

DOI:10.19651/j.cnki.emt.2519483

基于改进 YOLOv11 光伏热斑小目标缺陷检测

牟涛¹ 王代强²

(1. 贵州大学大数据与信息工程学院 贵阳 550025; 2. 贵州民族大学物理与机电工程学院 贵阳 550025)

摘要: 针对光伏红外图像中缺陷区域存在目标尺寸小、边缘模糊、易受噪声与背景干扰等问题,提出一种基于 YOLOv11 的改进算法。首先设计一种引导式局部-全局空间注意力 GLGSA 模块,用于有效融合局部显著区域信息与全局上下文语义,提升特征表达的判别性。其次,将 GLGSA 模块与双向特征融合结构 BiFPN 结合,构建出 GLGSA-BiFPN 结构提升多尺度特征融合的效果。新增 P2 检测层以增强对极小目标的检测能力。最后引入 NWD 损失函数替换原损失函数,增强对小目标的定位精度。在 PV-HSD-2025 光伏热斑数据集上进行实验验证,结果表明改进算法的检测精度 mAP_{50} 和 mAP_{50-95} 相比于 YOLOv11n 分别提高 9.1% 和 5.6%。有效提高光伏小目标缺陷检测的精度。

关键词: YOLOv11; 热斑缺陷; 小目标检测; 光伏板

中图分类号: TN219 **文献标识码:** A **国家标准学科分类代码:** 510.4030

Defect detection of small targets based on the improved YOLOv11 photovoltaic hot spot

Mou Tao¹ Wang Daiqiang²

(1. College of Big Data and Information Engineering, Guizhou University, Guiyang 550025, China;

2. School of Physics and Mechatronic Engineering, Guizhou Minzu University, Guiyang 550025, China)

Abstract: Aiming at the problems of small target size, fuzzy edges, and vulnerability to noise and background interference in defect areas of photovoltaic infrared images, an improved algorithm based on YOLOv11 was proposed. Firstly, a guided local-global spatial attention (GLGSA) module is designed to effectively integrate Local salient region information and Global context semantics to improve the discrimination of feature representation. Secondly, the GLGSA module was combined with the bidirectional feature fusion structure BiFPN to construct the GLGSA-BiFPN structure to improve the effect of multi-scale feature fusion. The P2 detection layer was added to enhance the detection ability of minimal targets. Finally, the NWD loss function is introduced to replace the original loss function to enhance the positioning accuracy of small targets. Experimental verification is carried out on the PV-HSD-2025 photovoltaic hot spot data set. The results show that the detection accuracy of the improved algorithm mAP_{50} and mAP_{50-95} is 9.1% and 5.6% higher than that of YOLOv11n. Effectively improve the accuracy of photovoltaic small target defect detection.

Keywords: YOLOv11; hot spot defects; small target detection; photovoltaic panel

0 引言

在全球能源结构转型和碳中和战略目标推动下,近年来,太阳能因其成本低、可再生、无污染等特性,受到人们的广泛青睐^[1]。然而在制造、运输和使用全周期中,电池板易产生隐裂、断栅等缺陷,导致光电转换效率衰减与使用寿命缩短,引发安全风险^[2-3]。其中热斑故障最为突出,轻则使光伏板发电效率大大降低,重则损坏整片光伏板甚至存在火灾隐患^[4]。因此对光伏组件进行快速、精准、高效的缺陷

检测已成为提高光伏电站运维效率的关键技术之一。

传统的太阳能电池板缺陷检测方法依赖人工检查,检测效率低、准确性差^[5]。已无法满足光伏产业对智能化运维需求。随着近年来深度学习技术的快速发展,以卷积神经网络为代表的智能检测算法逐渐成为主流,并成功应用于各种缺陷自动检测领域。当前,基于深度学习的目标检测方法主要可分为 3 大类:两阶段检测算法和一阶段检测算法以及基于 Transformer 的检测算法。在光伏热斑小目标检测任务中,两阶段检测方法如区域神经网络(region-

based convolution neural network, R-CNN)^[6]系列,它通常先生成一系列候选区域,然后再对这些区域进行分类和回归,这类方法的优点是检测精度较高,尤其在复杂背景下具有较好的定位稳定性。但是由于其网络结构的复杂性,推理速度相对较慢,难以满足无人机光伏热斑检测实时性的需求^[7]。基于 Transformer 检测模型如 DETR (detection transformer, DETR)通过自注意力机制建模全局上下文关系,具备端到端无需非极大值抑制的优势,可避免相邻热斑被误删。对微小局部光伏热斑缺陷的识别更具优势。然而,其缺点是计算开销大(标准 DETR 参数量超过 40 M)、训练收敛慢,显存消耗高,在光伏巡检的边缘设备上部署困难。相比之下,一阶段检测方法如单次多边框检测(single shot multibox detector, SSD)^[8]算法和 YOLO(you only look once, YOLO)^[9]系列在保持与前两者相近的检测精度前提下显著提升了推理速度,因而更适用于工业场景下实时检测的需求^[10]。特别是 YOLO 系列凭借其在检测精度与推理效率之间取得了良好平衡,为光伏缺陷检测提供了新的解决方案。近年来,针对光伏小目标缺陷检测任务,学者们对 YOLO 系列模型开展了多种改进探索。孔松涛等^[11]提出基于 YOLOv5 的光伏板红外热斑检测方法,在 Neck 层引入卷积块注意力,增强特征融合过程中的语义提取能力,并使用完全交并比损失函数(complete intersection over union, CIoU)以更好地区分长条形与小型热斑目标,提高了检测精度,但在处理多尺度小目标时仍表现不佳。彭自然等^[12]针对无人机航拍红外图像的微弱光伏热斑检测场景提出了 YOLOv5s-HWHA 改进算法。通过在主干网络后端引入层级聚合注入网络,以增强特征表达能力,并结合 WIoU 损失函数改善了边界框回归质量,提升了对困难样本的检测能力,但局部特征模块仅用两个 7×7 卷积,固定感受野,难以适应红外图像中多尺度、微弱的热斑目标,存在局部信息利用不足的问题。李冰等^[13]针对光伏电站航拍光伏红外图像复杂背景下小目标难检测的问题,提出一种基于 YOLOv7-EPAN 的红外图像缺陷检测方法,引入高效路径特征聚合网络(efficient layer aggregation network, ELAN)和归一化高斯 Wasserstein 距离损失函数(normalized gaussian wassertein distance, NWD),有效缓解了小目标匹配困难与尺度失衡问题,但其参数量大接近 36 M,在边缘设备上部署不理想。高航等^[14]针对 YOLOv8n 太阳能电池板检测微小细微缺陷方面表现不佳的问题,借鉴 ASF-YOLO 中的思想,和引入 Slide Loss 损失函数改进算法,提高模型对易混淆样本的识别能力和检测精度。但在轻量化与检测精度之间仍存在平衡挑战。刘玉萍等^[15]针对无人机视角下的小目标检测场景提出改进 YOLOv11 的小目标检测算法。通过引入双卷积结构(dual convolution, DualConv)^[16]构建 ELAN-DC 模块,提升了浅层特征提取能力,同时设计局部细节和全局特征的双向特征融合结构,并引入动态检测头。该方法显著提升了小目

