

控制与决策

Control and Decision

面向复杂背景光伏电池红外图像的小目标缺陷检测研究

彭道刚, 邓玉澳, 王丹豪, 潘俊臻

引用本文:

彭道刚, 邓玉澳, 王丹豪, 等. 面向复杂背景光伏电池红外图像的小目标缺陷检测研究[J]. *控制与决策*, 2026, 41(1): 186-200.

在线阅读 View online: <https://doi.org/10.13195/j.kzyjc.2025.0774>

您可能感兴趣的其他文章

Articles you may be interested in

基于低秩矩阵恢复的视觉显著性目标检测与细化

Saliency object detection and refinement based on low rank matrix recovery

控制与决策. 2021, 36(7): 1707-1713 <https://doi.org/10.13195/j.kzyjc.2019.1795>

基于局部信息增量与MPLS的质量相关故障检测方法

Quality-related fault detection method based on local information increment and MPLS

控制与决策. 2021, 36(7): 1647-1654 <https://doi.org/10.13195/j.kzyjc.2019.1402>

基于生成对抗网络学习被遮挡特征的目标检测方法

Object detection via learning occluded features based on generative adversarial networks

控制与决策. 2021, 36(5): 1199-1205 <https://doi.org/10.13195/j.kzyjc.2019.1319>

基于广义主成分分析的重构故障子空间建模方法

Reconstructed fault subspace modelling method based on generalized principal component analysis

控制与决策. 2021, 36(4): 808-814 <https://doi.org/10.13195/j.kzyjc.2019.0818>

基于向量集逆区间滤波的故障观测器设计

Vector set inversion interval filtering based fault observer design

控制与决策. 2021, 36(12): 2973-2981 <https://doi.org/10.13195/j.kzyjc.2020.0212>

面向复杂背景光伏电池红外图像的小目标缺陷检测研究

彭道刚[†], 邓玉澳, 王丹豪, 潘俊臻

(上海电力大学 自动化工程学院, 上海 200090)

摘要: 光伏电池的高效发电对于实现碳达峰与碳中和目标具有重要意义, 针对光伏电池红外图像缺陷中存在背景复杂和目标尺寸较小等问题, 提出一种改进 RT-DETR-R18 的缺陷检测模型——FSC-Net. 首先, 在骨干网络提出部分重参数化快速残差块 (PRFRB), 该模块通过设计部分重参数化卷积, 在实现轻量化的同时能够增强对目标特征的提取能力; 然后, 在颈部网络中构建小目标感知特征金字塔 (SOAFP), 利用 P_2 检测层、空间到深度卷积和浅层特征融合模块, 重构特征图的同时保留细节特征, 强化对小目标特征的表达; 接着, 引入多分支特征融合模块 Conv3XCC3, 通过差异化结构分支进一步提升颈部网络特征融合的效率和多尺度信息的整合能力; 最后, 在损失函数方面, 采用 Inner-WIoUv3 替代传统 GIoU 来提升边界框回归定位的精度. 实验结果表明, FSC-Net 与 RT-DETR-R18 模型相比, 参数量减少了 2.47 M, 模型尺寸压缩了 9.8% 的同时平均检测精度 $mAP@0.5$ 和 $mAP@0.5 : 0.95$ 分别达到 90.0%、75.8%, 相较于基准模型分别提升了 5.7%、4.1%, 实现了对光伏电池红外图像缺陷的高效检测.

关键词: 光伏电池缺陷检测; 目标检测; 部分重参数化快速残差块; 小目标感知特征金字塔; Inner-WIoUv3 损失函数; 空间到深度卷积

中图分类号: TP391 文献标志码: A

DOI: 10.13195/j.kzyjc.2025.0774

引用格式: 彭道刚, 邓玉澳, 王丹豪, 等. 面向复杂背景光伏电池红外图像的小目标缺陷检测研究 [J]. 控制与决策, 2026, 41(1): 186-200.

Research on small-target defect detection for photovoltaic infrared images under complex backgrounds

PENG Dao-gang[†], DENG Yu-ao, WANG Dan-hao, PAN Jun-zhen

(College of Automation Engineering, Shanghai University of Electric Power, Shanghai 200090, China)

Abstract: The high-efficiency power generation of photovoltaic cells is of great significance for achieving the goals of carbon peaking and carbon neutrality. Aiming at the problems of complex backgrounds and small target sizes in infrared image defects of photovoltaic cells, an improved RT-DETR-R18 defect detection model — FSC-Net is proposed. First, in the backbone network, a partial reparameterized faster residual block (PRFRB) is introduced, which realizes lightweight computation while enhancing the ability to extract target features by designing partially reparameterized convolution. In the neck network, a small object aware feature pyramid (SOAFP) is constructed, which employs a P_2 detection layer, space-to-depth convolution (SPDConv), and a shallow feature fusion module to reconstruct feature maps while retaining detailed features, thereby strengthening the expression of small target features. In addition, a multi-branch feature fusion module Conv3XCC3 is introduced, which further improves the efficiency of feature fusion in the neck network and the integration of multi-scale information through differentiated structural branches. Regarding the loss function, Inner-WIoUv3 is adopted to replace the traditional GIoU, in order to improve the accuracy of bounding box regression and localization. Experimental results show that compared with the RT-DETR-R18 model, the FSC-Net reduces the number of parameters by 2.47M, compresses the model size by 9.8%, and achieves average detection precisions of $mAP@0.5$ and $mAP@0.5 : 0.95$ of 90.0% and 75.8%, respectively, which are improvements of 5.7% and 4.1% over the baseline model, thereby realizing efficient detection of infrared image defects in photovoltaic cells.

Keywords: photovoltaic cell defect detection; object detection; PRFRB; SOAFP; Inner-WIoUv3 loss function; SPDConv

收稿日期: 2025-07-22; 录用日期: 2025-10-14.

责任编辑: 王琦.

[†]通信作者. E-mail: pengdaogang@126.com.

0 引言

随着全球能源结构加速向低碳转型,光伏电池凭借其清洁、高效和可再生的优势,正日益成为新型电力系统的关键支撑。在“双碳”目标引导下,光伏产业迅猛发展,大规模部署光伏组件已成常态。然而,由于生产、运输及长期运行中的机械应力、热应力以及环境腐蚀等因素,光伏电池易出现裂纹、断栅、热斑以及电势诱导衰减等缺陷,这会显著降低组件的光电转换效率并缩短其服役寿命,进而影响电站系统的整体稳定。因此,构建一套精度高、鲁棒性强、兼备实时性的缺陷检测体系,对于实现智能制造、保障光伏电站高效运行具有重要意义^[1]。

为应对上述挑战,研究者们主要沿两条路径开展探索。一类研究着眼于增强图像特征建模能力,以提升复杂背景下缺陷的可分辨性。如:Ma等^[2]基于傅里叶单像素成像构建检测方案,在强反射背景下实现了多类缺陷精准提取,但是,仍然依赖于复杂的光学系统,难以满足在线检测需求;Liu等^[3]提出了MNFI-Net网络,融合Ghost模块与多尺度邻域交互机制,有效提升了复杂纹理下的识别性能,但是,对于小尺寸目标缺陷仍然有待提高;Duan等^[4]则从分割任务出发,构建了CAAK-Net模型,首次将检测误差与系统效率关联,为产业化评估提供了依据,但是其模型结构较重,不利于边缘部署;Chen等^[5]、Bu等^[6]和Zhang等^[7]分别从结构设计、热成像融合和无监督学习角度切入,构建了CSPD-DETR、CNN分类器和SSP_CAE模型,在特定场景下取得了优异检测效果,但是,存在模型复杂度高、推理速度慢或训练依赖性强等问题,限制了其实际部署价值。

为进一步平衡检测性能与推理效率,另一类研究聚焦于深度学习^[8]模型的轻量化和鲁棒性。Zhu等^[9]提出了C2DEM-YOLO,引入C2Dense模块和EMA注意力机制,提升了多尺度特征提取和边界回归能力,但是其推理延迟较高;Bao等^[10]通过设计坐标通道感知模块和多尺度定位机制,提出了CCA-YOLO模型,显著增强了缺陷检测的精度和稳定性,但是,在资源受限设备上仍然存在部署瓶颈;Li等^[11]、Pan等^[12]分别提出了YOLO-LitePV和YOLO-ACF模型,前者结合SimAM注意力、通道剪枝和知识蒸馏实现了检测性能和推理速度双提升,后者则通过引入自适应互补融合机制兼顾了模型轻量化与表达能力,两者均适用于嵌入式部署,但是,在极端缺陷识别和多类适应性方面仍然有提升空间;此外,Zhang等^[13]通过融合部分卷积与可切换空洞卷积,优化了YOLOv7以增强其对于异常尺度缺陷的适应

能力;Zhou等^[14]提出了语义信息交互融合网络,引入MLPS和NAF模块,有效提升了跨层语义聚合能力,但是,其对极复杂背景下的目标定位仍然需要进一步优化。

在现有研究取得重要进展的同时,光伏电池缺陷检测仍然面临多重挑战:一方面,缺陷目标往往尺度较小、形态复杂,易被复杂背景信息所干扰,造成漏检或误检;另一方面,在实际应用中,对检测算法提出了更高的精度和实时性要求,尤其是在高分辨率图像下,传统方法在目标定位和分类精度方面难以兼顾计算效率与模型轻量化,这些问题制约了现有方法在实际场景中的应用价值。

在光伏电池红外图像缺陷检测任务中,复杂环境主要是指目标区域受到遮挡或邻近发热源干扰,从而导致识别结果出现偏差。如:输电线路与植被(如树木枝叶)遮挡,电气设备、墙体、道路路面以及路灯(因吸热发热)等因素,均可能在红外图像中引入干扰信号。这类复杂背景不仅增加了光伏电池缺陷特征与环境特征间的相似性,还可能造成误检或漏检,进而降低检测精度。

针对上述问题,本文提出一种改进RT-DETR-R18^[15]的光伏电池红外图像缺陷检测算法——FSC-Net。算法的改进如下:

1) 针对无人机航拍图像背景复杂、易受到环境特征干扰等问题,在骨干网络中提出部分重参数化快速残差块(PRFRB),在保持模型轻量化的同时提升对目标特征的感知能力;

2) 针对光伏电池红外图像缺陷尺寸较小等问题,在颈部网络设计小目标^[16]感知特征金字塔(SOAFP),融合 P_2 检测层、空间到深度卷积(SPDCConv^[17])和浅层特征融合模块(SFFM)来增强小目标细节特征的表达;

3) 为进一步优化颈部网络特征融合效率,引入多分支特征融合模块Conv3XCC3来提升多尺度语义信息的整合能力;

4) 为了更好地优化定位精度,损失函数采用Inner-WIoUv3,以实现更高质量的边界框回归。

FSC-Net能够提升对微小缺陷目标的感知能力,在轻量化的同时能够实现对复杂背景光伏电池红外图像缺陷的有效检测。

1 改进RT-DETR-R18的光伏电池红外图像缺陷检测算法

1.1 RT-DETR-R18网络结构

RT-DETR-R18是一种基于Transformer架构的

实时目标检测模型,其结构主要由骨干网络(Backbone)、高效混合编码器(Efficient Hybrid Encoder)以及检测头和解码器(Decoder)组成。

