

DOI:10.19651/j.cnki.emt.2416770

基于改进 YOLOv8 光伏板缺陷检测算法*

杨长春 贺轩轩 王睿 朱士竹 严昊

(常州大学王诤微电子学院集成电路产业学院 常州 213159)

摘要: 针对现有的分布式光伏电池板缺陷检测精度低、计算量高、参数量大以及复杂背景多变等因素导致的误检、漏检问题,提出了一种改进的轻量级 YOLOv8 分布式光伏电池板缺陷检测算法。采用高效轻量级的 StarNet 网络架构作为特征提取网络,减少计算成本和参数量,实现高效率和高性能之间平衡;设计 SPPF-AM 模块,增强了模型对空间信息的感知能力,有效应对不同尺度的目标;加入三元组注意力机制 Triplet,有效地提取多尺度目标特征,提升模型的表征能力和任务性能;设计 C2f_DSCConv2D 结合可变形卷积取代原网络中的 C2f,以较低的存储和较高的计算速度,从而提高缺陷检测模型的效率;在特征融合网络中引入空间上下文感知模块 SCAM,减少噪声影响,有效抑制图像中的噪声,抑制无关背景信息的干扰。设计 ECIOU 替换 CIOU,增强边界框损失的拟合能力,加快网络模型的收敛速度。实验结果表明:改进后的 YOLOv8 模型参数量下降 35%,计算量下降 29.6%,检测精度达到了 90.1%,mAP@50 从 85.9% 提升到了 89.7%,提升了 4.2%。改进后模型在参数量、计算量下降的情况下检测精度也有一定的提升。所提出的改进算法在缺陷检测任务中表现出了较好的性能,有效增强了光伏电池板缺陷检测模型的检测性能。

关键词: 缺陷检测;YOLOv8;轻量级;移位卷积;SCAM;损失函数

中图分类号: TK514;TN98;TP391.4 **文献标识码:** A **国家标准学科分类代码:** 520.6040

Based on the improved YOLOv8 photovoltaic panel defect detection algorithm

Yang Changchun He Xuanxuan Wang Rui Zhu Shizhu Yan Hao

(Wang Zheng Microelectronics College, Integrated Circuit Industry College, Changzhou University, Changzhou 213159, China)

Abstract: To address the issues of low detection accuracy, high computational load, large parameter size, and complex variable backgrounds in existing distributed photovoltaic panel defect detection, we propose an improved lightweight YOLOv8 defect detection algorithm for distributed photovoltaic panels. We adopt the efficient lightweight StarNet architecture as the feature extraction network to reduce computational costs and parameter size, achieving a balance between high efficiency and high performance. The SPPF-AM module is designed to enhance the model's spatial information perception capability, effectively handling targets of different scales. We incorporate the Triplet attention mechanism to effectively extract multi-scale target features, improving the model's representation ability and task performance. The C2f_DSCConv2D, which combines deformable convolution, replaces the original C2f in the network, improving defect detection efficiency with lower storage and higher computation speed. A spatial context-aware module (SCAM) is introduced in the feature fusion network to reduce noise impact and effectively suppress irrelevant background interference. We design ECIOU to replace CIOU, enhancing the fitting ability of the bounding box loss and accelerating the network's convergence speed. Experimental results show that the improved YOLOv8 model reduces parameter size by 35% and computational load by 29.6%, achieving a detection accuracy of 90.1%, with mAP@50 increasing from 85.9% to 89.7%, an improvement of 4.2%. The improved model demonstrates a certain enhancement in detection accuracy while reducing parameter size and computational load. The proposed improved algorithm demonstrated good performance in defect detection tasks, effectively enhancing the detection capability of the photovoltaic panel defect detection model.

Keywords: defect detection;YOLOv8;lightweight;shifted convolution;SCAM;loss function

0 引言

光伏电池板作为可再生能源领域的重要组成部分,其

性能和质量直接影响光伏系统的效率和可靠性。然而,由于光伏电池板通常放置于恶劣的室外环境下,常年受到风吹日晒,可能会出现各种缺陷,如脏污、电气损坏、物理损

收稿日期:2024-08-31

* 基金项目:国家自然科学基金(62301086)项目资助

坏、异物覆盖等。这些缺陷可能导致能量损失、寿命缩短和系统故障等问题,尽早准确检测这些缺陷对于提高光伏电池板的整体性能至关重要。然而,训练和测试深度学习模型所需的计算资源相对庞大,而进行光伏电池板缺陷检测的终端设备计算能力有限,这给实际应用带来了挑战。因此,在确保检测精度的前提下,如何尽量降低深度学习模型的复杂度和计算量,是该领域亟待解决的重要问题之一。

为此,许多学者致力于此进行研究。赵强等^[1]提出了一种基于 HSV 空间模型的 K 均值聚类图像处理方法,用于处理光伏板红外图像。管宽岐等^[2]设计了光伏板识别模型去除地面背景噪声对热斑识别的影响,将光伏板从红外图像中截取出来,从而快速检测热斑。Prabhakaran 等^[3]提出了一种实时多变体深度学习模型(RMVDM)来处理这个问题。张猛等^[4]提出了一种基于 SimAM-AdaYOLOv5 算法的太阳能电池缺陷检测方法,实现自适应学习特征尺度和感受野的大小;通过引入 SimAM^[5]注意机制,进一步提高模型的特征提取能力。然而,上述方法需要更多的计算资源和时间进行训练和推理,且检测速度较慢,捕捉位置信息的能力也有所欠缺。

近年来,深度学习模型在目标检测、图像分类以及语义分割等领域广泛应用。与传统方法相比,深度学习通过学习大量样本获得数据集的特征表示,因此具备更好的泛化性、更高的精确度和更强的鲁棒性。基于深度学习的目标检测方法目前主要分为两类:一类是单阶段目标检测算法,如 YOLO、SSD^[6]等;另一类是二阶段目标检测算法,例如 Faster-R-CNN^[7]等。刘承毅等^[8]提出了改进 Dn-YOLOv7 算法。该算法结合去噪卷积神经网络(DnCNN),提出一个降噪模块(DnBlock)利用空间坐标卷积(CoordConv)对噪声通道进行卷积整合,增强了网络的降噪能力提高了模型对小目标的检测能力。周颖等^[9]提出的光伏电池缺陷检测方法 YOLOv8-EL,使用 GauGAN 进行数据增强解决了数据集中各类缺陷不平衡问题,在骨干网络和特征提取网络中引入上下文聚合模块 CAM 来抑制背景噪声,最后构建多注意力检测头 MADH 有效提升了检测精度。Phan 等^[10]提出一种检测光伏电池故障的新方法,在 YOLOv8 中引入了粒子群优化算法来优化模型参数,使模型达到最佳的检测精度。Masita 等^[11]提出一种通过融合卷积神经网络(CNN)与残差网络的方法对红外图像热差进行缺陷检测;赵晓雨等^[12]提出了一种融合注意力机制的 YOLOv5 改进算法,将有效通道注意力 ECA^[13]与 YOLOv5 模型主干网络中的 C3 模块相融合形成 C3-ECA 模块,有效提升了模型的检测精度;焦思韬等^[14]提出了一种改进的 SSD 算法,使用 Involution^[15]算子替换原 SSD 中的 Conv7 卷积层以及使用 FocalLoss^[16]损失函数,提高了准确度。上述算法的检测精度虽有提升,但仍存在模型参数多、计算量大等问题,对于检测模型的轻量化仍有较大改进空间。

基于上述研究存在的局限性,本文针对计算资源受限

的条件,兼顾检测精度和模型体积的问题,从模型轻量化角度出发,提出了一种基于 YOLOv8 的轻量级分布式光伏电池板缺陷检测算法。在保证检测精度的同时,降低计算量、参数量,减小模型体积,减少背景噪声对缺陷检测的干扰,本文主要贡献有:

1)采用轻量化网络结构 StarNet 作为特征提取网络,去除特征提取网络中大量的分支结构,有效降低了模型的复杂度。

2)在特征提取网络中引入注意力机制 Triplet^[17]注意力机制,提高了网络对多尺度目标特征的提取能力。

3)设计 SPPF-AM 模块,提升模型在空间信息方面的感知能力,并增强检测模型对缺陷目标的空间定位能力,从而更好地适应不同尺度的目标。

4)在特征融合网络中引入 C2f_DSCConv2D 结合可变形卷积取代原网络中的 C2f,在确保模型性能的同时,降低了模型的计算复杂度和参数量,提升了模型的运行效率。

5)在特征融合网络中引入空间上下文感知模块 SCAM,自适应地融合不同层次的语义信息,对齐局部特征,减少微小微弱缺陷信息的丢失,有效抑制图像中的噪声,抑制无关背景信息的干扰。

6)设计 ECIoU 替换 CIoU,增强边界框损失的拟合能力,并加速网络模型的收敛速度。

1 改进 YOLOv8 算法

YOLOv8 目标检测算法作为深度学习领域经典的一阶段目标检测方法,因其简洁高效、实时性强和多尺度检测等特点而脱颖而出。其独特之处在于能够充分利用全局上下文信息和多任务学习,从而在快速目标检测和实时场景中表现出色。通过优化网络结构和算法设计,提高了检测精度,加快了处理速度,为目标识别提供了可靠解决方案。

本文通过改进基于 YOLOv8n 模型的结构,以提升光伏电池板缺陷检测的效果和性能。首先,在主干网络中,引入轻量化 StarNet 作为特征提取网络,增强对特征信息的提取能力。StarNet 有助于在确保检测精度的前提下,降低模型的计算复杂度,进而提升整体运行效率。为了进一步提升网络对多尺度目标特征的提取能力,引入注意力机制 Triplet。调控特征图中的通道交互,增强了模型在不同尺度缺陷检测中的适应性和准确性。在特征提取网络中设计 SPPF-AM 模块,增强模型对空间信息的感知能力。有效提升检测模型对光伏电池板缺陷目标的空间定位精度,尤其是在处理不同尺度目标表现出色。在特征融合网络中设计 C2f_DSCConv2D 取代 C2f,在降低模型的计算复杂度和参数量的同时,也确保模型性能和运行效率。在特征融合网络中引入空间上下文感知模块 SCAM,该模块自适应地融合不同层次的语义信息,减少微小缺陷信息的丢失,同时有效抑制图像中的噪声,降低无关背景信息的干扰。最后,

为提高边界框损失的拟合能力和加速网络模型的收敛速度,设计 ECIoU 替换 CIoU。有效提升了模型在边界框预

测和目标定位方面的精度和稳定性,有效加速了整体训练过程。改进后的 YOLOv8 网络结构如图 1 所示。

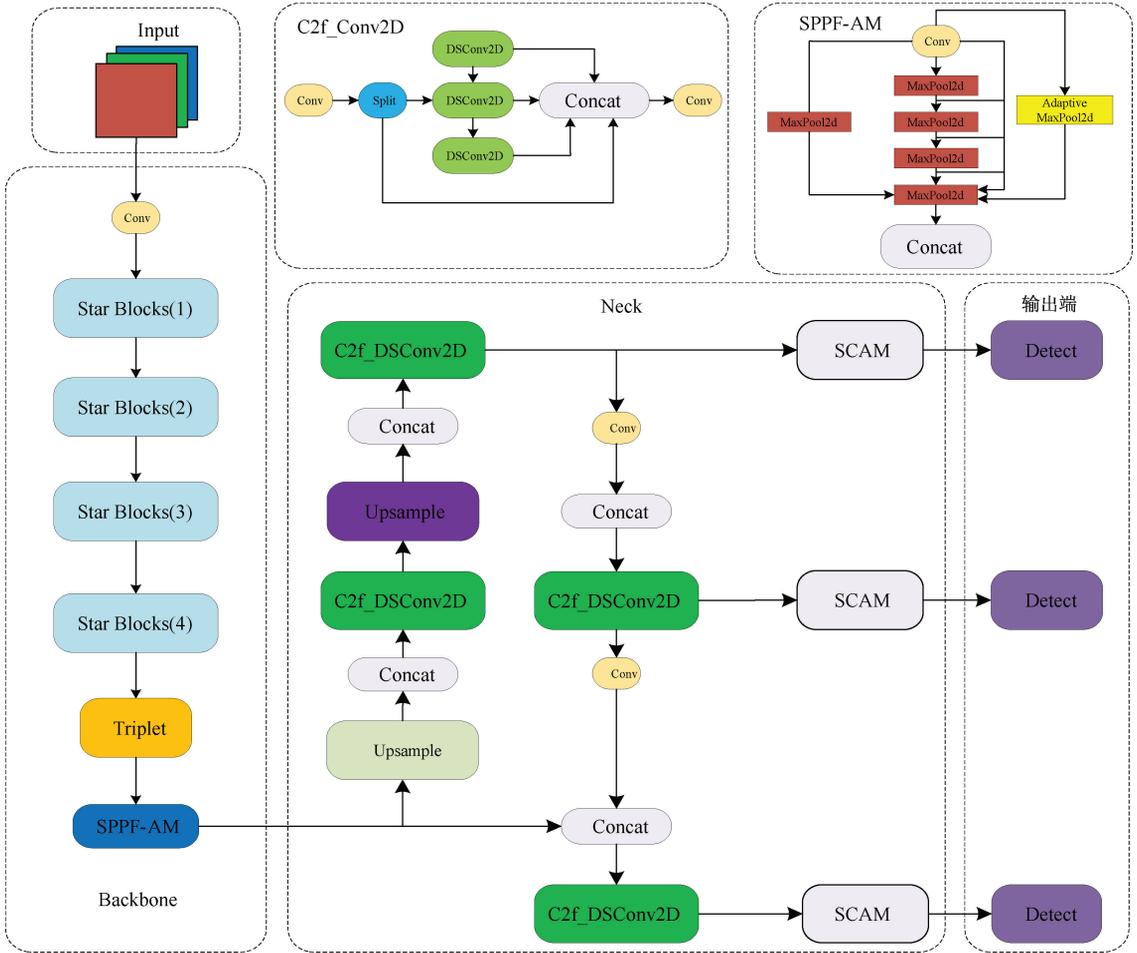


图 1 SN-YOLOv8 网络结构

Fig. 1 SN-YOLOv8 network structure

1.1 StarNet

在缺陷检测模型中,通常需要对主干网络进行轻量化操作。然而,现有的轻量化方法并未显著减少模型的复杂度、体积以及计算资源的使用,因此难以达到最优的轻量化效果。为了解决这一问题,本文利用快速特征提取网络 StarNet 替换 YOLOv8n 的主干网络。通过采用 StarNet,旨在简化主干网络的结构,同时有效减少计算资源的消耗,从而实现模型的轻量化。这样的设计旨在保证缺陷检测模型的精度不受影响,同时解决了主干网络复杂和计算资源受限的问题。

本文以 StarNet 替换原 YOLOv8n 中的主干网络,StarNet 由一个 4 阶层次结构构建而成,利用卷积层进行下采样,并利用改进的演示块 Demo 进行特征提取。为了满足效率的要求,用 BatchNormalization 替换 LayerNormalization,并将其放置在深度卷积之后,以便在推断期间可以融合。结合 MobileNeXt 的设计,在每个块的末尾加入了深度卷积 DW-Conv。通道扩展因子始终设

置为 4,网络宽度在每个阶段加倍。在 demo 演示块中,GELU 激活被替换为 ReLU6,遵循 MobileNetv2 的设计。由此构成本算法的轻量化特征提取网络,用于快速且有效地提取光伏板缺陷特征。快速特征提取网络 StarNet 整体架构如图 2 所示。

结构图中 Element-wise mul 为元素级乘法,通过逐元素相乘从而融合不同的子空间特征。元素级乘法符号类似于星星,将其命名为星操作(star operation)。当星操作被纳入神经网络并堆叠多个层时,每一层都会导致隐式维度复杂性的指数增长。仅仅通过几层,星操作就能在一个紧凑的特征空间内实现几乎无限的维度。

