	0.421
LED-YOLOv10 ^[22]	2.70	8.7	161	0.757	0.689	0.691	0.366
YOLOv10-s ^[5]	7.20	21.6	123	0.846	0.747	0.798	0.415
TLDDet(本文)	4.61	15.2	143	0.889	0.771	0.835	0.437

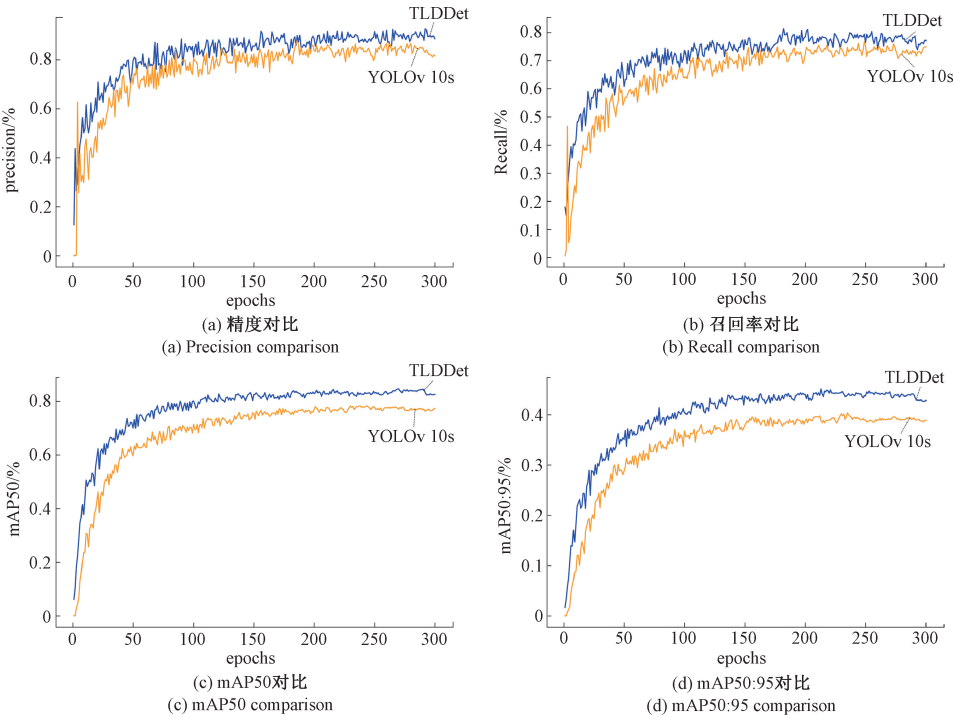


图 7 TLDDet 与原始模型 YOLOv10s 的精度、召回率、平均精度均值 mAP50,mAP50:95 对比
Fig. 7 Comparison of precision, recall, mAP50, and mAP50:95 between TLDDet and the YOLOv10s model

从图 7 中可以明显看出,TLDDet 的各项指标均高于 YOLOv10s,召回率的提升也表明模型能够有效减少漏检现象。此外,在 mAP50 和 mAP50:95 的评估中,TLDDet 同样展现出更优的性能。实验结果验证了本文提出的改进方法在实际应用中的有效性。

3.4 不同缺陷类别检测精度对比

为了验证 TLDDet 模型的缺陷检测性能,本文对比了 TLDD 数据集中 4 种缺陷类别在多个模型上的检测精度差异,实验结果如表 3 所示,TLDDet 在防震锤、蜂巢和鸟巢 3 种类别的检测效果均优于原始 YOLOv10s 模型,三者 mAP50 分别提高了 3.4%、2.5% 和 4.5%,Recall 也有一定提高,而在绝缘子类别上的检测精度略微下降了 0.3%。相比其他实时检测模型,如 YOLOv8, YOLOv9 等,