标检测精度,但在处理目标密集且背景复杂的场景中仍存在一定误检。

尽管现有方法在光伏小目标缺陷检测领域进行了多方面的优化,但无人机或手持红外设备拍摄的光伏红外图像往往存在缺陷目标尺寸较小、对比度低、背景复杂且目标密集分布的特点^[17],检测过程中仍然存在显著的漏检和误检问题。针对上述问题,本文在 YOLOv11 基础上提出了一种新的光伏小目标缺陷检测改进方法。研究主要贡献为:

1) 设计引导式局部-全局空间注意力模块(guided local-global spatial attention, GLGSA),融合引导注意力机制、全局上下文建模与局部细节建模。提升特征表达的判别性,增强网络对小目标关键区域的感知能力。

2) 构建了改进的多尺度特征融合结构 GLGSA-BiFPN,结合双向特征金字塔网络(bi-directional feature pyramid network, BiFPN)^[18]的高效融合策略与 GLGSA 的特征增强能力,实现对不同尺度目标的高效建模,提升小目标检测的鲁棒性与准确性。

3) 增加 P2 检测头,提升对浅层细粒度特征的利用能力和对极小目标的检测性能,特别适用于密集小目标场景。

4) 引入加 NWD 损失函数,进一步提高小目标定位精度,提高了检测框的回归质量。

1 YOLOv11 算法概述

YOLOv11 是 Ultralytics 公司在 2024 年推出的一种新型目标检测算法,网络结构主要由主干网络、颈部网络以及检测头这 3 部分组成,主干网络承担着图像特征提取的工作,是由卷积层、C3k2 模块、和跨通道金字塔显著注意力模块(cross stage partial with pyramid squeeze attention, C2PSA)及快速空间金字塔池化模块(spatial pyramid pooling fast, SPPF)所构成。其中该模型采用 C3k2 模块替代了原先的 C2f 模块,显著提升了模块的适应性,SPPF 模块用于整合不同尺度的上下文信息,把任意大小的特征图压缩成固定维度特征,提升了模型对多尺度目标的适应能力,YOLOv11 在主干后部引入了 C2PSA 模块,该模块融合了轻量化的跨阶段结构和多头自注意力机制,在保持计算效率的情况下提高了模型对关键区域和目标语义的感知能力。颈部网络采用路径聚合网络(path aggregation network, PANet)^[19]结构作为特征融合路径,它结合了自顶向下的特征金字塔网络和自底向上的路径聚合机制,让深层语义信息与浅层空间信息充分融合,为小目标检测提供更丰富的上下文支持。检测头部分采用解耦式设计,将分类与回归任务分开,可提升模型的泛化能力与收敛速度。YOLOv11 支持多尺度输出,有良好的适应性和扩展性,适合部署在多种工业与边缘计算场景中。

2 改进 YOLOv11 网络模型

针对光伏红外图像中存在的目标尺寸小、缺陷边界模

糊以及复杂纹理背景干扰等因素所导致的漏检与误检问题,本文提出一种基于 YOLOv11 的改进检测模型,其结构图如图 1 所示。首先设计一种引导式局部-全局空间注意力机制模块 GLGSA,用于对局部和全局信息进行统一引导,提升整体特征表达的判别性与对关键区域的聚焦能力,提高对小尺度缺陷的感知能力,接着运用 GLGSA 改进模

块来改进 BiFPN 结构,构建特征融合结构 GLGSA-BiFPN,以此优化 BiFPN 里多尺度信息的交互方式,达成语义与空间特征的协同表达。再引入 P2 检测头,来保留高分辨率特征图中的细节信息,提升微小缺陷目标的检测精度,并且引用 NWD 损失函数,提高模型在目标尺度变化时的稳定性。下面对改进网络中的每一部分进行详细说明。

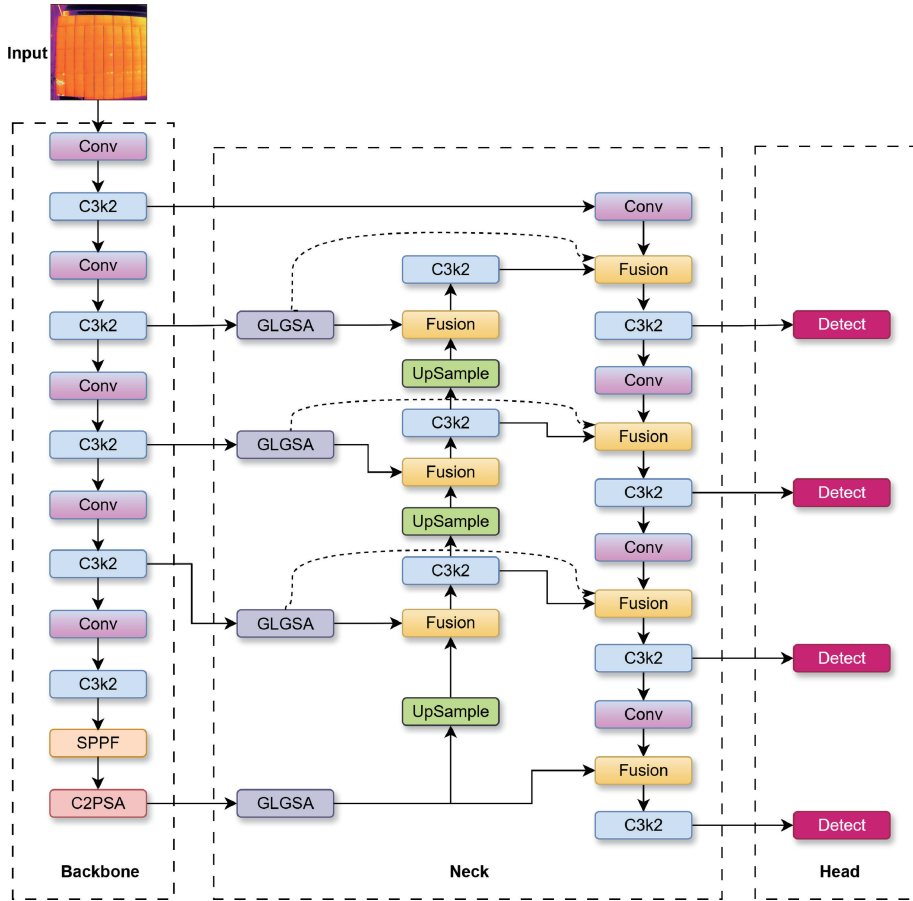


图 1 改进后的模型结构
Fig.1 Structure of improved model

2.1 Neck 结构优化

1) GLGSA 模块的设计

在小目标检测任务中,尤其是如无人机视角下的光伏板热斑缺陷检测这类红外图像场景下,目标往往具有尺寸小、纹理弱、边界模糊、背景干扰强等特征,直接导致模型容易出现漏检和误检。为了强化特征表征能力并提升对小目标的感知能力,受全局局部空间注意力(global-local spatial attention, GLSA)模块的启发,其结构如图 2 所示,设计出 GLGSA 引导式局部-全局空间注意力机制模块,其结构如图 3 所示。

