

基于改进 YOLOv8 的光伏电池缺陷检测

周颖^{1,2}, 颜毓泽¹, 陈海永^{1,2*}, 裴盛虎¹

¹河北工业大学人工智能与数据科学学院, 天津 300130;

²河北省控制工程技术研究中心, 天津 300130

摘要 针对光伏(PV)电池缺陷检测中存在的数​​据不均衡、缺陷尺度不一和背景纹理复杂多变等因素导致的误检、漏检问题,提出一种基于 YOLOv8 的缺陷检测算法 YOLOv8-EL。首先,使用 GauGAN 进行数据增强,处理数据集的类内和类间不均衡的问题,提高模型泛化能力,降低模型过拟合的风险;其次,在主干网络和特征融合网络之间嵌入上下文聚合模块,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,抑制无关背景信息的干扰;最后,构建多注意力检测头以替换解耦头,引入不同的注意力机制细化分类和定位任务,提取空间和通道层面上的关键信息,减少特征混淆。实验结果表明,该模型在扩充后的 PV 电池 EL 数据集上的平均精度达到 89.90%,模型参数量为 13.13×10^6 ,在提升精度的基础上兼顾了后续部署时轻量化的要求。在 PASCAL VOC 数据集上进行泛化性实验,证明了改进算法的泛化性能。

关键词 缺陷检测; YOLOv8; 生成对抗网络; 特征融合; 注意力机制; 解耦头

中图分类号 TP319.4

文献标志码 A

DOI: 10.3788/LOP231622

Defect Detection of Photovoltaic Cells Based on Improved YOLOv8

Zhou Ying^{1,2}, Yan Yuze¹, Chen Haiyong^{1,2*}, Pei Shenghu¹

¹School of Artificial Intelligence and Data Science, Hebei University of Technology, Tianjin 300130, China;

²Hebei Control Engineering Technology Research Center, Tianjin 300130, China

Abstract A YOLOv8-based defect detection algorithm, YOLOv8-EL, is proposed to address the problems of false detection and missing detection caused by data imbalance, varied defect scales, and complex background textures in photovoltaic (PV) cell defect detection. First, GauGAN is used for data augmentation to address the issue of intra-class and inter-class imbalance, improve model generalization ability, and reduce the risk of overfitting. Second, a context aggregation module is embedded between the backbone network and the feature fusion network to adaptively fuse semantic information from different levels, align local features, reduce the loss of minor defect information, and suppress irrelevant background interference. Finally, a multi-attention detection head is constructed to replace the decoupling head, introducing different attention mechanisms to refine classification and localization tasks, extract key information at the spatial and channel levels, and reduce feature confusion. Experimental results show that the proposed model achieves an average precision of 89.90% on the expanded PV cell EL dataset with a parameter count of 13.13×10^6 , achieving both precision improvement and lightweight deployment requirements. Generalization experiments on the PASCAL VOC dataset demonstrate the improved algorithm's generalization performance.

Key words defect detection; YOLOv8; generative adversarial network; feature fusion; attention mechanism; decoupled head

1 引言

光伏(PV)发电具有无污染、可再生等特点,被广泛应用于各种工业、生活场景中。PV 电池在生产、使

用和维护的过程中难免会产生隐裂、断栅等缺陷,影响整体阵列的寿命和性能,因此,对其进行缺陷检测尤为重要^[1]。某些种类的缺陷在电池内部产生,摄像机无法直接进行拍摄,故使用电致发光(EL)成像技术采集

收稿日期: 2023-06-28; 修回日期: 2023-07-08; 录用日期: 2023-07-31; 网络首发日期: 2023-08-11

基金项目: 国家自然科学基金(U21A20482,62073117)

通信作者: haiyong.chen@hebut.edu.cn

PV 电池缺陷图像。EL 成像技术可以不受外部环境或制造过程中电池表面不光滑等因素的影响,有效识别 PV 电池缺陷^[2]。但是 PV 电池的 EL 图像仍然存在数据不平衡、缺陷与背景耦合、缺陷尺度不一等问题,给缺陷检测任务带来极大困难。

在 PV 电池的生产过程中,部分类型的缺陷出现的概率较低,获取费时费力,导致样本数量稀少,模型无法充分学习缺陷特征。传统的数据增强方式如翻转、添加噪声、裁剪等并没有利用原始图像的深层次内在信息,导致模型检测精度的提升度有限,甚至出现过拟合的现象。生成对抗网络(GAN)能在学习真实样本分布规律的基础上生成具有缺陷类别一致性和属性多样性的样本,在目标检测领域已经被用于数据增强^[3]。文献[4]将 DCGAN(deep convolutional GAN)生成的 PV 缺陷图片与真实图片随机融合进行数据增强,提高了模型分类准确率。文献[5]和[6]通过改进的 WGAN(Wasserstein GAN)生成少数类别样本,以缓解数据集类间不平衡的问题。文献[7]提出了一种新型的 GAN——GP-GAN,并用其进行数据增强以提高 YOLOv8 网络的检测精度。