具体而言,该模型采用 ResNet-18 作为骨干网络,输入图像首先经卷积层和池化层处理,以提取丰富的特征信息,然后通过 BasicBlock 模块实现更深层的语义特征提取以及空间尺度的下采样.每个 BasicBlock 由若干个卷积层、批归一化(BN)层以及 ReLU 激活函数构成,用于增强特征的表达能力。

混合编码器部分由两个关键模块组成:自适应交互特征融合模块(AIFI)和跨尺度交叉特征融合模块(CCFF).前者通过引入动态权重机制和特征交互策略,对骨干网络末3层(S_3 、 S_4 、 S_5)输出的多尺

度信息进行融合,增强模型对目标细节的建模能力;后者则由多个融合单元构成,重点实现不同尺度和层次间特征的整合,充分发挥各层特征的互补优势,以提升模型在多尺度目标检测任务中的鲁棒性和准确性。

此外,RT-DETR-R18 引入了基于 IoU 感知的目标查询初始化机制,在生成查询向量时融入 IoU 相关信息,使得模型更加聚焦于高置信度目标区域,从而提升其整体检测性能.最后,设计可调节的解码器结构,用户可根据具体场景对推理速度与精度进行动态平衡,无需重新训练模型.这种设计提升了 RT-DETR-R18 在实时性要求高的应用场景中的适应能力.RT-DETR-R18 网络结构如图 1 所示。

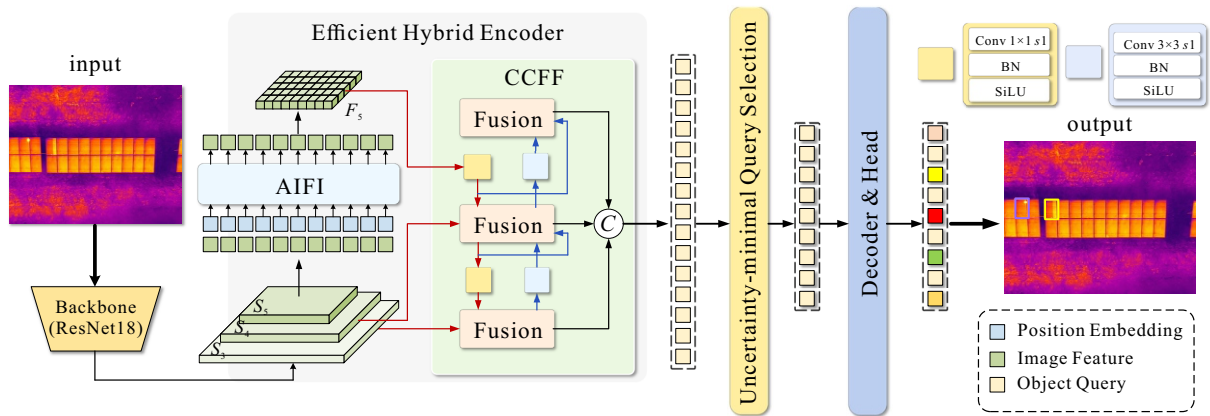


图1 RT-DETR-R18 网络结构

1.2 FSC-Net 模型

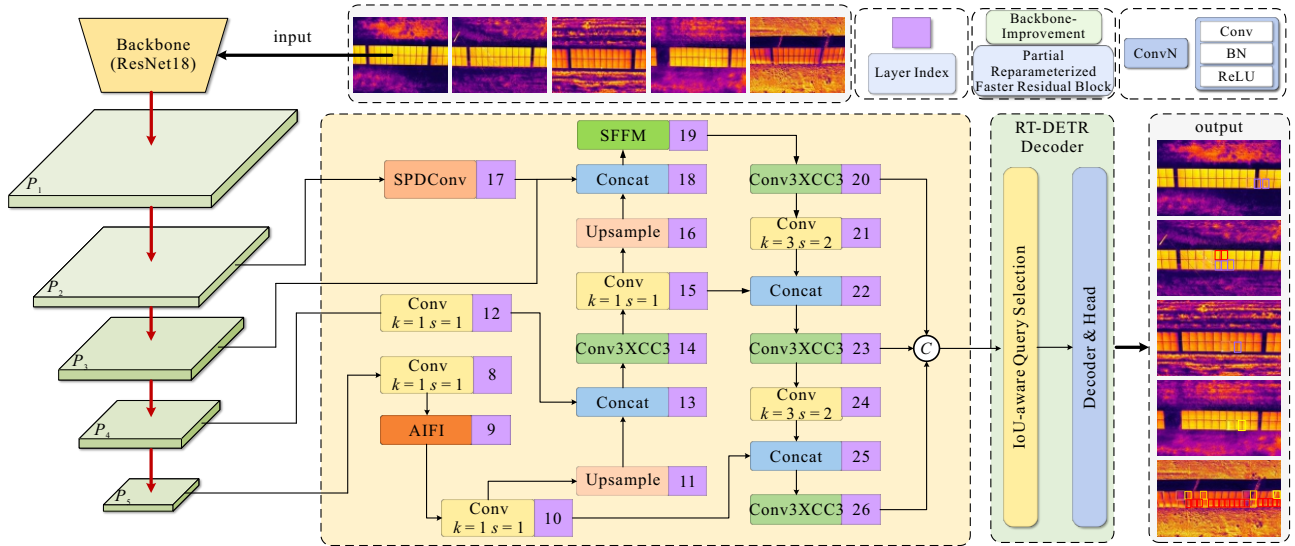
针对光伏电池红外图像中背景干扰强、缺陷目标小等问题,RT-DETR-R18 在该类任务中的检测性能仍有提升空间.本文在兼顾模型轻量化与检测精度平衡的基础上,对 RT-DETR-R18 网络进行多方面优化,旨在提升其在缺陷检测中的精度和效率.具体工作如下:首先,在骨干网络中提取 PRFRB 模块,通过设计部分重参数化卷积(PRConv)来提升模型对目标特征的提取能力,同时降低参数数量和计算量;然后,构建小目标感知特征金字塔,该模块融合骨干网络中 P_2 层的特征表达,并结合 SPDCConv 和 SFM 模块,可有效增强模型对小尺度目标的表达能力,进而增强颈部网络的特征融合效果;此外,在颈部网络中引入 Conv3XCC3 模块,通过将原最右分支中 RepConv 替换为 Conv3XC 卷积,提升特征融合的效率和多尺度语义信息的整合能力;最后,在损失函数方面,采用 Inner-WIoUv3 替代传统的 GIoU (Generalized IoU),可有效提升边界框回归定位的精度,改善整体的检测性能.FSC-Net 网络结构如图 2 所示。

1.2.1 部分重参数化快速残差块

在光伏电池红外图像缺陷检测任务中,模型对多尺度特征的有效提取在很大程度上决定其对缺陷目标的感知效果.BasicBlock 模块作为核心特征提取模块,采用多分支设计来增强特征表达能力.该模块依托传统卷积神经网络架构,通过堆叠多个 3×3 卷积核并引入残差连接机制来提升特征学习效果.然而,由于需要在网络深度与计算效率间权衡,其整体参数量较大,计算开销较高,且在检测性能方面仍然存在一定局限性。

为解决上述问题,本文提出一种部分重参数化快速残差块 PRFRB,如图 3 所示.该结构在主干网络分支中使用部分重参数化融合块(PRFB)替代原有的第 2 个 3×3 卷积层.PRFB 借鉴 Faster-Net 的结构设计,由部分重参数化卷积模块(PRConv)与 Faster-Net 融合而得到.PRConv 是在部分卷积(PConv)^[18]的基础上进行改进,PConv 本身旨在通过削减冗余计算和减少内存访问来实现对空间特征的高效提取,从而提升计算效率。

在 FasterNet 结构中,引入部分卷积 PConv,仅



注: 紫色方框数字表示颈部网络的层级索引, 如17表示FSC-Net.yaml文件的第17层深度

图2 FSC-Net 网络结构

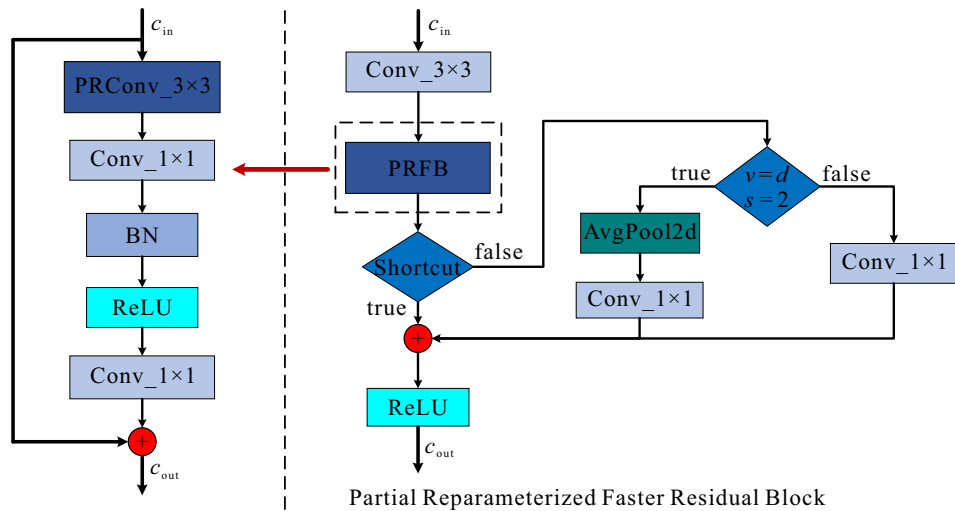


图3 PRFRB 网络结构

对输入特征图中部分通道 c_p 进行空间特征提取. PConv 操作的计算量和内存访问量如下所示:

$$FLOP_{SPConv} = h \times w \times k^2 \times c_p^2, \quad (1)$$

$$MAC = h \times w \times 2c_p + k^2 \times c_p^2. \quad (2)$$

其中: h 为输入/输出数据的高度, w 为输入/输出数据的宽度, c_p 为部分通道, k 为卷积核的大小. 式 (1) 和 (2) 的参数解释同样适用于图 4 中的参数. 当通道分离比设定为 $r = c_p/c = 1/4$ 时 (c 为输入数据的总通道数), PConv 的计算量仅为传统卷积的 1/16, 内存访问量为常规卷积的 1/4.

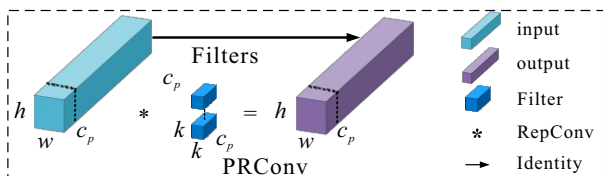


图4 PRConv 网络结构

然而, PConv 在信息交互和表达能力方面较弱. 为此, 本文进一步提出一种部分重参数化卷积结构——PRConv. 该结构以 RepConv^[19] 模块替代常规卷积操作, 增强了模型的特征提取能力. 经实验验证, 将 PRConv 融入 FasterNet 并集成至原始 BasicBlock 结构中, 可在提高检测精度的同时, 有效降低模型的参数量和计算量. PRConv 网络结构如图 4 所示.