TLDDet 整体检测水平平均更为优异;与输电线路异物检测算法 SC-YOLO 相比,TLDDet 在检测蜂巢和绝缘子时召回率分别少 0.05% 和 0.06%,但是防震锤和鸟巢缺陷的检测精度分别高 0.08% 和 2.1%,且 TLDDet 的参数量和计算量仅为 SC-YOLO 的 1/2;与无人机检测方法 LED-YOLOv10 相比,TLDDet 的检测速度不如前者,但是所有缺陷类别的检测精度都更高;而与基于 Transformer 的实时端到端模型 RT-DETR 相比,TLDDet 在防震锤和蜂巢类别的检测效果表现更好,虽然在绝缘子和鸟巢类别上的检测精度略低于 RT-DETR 模型,但是 TLDDet 的参数量和计算量仅为 RT-DETR 的 1/4。这表明了 TLDDet 在实际巡检中能够准确快速地检测输电线路缺陷,更加适合无人机巡检任务。

表 3 不同模型在多个缺陷类别检测效果的对比实验结果

Table 3 Comparison of detection performance across multiple defect categories for different models %

模型	Shockproof hammer		Honeycomb		Insulator		Nest	
	mAP50	Recall	mAP50	Recall	mAP50	Recall	mAP50	Recall
RTMDet-s ^[6]	0.599	0.552	0.881	0.780	0.902	0.871	0.776	0.707
RT-DETR-R18 ^[9]	0.671	0.601	0.886	0.830	0.919	0.885	0.841	0.751
YOLOX-s ^[2]	0.595	0.568	0.883	0.801	0.911	0.855	0.735	0.757
YOLOv8-s ^[3]	0.649	0.597	0.878	0.907	0.909	0.915	0.775	0.730
YOLOv9-s ^[4]	0.645	0.525	0.880	0.815	0.914	0.896	0.788	0.713
SC-YOLO ^[14]	0.679	0.581	0.901	0.839	0.917	0.891	0.817	0.744
LED-YOLOv10 ^[22]	0.571	0.522	0.855	0.800	0.891	0.870	0.624	0.598
YOLOv10-s ^[5]	0.653	0.571	0.879	0.833	0.915	0.893	0.793	0.737
TLDDet(本文)	0.687	0.598	0.904	0.834	0.912	0.885	0.838	0.767

3.5 FC2FA 模块消融实验

FC2FA 模块中的关键部分为 Efficient Fusion Block,其主要由 PConv 和 MLP 组成的模块(下文称为 Faster block)以及 CAA 组成。为了验证这两个模块的有效性,对其进行拆解消融实验,实验分别顺序将 Faster block 和 CAA 模块添加到原始模型的 C2F 模块中,实验结果如表 4 所示。

表 4 FC2FA 模块消融实验结果

Table 4 Ablation experiment results of the FC2FA

模块	Param/M	mAP50/%	Recall/%
C2F	7.20	0.798	0.747
+Faster block	4.43	0.792	0.733
+CAA	4.79	0.818	0.751

由实验结果可知,仅使用 Faster block 时,得益于部分卷积的作用,模型的参数量减少了 2.77 M,但 mAP50 下降了 0.6%。当同时包含 Faster block 和 CAA 时参数量比原始 C2F 模块降低了 2.41 M,同时 mAP50 提高了 2%。

这表明,Faster block 在减少模型参数量时保留了目标的主要特征,而 CAA 能够更好地捕捉上下文语义信息,从而增强缺陷特征的检测效果。

3.6 整体消融实验

为了验证本文多个改进方法的有效性,针对 FC2FA、AIFI^[9] 和 SEAM-Head^[23] 模块进行了消融实验,实验以 YOLOv10 s 作为基准模型,实验过程严格控制模型参数一致。消融实验结果如表 5 所示。实验结果表明,对于每个单独模块的改进均一定程度上提高了模型的精度与性能,而且 FC2FA 模块的改进效果最为明显,不仅参数量有所降低,mAP50 也提高了 2%。在添加 AIFI 模块和 SEAM-Head 检测头时,本文并不是直接添加,而是在使用 AIFI 模块时去掉原模型中主干网络的 SPPF 和 PSA 模块,在使用 SEAM-Head 检测头时,重新设计原始检测头中 CV2 和 CV3 操作序列,将 CV2 和 CV3 操作系列中的 CBS 操作序列替换为遮挡感知注意网络 SEAM 模块,经实验证明,这些改进不仅能够保证检测效果的同时能够降低模型的参数量和计算量。在 3 种改进中两两组合的情况下 FC2FA 和 AIFI 的组合效果最好,不仅参数量只有 4.79 M,

mAP50 也提高了 2.6%,而在 3 种改进同时使用时,模型的性能得到了更大的提升,对输电线路多种缺陷的平均检测精度达到了 83.5%,整体检测效果提高了 3.7%。参数量降低了 2.59 M。综上所述,FC2FA, AIFI 和 SEAM-Head 模块均对模型的检测精度提升起到了正向促进作用,且三者协同作用下,模型在降低计算量的同时有效提升了检测精度,证明了改进方法的有效性。