目前,基于机器视觉的缺陷检测效率更高,成本更低,逐渐取代了传统的图像处理技术和人工检测。YOLO 系列网络因其综合性能较好,成为了目标检测领域广泛使用的框架^[8]。文献[9]在 YOLOv3 算法中加入空间注意力检测头,降低了冗余预测框的数量,减少了计算量。文献[10]在 YOLOv4 算法的主干网络中加入密集块和非极大值抑制算法,提高了模型的检测精准度和检测速度。文献[11]将两种注意力机制级联,提高了模型对小目标的检测精度。文献[12]在 YOLOv5 中加入不同的损失函数,提升了算法的检测精度。文献[13]使用跨阶段局部瓶颈模块和一个结合 YOLOv5 的微小目标预测头来检测 PV 电池板的缺陷,模型在多尺度目标检测中获得了较好的精度。文献[14]基于 YOLOv5 模型,引入了一种新型的空间金字塔池化操作和通道注意力机制来定位电池 EL 图像中的裂纹和碎裂缺陷。文献[15]为了提高小目标检测精度,在 YOLOX 的基础上加入分支注意力模块,分别提取空间和通道层面的关键信息,提高了分类和定位任务的准确率。文献[16]在 YOLOX 的基础上设计了一种新的空间金字塔池化操作以充分提取目标空间特征,此外,在解耦头中加入角度检测分支提高了模型对完整目标的检测能力。文献[17]提出了一种动态检测头,结合三种不同的注意力机制提高了模型的检测性能。文献[18]提出了一种融合金字塔和注意力机制的缺陷检测算法,使用具有跳跃结构的加权双向特征金字塔,通过改进 YOLOv7 融合网络加强了多尺度目标特征融合能力。文献[19]提出了三方特征增强金字塔网络,该网络通过引入特征校准模块对上采样的特征进行校准,实现了特征融合的准确对应。

YOLOv8 算法在以往 YOLO 版本的基础上引入新的功能和改进,进一步提高了模型的结构灵活性和检测精度。文献[20]针对特殊场景的小目标检测,在 YOLOv8 的基础上提出了一种新的下采样方法和特征融合网络,能够在保留背景特征信息的同时兼顾浅层和深层信息。文献[21]结合原始样本生成新的样本以缓解数据集类不平衡的问题,提高了 YOLOv8 的检测精度。文献[22]针对 PV 电池 EL 图像的缺陷检测问题设计了一种新的算法,以优化学习率和模型参数,提高 YOLOv8 的各项性能;但是针对不同类别和尺度的缺陷检测,模型的精度尚待提高。

目前,部分现有的检测算法在解决 EL 图像缺陷和背景耦合、提高微小微弱缺陷检测精度上存在一定的局限性。YOLOv8 目标检测算法作为 YOLO 系列的最新成果,具有更高的检测精度。因此,在此基础上,提出了一种基于 YOLOv8 的缺陷检测算法——YOLOv8-EL,即:首先,使用 GauGAN 进行数据增强,缓解数据不均衡的问题;其次,在主干网络和特征融合网络之间添加上下文聚合模块(CAM),该模块可以针对性地融合深层和浅层特征图的空间信息,自适应地对齐局部特征间的空间关系;最后,构建多注意力检测头(MADH),根据分类和定位任务的差异插入不同的注意力模块,提升模型分类和检测性能。

2 基于 GauGAN 进行数据增强

PV 电池缺陷 EL 图片背景复杂,缺陷特征信息占比较低,使用普通的 GAN 网络很难学习到其有效的数据分布,故选用 GauGAN^[23]对其进行数据增强。相较于传统的 GAN 网络,GauGAN 在生成过程中通过向空间自适应归一化(SPADE)模块中嵌入语义分割掩码(mask)来控制高质量样本的生成,具体过程如图 1 所示,其中, \mathbf{Z} 为一维随机向量。传统的归一化过程会破坏语义布局图像的语义信息,故 GauGAN 使用 SPADE 模块替代传统的归一化模块。SPADE 模块中输入的语义分割掩码经过重构后与各层网络的分辨率相适应,并通过卷积层扩充至相应的通道数,最后分别通过两个卷积层得到归一化层的缩放系数与偏置项,促使 GAN 网络在归一化过程中保留完整的语义信息。