在 PRFB 模块中, PRConv 操作后接入逐点卷积 (PWConv), 这两者结合后, 在输入特征图上的有效感受野看起来像一个 T 型卷积 (T-shaped Conv). 具体而言, 如图 5(b) 所示: 该结构的中心区域权重较大, 周围区域权重较小, 相比于普通卷积, 它更加关注中心区域. 这种设计有助于充分挖掘各通道的特征信息, 提升模型对关键区域的感知和响应能力. 然而, 在实际应用中, T 型结构的分解布局效果更佳, 因为这种分解有助于进一步减少计算量 (FLOPs). 如

图 5(a) 所示, 对于相同的输入 $I \in \mathbb{R}^{c \times h \times w}$ 和输出 $O \in \mathbb{R}^{c \times h \times w}$, T 型卷积的 FLOPs 公式如下所示:

$$\text{FLOPs}_{\text{T-shaped-Conv}} = h \times w \times (k^2 \times c_p \times c + c \times (c - c_p)). \quad (3)$$

分解后卷积模块的 FLOPs 公式如下所示:

$$\text{FLOPs}_{\text{PRConv+PWConv}} = h \times w \times (k^2 \times c_p^2 + c^2). \quad (4)$$

其中: 当 $(k^2 - 1)c > k^2 c_p$, 如 $k = 3$, $c_p/c = 1/4$ 时, 分解后的卷积计算量更低, 这种两步实现方式可更方便地利用常规卷积 (regular Conv).

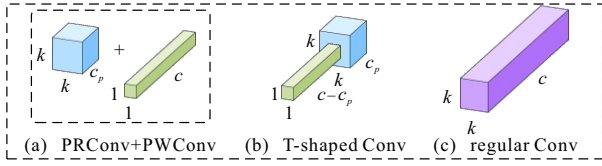


图5 T型卷积分解、合并和常规卷积网络结构

1.2.2 小目标感知特征金字塔

由于无人机采集到的光伏电池红外图像中缺陷

目标普遍较小, P_3 、 P_4 和 P_5 检测层在识别小尺寸目标时存在一定局限性. 通常, 为了增强小目标检测能力, 会增加 P_2 检测层, 但是, 这会造成计算复杂度的提升以及后处理阶段耗时增加, 影响整体系统的效率. 具体而言, 直接增加 P_2 检测层的计算量为 78.8 GLOPs, 明显增加了计算负担. 鉴于此, 设计一种能够更高效处理小目标的特征金字塔结构显得尤为重要. 针对这一需求, 本文提出一种小目标感知特征金字塔 SOAFP, 具体网络结构如图 6 所示. 该模块旨在增强小目标特征的提取质量, 同时, 兼顾模型体积与计算开销间的平衡. 与直接增加 P_2 检测层的方法不同, 本文通过对 P_2 特征层应用空间到深度卷积 (SPDConv) 进行处理, 充分提取红外缺陷图像中小目标的丰富特征, 并将其与 P_3 层特征实现融合. 这种设计不仅有效避免了因引入额外检测分支而导致的计算量增加 (SOAFP 的计算量为 65.2 GLOPs), 还提升了特征在不同尺度间的表达和交互效率.

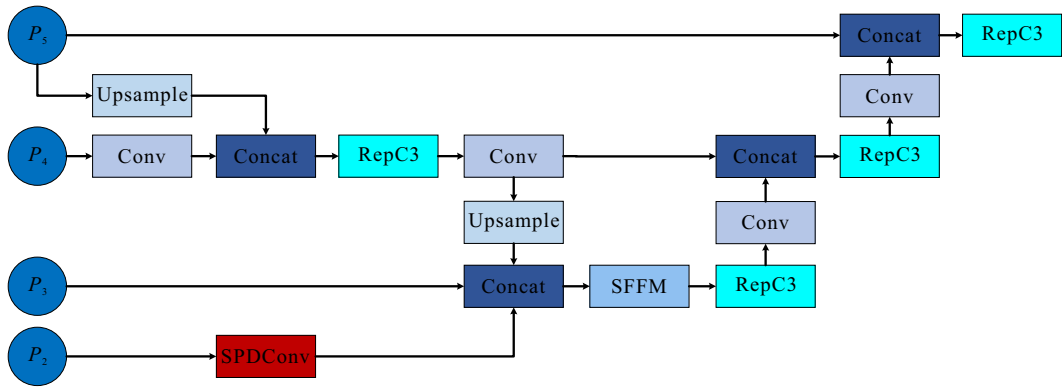


图6 SOAFP 网络结构

进一步来看, SPDConv 模块不同于传统的跨步卷积与池化操作, 它通过将空间维度信息映射至深度维度, 实现了对特征图的高效采样和重构, 有效避免了传统下采样过程中常见的信息丢失问题, 并保持了关键细节特征的完整性. 具体而言, SPD 模块首先将输入特征图划分为若干子区域, 如下所示:

$$\begin{cases} f_{0,0} = X[0 : S : \text{scale}, 0 : S : \text{scale}], \\ f_{1,0} = X[1 : S : \text{scale}, 0 : S : \text{scale}], \dots, \\ f_{\text{scale}-1,0} = X[\text{scale} - 1 : S : \text{scale}, \\ \quad 0 : S : \text{scale}]; \\ f_{0,1} = X[0 : S : \text{scale}, 1 : S : \text{scale}], f_{1,1}, \dots, \\ f_{\text{scale}-1,1} = X[\text{scale} - 1 : S : \text{scale}, 1 : S : \text{scale}]; \\ \vdots \\ f_{0,\text{scale}-1} = \\ X[0 : S : \text{scale}, \text{scale} - 1 : S : \text{scale}], f_{1,\text{scale}-1}, \dots, \\ f_{\text{scale}-1,\text{scale}-1} = X[\text{scale} - 1 : S : \text{scale}, \\ \quad \text{scale} - 1 : S : \text{scale}]. \end{cases} \quad (5)$$

其中: 对于任意 (原始) 特征图 X , 一个子图 $f_{x,y}$ 由所有满足 $i + x$ 和 $j + y$, 均能够整除 scale 的 $X(i, j)$ 组成.

以原始特征图 X (尺寸为 $S \times S \times C_1$) 为例, 处理后得到新特征图 X' 在空间尺寸上缩小为原来的 $1/\text{scale}$, 通道数扩展为 scale^2 倍. 如图 7 所示: 当 $\text{scale} = 2$ 时, 输入特征图被划分为 4 个大小为 $(S/2 \times S/2 \times C_1)$ 的子特征图 $f_{0,0}$ 、 $f_{1,0}$ 、 $f_{0,1}$ 、 $f_{1,1}$, 然后在通道维度拼接为 $(S/2 \times S/2 \times 4C_1)$ 的新特征图 X' . 接着, X' 通过一个非跨步 (步长为 1) 卷积层处理, 该层具有 C_2 个过滤器, 其中 $C_2 < \text{scale}^2 C_1$, 进一步将 $X'(S/\text{scale}, S/\text{scale}, \text{scale}^2 C_1)$ 输出特征图的尺寸调整为 $X''(S/\text{scale}, S/\text{scale}, C_2)$. 采用非跨步卷积有助于充分保留与光伏电池红外图像缺陷相关的判别性特征, 避免重要特征的丢失, 从而提升检测的准确性. 将 SPDConv 模块集成至 SOAFP 网络结构

中,不仅保留了光伏电池红外图像缺陷中关键空间特征,还提升了整体检测性能。

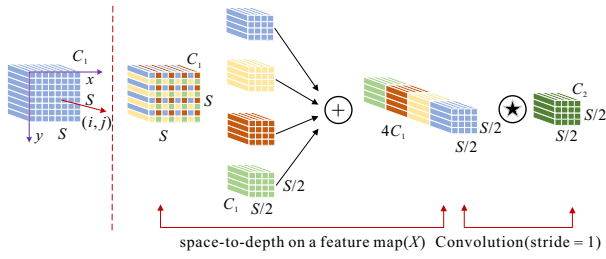


图7 SPDCConv 网络结构

为进一步加强模型对小目标缺陷的多尺度表达能力,SOAFP 模块还引入了浅层特征融合模块 SFFM.考虑直接使用 Omni-Kernel 结构将带来参数量和计算量的大幅增长.为此,本文基于 CSP (cross stage partial) 思想,在保持特征传递完整性的基础上对通道进行处理.具体而言,仅对 25% 的通道进行 Omni-Kernel 计算,其余 75% 的通道则通过直接拼接方式保留,有效控制了模型复杂度。

此外,为进一步减轻参数冗余,Omni-Kernel 中大分支卷积核的尺寸由原始的 63 调整为 31,在不牺牲建模能力的前提下降低了计算负担. Omni-Kernel 模块由 3 个并行分支构成,分别为全局分支、大分支和局部分支,旨在实现从全局到局部的特征学习,进而提升对小目标的检测效果,其中全局分支结合双域通道注意力模块 (DCAM) 和基于频率的空间注意力模块 (FSAM),用于捕获广泛的上下文信息,增强网络的全局建模能力.具体而言,给定输入特征 $X_{\text{Global}} \in \mathbb{R}^{C \times H \times W}$, DCAM 首先对 X_{Global} 应用频率通道注意力 (FCA),如下所示:

$$X_{\text{FCA}} = \mathcal{I}\mathcal{F}(\mathcal{F}(X_{\text{Global}}) \otimes W_{1 \times 1}^{\text{FCA}}(\text{GAP}(X_{\text{Global}}))). \quad (6)$$

其中: \mathcal{F} 和 $\mathcal{I}\mathcal{F}$ 分别为快速傅里叶变换及其逆变换, X_{FCA} 、 $W_{1 \times 1}$ 和 GAP 分别为 FCA 的输出、一个 1×1 卷积层和全局平均池化, \otimes 表示逐元素乘法运算.通过傅里叶处理,根据谱卷积定理有效地提炼全局特征.在谱域全局调制后,将得到的特征进一步输入至空间通道注意力模块 (SCA),其公式如下所示:

$$X_{\text{DCAM}} = X_{\text{FCA}} \otimes W_{1 \times 1}^{\text{SCA}}(\text{GAP}(X_{\text{FCA}})), \quad (7)$$

其中 X_{DCAM} 为 DCAM 的输出. DCAM 仅在通道级粗粒度上增强双域特征;在空间维度上应用基于频率的注意力模块,以在细粒度层面对频谱进行优化,公式表示为

$$X_{\text{FSAM}} = \mathcal{I}\mathcal{F}(\mathcal{F}(W_{1 \times 1}^1(X_{\text{DCAM}})) \otimes W_{1 \times 1}^2(X_{\text{DCAM}})), \quad (8)$$