在神经网络的单层中,星操作通常被写为 $(W_1^T X + B_1) * (W_2^T X + B_2)$,这表示通过逐元素乘法融合两个线性变换的特征。为了方便起见,将权重矩阵和偏置合并为一个实体,记作 $W = [W, B]^T$,同样地 $X = [X, 1]^T$,从而简化星操作为 $(W_1^T X) * (W_2^T X)$,为了简化,关注涉及一个输出通道变换和单元元素输入的情况。具体来说,定义 ω_1 ,

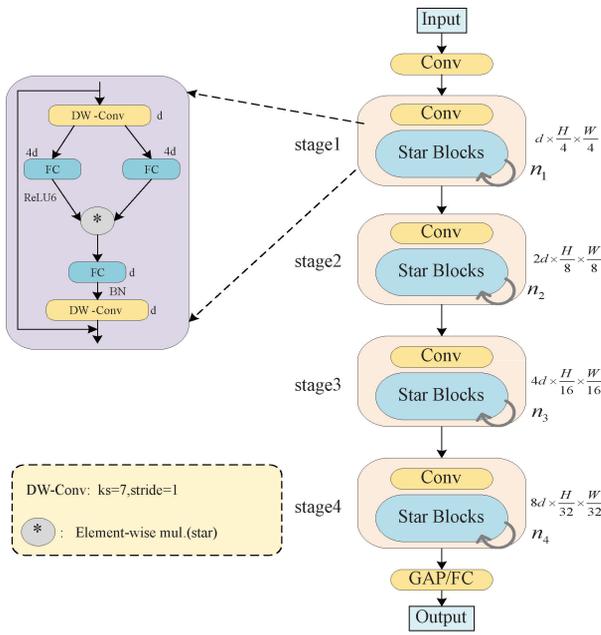


图 2 StarNet 结构图

Fig. 2 StarNet structure diagram

$\omega_2, x \in \mathbf{R}^{(d+1) \times 1}$, 其中 d 是输入通道数。它可以很容易地扩展到多个输出通道 $\mathbf{W}_1, \mathbf{W}_2 \in \mathbf{R}^{(d+1) \times (d+1)}$ 和处理多个特征元素, 其中 $\mathbf{X} \in \mathbf{R}^{(d+1) \times n}$, 一般来说, 可以将星操作重写为:

$$\omega_1^T x * \omega_2^T x \tag{1}$$

$$= \left(\sum_{i=1}^{d+1} \omega_1^i x^i \right) * \left(\sum_{j=1}^{d+1} \omega_2^j x^j \right) \tag{2}$$

$$= \sum_{i=1}^{d+1} \sum_{j=1}^{d+1} \omega_1^i \omega_2^j x^i x^j \tag{3}$$

$$= \alpha_1 x^1 + \dots + \alpha_{(4,5)} x^4 x^5 + \dots + \alpha_{(d+1, d+1)} x^{d+1} x^{d+1} \tag{4}$$

式中: 使用 i, j 来索引通道, α 是每个项的系数:

$$\alpha_{(i,j)} = \begin{cases} \omega_1^i \omega_2^j, & i = j \\ \omega_1^i \omega_2^j + \omega_1^j \omega_2^i, & i \neq j \end{cases} \tag{5}$$

在将式(1)中的星操作重写后, 将其扩展为一个由 $(d+2)(d+1)/2$ 个不同项组成的组合, 如式(4)所示。值得注意的是, 除了 $\alpha_{(d+1, d+1)} x^{d+1} x$ 项外, 每个项与 x 都表现出非线性关联, 表明它们是独立且隐式的维度。因此, 使用计算效率高的星操作在 d 维空间中执行计算, 却能在 $(d+2)(d+1)/2 \approx (d/\sqrt{2})^2$ (考虑 $d \gg 2$) 隐式维特征空间中实现表示, 显著放大了特征维度, 而无需在单个层中增加任何额外的计算开销。

接下来, 展示通过堆叠多层, 以递归的方式将隐式维度几乎无限地指数增长。考虑一个初始网络层, 其宽度 d , 应用一个星操作会得到表达式 $\sum_{i=1}^{d+1} \sum_{j=1}^{d+1} \omega_1^i \omega_2^j x^i x^j$, 如式(3)中所述。这导致在隐式特征空间 $\mathbf{R}(d/\sqrt{2})^{2^l}$ 中的表示。

令 Q_l 表示第 l 个星操作的输出, 得到:

$$O_1 = \sum_{i=1}^{d+1} \sum_{j=1}^{d+1} \omega_{(1,1)}^i \omega_{(1,2)}^j x^i x^j \in \mathbf{R}^{2^1} \tag{6}$$

$$O_2 = \mathbf{W}_{2,1}^T O_1 * \mathbf{W}_{2,2}^T O_1 \in \mathbf{R}(d/\sqrt{2})^{2^2} \tag{7}$$

$$O_3 = \mathbf{W}_{3,1}^T O_2 * \mathbf{W}_{3,2}^T O_2 \in \mathbf{R}(d/\sqrt{2})^{2^3} \tag{8}$$

⋮

$$O_l = \mathbf{W}_{l,1}^T O_{l-1} * \mathbf{W}_{l,2}^T O_{l-1} \in \mathbf{R}(d/\sqrt{2})^{2^l} \tag{9}$$

也就是说, 通过 l 层, 可以隐式地获得一个属于 $\mathbf{R}(d/\sqrt{2})^{2^l}$ 的特征空间。例如, 给定一个宽度为 128 的 10 层等宽网络, 通过星操作实现的隐式特征维度数量近似为 90^{1024} , 这可以合理地近似为无限维度。因此, 通过堆叠多层, 即使只有几层, 星操作也可以以指数方式显著放大隐式维度。StarNet 通过星操作 (star operation) 实现高维和非线性特征空间的映射, 而无需增加计算复杂度。

1.2 Triplet

为了增强对小目标的特征提取能力, 引入了 Triplet 注意力机制到主干网络的末端。Triplet 能够有效捕获不同维度之间的交互信息, 减少无关信息的干扰, 从而使模型更加集中于提取目标特征, 提升网络模型对密集分布小目标的检测能力^[18], 提高识别准确率。Triplet 的结构如图 3 所示。

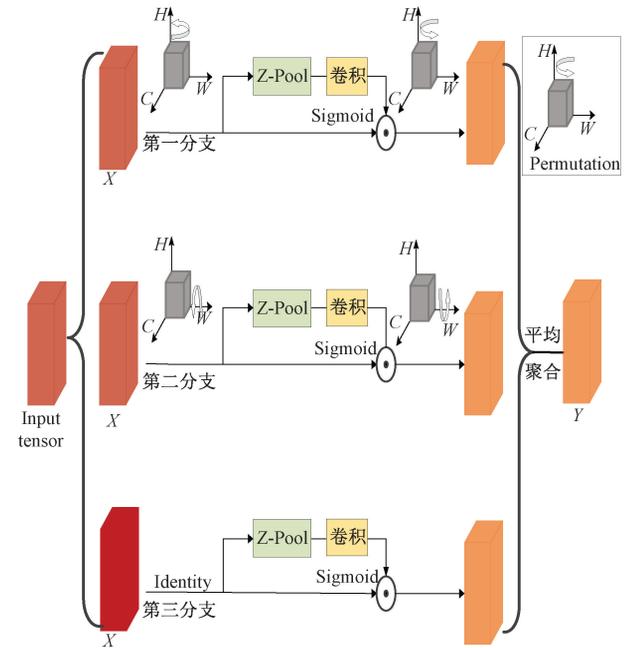


图 3 Triplet 结构图

Fig. 3 Triplet structure diagram

Triplet 由 3 个平行的分支组成, 每个分支接受输入张量 $\mathbf{X} \in \mathbf{R}^{C \times H \times W}$ 。在第 1 个分支中, 张量 \mathbf{X} 经过绕维度 H 逆时针旋转 90° 操作, 然后通过 Z-Pool 操作得到形状为 $2 \times H \times C$ 的张量, 经过卷积和 Sigmoid 激活函数生成注意力