GLGSA 模块主要包括局部分支、全局分支及引导注意力机制。首先将输入特征按通道均分为 F_1^l, F_1^g , 一部分送入局部分支进行轻量卷积并融合通道注意力(CA)与空间注意力(SA)以增强局部目标感知;另一部分送入全局分

支通过建模长距离依赖关系,补充全局上下文信息。同时,设计引导注意力模块对输入特征生成统一的引导权重,用于调控局部与全局特征的响应强度。最终将局部与全局输出拼接和融合,获得更具判别性的多尺度融合特征,从而有效提升小目标检测性能。下面详细介绍 GLGSA 模块的各个组成部分。

引导注意力模块设计:原始的 GLSA 的全局注意力和局部注意力模块融合方法缺乏主动引导网络聚焦于目标区域的机制,易受背景干扰,造成漏检和误检。为此设计了引导机制(guided attention, GA)。主动引导局部与全局特征更关注关键的目标区域。其结构如图 4 所示。引导机制模块设计的初衷是生成一个高效且紧凑的全局通道级注意力图,以明确引导后续的局部和全局分支更加有效地关注关键目标区域,提升其检测精度。

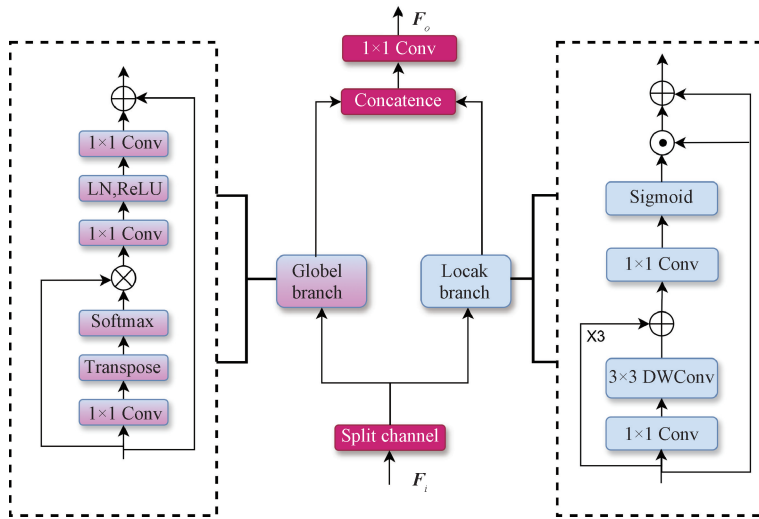


图 2 GLSA 结构

Fig. 2 Structure of GLSA

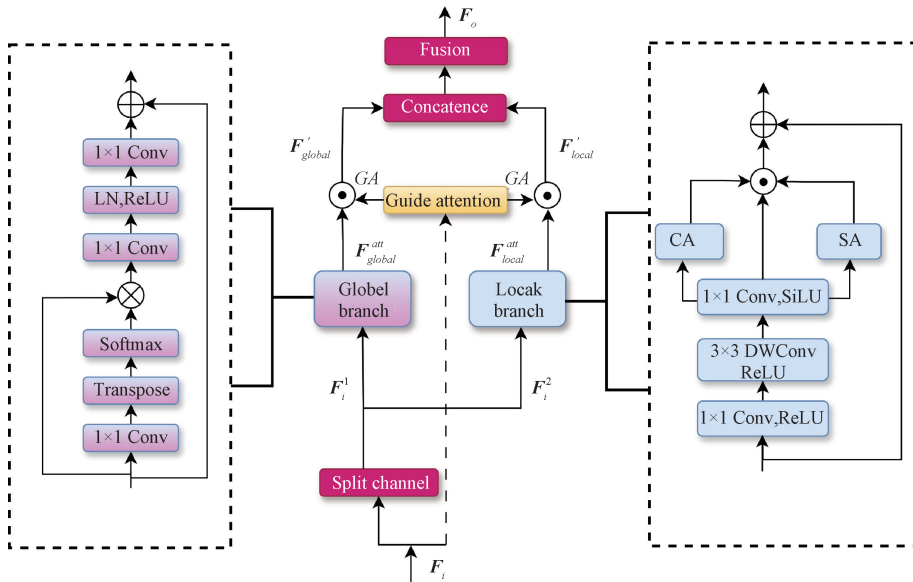


图 3 GLGSA 结构

Fig. 3 Structure of GLGSA

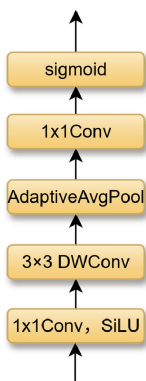


图 4 引导注意力结构

Fig. 4 Structure of guided attention

具体流程为：首先，通过 1×1 卷积降维，将输入特征从较大的通道降至更少的通道数，有效捕捉输入特征中最具代表性的关键信息。接着，通过深度卷积捕捉局部空间上下文关系，并使用自适应全局平均池化捕获全局信息，生成全局空间压缩表示。最后，利用 Sigmoid 函数得到一个介于 $[0, 1]$ 的注意力权重，该权重作用于后续局部和全局特征上，引导特征关注关键区域。其中 DWConv 的卷积核尺寸是 3×3 ，通道压缩率为 6.25%。具体的计算公式如式(1)所示。

$$GA = \sigma(\text{Conv}(\text{GAP}(\text{DWConv}(\text{SiLU}(\text{Conv}(x)))))) \quad (1)$$

其中， σ 为 Sigmoid 函数，Conv 是 1×1 的卷积，GAP

为全局自适应平均池化。 \mathbf{GA} 为引导注意力输出, DWConv 为深度卷积。

局部注意力分支设计:局部卷积分支的目的在于充分挖掘目标区域的局部细节特征,针对原始 GLSA 模块里的局部分支本文进行了轻量化重构工作,改进后的分支运用轻量化瓶颈结构,同时融合通道注意力以及空间注意力来提高关键区域的响应能力,与原始模块仅仅把注意力用作残差掩码的设计不一样,改进结构将注意力直接施加于主干路径特征上,切实提升了对边缘纹理和目标区域的较大性建模能力,特别适用于小目标检测场景。具体过程为:

使用 1×1 卷积来压缩通道,以此减轻计算复杂度,运用深度可分离卷积提高局部空间特征感知能力,然后凭借通道注意力和空间注意力突出小目标区域特征,抑制背景噪声。局部注意力分支计算公式如式(2)~(3)所示。

$$\mathbf{F}_{local} = \text{SiLU}(\text{Conv}(\text{DWConv}(\text{ReLU}(\text{Conv}(\mathbf{F}_i^1)))))) \quad (2)$$

$$\mathbf{F}'_{local} = \mathbf{F}_{local} * \mathbf{CA}(\mathbf{F}_{local}) * \mathbf{SA}(\mathbf{F}_{local}) + \mathbf{F}'_i \quad (3)$$