这里 X_{FSAM} 为 FSAM 的结果.通过这种方式,模型关注于有用的频率成分,用上标 1 和上标 2 对这两个卷积层进行标识。

大分支则通过大尺寸卷积核设计,扩展感受野,以实现多粒度特征的丰富表达.为避免由于大核卷积引入过大的计算开销,本文设置不同的卷积核大小 (15、31、63) 进行实验.实验结果表明,当卷积核大小为 31 时,计算量与平均精度间达到较好的平衡.此外,局部分支通过使用 1×1 深度卷积对局部特征信息进行调制,确保对微小缺陷的精确定位.本文通过合理设计通道比例和卷积核大小,结合多分支结构,有效提升了模型对光伏电池红外图像中小目标缺陷的检测性能,同时兼顾了计算效率. SFFM 模块网络结构如图 8 所示。

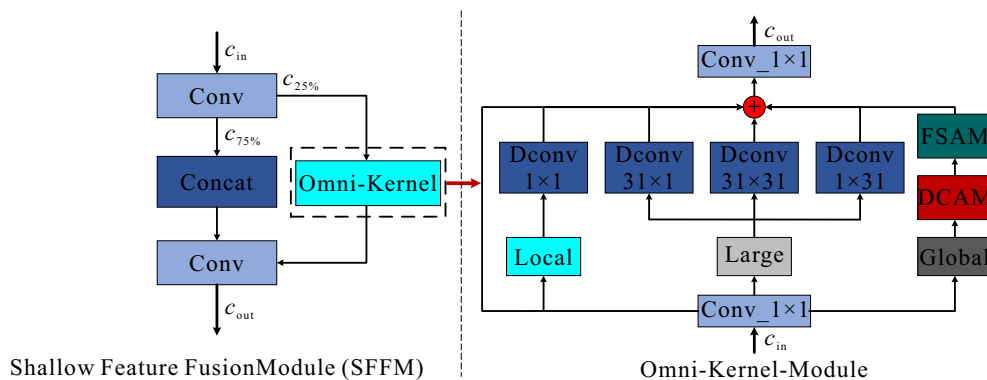


图8 SFFM 网络结构

1.2.3 特征融合 Conv3XCC3 模块

在光伏电池红外图像缺陷检测任务中,热斑、二极管短路等缺陷目标通常具有尺寸小、特征弱等特点,易造成误检和漏检.引入结构重参数化的多分支卷积进行多尺度特征提取成为一种有效的解决方案。

模块可在训练和推理阶段采用结构重参数化策略,通过差异化的分支结构,既增强了训练阶段的特征表达能力,又在推理阶段实现了网络结构的简化,从而提升模型的推理效率。

基于上述思想,本文引入多分支特征融合模块

Conv3XCC3, 如图9所示. 该模块在 RepC3 模块基础上集成了卷积 Conv3XC^[20], 通过主干分支结构实现多尺度信息的有效融合, 进一步提升后续检测头对语义特征的利用能力. Conv3XC 是 Conv3XCC3 模块的核心组成部分, 在训练阶段, 其右分支包含 3 层卷积结构, 通过与左分支输出的特征进行融合, 随后依次经批归一化 (Batch Normalization) 和 SiLU 激活函数, 形成最终输出. 具体而言, 输入特征图

$X \in \mathbb{R}^{c_{in} \times h \times w}$ 表示一个大小为 $h \times w$ 的二维特征图 (h 和 w 分别为特征图的高和宽), 具有 c_{in} 个通道. 通过零填充操作将其扩展为 X_p , 其尺寸为 $(h+2) \times (w+2)$, 即在每个边缘添加 1 个像素的零填充, 其公式如下所示:

$$X_p = \text{Padding}(X), \quad (9)$$

其中 X_p 为扩展后的特征图.

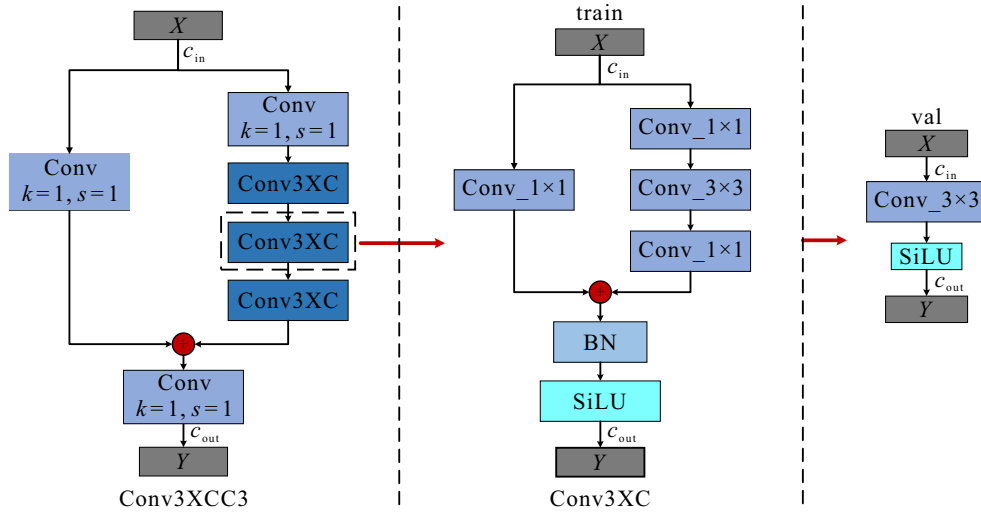


图9 Conv3XCC3 网络结构

接下来, 右分支部分使用 3 个卷积层处理特征图. 第 1 个 1×1 卷积通过对每个像素的通道进行线性变换来增加特征的代表能力, 第 2 个 3×3 卷积的目的是进行空间特征提取, 第 3 个 1×1 卷积主要用于调整通道数使其符合模型的最终输出要求. 通过这一系列操作, 右侧分支输出的特征图 X_r 被获得. 左侧分支通过 1×1 卷积调整输入特征的通道数, 确保其与右分支输出的特征图 X_r 的通道数一致. 经维度调整后, 合并操作如下所示:

$$X_s = X_r \oplus X_l. \quad (10)$$

其中: X_l 和 X_r 分别为左右侧分支输出的特征图, \oplus 表示每个位置的像素值将被相加.

合并后的特征图 X_s 经批量归一化 (BN) 和 SiLU 激活函数处理, 最终输出特征图 Y . 该结构设计使得网络对图像全局上下文建模方面具备更强的表达能力.

在推理阶段, 借助结构重参数化策略, Conv3XC 模块的多分支路径被转换为单一的 3×3 卷积结构, 在保证特征一致性的同时降低计算成本, 有效避免了冗余操作, 为模型推理提供了支持. 具体而言, 首先对卷积核和偏置进行融合, 减少冗余的卷积操作, 计算公式如下所示:

$$W_{\text{final}} = W_1 \oplus W_2, \quad (11)$$

$$b_{\text{final}} = b_1 \oplus b_2. \quad (12)$$

其中: W_1 、 W_2 、 b_1 、 b_2 为卷积层左右侧分支的卷积核和偏置, \oplus 表示在多个卷积层与偏置间进行加权和, W_{final} 和 b_{final} 为融合后的卷积核和偏置. 最后, 经 SiLU 激活函数得到的输出 Y 为

$$Y = \text{SiLU}(F_{\text{conv2d}}(X, W_{\text{final}}) + b_{\text{final}}), \quad (13)$$

这里 $F_{\text{conv2d}}(X, W_{\text{final}})$ 表示将输入特征图 X 和合并后的卷积核 W_{final} 进行卷积计算.

整体而言, Conv3XCC3 模块通过结构优化实现了特征层级的深度融合. 在训练阶段, 模型的参数量和计算量分别为 20.58 M 和 60.6 GFLOPs; 而在推理阶段, 参数量和计算量则降低至 19.88 M 和 57.0 GFLOPs. 由此可见, 该优化策略不仅有效提升了模型的推理效率, 同时, 也为光伏电池红外图像缺陷检测精度的提升提供了有力保障.

1.2.4 Inner-WIoUv3 损失函数的构建

传统的 IoU (intersection over union) 指标在边界框重叠评估中仅关注交并比, 难以准确刻画两个框之间的相对关系, 尤其是在无重叠或重叠较小的情况下, 其判别能力有限. 为克服这一局限, GIoU 在 IoU 的基础上引入预测框与真实框的最小外接矩形,

引入非重叠区域的信息, 有效提升了回归优化过程中对于目标定位偏差的敏感性. 在 RT-DETR-R18 模型中选用 GIoU 作为损失函数. IoU 的计算是用预测框 (A) 和真实框 (B) 的交集除以其并集, 公式为

$$\text{IoU} = \frac{|A \cap B|}{|A \cup B|}. \quad (14)$$

进一步地, 可得到 IoU 的损失函数为

$$L_{\text{IoU}} = 1 - \text{IoU}. \quad (15)$$

其中: $A \cap B$ 为预测框与真实框的交集面积, $A \cup B$ 为并集面积. C 为预测框与真实框的最小外接矩形, 用于表征两个框之间的距离, 其公式为

$$\text{GIoU} = \text{IoU} - \frac{|C - (A \cap B)|}{|C|}. \quad (16)$$

GIoU 作为损失函数时, 其公式如下所示:

$$L_{\text{GIoU}} = 1 - \text{IoU} + \frac{|C - (A \cap B)|}{|C|}. \quad (17)$$

在光伏电池红外图像缺陷检测的实际场景中, 采集到的图像数据通常背景复杂、分布多样, 且标注质量存在一定差异. GIoU 损失函数在边界框回归过程中未能充分体现样本质量差异对目标的影响, 使得低质量样本在训练中对模型性能造成干扰, 增加了缺陷识别的难度. 为进一步提升回归任务的鲁棒性, 本文引入 Inner-WIoUv3 作为损失函数.

WIoUv1 通过引入距离注意力机制, 该机制会随着预测框与真实框重合度的提高减弱对边界的关注, 集中优化高质量样本. 其公式如下所示:

$$L_{\text{WIoUv1}} = R_{\text{WIoU}} L_{\text{IoU}}, \quad (18)$$

$$R_{\text{WIoU}} = \exp\left(\frac{(x - x_{gt})^2 + (y - y_{gt})^2}{(W_g^2 + H_g^2)^*}\right). \quad (19)$$

其中: (x, y) 和 (x_{gt}, y_{gt}) 分别为预测边界框和真实边界框的中心点坐标; W_g 和 H_g 为最小封闭框的尺寸

大小; * 表示分离操作, 目的是防止 R_{WIoU} 产生阻碍收敛的梯度.

WIoUv3 在权重分配机制上实现了更高的动态适应性和收敛稳定性, 引入了动态调整因子 r 来增强 WIoUv1 的检测性能, 如下所示:

$$L_{\text{WIoUv3}} = r L_{\text{WIoUv1}}. \quad (20)$$

式 (20) 可对低质量样本赋予较小的梯度增益, 从而避免低质量检测框产生有害的梯度反向传播. 动态调整因子 r 基于离群度 β 计算, β 用于量化预测框与真实边界框间的匹配程度. r 、 β 的计算公式为

$$r = \frac{\beta}{\delta \alpha^{m-\delta}}, \quad \beta = \frac{L_{\text{IoU}}^*}{L_{\text{IoU}}} \in [0, +\infty), \quad m = \sqrt[n]{0.05}. \quad (21)$$

这里: L_{IoU}^* 为优化后的交并比 (IoU) 值, \bar{L}_{IoU} 为动量 m 的运行平均值, t 为当前迭代次数, n 为批处理大小, α 和 δ 为模型的超参数.