训练前,选取 400 张缺陷图片使用 LabelImg 标注工具进行标注,获得对应的语义分割掩码。将语义分割掩码和原始图片成对输入到 GauGAN 中进行训练。在训练完成后,任意修改语义分割掩码可以控制生成缺陷的位置、形状、尺度等特征,同时还能在保证栅线位置和遮挡关系正确的前提下,随机生成复杂的背景纹理信息,增加缺陷和背景的多样性。GauGAN 的生成效果如图 2 所示,其中:图 2(a)为隐裂与星状缺陷的语义分割掩码;图 2(b)、(c)展示了在训练过程中,不同迭代次数下 GauGAN 生成的具有多样性的背景和缺陷。

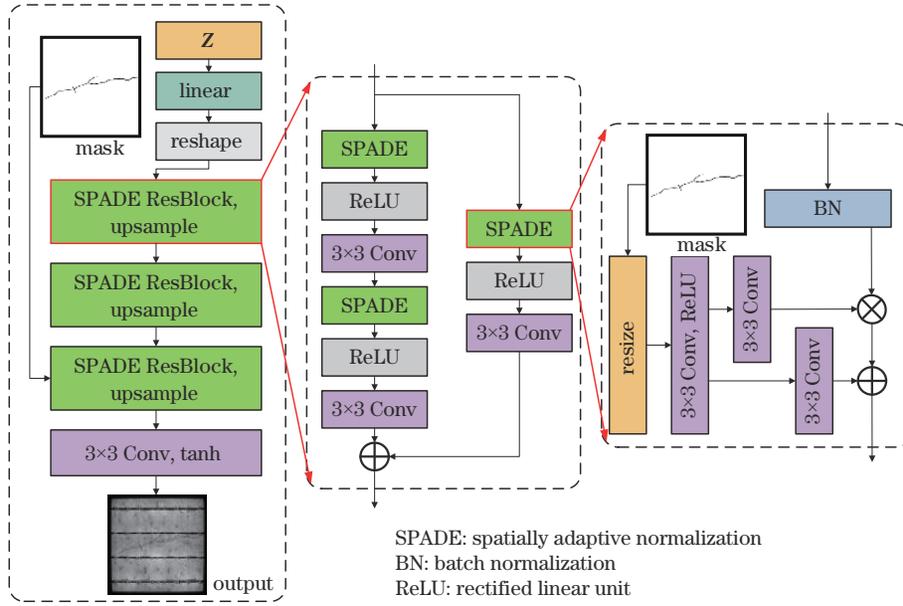


图 1 GauGAN 的结构

Fig. 1 Structure of GauGAN

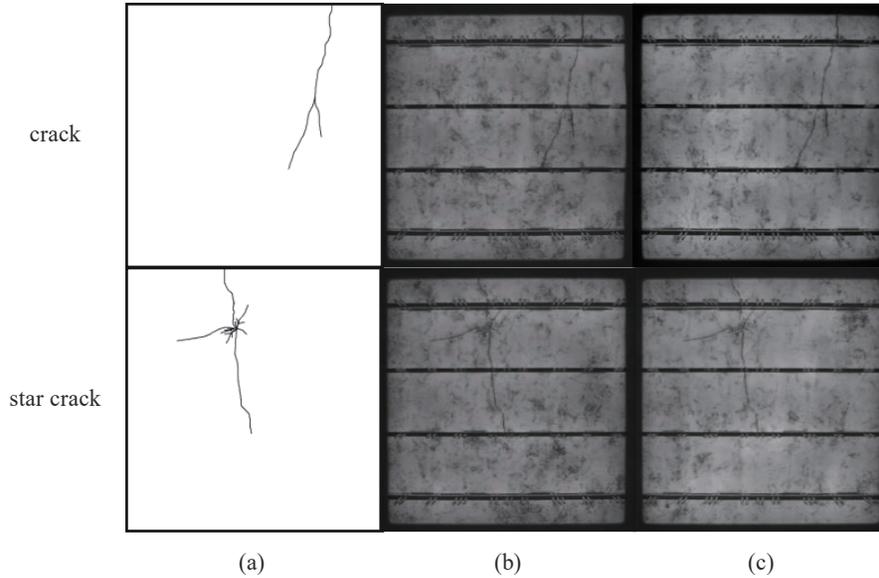


图 2 GauGAN 生成效果图。(a) 语义分割掩码；(b) 迭代次数为 100 时生成的图片；(c) 迭代次数为 200 时生成的图片

Fig. 2 Images generated by GauGAN. (a) Semantic segmentation mask; (b) images generated when epoch is 100; (c) images generated when epoch is 200