其中, \mathbf{F}'_{local} 为局部注意力分支输出, \mathbf{F}_{local} 为局部注意力中间特征, \mathbf{F}_i^1 为输入特征映射 \mathbf{F}_i 的通道均分的其第 1 组。

全局注意力分支设计:全局分支通过对输入特征进行上下文建模,利用注意力机制捕捉长距离依赖关系和全局语义信息,从而提升特征的判别能力,补充了局部空间注意力的不足。其计算公式如式(4)~(5)所示。

$$\mathbf{F}_{global} = \text{Softmax}(\text{Transpose}(\text{Conv}_{1 \times 1}(\mathbf{F}_i^2))) \quad (4)$$

$$\mathbf{F}'_{global} = \text{MLP}(\mathbf{F}_{global} \otimes \mathbf{F}_i^2) + \mathbf{F}_i^2 \quad (5)$$

其中, \mathbf{F}'_{global} 局部注意力分支输出, \mathbf{F}_{global} 为表示全局上下文注意力特征, \otimes 表示矩阵乘法。MLP(\cdot) 由两个全连接层和一个 ReLU 非线性和规范化层组成。MLP 的第 1 层将其输入转换为一个扩展比为 2 的高维空间,第 2 层将其恢复为与输入相同的维数。

特征融合设计:局部和全局分支分别经过引导注意力引导后,形成两路特征,这两路特征通过拼接和融合生成最终的增强特征图。其计算公式为:

$$\begin{cases} \mathbf{F}'_{local} = \mathbf{F}'_{local} * \mathbf{GA} \\ \mathbf{F}'_{global} = \mathbf{F}'_{global} * \mathbf{GA} \end{cases} \quad (6)$$

$$\mathbf{F}_o = \text{SiLU}(\text{BN}(\text{Conv}_{1 \times 1}(\text{ConCat}(\mathbf{F}'_{local}, \mathbf{F}'_{global})))) \quad (7)$$

其中, \mathbf{GA} 为引导注意力, \mathbf{F}'_{local} 为局部分支经过引导后的特征表示, \mathbf{F}'_{global} 为全局分支经过引导后的特征表示。 \mathbf{F}_o 融合后的最终输出特征,可用于下游模块输入。

2)GLGSA-BiFPN 结构设计

在目标检测任务中,多尺度特征融合结构对提升模型感知不同尺寸目标的能力起着关键作用, YOLOv11 所采用的 PANet 融合策略,虽实现了浅层与深层特征的连通,可融合过程主要依靠固定路径以及简单相加操作,缺少对特征关键性以及层级信息的自适应建模,在处理复杂背景

下的小目标检测时,会出现信息表达受到限制、细节特征容易丢失等状况。

为了处理上述问题,本文引入了 BiFPN 为特征融合的核心结构, BiFPN 依靠自上而下和自下而上的双向路径,有效地达成了多层特征的高效传递与整合,其可学习的加权融合机制可自动评估各尺度特征的贡献程度,优化特征选择与融合质量,该机制在提升中等尺寸与大目标检测性能方面表现出色,不过在小目标检测中依旧存在局限,因为小目标自身在高层语义特征图中已经被严重压缩,直接参与加权融合容易致使信息被淹没。

为此本文提出 GLGSA-BiFPN 结构,其结构和 GLGSA 模块插入位置如图 5 所示。其中 GLGSA 模块插入于 BiFPN 的每一融合节点之前,输入为主干网络输出的 P2、P3、P4、P5 特征图。该模块通过局部-全局双分支结构建模空间注意力,局部分支提取纹理边缘等细粒度特征,全局分支捕捉上下文语义,判断潜在的目标分布。并由引导注意力输出一张“目标优先权重图”。该图在空间维度突出可能包含目标的区域。在进入 BiFPN 的加权求和前,该权重图作用于输入特征,能主动增强小目标所在位置的响应强度,提升其在融合过程中的相对权重,从而减少在后续融合和下采样中的信息被稀释,有效缓解了小目标特征在高层被稀释的问题。通过这一先引导再融合的机制 GLGSA-BiFPN 结构在保留 BiFPN 自适应通道加权优势的同时,增强了网络对小目标局部结构与上下文语义的联合建模能力,提高了对关键区域的特征聚焦能力,又有效地避免了直接加权带来的特征冲突问题。

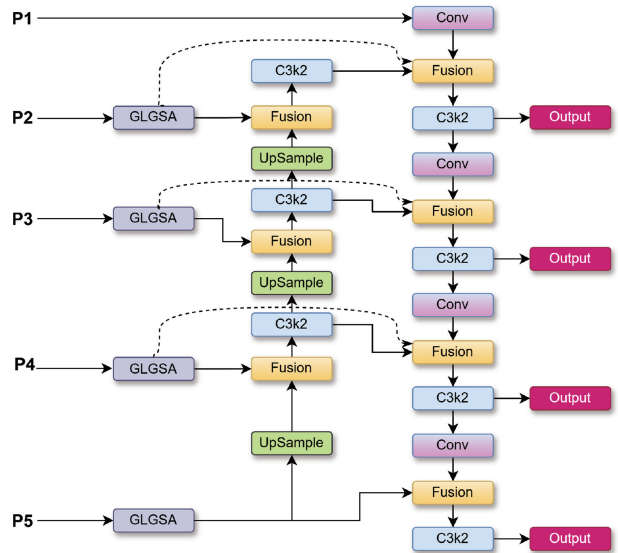


图 5 GLGSA-BiFPN 结构

Fig. 5 Structure of GLGSA-BiFPN

2.2 增加小目标检测层

YOLOv11 处理经过 n 次连续下采样操作生成的特征图的检测头被定义为 P_n 检测头。在基于 YOLOv11 的目

标检测框架中,有 P3、P4、P5 检测头。分别对应检测小中大目标。然而对于无人机视角下的光伏热斑小目标缺陷检测来说,即使下采样次数最少的 P3 检测头,其输入特征图尺寸也较原始图像缩小至 1/8。目标尺寸缩小后仅为几个像素或不足一像素,其显著特征极易被压缩抹除,严重影响检测精度。为此,本文在 YOLOv11 基础结构上增加了 P2 检测头,对应原始图像的 4 倍下采样。该层能够直接利用浅层的细粒度特征信息,从而显著提升模型对小目标的感知与区分能力。此外本文设计的 GLGSA-BiFPN 特征融合结构,也为 P2 层提供更加判别性强、信息量丰富的细粒度特征。