权重,最后绕维度 H 顺时针旋转 90° ,输出张量 \hat{X}_1^* ,实现通道 C 和维度 H 之间的交互。第2个分支中,张量 H 绕维度 W 逆时针旋转 90° ,经过 Z-Pool 操作、卷积和 Sigmoid 激活函数后,绕维度 W 顺时针旋转 90° ,输出张量 \hat{X}_2^* ,完成通道 C 与空间维度 W 的交互。第3个分支经过 Z-Pool 操作、卷积和 Sigmoid 激活函数后,直接输出张量 \hat{X}_3^* 。最后将这3个张量进行平均聚合,输出的张量为:

$$Y = \left(\frac{1}{3} \hat{X}_1^* \sigma(\psi_1(\hat{X}_1^*)) + \hat{X}_2^* \sigma(\psi_2(\hat{X}_2^*)) + X \sigma(\psi_3(\hat{X}_3^*)) \right) \quad (10)$$

式中: σ 为 Sigmoid 激活函数;而 ψ_1, ψ_2, ψ_3 则表示标准卷积层。

1.3 SPPF-AM

YOLOv8 中采用的 SPPF 结构是在 SPP 结构的基础上引入特征融合模块改进而来,提升了模型的感知能力和检测性能。具体而言,首先对输入特征图进行多种尺寸的池化操作,随后通过卷积操作将各个尺度的池化结果进行融合,最终输出融合后的特征图。SPPF 结构的特点在于能够自适应地整合不同尺度的特征信息,从而具备优秀的特征提取能力,但是 SPPF 更多地关注缺陷图像的边缘位置信息而忽略背景信息^[19],从而影响模型的准确率,可做出进一步的改进。通过加入最大池化层和自适应最大池化层来帮助网络更好的做出判断,从而降低无用信息的干扰,突出有效信息^[20]。

最大池化层 MaxPool2d 在深度学习模型中被广泛应用,通过滑动窗口操作在每个局部区域中选择最大值作为输出。这种操作不仅能有效提取输入数据的关键特征,还实现了数据的空间降采样,即减少数据的空间尺寸,从而减少模型的参数数量和计算量。这种降采样过程有助于在保持重要信息的同时,提高模型的计算效率和推理速度。最大池化层的选择机制确保了输出特征包含每个窗口区域的最显著特征,这对于图像处理和视觉识别任务尤为重要。通过选择最大值而非平均值或其他统计量,最大池化层能够更好地捕捉到图像中的显著纹理、形状和模式,从而增强模型对于视觉特征的感知能力。图4展示了最大池化层的计算过程。

自适应最大池化层 AdaptiveMaxPool2d 是池化层的一种进化形式。它能够在保持输入张量形状不变的情况下,自动调整池化窗口的大小,以便在每个通道中找到整个输入张量的最大值。这种池化方法对每个输入通道分别计算最大值,并将这些最大值用于生成每个通道的输出。在卷积神经网络中,自适应最大池化通过减小特征图的空间维度,这种方法显著减少了模型的计算复杂度,对处理大规模数据和提升模型效率至关重要。与传统的固定大小池化相比,自适应最大池化能够根据输入张量的大小和形状动态调整池化窗口,因此在不同尺度和分辨率的

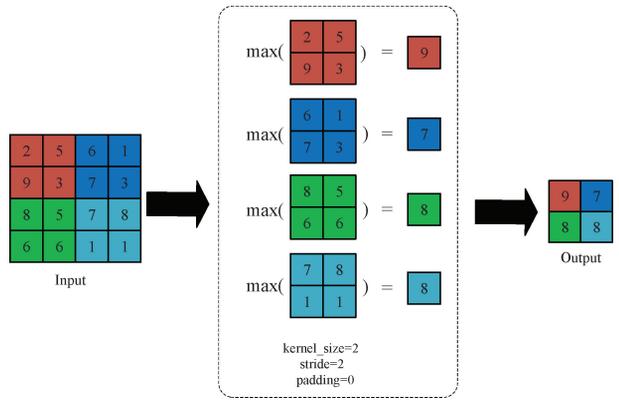


图4 最大池化层计算过程

Fig. 4 Calculation process of the max pooling layer

输入上表现更加灵活和适应性强。这种灵活性使得模型能够更好地捕捉到输入数据中的关键特征,从而提高了整体的特征提取能力和模型的识别精度。还能够有效避免信息丢失的问题。图5为自适应最大池化层的计算过程。

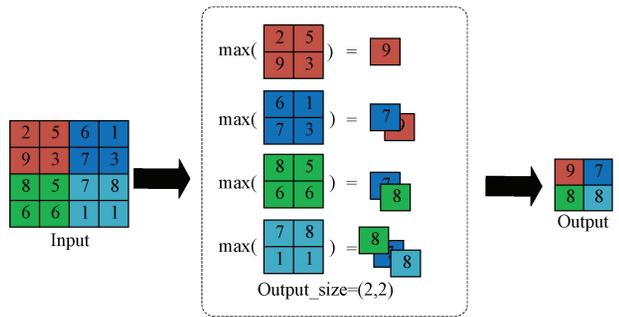


图5 自适应最大池化层计算过程

Fig. 5 Calculation process of the adaptive max pooling layer

为了确保在光伏板缺陷识别任务中达到实时性,本文在 SPPF 模块的基础上引入 AdaptiveMaxPool2 自适应最大池化层和 MaxPool2d 最大池化层,命名为 SPPF-AM。以增强模型的全局视角信息获取能力,并有效减轻不同尺度大小对模型表现的影响。具体来说,输入端的 Conv 层结合了 AdaptiveMaxPool2 和 MaxPool2d,这两种池化层可以分别根据输入的尺寸自适应地进行池化操作,从而有效地捕获不同尺度下的特征。最后,通过 Concat 将两种池化层的输出连接起来,使得模型能够综合利用池化后的全局特征信息。这种设计显著增强了模型的特征表达能力和感知能力。通过引入自适应最大池化层和最大池化层操作,SPPF-AM 模块不仅能够更好地处理光伏板缺陷识别任务中的多尺度特征,还能有效提升模型的计算效率和实时性,从而使得系统在实际场景中能够更快速、更精确地进行缺陷检测和识别。SPPF-AM 的结构如图6所示。

1.4 C2f_DSConv2D

由于深度卷积神经网络深度和参数数量的增加,导致缺陷检测模型存在内存占用大和运行速度慢的问题。因此,

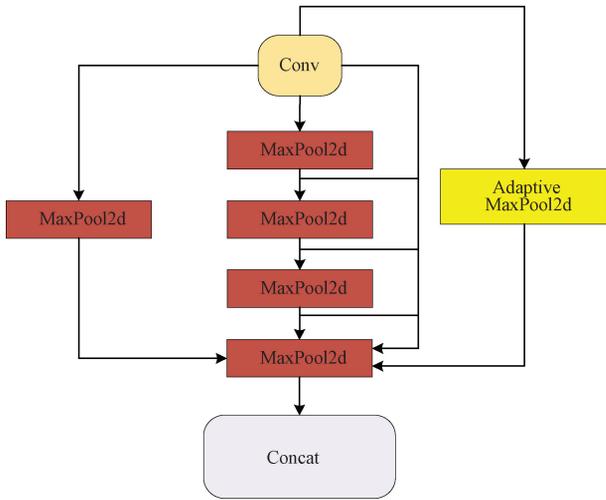


图 6 SPPF-AM 模块
Fig. 6 SPPF-AM module

引入一种深度可分离卷积变体-分布移位卷积(DSConv^[21]), 以实现使用较低的存储和较高的计算速度, 从而提高缺陷检测模型的效率。

DSConv 模块集成了可变量化核(VQK)与分布移位两大组件, DSConv 结构如图 7 所示。可变量化核是 DSConv 的量化组件, 它存储可变长度的整数值, 一旦设置即不可更改。这一设计旨在减少计算和存储需求, 特别适用于计算资源有限的环境。分布移位组件包括两个部分: 核分布移位器(KDS)和信道分布移位器(CDS)。核分布移位器中的每个值用于对 VQK 中的每个块(BLK)的深度进行量化, 实现卷积核的分布移位。信道分布移位器则负责调整每个块中通道的分布, 从而优化特征的表达。整个组件的目标是生成与原始卷积相同输出的新型卷积结构。这种结构不仅保持了与原始卷积相等的输出质量, 还显著减小了卷积核的大小, 从而进一步减少卷积层中的参数数量和计算复杂度^[22]。显著提高运算速度和降低内存使用量。