3 YOLOv8 算法结构与工作原理

YOLOv8 网络主要由 backbone、neck 和 head 部分组成，网络结构如图 3 所示，其中：box. loss 和 cls. loss 分别为回归和分类损失。

backbone 部分主要用于特征提取。YOLOv8 将 YOLOv5 中的跨阶段局部网络 (CSP) 模块替换为轻量级 C2f 模块，通过密集残差结构增强模型的特征表达能力；根据缩放系数，通过拆分和拼接操作改变通道数，降低模型的计算复杂度和容量；在尾部使用快速空间金字塔池化 (SPPF) 层增加感受野并捕获场景中不同层次的特征信息。neck 部分主要用于特征融合。它利

用路径聚合网络与 C2f 模块，针对 backbone 三个阶段输出的不同尺度的特征图进行融合，帮助浅层信息向深层特征聚合。head 部分使用解耦头结构，分为分类端和定位预测端，用于缓解分类和定位任务之间存在的冲突。采用 Anchor-Free 框架提高模型的检测性能，该框架对于检测长宽不规则的目标来说比较有优势。

损失计算包括分类损失 (varifocal loss)^[24] 和回归损失 [CIoU (complete intersection over unit) loss + DF (distribution focal) loss]^[25]。分类损失是一种基于焦点损失函数的目标检测损失函数，它可以更好地处理类别不平衡的情况，提高模型的检测准确率；CIoU 损失可以更好地处理预测框和真实框之间的重叠问