WIoUv3 在处理高度重合框时可能会对这些框产生过度惩罚, Inner-IoU 引入尺度因子 ratio (本文实验设置 ratio = 0.7) 和辅助边界框, 动态调整边界框的尺寸, 能够实现损失函数计算方式的改进. 将 Inner-IoU 与现有的 IoU 进行融合, 有助于提升边界框定位的精度和鲁棒性. Inner-IoU 如图 10 所示, 其中 w_{inner}^{gt} 、 h_{inner}^{gt} 和 w_{inner} 、 h_{inner} 为辅助边框、辅助预测框的宽和高. 尺度因子 ratio $\in [0.5, 1.5]$, 当尺度因子小于 1 时 Inner-IoU 如图 10(a) 所示, 当尺度因子大于 1 时 Inner-IoU 如图 10(b) 所示, 辅助边框大于实际边框.

Inner-IoU 的损失函数公式如下所示:

$$b_l^{gt} = x_c^{gt} - \frac{w^{gt} \times \text{ratio}}{2}, \quad b_r^{gt} = x_c^{gt} + \frac{w^{gt} \times \text{ratio}}{2}, \quad (22)$$

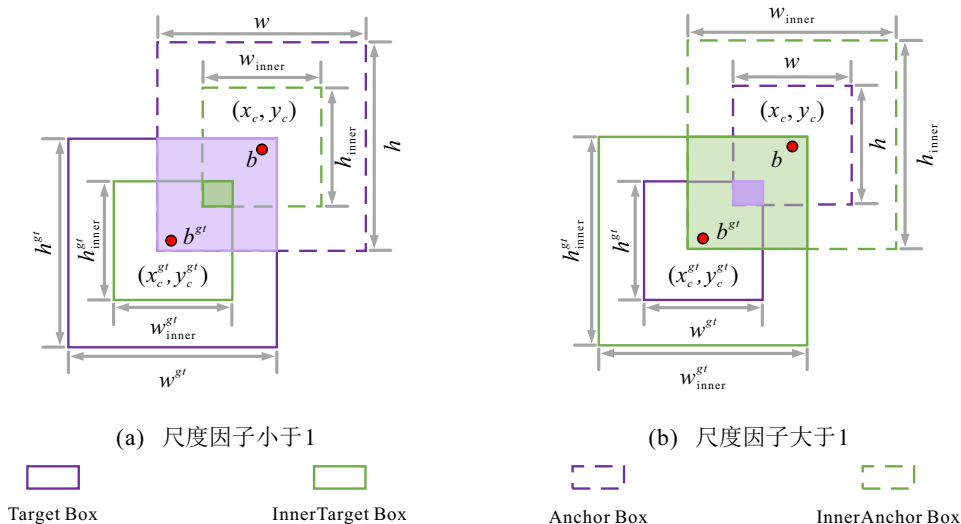


图10 Inner-IoU 损失函数

$$b_t^{gt} = y_c^{gt} - \frac{h^{gt} \times \text{ratio}}{2}, b_b^{gt} = y_c^{gt} + \frac{h^{gt} \times \text{ratio}}{2}, \quad (23)$$

$$b_l = x_c - \frac{w \times \text{ratio}}{2}, b_r = x_c + \frac{w \times \text{ratio}}{2}, \quad (24)$$

$$b_t = y_c - \frac{h \times \text{ratio}}{2}, b_b = y_c + \frac{h \times \text{ratio}}{2}, \quad (25)$$

$$\text{inter} = (\min(b_r^{gt}, b_r) - \max(b_l^{gt}, b_l)) \times (\min(b_b^{gt}, b_b) - \max(b_t^{gt}, b_t)), \quad (26)$$

$$\text{union} = (w^{gt} \times h^{gt}) \times (\text{ratio})^2 + (w \times h) \times (\text{ratio})^2 - \text{inter}, \quad (27)$$

$$\text{IoU}^{\text{inner}} = \frac{\text{inter}}{\text{union}}. \quad (28)$$

其中: b_t, b_b, b_l, b_r 和 $b_t^{gt}, b_b^{gt}, b_l^{gt}, b_r^{gt}$ 分别为预测框和真实框的上下左右 4 个边界坐标; (x_c^{gt}, y_c^{gt}) 和 (x_c, y_c) 分别为真实框和预测框中心点坐标; w^{gt} 和 h^{gt} 分别为真实框的宽和高; w 和 h 分别为预测框的宽和高; inter 为预测框与真实框的重叠区域; union 为预测框与真实框的联合区域. 最后得到 Inner-IoU 的损失函数 $L_{\text{Inner-IoU}}$ 为

$$L_{\text{Inner-IoU}} = 1 - \text{IoU}^{\text{inner}}. \quad (29)$$

将 Inner-IoU 与 WIoUv3 相结合, 构建形成改进的边界框回归损失函数——Inner-WIoUv3, 其中 Inner-WIoUv3 不仅能够借助辅助边界框计算 IoU, 还能有效缓解低质量样本与类别不平衡对检测性能

的干扰, 从而提升模型整体的泛化能力. 针对光伏电池红外图像缺陷检测任务, 该损失函数在保持检测精度的基础上, 进一步优化了回归过程中的样本权重分配策略. Inner-WIoUv3 公式如下所示:

$$L_{\text{Inner-WIoUv3}} = L_{\text{WIoUv3}} + \text{IoU} - \text{IoU}^{\text{inner}}. \quad (30)$$

2 实验与结果

2.1 实验环境与参数配置

本研究实验平台基于 Windows 11 操作系统, 硬件配置包括主频为 3.0 GHz 的 Intel i7-9700 处理器、16 GB 内存以及配备 8 GB 显存的 NVIDIA RTX 2080 显卡. 软件环境方面, 采用 Python 3.9.21 作为编程语言, 深度学习模型在 Pytorch 2.0.1 框架下构建和训练. 具体的实验环境和参数配置如表 1 所示.

表1 实验环境和参数

实验环境	版本	实验参数	参数值
CPU	Intel i7-9700	Epoch	150
GPU	RTX 2080 8 GB	Batch size	4
Python	3.9.21	Image size	640 × 640
Pytorch	2.0.1	Ratio	0.7
CUDA	11.7	Momentum	0.937

2.2 数据集

本文选用的数据集为某光伏电站无人机采集得到的图像数据, 数据集样例如图 11 所示.

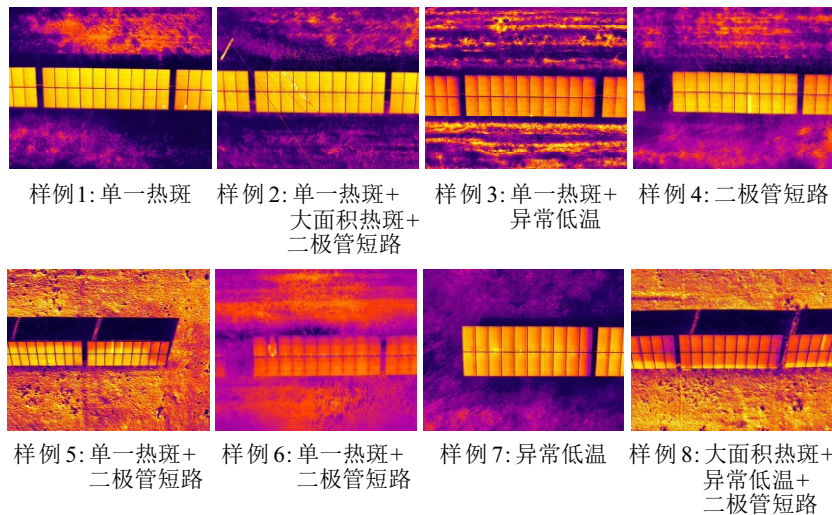


图11 数据集样例

数据集中包含 10 类光伏电池红外图像缺陷. 为聚焦典型特征, 本文选取其中 6 类具有代表性的缺陷作为研究对象, 包括大面积热斑 (dmjrb)、单一热斑 (dyrb)、光伏板正常热成像 (gfb_zc_rcx)、异常低温 (ycdw)、二极管短路 (ejgdl) 以及单一热斑伴随异常低温 (dyrb_ycdw). 大面积热斑表现为覆盖半板或

整板的连续高温区, 导致功率耗散和加速老化; 单一热斑多为由微裂或局部遮挡引起的孤立小型高温点; 光伏板正常热成像下温度分布均匀, 热-电耦合处于稳态, 功率输出正常; 异常低温则呈现暗区或阴影状轮廓, 可能会引起载流不足和功率下降; 二极管短路常在约 1/3 模块范围内形成规则异常温区, 一般呈长

条状,并存在过热风险;单一热斑伴随异常低温表现为点状高温与块状低温并置,形态表现为高温点与暗区并存.随后,数据集按照 8 : 1 : 1 的比例划分为训练集、验证集和测试集.考虑到各类样本在数量上存在明显差异,且整体类别分布较为不均,为缓解数据不平衡对于模型性能的影响,训练阶段引入多种数据增强策略,包括仿射变换、随机裁剪以及透视变换等方法,以提升样本的多样性并增强模型的泛化能力.在划分后的数据集中:训练集共包含 1962 张图像,验证集和测试集分别包含 166 张和 167 张图像.数据集的标签分布如表 2 所示.

表2 标签分布

Instance	dmjrb	dyrb	gfb_zc_rcx	ycdw	ejgd	dyrb_ycdw
Train	1058	2631	25536	4566	405	99
Val	147	355	3746	672	49	10
Test	248	280	3148	615	48	12
Total	1453	3266	32430	5853	502	121

2.3 评价指标

为全面评估 FSC-Net 模型的性能,本实验选取准确率 (P)、召回率 (R)、平均精度均值 (mAP)、模型参数量 (Parameters) 以及计算量 (FLOPs) 作为主要评价指标,通过 TP、FP 和 FN 可以计算出准确率 P 和召回率 R , 如下所示:

$$P = \frac{TP}{TP + FP} \times 100\%, \quad (31)$$

$$R = \frac{TP}{TP + FN} \times 100\%. \quad (32)$$

其中: TP 表示实际为正类且被正确预测为正类的样本数量, FP 表示实际为负类但是被错误识别为正类的样本数量, FN 表示实际为正类但是被误判为负类的样本数量.