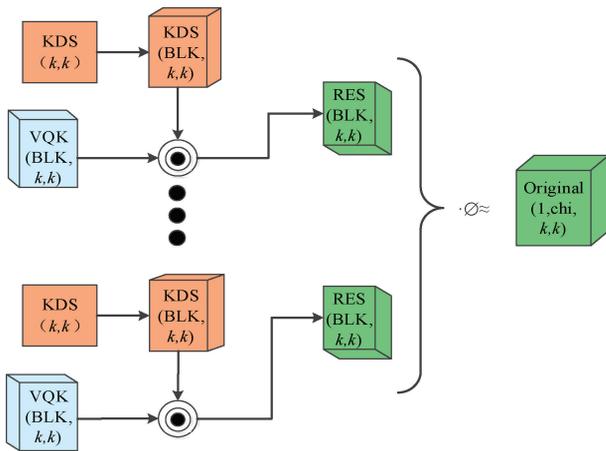


图 7 DSConv 结构图
Fig. 7 DSConv structure diagram

基于分布移位卷积与 C2f 模块相融合, 设计新的模块 C2f_DSConv2D, 如图所示。该模块在确保缺陷检测模型轻量化的同时, 能够获取更加丰富的梯度流信息, 并提高缺陷检测模型计算资源利用率, 从而实现缺陷检测模型整体性能的提升。C2f_DSConv2D 结构如图 8 所示。

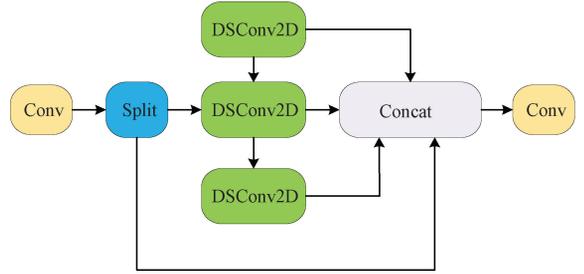


图 8 C2f_DSConv2D 模块
Fig. 8 C2f_DSConv2D module

1.5 SCAM

空间上下文感知模块 SCAM 利用全局上下文信息来表示像素之间的跨空间关系, 抑制图像中的噪声, 抑制无用背景, 增强目标和背景的区分能力, SCAM 受到 GCNet 与 SCP 的启发。SCAM 由 3 个分支组成, 第 1 个分支使用 GAP(全局平均池化)与 GMP(全局最大池化)来整合全局信息, 第 2 个分支使用 1×1 的卷积来生成特征映射的线性变换结果, 第 3 个分支使用 1×1 卷积来简化键值对的倍数。最后将第 1 分支与第 3 分支分别于第 2 分支进行矩阵相乘。得到的两个分支分别表示跨通道以及空间上的上下文信息, 随后利用 Hadamard 乘积在这两个分支上得到了 SCAM 的输出。SCAM 的结构如图 9 所示。

在每一层中, 像素空间上下文可以表示为:

$$Q_i^j = P_i^j + a_i^j \sum_{j=1}^{N_i} \left[\frac{\exp(\omega_{qk} P_i^j)}{\sum_{n=1}^{N_i} \exp(\omega_{qk} P_i^n)} \cdot \omega_v P_i^j \right] \quad (11)$$

$$a_i^j = \frac{\exp([\text{avg}(P_i); \max(P_i)] P_i^j)}{\sum_{n=1}^{N_i} \exp([\text{avg}(P_i); \max(P_i)] P_i^n)} \cdot \omega_v \quad (12)$$

式中: P_i^j 和 Q_i^j 分别表示第 i 级特征映射中第 j 个像素的输入和输出。 N_i 表示像素总数。 ω_{qk} 和 ω_v 是用于投影特征映射的线性变换矩阵, 它们通过 1×1 卷积进行简化。 $\text{avg}(\cdot)$ 和 $\text{max}(\cdot)$ 分别执行 GAP 和 GMP。 GAP 和 GMP 可以引导 feature map 选择具有重要信息的通道, 使 SCAM 能够了解通道尺寸的上下文信息。

1.6 ECIOU

在 YOLOv8 模型中, 使用了 DFL Loss 和 CIOU Loss 作为关键的边界框回归损失函数。 DFL Loss 主要有助于快速定位各种目标及其周围区域, 这种损失函数的设计使得检测网络能够更有效地捕捉到目标的位置信息。 而 CIOU Loss 则进一步评估边界框预测的精确性, 确保预测的边界框与实际目标的位置和形状更加贴合。 然而, 光伏

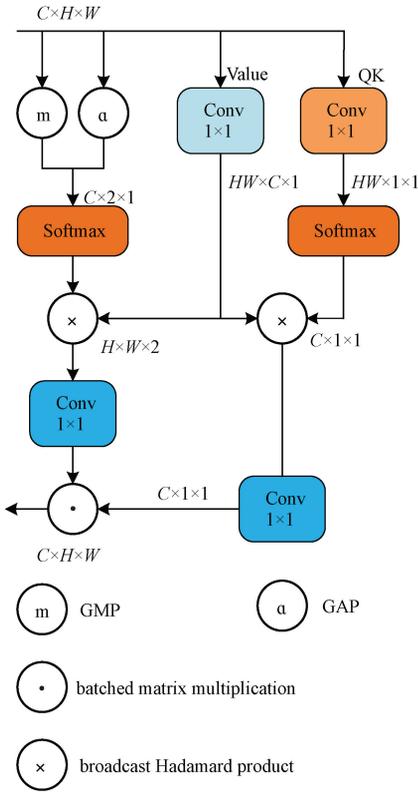


图9 SCAM模块
Fig.9 SCAM module

板表面缺陷数据集通常包含多样化的缺陷类型,这些缺陷可能具有不同的尺寸、形状和分布特征。这种多样性和复杂性会对模型的边界框预测造成挑战,因为模型需要能够有效地适应各种不同特征的缺陷目标。为了解决这个问题,设计ECIoU Loss来替代传统的CIoU Loss损失函数。ECIoU Loss进一步提高普通质量锚框的预测精度,并增强边界框损失函数的拟合能力。通过这种优化,网络模型能够更快地收敛,提高在复杂场景下的检测准确性和稳定性。ECIoU结合CIoU和EIoU两种损失函数,CIoU损失函数公式如下:

$$CIoU = 1 - IOU + \frac{\rho^2(b, b^{gt})}{c^2} + \alpha v \quad (13)$$

式中: b 和 b^{gt} 分别代表预测框和真实框的中心点坐标。 ρ 则表示这两个中心点之间的欧氏距离, c 指的是能够同时完全覆盖(即包含)预测框与真实框的最小闭合矩形的对角线长度, α 是惩罚项的控制参数,定义如下:

$$\alpha = \frac{v}{(1 - IOU) + v} \quad (14)$$

式中: v 是衡量长宽一致性的参数,定义如下:

$$v = \frac{4}{\pi^2} (\arctan \frac{\omega^{gt}}{h^{gt}} - \arctan \frac{\omega}{h})^2 \quad (15)$$

式中: (ω, h) 和 (ω^{gt}, h^{gt}) 代表预测框和真实框的尺寸。

EIoU损失函数的公式如下:

$$EIoU = 1 - IOU + \frac{\rho^2(b^{gt}, b)}{c^2} + \frac{\rho^2(h^{gt}, h)}{c_h^2} + \frac{\rho^2(\omega^{gt}, \omega)}{c_w^2} \quad (16)$$

式中: c_w 和 c_h 是覆盖两个框的最小外接框的宽度和高度。

将CIoU损失函数和EIoU损失函数结合为ECIoU损失函数,具体公式如下:

$$ECIoU_Loss = 1 - IoU + \alpha v + \frac{\rho^2(b^{gt}, b)}{c^2} + \frac{\rho^2(h^{gt}, h)}{c_h^2} + \frac{\rho^2(\omega^{gt}, \omega)}{c_w^2} \quad (17)$$