表 5 TLDDet 模型整体消融实验结果

Table 5 Ablation experiment results of the tlddet model					
FC2FA	AIFI	SEAM-Head	Param /M	mAP50 /%	Recall /%
—	—	—	7.20	0.798	0.747
✓	—	—	6.68	0.818	0.751
—	✓	—	6.55	0.805	0.755
—	—	✓	7.04	0.801	0.747
✓	✓	—	4.79	0.824	0.759
✓	—	✓	6.16	0.820	0.755
—	✓	✓	5.50	0.819	0.769
✓	✓	✓	4.61	0.835	0.771

3.7 检测效果对比分析

为了更加直观地展示改进模型 TLDDet 的检测效果, 本文将其与原始模型 YOLOv10 s 的检测结果进行了可视化对比。图 8 展示了两个模型的检测效果对比,其中左侧为改进模型 TLDDet 的检测结果,右侧为原始 YOLOv10 s 的结果,同时本研究使用 Grad-CAM^[26] 进行热力图可视化,它通过生成热力图来展示模型在做决策时关注的区域,从而提供模型决策过程的可视化解释。在图 8(a)中,场景为高空树木背景下检测较远距离的防震锤。可以明显看出,TLDDet 能够完整检测出具有缺陷的防震锤,而 YOLOv10 s 则出现了漏检的情况,从热力图的对比中发现,TLDDet 对防震锤缺陷有非常好的特征捕捉能力, YOLOv10 s 无法聚焦缺陷特征。在图 8(b)中,背景为高空铁塔,检测目标是蜂巢,TLDDet 成功检测到了蜂巢,而 YOLOv10 s 出现了漏检现象,图 8(c)展示了城镇背景下对绝缘子的检测效果。尽管两个模型都能检测出绝缘子缺陷,但 YOLOv10 s 出现了误检,将阴影处的铁塔部分误识别为鸟巢。最后在图 8(d)中,背景为房屋环境,目标为鸟巢。在背景颜色与鸟巢接近的情况下,TLDDet 依然能够准确检测出鸟巢,而 YOLOv10 s 未能检测到,热力图的对比也显示出了 TLDDet 对缺陷目标更好的聚焦能力,综合这些对比分析,TLDDet 在复杂背景下的输电线路缺陷检测任务中展现出了更优异的检测性能,为无人机输电线路巡检提供了更为可靠的技术支持。

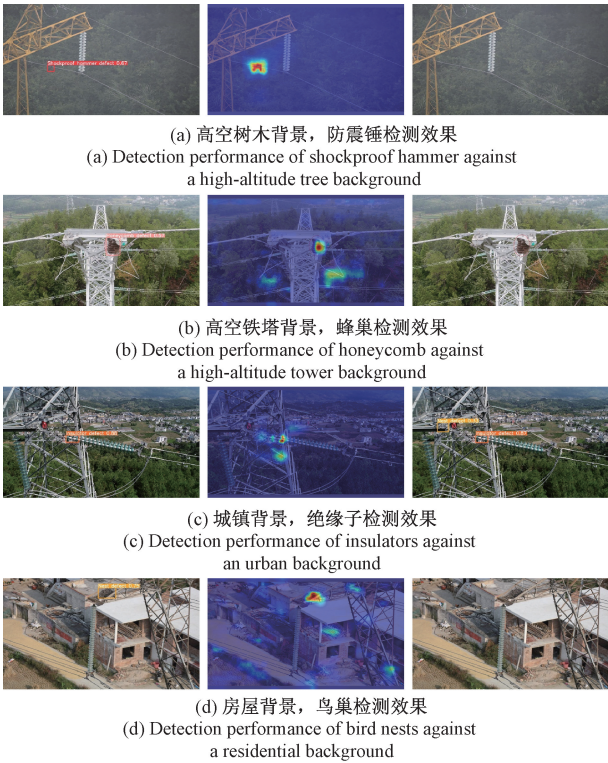


图 8 不同背景下 TLDDet 与 YOLOv10s 的检测效果对比可视化

Fig. 8 Comparison of detection performance between TLDDet and YOLOv10s under different backgrounds