通过计算每个类别的 P - R 曲线下方的面积可以确定单个类别的平均精度 AP, mAP 为对所有类别的平均精度, 如下所示:

$$AP = \int_0^1 P(R) dR, \quad (33)$$

$$mAP = \frac{1}{N} \sum_{i=1}^N AP. \quad (34)$$

其中: N 为检测的类别总数. mAP@0.5 表示 IoU 阈值为 0.5 的所有类别的平均精度; mAP@0.5 : 0.95 表示以 0.05 为步进的单位、IoU 阈值从 0.5 ~ 0.95 的所有类别的平均精度的平均值, 该值越高, 模型检测性能越好. mAP@0.5 和 mAP@0.5 : 0.95 公式如下所示:

$$mAP@0.5 = \frac{1}{N} \sum_{i=1}^N AP|_{IoU=0.5}, \quad (35)$$

$$mAP@0.5 : 0.95 = \frac{1}{10} \sum_{k=0}^9 mAP@(0.5 + 0.05k). \quad (36)$$

2.4 消融实验

2.4.1 FSC-Net 消融实验

为验证 FSC-Net 模型在光伏电池红外图像缺陷检测任务中的改进效果, 本文选用 RT-DETR-R18 作为基准模型, 依次加入 PRFRB 模块、SOAFP 模块、Conv3XCC3 模块以及 Inner-WIoUv3 损失函数 (ratio = 0.7, α = 1.7, δ = 2.7), 在本文数据集上开展消融实验. 实验以准确率、召回率、mAP@0.5、mAP@0.5 : 0.95、参数量、计算量以及模型尺寸等指标作为评价标准, 通过对比不同模块组合下的检测结果, 验证 FSC-Net 整体结构设计的有效性和合理性. 具体的消融实验如表 3 所示.

表3 消融实验结果

model	PRFRB	SOAFP	Conv3XCC3	Inner-WIoU	P (%↑)	R (%↑)	mAP@0.5 (%↑)	mAP@0.5 : 0.95 (%↑)	Params (M↓)	FLOPs (G↓)	model size (MB↓)
1					85.9	80.4	84.3	71.7	19.88	57.0	76.9
2	√				83.4	83.2	86.4	72.6	16.79	49.5	65.4
3		√			89.1	82.9	86.1	72.4	20.49	65.2	79.3
4			√		86.0	82.3	86.2	71.9	19.88	57.0	78.4
5	√	√			83.7	83.0	87.2	73.5	17.41	57.7	67.8
6	√		√		86.2	85.1	87.7	74.0	16.79	49.5	66.9
7		√	√		87.5	84.5	87.3	72.7	20.49	65.2	80.8
8	√	√	√		85.1	85.7	89.3	75.3	17.41	57.7	69.3
ours	√	√	√	√	91.5	84.5	90.0	75.8	17.41	57.7	69.3

由表 3 结果可知, 各模块的逐一引入均在不同程度上提升了 RT-DETR-R18 在光伏电池红外图像缺陷检测任务的检测性能. 首先, 引入 PRFRB 模块

后, 在增强骨干网络特征提取能力的同时实现了模型的轻量化. 具体表现为参数量减少了 3.09 M, 计算量降低了 7.5 G, 模型尺寸压缩了 14.9% 的同时

mAP@0.5 和 mAP@0.5 : 0.95 分别提升了 2.1% 和 0.9%。在此基础上, 加入 SOAFP 模块 (对应模型 5), 虽然相较于模型 2 参数与计算量分别增加了 0.62 M 和 8.2 G, 但是检测精度继续提升, mAP@0.5 和 mAP@0.5 : 0.95 分别增长了 0.8% 和 0.9%, 表明该模块在增强模型对小尺度缺陷目标感知方面具有显著效果。然后, 添加 ConvXCC3 模块 (对应模型 8), 进一步强化了颈部网络的多尺度特征融合能力。与模型 5 相比, 虽然模型尺寸增加了 2.2%, 但是, mAP@0.5 和 mAP@0.5 : 0.95 分别提升 2.1% 和 1.8%, 展现出更强的语义信息整合能力。最后, 采用 Inner-WIoUv3 损失函数替代传统 GIoU (对应模型 9), 在保持模型结构不变的前提下, 进一步提升了边界框回归的定位精度, 相较于模型 8, mAP@0.5 和 mAP@0.5 : 0.95 分别提升了 0.7% 和 0.5%。综上所述, 在参数量减少 2.47 M、模型尺寸压缩 9.8%、准确率提升 5.6% 以及召回率提升 4.1% 的前提下, FSC-Net 在 mAP@0.5 和 mAP@0.5 : 0.95 指标上分别实现了 5.7% 和 4.1% 的性能提升, 验证了 FSC-Net 在光伏电池红外图像缺陷检测任务中具有良好的检测性能。

2.4.2 SOAFP 参数消融实验

为深入分析 SFFM 模块在颈部网络中不同参数对于模型参数量、计算量以及检测精度的影响, 本文在 Omin-Kernel 模块中分别设置卷积核大小为 63、31 和 15 进行实验, 结果如表 4 所示。实验结果表明, 卷积核尺寸的增大并不一定带来更好的检测性能。如: 当 Kernel Size = 63 时, 相较于 31, 参数量和计算量分别增加了 0.3 M 和 3.8 G, 但是, 平均检测精度 mAP@0.5 和 mAP@0.5 : 0.95 分别下降了 0.3% 和 0.1%; 当 Kernel Size = 15 时, 虽然参数量和计算量相较于 31 分别减少了 0.07 M 和 1.0 G, 但是, 平均检测精度 mAP@0.5 和 mAP@0.5 : 0.95 却显著下降, 降幅分别达到 5.0% 和 3.4%; 当 Kernel Size = 31 时, 模型在精度与复杂度间实现了较好的平衡, 在参数量和计算量仅略微增加的情况下, 获得了最佳的检测性能。

表4 Kernel Size 调参实验

组别	Params (M↓)	FLOPs (G↓)	mAP@0.5 (%↑)	mAP@0.5 : 0.95 (%↑)	Kernel Size
1	20.79	69.0	85.8	72.3	63
2	20.49	65.2	86.1	72.4	31
3	20.42	64.2	81.1	69.0	15

为评估 SFFM 模块中将部分通道用于 Omni-Kernel 分支对于模型复杂度和检测性能的影响, 本文将通道占比设为 75%、50% 和 25% 进行实验, 实

验结果如表 5 所示。表 5 结果表明, 当 25% 的通道用于 Omni-Kernel 分支时, 模型参数量、计算量分别降至 20.49 M 和 65.2 G, mAP@0.5 达到 86.1%, mAP@0.5 : 0.95 达到 72.4%。与通道占比为 50% 相比, 参数量和计算量分别减少了 0.27 M 和 2.7 G, mAP@0.5 和 mAP@0.5 : 0.95 分别提升了 2.5% 和 1.3%; 与通道占比为 75% 相比, 参数量和计算量分别减少了 0.65 M 和 6.3 G, mAP@0.5 和 mAP@0.5 : 0.95 分别提升了 0.3% 和 1.1%。因此, 将 25% 通道应用 Omni-Kernel 模块操作可在检测精度与模型复杂度间实现更好的平衡。

表5 通道占比选择实验

组别	Params (M↓)	FLOPs (G↓)	mAP@0.5 (%↑)	mAP@0.5 : 0.95 (%↑)	通道占比/%
1	21.14	71.5	85.8	71.3	75
2	20.76	67.9	83.6	71.1	50
3	20.49	65.2	86.1	72.4	25

为验证 SOAFP 模块在 P_2 层应用 SPDCConv 操作以及 SFFM 模块在颈部网络特征融合的效果, 本文进行消融实验, 实验结果如表 6 所示。

表6 应用 SPDCConv 与 SFFM 模块消融实验

组别	Params (M↓)	FLOPs (G↓)	mAP@0.5 (%↑)	mAP@0.5 : 0.95 (%↑)
1	19.88	57.0	84.3	71.7
2	20.04	59.7	85.7	71.8
3	20.49	65.2	86.1	72.4

注: 1为基础模型RT-DETR-R18; 2为基础模型应用 P_2 层 + SPDCConv; 3为基础模型 + SOAFP模块。

表 6 实验结果表明, 加入 SFFM 后, 颈部网络的平均检测精度在 mAP@0.5 和 mAP@0.5 : 0.95 上分别达到 86.1% 和 72.4%, 相比于仅在 P_2 层应用 SPDCConv 模块的模型, 分别提高了 0.3% 和 0.6%。在参数量和计算量方面, 加入 SFFM 后, 模型的参数量和计算量略微增加, 分别为 0.45 M 和 5.5 G。此外, 在 P_2 层应用 SPDCConv 模块相比于基础网络模型 RT-DETR-R18, 模型的平均检测精度 mAP@0.5 和 mAP@0.5 : 0.95 分别达到 85.7% 和 71.8%, 相比于基础模型提高了 1.4% 和 0.1%。参数量和计算量分别增加了 0.16 M 和 2.7 G。因此, 在 P_2 层应用 SPDCConv 和 SFFM 模块有助于提升模型的检测精度。

2.5 不同模型对比实验

为了更好地验证 FSC-Net 的性能表现, 本文选取多个主流目标检测模型作为对比对象, 包括 YOLOv3t、YOLOv3-spp、YOLOv5n、YOLOv8n、

YOLOv9t、YOLO11n、YOLO12n、Hyper-YOLO11n 以及不同版本的 RT-DETR 模型,实验结果如表 7 所示.表 7 结果表明,RT-DETR-L 在光伏电池红外图像缺陷中的检测表现稳健,但是,其计算开销与处理速度的平衡在实际应用中可能会成为瓶颈.FSC-Net 相较于 RT-DETR-L 模型,在参数量和计算量分别减少 14.58 M 和 45.8 G 的情况下,mAP@0.5 和 mAP@0.5 : 0.95 分别比其高出 1.0% 和 0.6%.尽管 YOLOv5n 在对比模型中参数量和计算量最低,但是,其在 mAP@0.5 和 mAP@0.5 : 0.95 指标上分别低于 FSC-Net 20.5% 和 18.2%.这表明 YOLOv5n 无法提供足够的细粒度检测.此外,FSC-Net 与性能较优的 YOLOv8n 模型相比,尽管参数量和计算量有所上升,但是,mAP@0.5 和 mAP@0.5 : 0.95 分别提高了 4.9% 和 4.4%.表明 YOLOv8n 虽然能够较快地处理图像,但是,其未能满足高精度缺陷检测的需求.针对其他轻量化模型,如 YOLOv3t、YOLOv9t、YOLO11n 和 YOLO12n,FSC-Net 在平均检测精度方面与其相比均取得了明显提升.虽然 YOLO^[21] 系列轻量级模型在实时性上具有优势,但是,在特征表达上表现不足(检测精度相对较低),而 RT-DETR 系列模型在主干网络嵌套层级较深,对特征的提取和理解方面更细致,且改进后的模型通过优化特征表达策略,在融合其精度优势的同时降低了计算代价,有效提升了在光伏电池红外图像中对微小缺陷目标的感知能力.