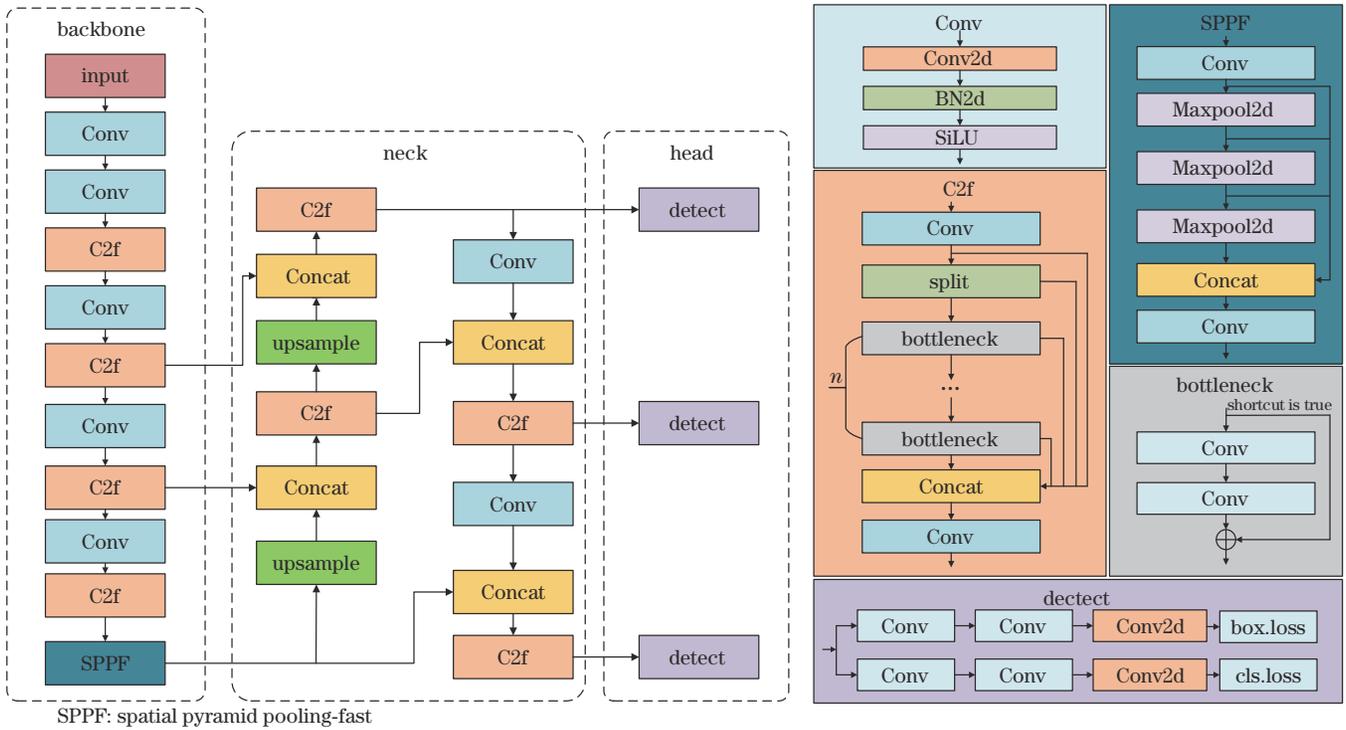


图 3 YOLOv8 的结构

Fig. 3 Structure of YOLOv8

题;DF 损失可以更好地处理类别不平衡和背景类别的问题,能够让网络快速地聚焦于目标附近的数据分布。

4 YOLOv8-EL PV 电池缺陷检测算法

4.1 CAM

不同层级的特征图信息存在差异,直接进行拼接 (Concat) 操作会增加冗余信息,降低网络的表达能力;

而直接添加像素增大图像分辨率的上采样方式会导致后续拼接时局部特征难以对齐。在复杂背景下对尺度变化巨大的 PV 电池缺陷特征进行特征融合时,浅层的细节特征会被深层语义特征所掩盖,使具有丰富空间细节的较低级特征丢失,进而影响模型对浅层边缘小目标和区域密集目标的检测。为提高模型的特征融合和空间交互能力,基于文献[26]设计了 CAM,其结构如图 4 所示。

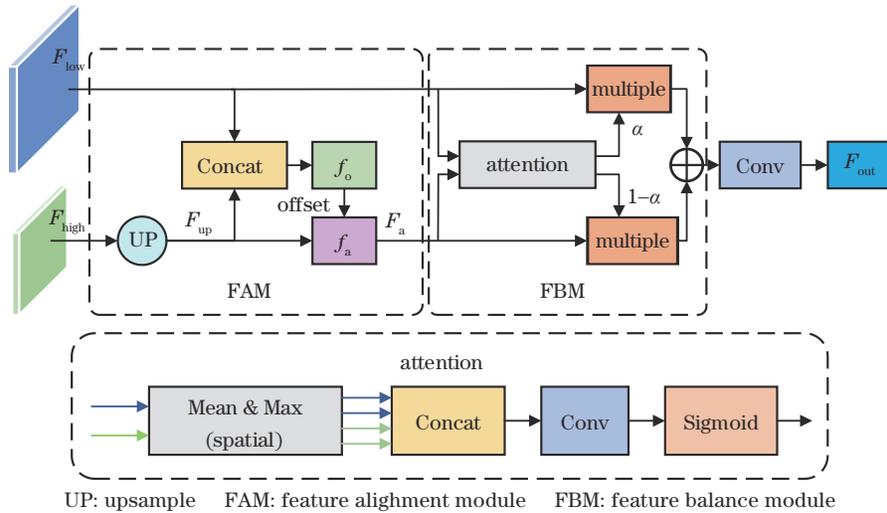


图 4 CAM 结构

Fig. 4 Structure of CAM

CAM 可以充分学习不同层级之间的特征差异,通过可变形卷积学习像素间的偏移关系,通过注意力机

制分配融合比例,从而提高特征信息的利用率。将 neck 中连续上采样后的两个 Concat 模块替换为 CAM,

能够自适应地校正局部特征之间的空间叠加关系,促进不同层次特征的交互融合。

CAM分为两部分:特征对齐模块(FAM)和特征平衡模块(FBM)。在FAM中,分辨率较低的高层语义信息特征图 F_{high} 经过上采样后变为 F_{up} 。 F_{up} 与分辨率较高的低层语义信息特征图 F_{low} 通过拼接操作级联,以获取特征之间的变换偏移量(offset)。使用可变形卷积将 F_{a} 与 F_{low} 的空间特征对齐,表示为

$$\Delta = f_o([F_{\text{up}}, F_{\text{low}}]), \quad (1)$$

$$\alpha = \text{Sigmoid}\left(\text{Conv}\left\{\text{Concat}\left[\text{Mean}(F_{\text{a}}), \text{Max}(F_{\text{a}}), \text{Mean}(F_{\text{low}}), \text{Max}(F_{\text{low}})\right]\right\}\right), \quad (3)$$

$$F_{\text{out}} = f[F_{\text{low}} \cdot \alpha + F_{\text{a}}(1 - \alpha)], \quad (4)$$

式中: Sigmoid(\cdot)为激活函数; Conv(\cdot)为卷积操作; Concat(\cdot)表示在通道维度进行合并操作; Mean(\cdot)和 Max(\cdot)表示在通道维度计算均值和最大值操作; $f(\cdot)$ 为 1×1 卷积,用于改变输出的通道数; F_{out} 为整个模块的最终输出。