4 结 论

针对输电线路多缺陷识别中检测精度低和检测速度慢的问题,本文提出了基于 YOLOv10 s 改进的实时端到端缺陷检测模型 TLDDet。首先,改进原始特征融合 C2F 模块设计出高效多尺度特征融合模块 FC2FA,此模块采用了部分卷积进行特征提取,以减少参数量和计算量,引入上下文锚点注意力(CAA)捕捉关键特征信息以提升检测精度。其次,使用 AIFI 模块对最小尺寸特征图进行尺度内特征交互,增强了模型对输电线路小目标缺陷的检测能力。最后,设计了遮挡感知注意检测头,通过 SEAM-Head 检测头有效减少了因目标遮挡导致的特征丢失问题。实验结果表明 TLDDet 具有更好的检测效果与模型性能,在低算力平台下的无人机缺陷检测场景中能够实时检测,为后续的输电线路多缺陷检测任务提供了一定的方法。

参考文献

[1] 柳方圆,任东,王露,等. 高分辨率的航拍输电线路绝缘子检测[J]. 电子测量技术, 2023, 46(13): 102-109.
LIU F Y, REN D, WANG L, et al. Transmission line insulator detection based on high resolution uav image[J]. Electronic Measurement Technology, 2023, 46 (13): 102-109.