表7 光伏电池缺陷数据集不同模型对比结果

model	Parameters (M)	FLOPs (G)	mAP@0.5 (%↑)	mAP@0.5 : 0.95 (%↑)
YOLOv3t	8.68	12.9	80.0	62.4
YOLOv3-spp	62.57	155.5	83.4	69.4
YOLOv5n	1.77	4.2	69.5	57.6
YOLOv8n	3.01	8.1	85.1	71.4
YOLOv9t	2.62	10.7	82.4	68.9
YOLO11n	2.58	6.3	83.8	70.5
YOLO12n	2.51	5.8	82.6	70.2
Hyper-YOLOn	3.62	9.5	88.7	74.5
RT-DETR-R18	19.88	57.0	84.3	71.7
RT-DETR-R34	31.11	88.8	82.1	69.2
RT-DETR-R50	41.97	129.6	88.6	74.7
RT-DETR-R101	74.66	247.1	88.0	74.1
RT-DETR-L	31.99	103.5	89.0	75.2
RT-DETR-X	65.48	222.5	83.1	69.4
FSC-Net (ours)	17.41	57.7	90.0	75.8

2.6 不同损失函数对比实验

为验证 Inner-WIoUv3 损失函数在边界框回归任务中的效果,本文在引入前 3 个模块的基础上进行实验,将 WIoUv3 与多种典型 IoU 损失函数(包

括 Inner-IoU、EIoU、GIoU、DIoU、CIoU、SIoU、MPDIoU 以及 ShapeIoU) 进行组合测试,以评估性能表现,如表 8 所示.由表 8 实验结果可知,Inner-IoU 通过引入尺度因子和辅助边界框计算损失,在多数情况下能够有效提升模型性能.如 Wise-Inner-EIoU、Wise-Inner-GIoU 与 Wise-Inner-CIoU 相较于 Wise-EIoU、Wise-GIoU 与 Wise-CIoU,分别在召回率、mAP@0.5 和 mAP@0.5 : 0.95 指标上提高了 2.5%、6.1%、5.3%、3.6%、1.7%、1.3% 以及 2.6%、4.7%、3.9%.这种性能提升表明,引入尺度因子和辅助边界框计算能够帮助更准确地捕捉到目标位置,尤其是在光伏电池红外图像缺陷检测需求任务中,提升了检测精度.尽管 GIoU 在召回率方面表现最佳,但是,改进后的 Inner-WIoUv3 损失函数在准确率、mAP@0.5 和 mAP@0.5 : 0.95 指标上相较于 GIoU 分别提升了 6.4%、0.7% 和 0.5%.其他损失函数(如 Wise-DIoU 和 Wise-Inner-GIoU)虽然在召回率上分别相较于 Inner-WIoU 略高 0.6% 和 0.3%,但是,在准确率、mAP@0.5 以及 mAP@0.5 : 0.95 指标上均表现不及,分别低于 Inner-WIoU 5.9%、2.5%、2.2% 和 10.6%、4.4%、3.9%.WIoUv3 结合 DIoU 和 MPDIoU 在处理复杂背景目标时表现出较好的检测效果,其中 mAP@0.5 达到了 87.0% 以上;GIoU 在处理小目标和遮挡目标时也表现优异,但是在本文数据集上,上述损失函数相比于 Inner-WIoUv3 均表现不及.

表8 不同损失函数对比实验

Loss Function	P (%↑)	R (%↑)	mAP@0.5 (%↑)	mAP@0.5 : 0.95 (%↑)
GIoU (base)	85.1	85.7	89.3	75.3
WIoUv3	82.2	80.4	82.5	69.8
Wise-EIoU	79.6	81.5	81.1	68.4
Wise-GIoU	82.1	81.2	83.9	70.6
Wise-DIoU	85.6	85.1	87.5	73.6
Wise-CIoU	79.9	79.3	80.6	68.0
Wise-SIoU	81.3	84.9	85.8	72.7
Wise-MPDIoU	85.6	84.4	87.5	73.8
Wise-ShapeIoU	79.1	79.2	80.7	68.4
Wise-Inner-EIoU	85.1	84.0	87.2	73.7
Wise-Inner-GIoU	80.9	84.8	85.6	71.9
Wise-Inner-DIoU	83.1	80.9	82.3	68.6
Wise-Inner-CIoU	84.9	81.9	85.3	71.9
Wise-Inner-SIoU	79.3	81.2	82.1	69.5
Inner-WIoUv3	91.5	84.5	90.0	75.8

2.7 Inner-WIoUv3 参数对比实验

在 Inner-WIoUv3 中: α 为加权强度,值越大,加权影响越强; δ 为边界框的调整灵敏度, δ 值越小,模型对于边界框的要求越严格.为验证在 Inner-WIoUv3

中设置 $\alpha = 1.7$ 和 $\delta = 2.7$ 时效果最佳, 本文进行不同值的对比实验, 实验结果如表 9 所示. 由表 9 实验结果可见, 当 $\delta = 5.0$ 时, 模型在小目标的特征学习方面存在不足, 导致小目标区域被忽略或无法有效学习. 一般而言, 较小的 δ 值有助于提高小目标检测效果, 这一点在实验中得到了验证. 与此同时, 当设置 $\alpha = 1.7$ 且 $\delta = 2.7$ 时, 平均精度取得了最佳的实验效果.

表9 Inner-WIoUv3 损失函数超参数对比实验

组别	α	δ	mAP@0.5% \uparrow	mAP@0.5 : 0.95% \uparrow
1	1.7	2.7	90.0	75.8
2	1.7	3.0	85.9	72.3
3	1.7	4.0	79.7	66.7
4	1.7	5.0	84.1	71.1
5	1.9	2.7	88.0	74.2
6	1.9	3.0	80.5	67.5
7	1.9	4.0	88.5	73.5
8	1.9	5.0	85.9	72.0

2.8 热力图分析

为深入评估所提出模型在光伏电池缺陷检测任务中的性能, 本文从测试集中选取 8 张具有代表性的光伏电池红外图像, 并采用 GradCAM++ (gradient-weighted class activation mapping plus plus) 技术生成类别激活热力图进行可视化分析. 与传统 Grad-

CAM 方法相比, GradCAM++ 通过引入梯度的高阶信息和空间加权策略来提升激活图的语义精准度, 适用于目标轮廓模糊、特征表达不明显的复杂图像场景. 随后, 选取多个主流检测模型进行可视化效果对比分析, 包括 YOLOv8n^[22]、YOLO11n^[23]、YOLO12n^[24] 以及 Transformer 架构下不同版本的 RT-DETR 模型 (RT-DETR-R18、RT-DETR-R34^[25]、RT-DETR-R50^[26] 等), 并与所提出 FSC-Net 模型在相同图像条件下进行热力图比较, 实验结果如图 12 所示.

图 12 结果表明, FSC-Net 在目标区域的激活响应更加集中, 具备更强的目标聚焦能力, 表现出更优异的可视化效果和识别性能.

为进一步分析热力图中各检测框的激活效果, 本文对热力图的激活像素占比进行定量分析. 具体而言, 统计每个检测框内激活区域的平均像素占比, 并结合红、黄、绿 3 种颜色高亮区域来区分激活强度. 为了清晰地区分背景与有效检测区域, 设定 75% 的阈值. 此外, 除关注激活区域的像素占比, 还应考虑聚焦程度的好坏, 聚焦程度可通过直观地观察热力图来评估. 为全面评判模型性能, 还应考虑误检、漏检情况的影响. 实验结果如表 10 所示. 由表 10 实验结果可见: RT-DETR-R18、RT-DETR-R34 和 RT-DETR-R50 的热力图激活像素平均占比分别为

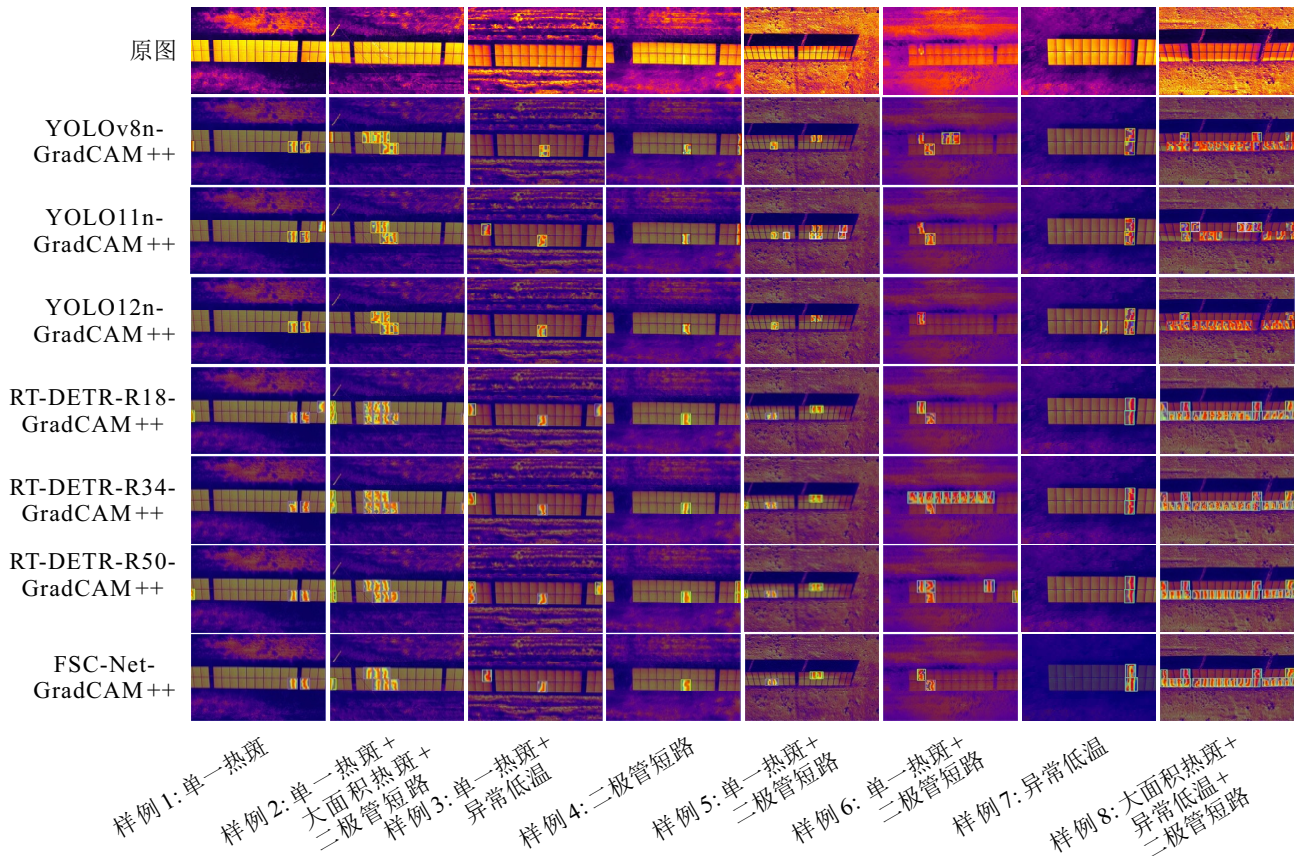


图12 不同模型的热力图效果对比

42.7%、44.0%和 47.9%; 相比之下, YOLOv8n、YOLO11n 与 YOLO12n 的热力图激活像素平均占比分别为 36.1%、37.9% 和 38.4%。结果表明, 在本文数据集上, RT-DETR 系列模型在热力图激活实验中表现优于 YOLO 系列模型。此外, FSC-Net 的热力图