2 实验结果与分析

2.1 实验配置

实验在GPU为GTX4060、内存16GB的计算机上运行。操作系统为Windows 10, Python版本为3.8, PyTorch版本为1.12, CUDA版本为12.2。训练时数据集的Batch size设置为16,总共进行了200个epochs的训练。优化器选用SGD,初始学习率设为0.01,动量为0.9,权重衰减为0.001。随着迭代的进行,学习率线性衰减至0。

实验环境如表1所示。

表1 实验环境

Table 1 Experimental environment

名称	环境参数
操作系统	Windows 10
CPU	Intel Core i5-12000F
GPU	Nvidia GeForce RTX 4060
内存	16 GB
Python	3.8
Pythorch	1.12
Cuda	12.2

2.2 实验数据集

基于没有公开的光伏电池板缺陷数据集,本文利用大疆M300型无人机搭载禅思H20T相机进行光伏电池板图像采集,数据集使用Labelimg进行手动标记,生成对应的标签用于训练。该数据集涵热斑、物理损坏、电气损坏、异物覆盖、脏污5类缺陷,共计12533张图片,数据集详细信息如表2所示。考虑到标签与数据的对应关系,为保证数据集分布均匀,将数据集按70%、20%、10%的比例,分为训练集、测试集和验证集。

2.3 评价指标

本次实验评估指标包括精确度(precision, P)、召回率(recall, R)、平均精度(mean average precision, mAP)、模型参数量(Param)、以及模型计算量(GFLOPs)。精确度(Precision)是在所有被模型预测为正类别的样本中,属

表 2 数据集详细信息

Table 2 Detailed information of the dataset

缺陷名称	数量
热斑	1 506
物理损坏	2 200
电气损坏	1 880
异物覆盖	3 900
脏污	3 047

于正类别的样本所占的百分比。召回率(Recall)是指模型正确找出所有正类别样本的能力,公式如下:

$$P = \frac{TP}{TP + FT} \quad (18)$$

$$R = \frac{TP}{TP + FN} \quad (19)$$

式中: TP 表示模型正确地将真实的正例样本分类为正例, FP 表示模型错误地将真实的负例样本分类为正例, FN 表示模型错误地将真实的正例样本分类为负例。 AP 指单个类别在不同召回率下的平均精确率,其计算公式如下:

$$P_{AP} = \int_0^1 P(R) dr \quad (20)$$

mAP 是所有类别的平均精确率的总和除以所有类别的总数,其计算公式如下:

$$P_{mAP} = \frac{1}{n} \sum_{i=1}^n P_{AP} \quad (21)$$

模型计算量和模型参数量分别是评估模型计算效率和模型容量的关键指标。

2.4 迁移学习

迁移学习(transfer learning)是一种机器学习方法,利用已经在一个任务上训练好的模型作为起点,来解决另一个类似但不完全相同的任务。在深度学习领域,特别是计

算机视觉任务中,迁移学习已经成为一种非常有效的技术。卷积神经网络往往需要大量的标注数据和计算资源,迁移学习可利用预训练模型学到的特征,减少对大量标注数据的需求,加快训练速度并提高模型性能。

首先构建 YOLOv8n 缺陷检测模型,将 COCO2017 数据集用于模型预训练,为识别光伏电池板缺陷检测提供最优化的初始化模型权重参数。之后将预训练得到的模型权重参数迁移至光伏电池板缺陷检测任务中,实现相似领域的知识迁移;最后使用自建数据集样本对 SN-YOLOv8 缺陷检测模型进行微调,并通过测试集样本进行验证,获取最终的光伏电池板缺陷检测模型。

2.5 消融实验

为了评估引入的模块及其不同组合顺序对算法性能优化的影响程度,设计了一系列消融实验进行分析。每个实验的环境和参数设置均保持统一。消融实验结果如表 3 所示。实验 1 使用未经修改的 YOLOv8n 算法,其 mAP 值为 85.9%。实验 2 则采用了 StarNet(SN)轻量化主干网络,替代了原有的 YOLOv8n 结构网络。这种改变使得模型参数量和计算量分别减少了 26.3%、19.8%,mAP 达到 85.6%。实验 3 和 4 分别加入 Triplet、SPPF-AM 模块,虽然这导致了 Params 和 GFLOPs 的增加,但 mAP 值相较于 YOLOv8n 算法提高了 0.5%。实验 5 引入 C2f_DSCConv2D 模块取代原网络中的 C2f,使得模型参数量和计算量分别减少了 9.9%、5.6%。mAP 值达到 86.4%。实验 6 引入 SCAM 模块,mAP 值达到 86.9%。实验 7 将损失函数 ECIoU 应用于 YOLOv8n 网络结构,加快了网络模型的收敛速度,mAP 值达到 87.6%。最终的实验 12 结果表明,通过加入 StarNet 轻量化主干网络,SPPF-AM 模块,Triplet 注意力机制、C2f_DSCConv2D 模块,SCAM 模块,损失函数 ECIoU,累积效应显著,mAP 值提高到了

表 3 消融实验

Table 3 Ablation experiment

序号	SN	Triplet	SPPF-AM	C2f_DSCConv2D	SCAM	ECIoU	FLOPs/G	R/%	mAP@0.5/%	Params/M
1	×	×	×	×	×	×	8.1	83.9	85.9	5.7
2	√	×	×	×	×	×	6.5	84.2	85.6	4.2
3	×	√	×	×	×	×	21.5	84.5	86.4	11.7
4	×	×	√	×	×	×	8.3	84.4	85.1	5.9
5	×	×	×	√	×	×	7.3	84.4	86.4	5.4
6	×	×	×	×	√	×	8.0	85.2	86.9	5.8
7	×	×	×	×	×	√	8.2	85.6	87.6	5.7
8	√	√	×	×	×	×	15.4	82.1	84.1	8.0
9	√	√	√	×	×	×	15.5	82.0	84.5	8.0
10	√	√	√	√	×	×	5.7	83.2	85.0	3.9
11	√	√	√	√	√	×	5.7	86.5	88.6	3.9
12	√	√	√	√	√	√	5.7	87.1	89.7	3.7

89.7%,模型参数数量和计算量分别减少了35%、29.6%。通过一系列的消融实验,可以清楚地看到每个改进点对整体检测效果的贡献,改进后的算法既降低了参数量,还提升了检测性能,具有一定的可行性。

2.6 对比实验

为了进一步验证改进模型的检测性能,本文与其他深度学习模型在同一数据集上进行对比,结果如表4所示。

表4 模型对比实验结果

Table 4 Results of model comparison experiments

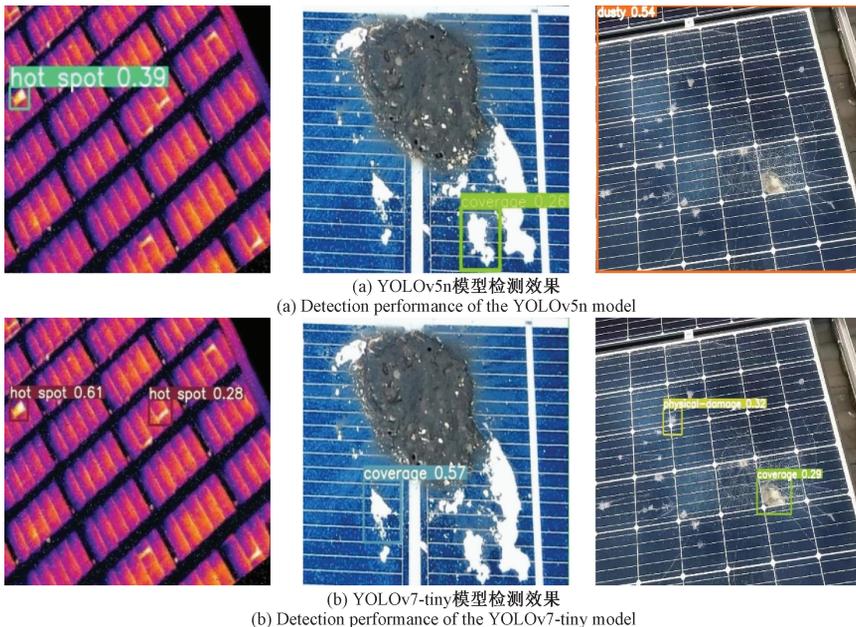
算法	P	R	mAP@0.5	Params/M	GFLOPs
Faster R-CNN	56.3	74.1	76.3	108.2	302.8
SSD	68.5	73.8	78.5	93.1	150.5
EfficientDet ^[23]	77.3	75.2	80.4	70.7	90.8
YOLOv5n	80.6	85.7	87.7	4.0	8.6
YOLOv7-tiny	79.3	65.4	72.3	11.5	13.2
YOLOv8n	80.6	83.9	85.9	5.7	8.1
YOLOv8s	81.2	84.7	89.0	21.2	28.4
YOLOv9t	82.2	85.3	86.4	5.0	10.7
RE-DETR	83.3	84.2	87.5	38.3	230.6
本文	90.1	87.1	89.7	3.7	5.7