2.3 改进损失函数

为进一步提高检测精度,引入 NWD 损失函数替代原有 CIoU 损失函数。原始的 CIoU 对于不同尺度物体的敏感度差异较大,尤其对于微小目标,微小位置的偏差会导致 CIoU 明显下降,从而不能准确的进行预测。相较于 CIoU 损失,NWD 对于目标的不同尺度并不敏感,因此更适合测量小物体之间的相似性。

NWD 是一种使用归一化 Wasserstein 距离来衡量高斯分布相似性的算法。该算法首先将边界框建模为二维高斯分布,并提出归一化 Wasserstein 距离,通过相应高斯分布计算其之间的相似性。对于微小目标,大部分真实物体都不是严格意义上的矩形,其边界框内往往存在背景要素,前景与背景要素分别集中分布于边界框的中心与边界。将边界框建模为高斯分布,对其中心赋予最高权重,像素重要性从中心向边界递减。因此,两个边框之间的相似度可以转换为两个高斯分布之间的分布距离。对从边界框 $A = (x_a, c y_a, w_a, h_a)$ 和 $B = (x_b, c y_b, w_b, h_b)$ 建模的高斯分布 N_a 和 N_b , 该方法的计算公式如式(8)~(10)所示。

$$W_2^2(N_a, N_b) = \left\| \left(\left[c x_a, c y_a, \frac{w_a}{2}, \frac{h_a}{2} \right]^T, \left[c x_b, c y_b, \frac{w_b}{2}, \frac{h_b}{2} \right]^T \right) \right\|_2^2 \quad (8)$$

$$\text{NWD}(N_a, N_b) = \exp\left(-\frac{\sqrt{W_2^2(N_a, N_b)}}{C}\right) \quad (9)$$

$$\text{Loss}_{\text{NWD}} = 1 - \text{NWD}(N_a, N_b) \quad (10)$$

其中, c 是与数据集密切相关的常数; $W_2^2(N_a, N_b)$ 表示距离度量。

3 实验及结果分析

3.1 实验环境与参数配置

为了确保本次实验数据的可靠性与可重复性,所有的消融实验和对比实验均在相同的计算环境和相同的训练设置中完成。本次实验使用的计算机操作系统 Windows11, 配备 NVIDIA GeForce RTX 3080 Ti GPU 及 Intel Core i7-10700k 处理器,主频为 2.90 GHz。实验环境方面,采用了 CUDA 11.8 版本,Python 3.11.5,深度学习

框架使用了 Pytorch2.0.0。输入图像尺寸统一为 640×640 ,模型训练的初始学习率设定为 0.01,优化器为 adamw, batchsize 为 64,训练周期为 200 轮。

3.2 数据集

本文使用最近开源的光伏热斑数据集 PV-HSD-2025^[20]进行模型的训练、验证以及测试。其光伏热斑缺陷如图 6 所示。该数据集包含 4 025 张带有缺陷的图像,将其按照 7:2:1 的比例随机划分为训练集、验证集和测试集。共计 25 181 个缺陷标注框,涵盖热斑单一类别的缺陷。该数据集中绝大多数标注框的面积平方根均小于 20 pixel,标注框的面积平方根小于 20 pixel 占全部标注框比的 92%。数据集标注情况如图 7 所示。

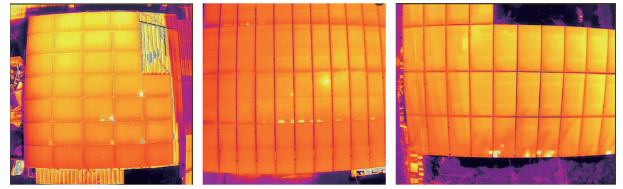


图 6 光伏板热斑缺陷图

Fig. 6 Diagram of hot spot defects on photovoltaic panels

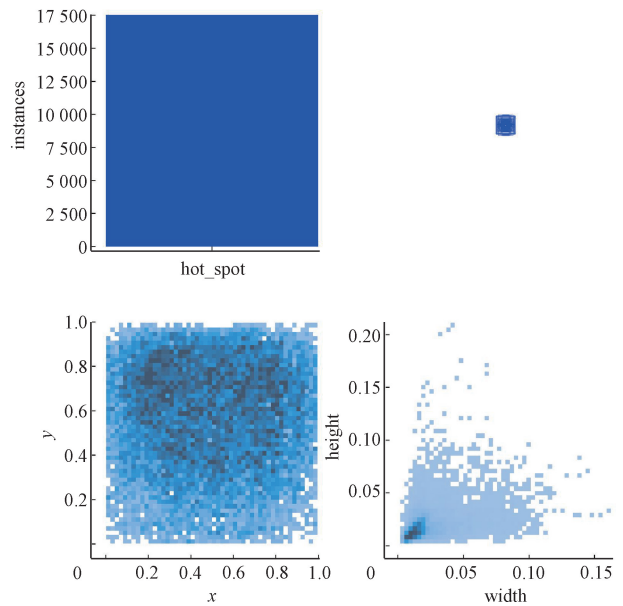


图 7 标签分布

Fig. 7 Label distribution

3.3 算法评价指标

采用精确率 (precision, P)、召回率 (recall, R) 平均精度均值 (mean average precision, mAP)、参数量 (Params) 和浮点运算数 (GFLOPs) 5 项指标评估模型性能。 $mAP_{0.5}$ 表示当 IoU 阈值为 0.5 时的下的平均精度值。 $mAP_{0.5-0.95}$ 表示当 IoU 阈值从 0.5 到 0.95, 每隔 0.05 计算一次该阈值下的 mAP 值最后取平均。参数量用于衡量模型的规模和复杂度,通过累加各层权重参数的数量计算得出。

GFLOPs 用于评估模型的计算复杂度和运行效率。相应计算公式为:

$$P = \frac{TP}{TP + FP} \quad (11)$$

$$R = \frac{TP}{TP + FN} \quad (12)$$

$$AP = \int_0^1 P dR \quad (13)$$

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

3.4 消融实验

为了验证每个改进策略的有效性,以 YOLOv11n 为基础进行消融实验,实验结果如表 1 所示。实验表明各改进模块均显著提高光伏缺陷检测性能。其中将颈部结构

替换为 GLGSA-BiFPN 结构后,模型的 mAP_{50} 和 mAP_{50-95} 分别提升了 1.9% 和 1.6%。实验表明, GLGSA-BiFPN 结构中的引导注意力机制和多尺度特征融合策略有效增强了局部与全局语义信息的建模能力,缓解了原有 PANet 结构中特征融合方式单一、跨层信息交互不足等问题,从而提升了多尺度目标的检测性能,尤其对小目标更加敏感。增加 P2 检测头后,模型的 mAP_{50} 显著增加了 6.3%, mAP_{50-95} 增加了 2%。实验表明增加 p2 检测头增强了网络对浅层细节特征的利用能力,提升了对微小尺寸目标检测能力。引入 NWD 损失函数后模型的 mAP_{50} 增加了 0.5%, mAP_{50-95} 增加了 0.2%。这说明 NWD 能有效缓解小目标检测任务中正负样本匹配不均和尺度失衡问题,提升了定位精度与模型在复杂样本下的稳定性,进一步增强了整体检测性能。