- [2] 黄力,万旭东,王凌云,等. 基于图像增广与迁移学习的输电线路金具多目标实时检测方法[J]. 电子测量技术,2022,45(20):135-142.
HUANG L, WAN X D, WANG L Y, et al. Multi-objective real-time detection method of transmission line fittings based on image augmentation and transfer learning [J]. Electronic Measurement Technology, 2022,45(20):135-142.
- [3] VARGHESE R, SAMBATH M. YOLOv8: A novel object detection algorithm with enhanced performance and robustness[C]. 2024 International Conference on Advances in Data Engineering and Intelligent Computing Systems. IEEE, 2024: 1-6.
- [4] WANG C Y, YE H I, MARK LIAO H Y. YOLOV9: Learning what you want to learn using programmable gradient information [C]. European Conference on Computer Vision. Cham: Springer, 2025: 1-21.
- [5] WANG AO, CHEN H, LIU L H, et al. Yolov10: Real-time end-to-end object detection [J]. ArXiv preprint arXiv:2405.14458,2024.
- [6] LYU CH Q, ZHANG W W, HUANG H AN, et al. RTMDet: An empirical study of designing real-time object detectors [J]. ArXiv preprint arXiv: 2212.07784,2022.
- [7] NEUBECK A, GOOL L V. Efficient non-maximum suppression [C]. 18th International Conference on Pattern Recognition. IEEE, 2006,3: 850-855.
- [8] REN SH Q, HE K M, 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.
- [9] ZHAO Y AN, LYU W Y, XU SH L, et al. DETRs beat YOLOs on real-time object detection[C]. IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2024: 16965-16974.
- [10] 赵霖,王素珍,邵明伟,等. 基于改进 YOLOv5 的输电线路鸟巢缺陷检测方法[J]. 电子测量技术,2023,46 (3):157-165.
ZHAO L, WANG S ZH, SHAO M W, et al. Improved YOLOv5-based bird's nest defect detection method for transmission lines [J]. Electronic Measurement Technology, 2023,46(3):157-165.
- [11] 肖黎俊,潘睿志,李超,等. 基于改进 YOLOv5s 绝缘子缺陷检测技术研究[J]. 电子测量技术, 2022,45(24): 137-144.
XIAO C J, PAN R ZH, LI CH, et al. Research on defect detection technology based on improved YOLOv5s insulator [J]. Electronic Measurement Technology, 2022,45(24):137-144.
- [12] 刘兰兰,万旭东,汪志刚,等. 基于超分辨率重建与多尺度特征融合的输电线路缺陷检测方法[J]. 电子测量与仪器学报,2023,37(1):130-139.
LIU L L, WAN X D, WANG ZH G, et al. Transmission line defect detection method based on super-resolution reconstruction and multi-scale feature fusion [J]. Journal of Electronic Measurement and Instrumentation, 2023,37(1):130-139.
- [13] 李鑫,刘帅男,杨桢,等. 基于改进 Cascade R-CNN 的输电线路多目标检测[J]. 电子测量与仪器学报, 2021, 35(10):24-32.
LI X, LIU SH N, YANG ZH, et al. Multi-target detection of transmission lines based on improved cascade R-CNN[J]. Journal of Electronic Measurement and Instrumentation, 2021,35(10):24-32.
- [14] 易磊,黄哲玮,易雅雯. 改进 YOLOv8 的输电线路异物检测方法[J]. 电子测量技术,2024, 47(15):125-134.
YI L, HUANG ZH W, YI Y W. Improved YOLOv8 foreign object detection method for transmission lines[J]. Electronic Measurement Technology, 2024, 47(15):125-134.
- [15] 刘传洋,吴一全,刘景景. 基于视觉的输电线路金具锈蚀缺陷检测方法研究进展[J]. 仪器仪表学报,2024,45 (3):286-305.
LIU CH Y, WU Y Q, LIU J J. Research progress of vision-based rust defect detection methods for metal fittings in transmission lines[J]. Chinese Journal of Scientific Instrument, 2024,45(3):286-305.
- [16] LIN T Y, DOLLÁR P, GIRSHICK R, et al. Feature pyramid networks for object detection [C]. IEEE Conference on Computer Vision and Pattern Recognition, 2017: 2117-2125.
- [17] 徐业东,蔡亚恒,李严,等. 基于 YOLOv5s 的轻量化架空输电线路鸟巢检测网络[J]. 电子测量技术,2024,47 (7):138-148.
XU Y D, CAI Y H, LI Y, et al. Lightweight overhead transmission line bird's nest detection network based on YOLOv5s [J]. Electronic Measurement Technology, 2024,47(7):138-148.
- [18] HU J, SHEN L, SUN G. Squeeze-and-excitation networks[C]. IEEE Conference on Computer Vision and Pattern Recognition, 2018: 7132-7141.
- [19] ZHU X ZH, CHENG D ZH, ZHANG ZH, et al. An empirical study of spatial attention mechanisms in deep networks[C]. IEEE/CVF International Conference on

Computer Vision, 2019: 6688-6697.

[20] HAN K, WANG Y H, CHEN H T, et al. A survey on vision transformer[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2022, 45(1): 87-110.

[21] WOO S, PARK J, LEE J Y, et al. CBAM: Convolutional block attention module[C]. European Conference on Computer Vision, 2018: 3-19.

[22] 黄毅,周纯,刘欣军,等. 基于 YOLOv10 的无人机复杂背景下多尺度检测模型[J/OL]. 光通信研究, 1-8 [2024-12-26]. <http://kns.cnki.net/kcms/detail/42.1266.TN.20240822.1307.002.html>.
HUANG Y, ZHOU CH, LIU X J, et al. Multiscale detection model for complex backgrounds in UAV images based on YOLOv10 [J/OL]. Study on Optical Communications, 1-8[2024-12-26]. <http://kns.cnki.net/kcms/detail/42.1266.TN.20240822.1307.002.html>.

[23] 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.

[24] CHEN J R, KAO S H, 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.

[25] CAI X H, LAI Q X, WANG Y W, et al. Poly kernel inception network for remote sensing detection[C]. IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2024: 27706-27716.

[26] SELVARAJU R R, COGSWELL M, DAS A, et al. Grad-CAM: Visual explanations from deep networks via gradient-based localization[C]. IEEE International Conference on Computer Vision, 2017: 618-626.

作者简介

李坤祥(通信作者), 硕士研究生, 主要研究方向为计算机视觉、目标检测。
E-mail:likunxiang@mail.shiep.edu.cn

刘大明, 副教授, 硕士生导师, 主要研究方向为人工智能、嵌入式系统。
E-mail:ldm@shiep.edu.cn