激活占比为 51.1%, 表现出最高的平均激活占比。在聚焦程度方面, FSC-Net 的检测框明显更加聚焦, 且在检测精度上也表现最为突出, 表明其漏检和误检率较低。综合来看, FSC-Net 在激活占比、聚焦程度以及检测精度等方面均展现出较为优异的性能。

表10 热力图激活区域分析实验

model	样例1/%↑	样例2/%↑	样例3/%↑	样例4/%↑	样例5/%↑	样例6/%↑	样例7/%↑	样例8/%↑	平均比例/%↑	聚焦程度
YOLOv8n	32.9	47.0	36.3	47.0	34.7	28.4	28.5	34.3	36.1	差
YOLO11n	35.6	38.5	43.2	46.4	30.5	44.4	34.4	30.1	37.9	差
YOLO12n	40.9	40.2	42.2	42.5	32.3	44.1	32.9	32.3	38.4	差
RT-DETR-R18	40.4	40.8	38.3	49.2	38.7	48.6	45.4	40.2	42.7	差
RT-DETR-R34	45.7	44.9	40.2	43.2	40.7	43.1	43.9	50.1	44.0	差
RT-DETR-R50	49.9	55.1	46.1	45.3	44.5	50.8	48.2	43.2	47.9	较好
FSC-Net	49.2	51.0	52.7	48.7	48.6	54.4	53.6	50.4	51.1	较好

2.9 模型部署

为验证所提出改进网络模型在算法服务器上部署的可行性, 本文选取实验室另外一台搭载 NVIDIA GeForce RTX 2080 算法的服务器作为模型部署设备, 通过将 Windows 系统中训练得到的模型部署于 Ubuntu 系统上, 以评估改进后模型的检测效果。具体实验结果如表 11 所示。

表11 模型部署实验

model	FPS (f/s)↑	Weight (MB)↓	mAP@0.5 (%↑)	mAP@0.5 : 0.95 (%↑)
RT-DETR-R18	49.0	76.9	84.3	71.7
FSC-Net	40.7	69.3	90.0	75.8

注: FPS是在3次检测结果取均值后得到。

由表 11 可知, 改进后模型的平均检测精度 mAP@0.5 和 mAP@0.5 : 0.95 分别达到了 90.0% 和 75.8%, 相较于改进前的模型提高了 5.7% 和 4.1%。同时, 模型尺寸压缩了 9.8%, FPS 相较于原模型略微下降 8.3/(f/s), 达到了 40.7/(f/s), 满足一定的部署需求。

3 结论

针对光伏电池红外图像缺陷检测中因背景复杂和目标尺寸较小所引发的漏检、误检等问题, 本文提出了一种网络模型——FSC-Net。该模型在主干网络中提出了 PRFRB 模块, 使得模型在轻量化的同时增强了对目标特征的感知能力; 在颈部网络中设计了小目标感知特征金字塔, 融合 P_2 检测层、空间到深度卷积和 SFFM 模块, 有效保留了细节信息, 提升了模型对小目标缺陷的检测精度。此外, 为进一步改善特征融合效果, 模型中还引入了 Conv3XCC3 模块, 以提升颈部网络多尺度信息的整合能力。在边界框回归阶段, 采用 Inner-WIoUv3 作为损失函数, 有

效提升了目标定位精度。实验结果表明, FSC-Net 在复杂背景下的光伏电池红外图像缺陷检测任务中表现优异, 验证了其出色的检测性能。

尽管 FSC-Net 在本文数据集缺陷目标检测中表现良好, 但是, 在应对更多样化、复杂化的缺陷类型时, 模型的泛化能力仍有提升空间, 未来的研究工作将聚焦于扩展更多类型的光伏电池缺陷。

参考文献 (References)

- [1] Hassan S, Dhimish M. Broad-scale electroluminescence analysis of 5 million+ photovoltaic cells for defect detection and degradation assessment[J]. *Renewable Energy*, 2024, 237: 121868.
- [2] Ma J, Gao X C, Wang J Q, et al. Surface defect detection of solar cells based on Fourier single-pixel imaging for removal of substrates interference[J]. *Optics and Lasers in Engineering*, 2024, 176: 108073.
- [3] Liu Y C, Hua Q, Chen L L, et al. A multi-scale neighbourhood feature interaction network for photovoltaic cell defect detection[J]. *Knowledge-Based Systems*, 2025, 309: 112882.
- [4] Duan R Y, Wang Y J, Chen X S, et al. An enhanced algorithm for cell-level anomaly segmentation in photovoltaic solar panels using electroluminescence imaging[J]. *Energy*, 2025, 331: 136711.
- [5] Chen S H, Lu Y, Qin G X, et al. CSPD-DETR: Real-time silicon crystalline photovoltaic cell surface defect detection transformer for building photovoltaic systems[J]. *Journal of Building Engineering*, 2025, 108: 112810.
- [6] Bu C W, Shen R H, Bai W L, et al. CNN-based defect detection and classification of PV cells by infrared thermography method[J]. *Nondestructive Testing and Evaluation*, 2025, 40(5): 1752-1769.
- [7] Zhang Y F, Zhang X, Tu D W. Unsupervised defect detection for solar photovoltaic cells based on convolutional autoencoder[J]. *Nondestructive Testing*

- and Evaluation, 2025: 1-17.
- [8] 艾明曦, 许庆, 张进, 等. 基于局部特征增强的浮选过程改进半监督工况识别方法[J]. 控制与决策, 2025, 40(9): 2891-2900.
(Ai M X, Xu Q, Zhang J, et al. Local feature enhancement-based improved semi-supervised condition recognition method for flotation process[J]. Control and Decision, 2025, 40(9): 2891-2900.)
- [9] Zhu J H, Zhou D Q, Lu R S, et al. C2DEM-YOLO: Improved YOLOv8 for defect detection of photovoltaic cell modules in electroluminescence image[J]. *Nondestructive Testing and Evaluation*, 2025, 40(1): 309-331.
- [10] Bao J Q, Yuan X C, Wu Q Y, et al. CCA-YOLO: Channel and coordinate aware-based YOLO for photovoltaic cell defect detection in electroluminescence images[J]. *IEEE Transactions on Instrumentation and Measurement*, 2025, 74: 5009412.
- [11] Li X R, Hou H S, Yu Y Y, et al. YOLO-LitePV: A lightweight detection algorithm for photovoltaic panel defects[J]. *Nondestructive Testing and Evaluation*, 2025: 1-28.
- [12] Pan W W, Sun X F, Wang Y L, et al. Enhanced photovoltaic panel defect detection via adaptive complementary fusion in YOLO-ACF[J]. *Scientific Reports*, 2024, 14(1): 26425.
- [13] Zhang J L, Yang W J, Chen Y M, et al. Fast object detection of anomaly photovoltaic (PV) cells using deep neural networks[J]. *Applied Energy*, 2024, 372: 123759.
- [14] Zhou P, Wang R, Wang C H, et al. SIIF: Semantic information interactive fusion network for photovoltaic defect segmentation[J]. *Applied Energy*, 2024, 371: 123643.
- [15] Liu G Y, Yang D R, Ye J, et al. A real-time welding defect detection framework based on RT-DETR deep neural network[J]. *Advanced Engineering Informatics*, 2025, 65: 103318.
- [16] 周葳楠, 吴治海, 张正道, 等. 基于弱特征增强的轻量化小目标检测方法[J]. 控制与决策, 2024, 39(2): 381-390.
(Zhou W N, Wu Z H, Zhang Z D, et al. Lightweight small target detection method based on weak feature enhancement[J]. Control and Decision, 2024, 39(2): 381-390.)
- [17] Pan J Z, Peng D G, Wang D H, et al. SRHS-net: A hybrid attentive architecture for real-time steam leak detection in power plants[J]. *IEEE Transactions on Industrial Informatics*, 2025, 21(9): 6879-6890.
- [18] Singh D, Choi Y, Park J, et al. Deep-BCSI: A deep learning-based framework for bias correction and spatial imputation of PM_{2.5} concentrations in South Korea[J]. *Atmospheric Research*, 2024, 301: 107283.
- [19] Cui X Z, Zhu L C, Zhao B, et al. Parallel RepConv network: Efficient vineyard obstacle detection with adaptability to multi-illumination conditions[J]. *Computers and Electronics in Agriculture*, 2025, 230: 109901.
- [20] Zuo C H, Huang N X, Yuan C, et al. Pavement-DETR: A high-precision real-time detection transformer for pavement defect detection[J]. *Sensors*, 2025, 25(8): 2426.
- [21] 丁浩晗, 贺万程, 万俊, 等. DI-YOLO: 一种面向无人机航拍图像的高效小目标检测框架[J]. 控制与决策, 2025, 40(10): 3106-3116.
(Ding H H, He W C, Wan J, et al. DI-YOLO: An efficient small object detection framework for UAV aerial imagery[J]. Control and Decision, 2025, 40(10): 3106-3116.)
- [22] 卢迪, 赵庆. 空间分组内卷积轻量级目标检测算法[J]. 控制与决策, 2025, 40(10): 3127-3135.
(Lu D, Zhao Q. Lightweight object detection algorithm based on SGWInvo[J]. Control and Decision, 2025, 40(10): 3127-3135.)
- [23] Akdoğan S, Öziç M Ü, Tassoker M. Development of an AI-supported clinical tool for assessing mandibular third molar tooth extraction difficulty using panoramic radiographs and YOLO11 sub-models[J]. *Diagnostics*, 2025, 15(4): 462.
- [24] Cui B Y, Liu Z, Yang Q F. UAV-YOLO12: A multi-scale road segmentation model for UAV remote sensing imagery[J]. *Drones*, 2025, 9(8): 533.
- [25] 武小兰, 孙奔奔, 白志峰, 等. 基于多尺度特征融合和组合结构损失优化的深度学习弯道车道线检测方法研究[J]. 控制与决策, 2025, 40(8): 2467-2472.
(Wu X L, Sun B B, Bai Z F, et al. Research on deep learning-based curved lane detection method using multi-scale feature fusion and composite structural loss optimization[J]. Control and Decision, 2025, 40(8): 2467-2472.)
- [26] Li C Y, Xu Q Y, Lu Y, et al. A new method for tomicus classification of forest pests based on improved ResNet50 algorithm[J]. *Scientific Reports*, 2025, 15(1): 9665.

作者简介

彭道刚 (1977-), 男, 教授, 博士, 博士生导师, 主要研究方向为低碳智能发电、综合智慧能源、电力巡检机器人, E-mail: pengdaogang@126.com;

邓玉澳 (2000-), 男, 硕士生, 主要研究方向为电力巡检机器人、目标检测, E-mail: deng_yuao@163.com;

王丹豪 (1992-), 男, 讲师, 博士, 主要研究方向为综合能源系统、能源互联网, E-mail: danhao.wang@shiep.edu.cn;

潘俊臻 (2000-), 博士生, 主要研究方向为电力巡检机器人、机器视觉, panjunzhen0712@163.com.