表 1 消融实验

Table 1 Ablation experiments

Baseline	GLGSA-BiFPN	P2	NWD LOSS	P/%	R/%	mAP_{50} /%	mAP_{50-95} /%	Param/M	GFLOPs
✓				68.6	63.2	67.0	27.6	2.5	6.3
✓	✓			67.7	66.6	68.9	29.2	2.0	6.7
✓		✓		70.6	69.1	73.3	31.2	2.6	10.3
✓			✓	68.5	64.2	67.5	27.8	2.5	6.3
✓	✓	✓		74.1	70.4	75.8	32.8	2.2	15.6
✓		✓	✓	72.4	69.5	74.4	31.5	2.6	10.3
✓	✓		✓	69.5	65.7	69.4	29.1	2.0	6.7
✓	✓	✓	✓	74.1	71.6	76.1	33.2	2.2	15.6

最终,结合 GLGSA-BiFPN 和 P2 检测层以及 NWD 损失函数后个模块后,改进模型的 mAP_{50} 和 mAP_{50-95} 分别较基线模型提升了 9.1% 和 5.6%。尽管改进模型的计算量有所增加,但其参数量仅 YOLOv11s 的 24%。但其参数量仅 YOLOv11s 的 24%。结果来看,提出的改进策略有效提升了模型对光伏小目标缺陷的检测性能,验证了本文所提方法的有效性。

3.5 对比实验

1) 引导注意力对比实验

为验证设计的 GLGSA 模块中引导注意力机制的必要性和先进性。选取 YOLO11 加 GLGSA-BiFPN 结构为实验模型,分别比较 GLGSA 模块加引导注意力和不加引导注意力的对比验证实验。以及将引导注意力替换为经典注意力机制 SE、CBAM、SimAM 注意力进行对比实验。实验结果如表 2 所示。实验表明加了引导注意力模型的 mAP_{50} 和 mAP_{50-95} 分别提升了 1.0% 和 0.7%。证明了引导注意力能主动引导局部与全局特征更关注关键的目标区域,提高检测精度。相比于现有的其他注意力机制其 mAP_{50} 和 mAP_{50-95} 指标也达到最优,证明了引导注意力机制的有效性和先进性。

表 2 引导注意力对比实验

Table 2 Guided attention contrast experiment

注意力机制	P/%	R/%	mAP_{50} /%	mAP_{50-95} /%
SE	68.9	63.9	67.7	28.7
CBAM	68.5	65.6	68.4	28.4
SimAM	70.5	63.5	68.2	28.6
无引导注意力	68.8	65.3	67.9	28.5
引导注意力	68.7	66.6	68.9	29.2

2) 颈部网络改进对比

为验证提出 GLGSA-BiFPN 结构在小目检测任务中的性能提升效果,在 VisDrone2019 数据集上以及 PV-HSD-2025 光伏热斑数据集。选取 YOLOv11n 为实验模型。分别对比 YOLOv11n 的原始颈部 PANet 网络、替换为 GLSA-BiFPN 结构和本文设计的 GLGSA-BiFPN 结构进行了对比实验评估,具体实验结果如表 3 和 4 所示。

实验表明相较于 YOLOv11 的原始 PANet 颈部网络,在 PV-HSD-2025 公开数据集上,替换为 GLGSA-BiFPN 结构后,模型的 mAP_{50} 和 mAP_{50-95} 分别提升了 1.9% 和 2.6%。且较 GLSA-BiFPN 结构,也提升了 0.3% 和 0.3%。而

表 3 PV-HSD-2025 上不同颈部网络对比实验

Table 3 Comparative experiments of different cervical networks on PV-HSD-2025

颈部网络	P/%	R/%	mAP_{50} /%	mAP_{50-95} /%
PANet	68.6	63.2	67.0	27.6
GLSA-BiFPN	69.0	64.4	68.6	28.9
GLGSA-BiFPN	68.7	66.6	68.9	29.2

表 4 VisDrone2019 上不同颈部网络对比实验

Table 4 Comparative experiments of different cervical networks on VisDrone2019

颈部网络	P/%	R/%	mAP_{50} /%	mAP_{50-95} /%
PANet	43.8	33.3	33.3	19.4
GLSA-BiFPN	45.5	34.1	34.3	20.2
GLGSA-BiFPN	45.4	34.9	34.7	20.4

在 VisDrone2019 公开数据集上, 替换为 GLGSA-BiFPN 结构后, 模型的 mAP_{50} 和 mAP_{50-95} 也分别提升了 1.4% 和 0.9%。且较 GLSA-BiFPN 结构, 也提升了 0.4% 和 0.2%。

对比实验结果表明 GLGSA-BiFPN 结构在多个小目标检测数据集上均展现出优于原始 PANet 和已有 GLSA-BiFPN 结构的性能表现。说明设计的网络结构在特征提取与多尺度融合方面具有显著优势。GLGSA-BiFPN 在融合策略上进一步优化, 有效增强关键区域的响应程度提升检测精度。

3) 损失函数对比实验

为验证 NWD 损失函数的有效性。本文使用 NWD 和几种主流损失函数进行对比实验, 结果如表 5 所示。实验表明引入 NWD 后, 模型在 mAP_{50} 上有显著提高, 从基线 YOLOv11 模型的 67.0 提高到 67.5。召回率达到 64.2%, 明显高于 CIoU 和 MPDIoU 说明模型在保证精度的同时更能捕捉到更多有效目标。此外, NWD 在保持 P 值稳定 68.5% 的前提下, 优化了正负样本匹配策略, 提高了模型对小目标和边界不确定区域的检测能力, 在光伏热斑检测等对目标定位鲁棒性要求较高的场景中具有良好的实用价值。

表 5 损失函数对比实验结果

Table 5 Comparative experimental results of loss function

损失函数	P/%	R/%	mAP_{50} /%	mAP_{50-95} /%
CIoU	68.6	63.2	67.0	27.6
WIoU	68.4	64.0	67.3	27.7
GIoU	67.9	63.1	66.7	27.2
EIoU	67.6	62.9	66.4	27.3
MPDIoU	67.4	63.3	66.2	27.6
NWD	68.5	64.2	67.5	27.8

4) 主流网络对比实验

为系统评估模型的改进效果, 本研究在 PV-HSD-2025 光伏热斑数据集上开展多模型对比实验。所有对比模型均采用相同的训练设置以确保公平性。对比模型包括 YOLO 系列 (YOLOv5s, YOLOv8n, YOLOv11n, YOLOv11s, YOLOv12n) RTDETR-l 以及文献[17] (数据集提出者的方法) 和加 P1 及 P2 检测层的 5 检测头 YOLOv11-P1 的改进方法其结构如图 8 所示。具体实验结果如表 6 所示。