在众多模型中,Faster R-CNN模型和SSD模型的检测精度相对较低,模型参数数量和计算量较大,需要更多的计算资源和时间进行训练和推理。EfficientDet和RE-DETR相较于Faster R-CNN和SSD,检测精度提高,但模型过于冗余,由于计算量庞大,不利于在计算资源受限的设备上进行部署。YOLOv5n、YOLOv8n、YOLOv9t模型具有较少的模型参数,计算复杂度低,但检测精度有待提高;而本文提出的模型相较以上算法相比,在确保小模型、低计算量的前提下,检测精度取得了明显提升,达到了90.1%,mAP@.5达到了89.7%,模型参数数量和计算量分别为3.9M、5.7G。在保持较高检测精度的同时,也实现了轻量化。在满足分布式光伏板缺陷检测需求的同时,也考虑嵌入式设备中实际应用,增强了其实用性。符合本

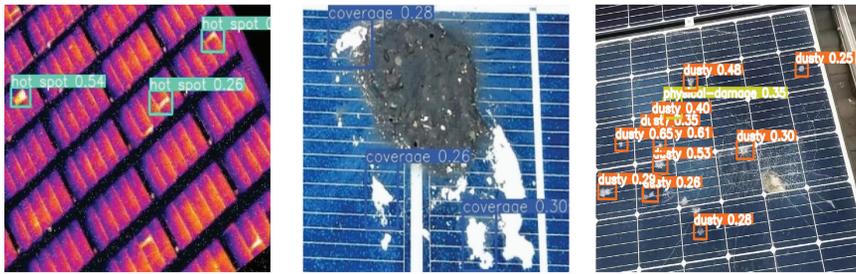
次改进的预期效果和实际应用的需求。

2.7 光伏板缺陷检测可视化及分析

为了评估改进模型的具体检测性能,将本文改进的模型与其他几种主流模型检测效果进行对比。各模型检测效果如图10所示。YOLOv5n、YOLOv7-tiny、YOLOv8n、YOLOv8s、YOLOv9t模型都出现缺陷目标漏检的问题,其中YOLOv5n检测效果差,YOLOv7-tiny、YOLOv8s与YOLOv9t模型出现了对同一目标多次检测导致检测框重叠的现象以及对缺陷目标的误检现象。本文提出的模型成功检测到了所有的目标,无漏检和误检现象的出现。改进后的算法在感知能力方面具有明显优势,能够检测出更多目标,具有更低的误检和漏检概率;在检测性能方面同样出色,具有更高的检测精度,同时具有较强的泛化能力。

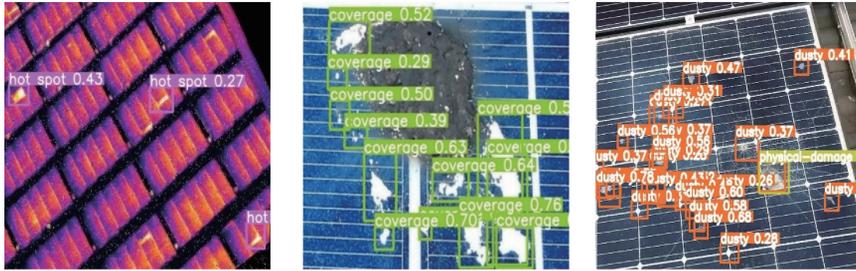


(b) YOLOv7-tiny模型检测效果
(b) Detection performance of the YOLOv7-tiny model



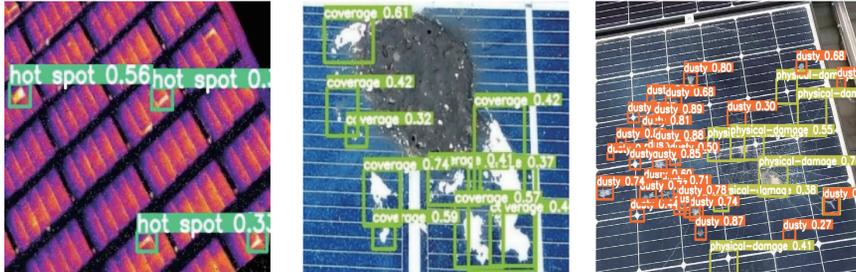
(c) YOLOv8n模型检测效果

(c) Detection performance of the YOLOv8n model



(d) YOLOv8s模型检测效果

(d) Detection performance of the YOLOv8s model



(e) YOLOv9t模型检测效果

(e) Detection performance of the YOLOv9t model



(f) 本文模型检测效果

(f) Detection performance of the model in this paper

图 10 各模型检测效果图

Fig. 10 Detection performance of various models

2.8 泛化性实验

为了验证 SN-YOLOv8 算法的泛化能力,本文选择采用通用目标检测公开数据集 PVELAD 对 SN-YOLOv8 的检测性能进行泛化性验证,PVELAD 包含 36 543 张具有各种内部缺陷和异构背景的近红外图像,其中包含 1 类无异常图像和具有 12 个不同类别的异常缺陷图像。运行环境保持一致的情况下,结果如表 5 所示。通过对比 YOLOv8n 和 SN-YOLOv8 的性能,可以观察到,SN-YOLOv8 在有效降低模型的参数量和计算量的同时,能够最大限度地保留其精度。这一结果不仅验证了所提出算

法的有效性,还展示了 SN-YOLOv8 在实际应用中的卓越泛化能力。

表 5 PVELAD 数据集实验结果对比

Table 5 Comparison of experimental results on the PVELAD dataset

模型	P	R	mAP@0.5/%	Params/M
YOLOv8n	81.2	84.0	85.0	5.7
本文	87.6	86.4	88.9	3.9

3 结 论

本文提出一种轻量级分布式光伏电池板缺陷检测算法 SN-YOLOv8。以 YOLOv8n 为基线模型,主干网络部分,采用轻量化网络结构 StarNet 作为特征提取网络,减少计算量和参数量,实现高效率和高性能之间平衡;引入三元组注意力机制 Triplet,提高了网络对多尺度目标特征的提取能力,有助于在检测过程中更准确地识别光伏电池板上的缺陷,从而提升检测精度;特征提取网络最后一层设计 SPPF-AM 模块,获取全局视角信息并减轻不同尺度大小所带来的影响,从而提高了模型处理各种尺寸缺陷的精度;特征融合网络中设计 C2f_DSCConv2D 取代原网络中的 C2f,增强了模型捕捉复杂形状和不规则目标特征的能力;在特征融合网络中嵌入空间上下文感知模块 SCAM,融合来自不同层次的丰富语义信息,对齐局部特征,抑制图像噪声,提升图像质量,进一步降低无关背景信息对目标特征提取的干扰,从而确保了特征表示的准确性和鲁棒性。最后,设计 ECIoU 替换 CIoU,增强边界框损失的拟合能力,加快网络模型的收敛速度,提高训练效率。这种改进减少计算成本和参数量,实现高效率和高性能之间平衡。实验证明,改进模型提高了缺陷检测精度、检测速度以及计算资源利用效率。该模型不仅在整体性能上优于其他主流网络模型,还易于使用和部署,具有轻量化特性。其平均精度达到 90.1%,使其在实际应用中更加高效灵活,对缺陷检测技术的发展具有重要意义。然而,由于数据集中部分类别较少,导致该类目标学习精度较低,进而影响了整体检测精度。为了消除检测结果存在的偏差,还需要进一步完善数据集。改进算法在检测精度上依然存在提升空间。下一步的研究将致力于如何进一步提高热斑和电气损坏的光伏板缺陷的检测精度,同时提升整体缺陷的平均检测精度,以期达到更加精准的缺陷检测效果,并将其部署到双光无人机上,从而实现硬件加速设计,进一步提高其应用场景的覆盖范围和实用性。