4.2 MADH

传统的检测头将多个任务并行处理,容易导致特征混淆和相互干扰。YOLOv8采用解耦头的设计,考虑到了分类和定位所关注的内容差异性,可以解耦出单独的特征通道并将其分别用于边界框坐标回归和对象分类。首先,YOLOv8中的head部分使用 3×3 卷积将neck部分输出的三个有效特征层的通道数 $C_{i,x}$ (i 为第 i 个检测头, x 为对应的通道数)按照定位和分类任务变为 C_1 和 C_2 。其次,继续使用卷积模块提取特征。最终,定位任务的输出通道数为 $4 \times \text{reg_max}$,即每个特征点会输出数量为 $4 \times \text{reg_max}$ 的回归预测值,分别对应框的中心点(x, y)和预测框的宽高信息调整参数(W, H)。分类任务的输出通道数为 C_N ,表示检测缺陷的类别数。每一个特征点都将用于判断缺陷的类别信息。

式中: $[\cdot, \cdot]$ 为拼接操作,提供了特征之间的空间差异; σ 为激活函数ReLU; $f_o(\cdot)$ 是用于从空间差异中学习偏移 Δ 的函数; $f_a(\cdot)$ 是将特征与学习到的偏移 Δ 进行对齐(align)的可变形卷积,可以实现特征的精准定位。

在FBM中, F_{low} 与 F_{a} 通过拼接操作级联后输入到空间注意力(attention)模块中,获得注意力权重 α 。注意力权重分配两个特征图的融合比例,捕捉有效特征并抑制无关特征,表示为

上述过程仅靠堆叠卷积层进行特征提取,忽视了定位任务对空间信息的敏感性和分类任务对通道信息的敏感性。因此,在YOLOv8解耦头的基础上,构建MADH,将head中定位预测端的部分卷积层替换为坐标注意力(CA)^[27]模块。为适应分类预测端的通道变化,在压缩激励(SE)^[28]模块的基础上构建通道压缩注意力(CCA)模块。MADH结构如图5所示。