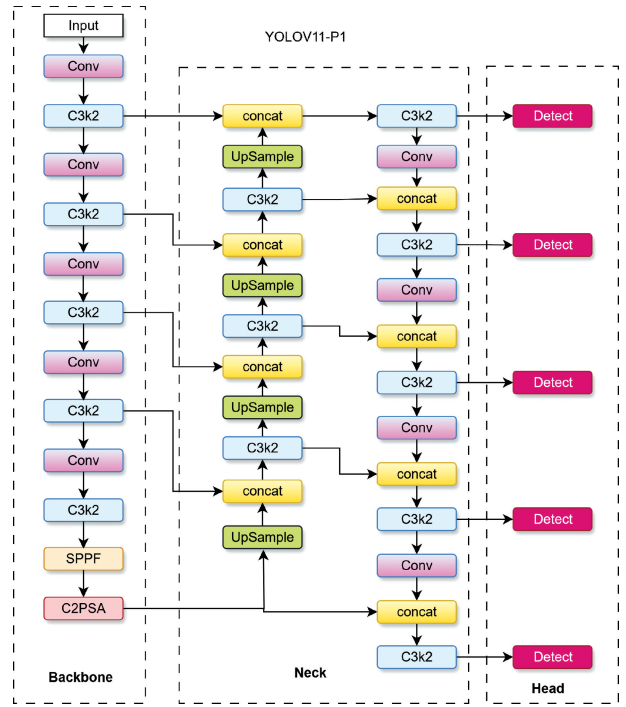


图 8 YOLOV11-P1 结构

Fig. 8 Structure of YOLOV11-P1

从表 6 可以看出, 本文提出的改进算法在多个性能指标上均取得了显著提升。以 YOLOv11n 为基线模型, 本文方法在 mAP_{50} 和 mAP_{50-95} 上分别提升了 5.1% 和 5.6%, 准确率和召回率分别提升了 5.5% 和 8.4%, 表明所提的改进策略有效增强了模型特征提取和检测精度。相比 YOLOv8n, 改进算法在 mAP_{50} 和 mAP_{50-95} 上分别提升 9.3% 和 5.0%, 准确率和召回率分别提升 5.4% 和 7.0%, 进一步验证了改进算法在复杂场景下的小目标检测任务中具有更强优势。同时与文献[18]的方法相比, mAP_{50} 和 mAP_{50-95} 分别提升了 2.3% 和 1.8%, 而参数量和 GFLOPs 降低了 0.7 M 和 7.7 G。表明改进算法使模型在保持轻量化的同时, 显著提升了小目标检测性能。与 YOLOv11s 相比, 改进算法不仅在检测精度有着提升, 其参数量也降低约 6.5 M 计算量减低 5.9 GFLOPs, 展现出更强的部署能力。改进模型与基线模型缺陷检测结果与对比图如图 9 所示, 改进后的模型成功检测到了 YOLOv11n 漏检的热斑缺陷。实验表明本文所提的改进算法, 不仅提升了小目标

表 6 主流模型对比实验
Table 6 Comparative experiments

算法	P/%	R/%	mAP_{50} /%	mAP_{50-95} /%	Param/M	GFLOPs
YOLOv5s	70.4	66.2	69.9	28.1	7.8	19.0
YOLOv8n	68.7	64.6	66.8	28.2	3.0	8.1
YOLOv11n	68.6	63.2	67.0	27.6	2.5	6.3
YOLOv11s	70.9	65.3	69.6	29.6	9.4	21.5
YOLOv12n	70.9	63.2	67.3	28.1	2.6	6.3
RTDETR-l	59.6	57.8	56.2	21.2	32.8	108.0
YOLOv11-P1	70.4	68.7	73.2	31.9	2.6	21.3
文献[17]	71.5	70.6	73.8	31.4	2.9	23.3
本文	74.1	71.6	76.1	33.2	2.2	15.6

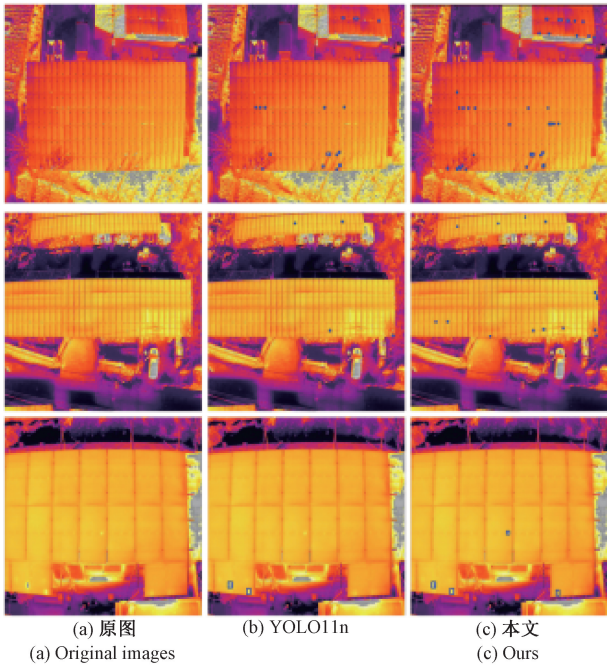


图 9 改进模型检测结果对比

Fig. 9 Comparison of detection results for improved

检测精度,也在模型轻量化与实际应用方面展现出良好的性能。

3.6 泛化性验证实验

为验证所提出模型的小目标检测的泛化能力,分别将基线算法 YOLOv11n 和改进算法在 DOTAv1.5 数据集和 VisDrone2019 数据集上进行实验。从表 7 可看出,改进算法在总体上比 YOLOv11n 提升较大, P 、 R 、 mAP_{50} 、

表 7 DOTAv1.5 效果对比实验

Table 7 Comparative experiment on the effects of DOTAv1.5

方法	P/%	R/%	mAP_{50} /%	mAP_{50-95} /%
YOLOv11n	69.2	32.3	36.2	21.7
本文	74.4	37.1	41.7	25.9

mAP_{50-95} 分别提升了 5.2%、4.8%、5.5%、4.2%。从表 8 可看出其 4 项指标也分别提升了 7.4%、6.7%、7.8%、5.4%,说明改进算法具有一定的泛化性,具备较好的跨数据集推广能力。

表 8 VisDrone2019 效果对比实验

Table 8 Comparative experiment on the effects of VisDrone2019

结构	P/%	R/%	mAP_{50} /%	mAP_{50-95} /%
YOLOv11n	43.8	33.3	33.3	19.4
本文	51.2	40.0	41.1	24.8

4 结 论

针对无人机视角下光伏场景小目标缺陷检测中存在的特征提取能力弱、目标尺寸小及背景干扰严重等问题,基于 YOLOv11 提出了一种改进检测算法。主要改进包括:设计引导注意力 GLGSA 模块,增强局部与全局特征建模能力,构建 GLGSA-BiFPN 结构,提升多尺度特征融合效果,增加 P2 检测头以提升模型微小目标检测能力,引入 NWD 损失函数进一步提升检测精度。在 PV-HSD-2025 光伏热斑缺陷数据集上表明,改进算法的检测精度大幅超越基准模型。在 DOTAv1.5 与 VisDrone2019 数据集上的测试也验证了模型的小目标检测泛化能力。未来将进一步优化模型性能,通过模型剪枝等技术做轻量化部署工作。

参考文献

[1] 陈光耀,陈田,高学海,等.跨层融合的轻量化太阳能电池片缺陷分割方法[J].电子测量技术,2024,47(24):118-127.
CHEN G Y, CHEN T, GAO X H, et al. C2LA-U2-Net: Lightweight defect segmentation method for solar cells with cross-layer fusion [J]. Electronic Measurement Technology, 2024, 47(24):118-127.