参考文献

- [1] 赵强,刘胜杰,韩东成,等.基于改进 K 均值聚类的光伏板缺陷检测方法[J].红外技术,2024,46(4):475-482.
ZHAO Q, LIU SH J, HAN D CH, et al. A defect detection method for photovoltaic panels based on improved K-means clustering [J]. Infrared Technology, 2024, 46(4): 475-482.
- [2] 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. Research on thermal spot detection method for aerial photovoltaic panel infrared images based on deep learning [J]. Electronic Measurement Technology, 2022, 45(22): 75-81.
- [3] PRABHAKARAN S, UTHRA R A, PREETHAROSELYN J. Deep learning-based model for defect detection and localization on photovoltaic panels [J]. Computer Systems Science & Engineering, 2023, 44 (3): 2683-2700.
- [4] 张猛,尹丽菊,周辉,等.基于 SimAM-Ada YOLOv5 的太阳能电池表面缺陷检测[J].电子测量技术,2023,46(22):17-25.
ZHANG M, YIN L J, ZHOU H, et al. Surface defect detection of solar cells based on SimAM-Ada YOLOv5 [J]. Electronic Measurement Technology, 2023, 46(22): 17-25.
- [5] YANG L, ZHANG R Y, LI L, et al. SimAM: A simple, parameter-free attention module for convolutional neural networks [C]. International Conference on Machine Learning, 2021: 11863-11874.
- [6] LIU W, ANGUELOV D, ERHAN D, et al. SSD: Singleshot MultiBox detector [C]. Computer Vision-ECCV 2016, 2016.
- [7] 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.
- [8] 刘承毅,董效杰,刘三军,等.基于改进 Dn-YOLOv7 算法的光伏板表面小目标缺陷检测[J].湖北民族大学学报(自然科学版),2024,42(2):212-218.
LIU CH Y, DONG X J, LIU S J, et al. Small target defect detection on photovoltaic panels based on the improved Dn-YOLOv7 algorithm [J]. Journal of Hubei University for Nationalities (Natural Science Edition), 2024, 42(2): 212-218.
- [9] 周颖,颜毓泽,陈海永,等.基于改进 YOLOv8 的光伏电池缺陷检测[J].激光与光电子学进展,2024,61(8):245-255.
ZHOU Y, YAN Y Z, CHEN H Y, et al. Defect detection of photovoltaic cells based on improved YOLOv8 [J]. Advances in Lasers and Optoelectronics, 2024, 61(8): 245-255.
- [10] PHAN Q B, NGUYEN T. A novel approach for PV cell fault detection using YOLOv8 and particle swarm optimization [C]. 2023 IEEE 66th International Midwest Symposium on Circuits and Systems (MWSCAS), 2023: 634-638.
- [11] MASITA K, HASAN A, SHONGWE T. 75 MW AC PV module field anomaly detection using drone-based IR orthogonal images with Res-CNN3 detector [J].

- IEEE Access, 2022, 10: 83711-83722.
- [12] 赵晓雨, 高林, 杨校李, 等. 融合注意力机制的 YOLOv5 光伏板电致发光图像缺陷检测算法[J]. 湖北民族大学学报(自然科学版), 2023, 41(1): 65-70.
ZHAO X Y, GAO L, YANG X L, et al. Defect detection algorithm for photovoltaic panel electroluminescence images based on an attention mechanism integrated YOLOv5[J]. Journal of Hubei University for Nationalities(Natural Science Edition), 2023, 41(1): 65-70.
- [13] WANG Q L, WU B G, ZHU P F, et al. ECA-Net: Efficient channel attention for deep convolutional neural networks [C]. IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2020: 11534-11542.
- [14] 焦思韬, 王可庆, 周奇, 等. 基于改进 SSD 算法的光伏板缺陷检测研究[J]. 软件, 2023, 44(12): 47-52.
JIAO S T, WANG K Q, ZHOU Q, et al. Research on photovoltaic panel defect detection based on improved SSD algorithm[J]. Software, 2023, 44(12): 47-52.
- [15] LI D, HU J, WANG CH H, et al. Involution: Inverting the inherence of convolution for visual recognition[C]. IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2021: 12321-12330.
- [16] LIN T Y, GOYAL P, GIRSHICK R, et al. Focal loss for dense object detection[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2020, 42(2): 318-327.
- [17] MISRA D, NALAMADA T, ARASANIPALAI A U, et al. Rotate to attend: Convolutional triplet attention module [C]. IEEE/CVF Winter Conference on Applications of Computer Vision, 2021: 3139-3148.
- [18] 翁俊辉, 成乐, 黄曼莉, 等. 基于 CS-YOLOv5s 的无人机航拍图像小目标检测[J]. 电子测量技术, 2024, 47(7): 157-162.
WENG J H, CHENG L, HUANG M L, et al. Small target detection in drone aerial images based on CS-YOLOv5s[J]. Electronic Measurement Technology, 2024, 47(7): 157-162.
- [19] 叶彦斐, 程立, 侯翔一. 基于改进 YOLOv5 的钢轨内部伤损 B 显图像识别与分类[J]. 国外电子测量技术, 2023, 42(12): 70-76.
YE Y F, CHENG L, HOU X Y. Identification and classification of internal rail damage B-scan images based on improved YOLOv5 [J]. Foreign Electronic Measurement Technology, 2023, 42(12): 70-76.
- [20] 高嘉, 刘涛, 王显峰, 等. TWRD-Net: 一种用于曳引钢丝绳缺陷的实时检测网络算法[J]. 仪器仪表学报, 2023, 44(6): 223-235.
GAO J, LIU T, WANG X F, et al. TWRD-Net: A real-time detection network algorithm for traction wire rope defects [J]. Chinese Journal of Scientific Instrument, 2023, 44(6): 223-235.
- [21] NASCIMENTO M G D, PRISACARIU V, FAWCETT R. DSConv: Efficient convolution operator[C]. IEEE/CVF International Conference on Computer Vision, 2019: 5148-5157.
- [22] 赵鸿亮, 郭佑民, 王建鑫, 等. 基于改进 YOLOv5 的轨道交通障碍物检测算法[J]. 电子测量技术, 2024, 47(1): 130-135.
ZHAO H L, GUO Y M, WANG J X, et al. Obstacle detection algorithm for rail transit based on improved YOLOv5 [J]. Electronic Measurement Technology, 2024, 47(1): 130-135.
- [23] TAN M X, PANG R M, LE Q V. EfficientDet: Scalable and efficient object detection[C]. IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2020: 10781-10790.

作者简介

杨长春, 教授, 博士, 主要研究方向为数据挖掘与决策支持系统。

E-mail: 2513676923@qq.com

贺轩轩(通信作者), 硕士研究生, 主要研究方向为深度学习与图像处理。

E-mail: 2269258075@qq.com

王睿, 讲师, 博士, 主要研究方向为声传感器信号处理与智能机器人。

E-mail: wangrui@cczu.edu.cn

朱士竹, 硕士研究生, 主要研究方向为机器视觉。

E-mail: s23060854018@smail.cczu.edu.cn

严昊, 硕士研究生, 主要研究方向为目标检测。

E-mail: S23060854044@smail.cczu.edu.cn