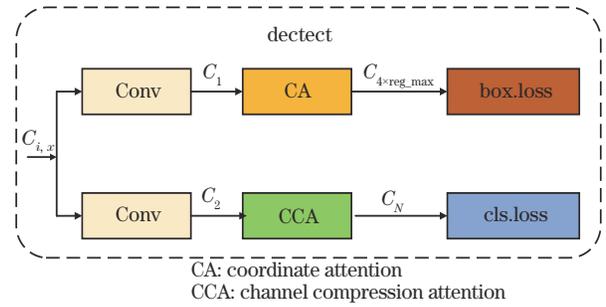


图5 MADH的结构

Fig. 5 Structure of MADH

两种注意力机制的结构如图6所示。CA模块具有如下作用: 1)沿着X和Y两个空间方向将输入特征

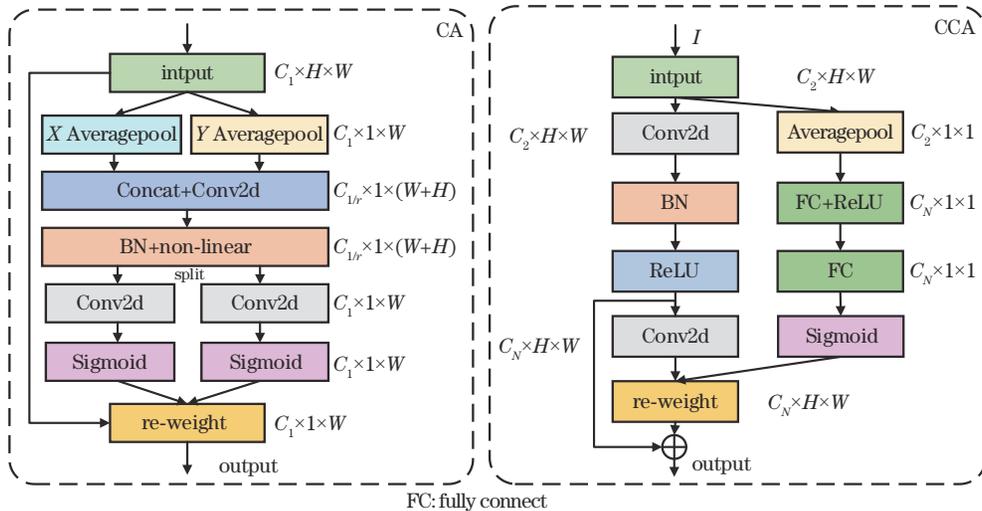


图6 两种注意力机制的结构

Fig. 6 Structure of two attention mechanisms

编码为一维特征,捕获输入特征图的长程依赖关系;2)将通道数压缩为 $1/r$ (r 为 CA 中的对通道维度进行压缩的系数),以降低计算量;3)生成对方向感知和空间敏感的注意力特征图,帮助模型在捕捉跨通道信息的同时获取精确的空间位置信息。CCA 模块中,输入特征图 I 的通道数为 C_2 , I 经过全局平均池化 (Averagepool) 生成一个 $C_2 \times 1 \times 1$ 的向量;为了匹配分类预测端的输出通道数 C_N , I 再经过两次全连接 (FC) 和 Sigmoid 激活函数生成一个 $C_N \times 1 \times 1$ 的权重向量,对应数量为 C_N 的缺陷类型。同时,为了获取具有相应通道数的特征图,将 I 通过两层卷积将通道

数降至 C_N 后与权重向量相乘,得到注意力加权特征图。最后引入一个残差连接,以提高模型的训练效率。CCA 模块通过压缩特征通道聚焦于类别特征紧密相关的数量为 C_N 的通道,提高了模型的分类能力。

5 实验结果与分析

为验证 YOLOv8-EL 的效果,在 EL 成像技术采集的 PV 电池缺陷图像制作的数据集上进行训练和测试,整体框架如图 7 所示。为评估模型的泛化性能,在公开数据集 PASCAL VOC 上进行训练和测试。

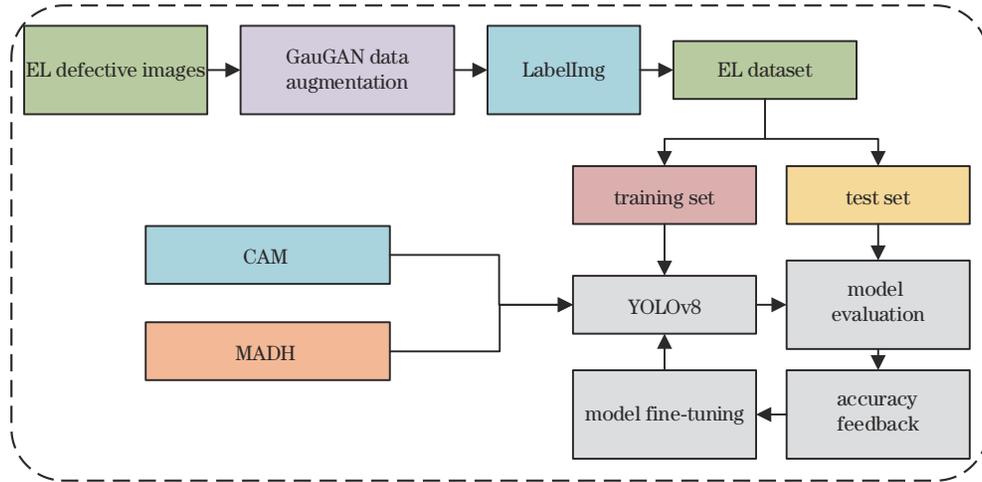


图 7 整体框架
Fig. 7 Overall frame

5.1 EL 数据集构建

EL 成像技术采集的 PV 电池缺陷数据集分为线状隐裂 (crack)、星状隐裂 (star crack)、断栅 (finger interruption) 和粗线 (thick line) 4 种类型,部分缺陷图片如图 8 所示。可以看出:隐裂缺陷形态多样,尺度变化大;断栅和粗线缺陷空间分布广,特征不明显。

对于微小微弱的缺陷目标,背景的栅线和各种不规则纹理的遮挡会给检测造成极大的困难。同时,线状隐裂存在类内不平衡的现象,大尺度缺陷数量稀少;星状隐裂也存在类间不平衡的现象,且相比于其他种类缺陷数量稀少。故针对上述两种缺陷进行数据增强。