[2] 王巍,余欣,缪佳欣,等.基于改进 YOLOv5s 太阳能电池片表面缺陷检测算法[J].电子测量技术,2025,

- 48(5):128-136.
WANG W, YU X, MIAO J X, et al. Solar cells surface defect detection algorithm based on improved YOLOv5s[J]. *Electronic Measurement Technology*, 2025, 48(5):128-136.
- [3] 王海群,武泽锴,于海峰.改进 YOLOv10n 的光伏电池缺陷检测算法[J]. *电子测量技术*,2025,48(15):52-62.
WANG H Q, WU Z K, YU H F. Improved defect detection algorithm for YOLOv10n photovoltaic cells[J]. *Electronic Measurement Technology*, 2025, 48(15): 52-62.
- [4] 管宽岐,蔺雨桐,赵雨薇,等.基于深度学习的航拍光伏板红外图像热斑检测方法研究[J]. *电子测量技术*, 2022,45(22):75-81.
GUAN K Q, LIN Y T, ZHAO Y W, et al. Photovoltaic hot spot detection of aerial infrared image based on deep learning[J]. *Electronic Measurement Technology*, 2022, 45(22): 75-81.
- [5] TANG W Q, YANG Q, DAI ZH, et al. Module defect detection and diagnosis for intelligent maintenance of solar photovoltaic plants: Techniques, systems and perspectives[J]. *Energy*, 2024, 297:20, DOI:10.1016/j.energy.2024.131222.
- [6] GIRSHICK R, DONAHUE J, DARRELL T, et al. Rich feature hierarchies for accurate object detection and semantic segmentation[C]. 2014 IEEE Conference on Computer Vision and Pattern Recognition, Columbus, 2014:580-587.
- [7] 杨丽,杨晨晨,杨耿煌,等.改进 YOLOv8 的光伏电池缺陷检测算法[J]. *电子测量技术*,2025,48(1):92-99.
YANG L, YANG CH CH, YANG G H, et al. Improved photovoltaic cell defect detection for YOLOv8[J]. *Electronic Measurement Technology*, 2025, 48(1):92-99.
- [8] LIU W, ANGUELOV D, ERHAN D, et al. SSD: Single shot multibox detector [C]. *European Conference on Computer Vision*, 2016: 21-37.
- [9] REDMON J, DIVVALA S, GIRSHICK R, et al. You only look once: Unified, real-time object detection[C]. 2016 Conference on Computer Vision and Pattern Recognition, 2016:779-788.
- [10] 杨长春,贺轩轩,王睿,等.基于改进 YOLOv8 光伏板缺陷检测算法[J]. *电子测量技术*,2024, 47(23): 181-192.
YANG CH CH, HE X X, WANG R, et al. Based on the improved YOLOv8 photovoltaic panel defect detection algorithm [J]. *Electronic Measurement Technology*, 2024, 47(23): 181-192.
- [11] 孔松涛,徐甄泽,林星宇,等.基于改进 YOLOv5 算法的光伏组件红外热成像缺陷检测[J]. *红外技术*, 2023, 45(9):974-981.
KONG S T, XU ZH Z, LIN X Y, et al. Infrared thermal imaging defect detection of photovoltaic module based on improved YOLOv5 algorithm [J]. *Infrared Technology*, 2023, 45(9):974-981.
- [12] 彭自然,王思远,张颖清,等.基于 YOLOv5 的无人机航拍红外图像的微弱光伏阵列热斑检测[J]. *太阳能学报*,2025,46(12):315-323.
PENG Z R, WANG S Y, ZHANG Y Q, et al. Weak hot spot detection in photovoltaic arrays using YOLOv5-based UAV aerial infrared images[J]. *Acta Energetica Solaris Sinica*, 2025, 46(12):315-323.
- [13] 李冰,赵宽,白云山,等.基于 YOLOv7-EPAN 的光伏板红外图像缺陷检测[J]. *红外技术*,2024, 46(11): 1315-1324.
LI B, ZHAO K, BAI Y SH, et al. Defect detection of photovoltaic panel infrared image based on YOLOv7-EPAN [J]. *Infrared Technology*, 2024, 46(11): 1315-1324.
- [14] 高航,祁云嵩.基于 YOLOv8n 改进的太阳能电池板缺陷检测[J]. *红外技术*,2025, 47(11):1444-1450.
GAO H, QI Y S. Improved defect detection of solar panels based on YOLOv8n[J]. *Infrared Technology*, 2025, 47(11):1444-1450.
- [15] 刘玉萍,尚翠娟,李明明.改进 YOLOv11 的无人机小目标检测算法[J]. *计算机工程与应用*,2025, 61(15): 124-131.
LIU Y P, SHANG C J, LI M M. Improved YOLOv11 algorithm for small target detection in UAVs[J]. *Computer Engineering and Applications*, 2025: 61(15):124-131.
- [16] ZHONG J CH, CHEN J Y, MIAN A. DualConv: Dual convolutional kernels for lightweight deep neural networks[J]. *IEEE Transactions on Neural Networks and Learning Systems*, 2023, 34(11): 9528-9535.
- [17] SHANG L, YANG Q, WANG J, et al. Small infrared target detection using attention-guided pyramid contextnetwork [J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2022, 60: 1-15.
- [18] TAN M X, PANG R M, LE Q V. Efficientdet: Scalable and efficient object detection[C]. 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2020: 10778-10787.
- [19] LIU SH, QI L, QIN H F, et al. Path aggregation network for instance segmentation[C]. 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2018: 8759-8768.
- [20] 刘沅林,金建辉,李盛洪,等.PV-HSD-2025:光伏热斑检测的红外基准数据集与 YOLOv8-P1 架构[J]. *计算机工程与应用*,2026,62(6):240-252.
LIU R L, JIN J H, LI SH H, et al. PV-HSD-2025: Infrared benchmark dataset for photovoltaic hotspot detection with YOLOv8-P1 architecture[J]. *Computer Engineering and Applications*,2026,62(6):240-252.

作者简介

牟涛,硕士,主要研究方向为目标检测、机器视觉。

E-mail:1870580536@qq.com

王代强(通信作者),教授,主要研究方向为电磁场与微波技术。

E-mail:wang_dai_qiang@163.com