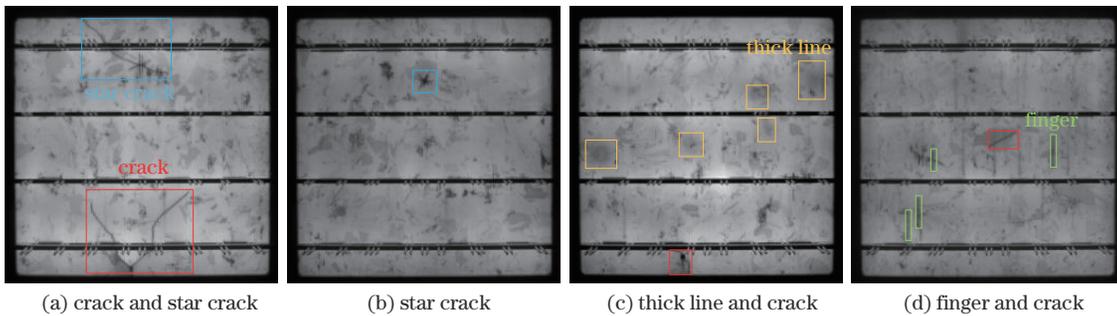


图 8 部分原始数据集
Fig. 8 Partial original dataset

数据集详细信息如表 1 所示,其中,大尺度线状隐裂缺陷图片仅有 104 张,星状隐裂图片仅有 219 张,存在类内和类间不平衡的问题。最终将线状隐裂图片扩充至 2000 张,星状隐裂图片扩充至 600 张。在完成扩充后,使用 LabelImg 标注软件对每类缺陷按照

PASCAL-VOC2007 数据集的格式进行标注。处理后的图片保存为 jpg 格式,分辨率为 640 pixel × 640 pixel,标注文件格式储存为 xml 文本。按照 8:1:1 的比例划分训练集、验证集和测试集,分别用于模型的训练和测试。

表 1 数据集详细信息
Table 1 Dataset details

Defect type		Crack		Star crack	Thick line	Finger	Total
		Small and medium scale	Large scale				
Before expansion	Label number	1028	104	252	1189	2918	5491
	Image number	907	104	219	775	1492	3497
After expansion	Label number	1821	300	633	1189	2918	6861
	Image number	1700	300	600	775	1492	4867

5.2 实验配置

本次实验采用 PyTorch 1.6.0 训练框架, 以 Windows 10 为系统环境, 所用显卡为 NVIDIA GeForce 2070Super, 8 GB 显存。通过 CUDA 10.1 对图形处理器(GPU)进行加速。模型训练参数 epoch 设置为 300, 优化器使用随机梯度下降法, 初始学习率设定为 0.01, 动量设定为 0.937, 使用余弦退火学习率调整算法。

5.3 评价指标

实验采用准确率(Precision)、召回率(Recall)、分类综合评价指标(F1)、平均精度[mAP(mean average precision)]、参数量(Params)、计算量(Flops)和每秒帧数(FPS)作为评估检测性能的指标。mAP@0.5 表

示当交并比(IoU)设为 0.5 时模型的平均精确度; mAP@0.5:0.95 表示 IoU 从 0.5 增加到 0.95(步长为 0.05)时模型的平均精确度。

5.4 数据增强效果

为验证 GauGAN 的数据增强效果, 分别使用扩充前后的数据集进行训练, 实验结果如表 2 所示。GauGAN 数据增强后, 线状隐裂和星状隐裂两种缺陷的 mAP@0.5 分别提升了 3.58 和 4.79 个百分点, 其余两种缺陷的 mAP@0.5 基本不变。分类指标均有所提高, 且平均 mAP@0.5:0.95 提高了 2.5 个百分点, 说明数据增强策略综合提升了模型的各项指标, 缓解了数据不平衡的问题, 同时证明了生成数据的有效性。后续实验皆在数据增强后的数据集上进行。

表 2 数据增强效果对比
Table 2 Comparison of data enhancement effects

Dataset	F1	Recall / %	Precision / %	mAP@0.5 / %				Average mAP@0.5 / %	Average mAP@0.5:0.95 / %
				Crack	Star crack	Thick line	Finger		
Before expansion	80.25	76.41	85.19	80.47	75.84	85.18	94.06	83.89	49.50
After expansion	82.75	77.29	86.93	84.05	80.63	85.81	94.34	86.21	52.00

5.5 CAM 改进实验

CAM 由特征对齐模块 FAM 和特征平衡模块 FBM 融合组成, 提高了模型的特征融合能力。为验证单个模块的效果和组合顺序的影响, 在保证模型其他结构参数相同的情况下进行实验, 结果如表 3 所示。分析可知: 分别加入两个模块都能提高模型的精度和对小目标的检

测能力, 但是对模型分类能力的提升效果不明显; 先加入 FBM 会使 FAM 对齐原始特征的效果不佳; 先加入 FAM 再分配融合权重比例能够较好地保留上下文信息; 与原始 YOLOv8 相比, YOLOv8+FAM+FBM 模型的检测精度 mAP@0.5 和 mAP@0.5:0.95 分别提高了 2.58 和 2.00 个百分点, 达到了 4 组实验中的最佳效果。

表 3 CAM 改进实验结果
Table 3 CAM improvement experimental results

Model	Params / 10 ⁶	Flops / 10 ⁹	F1	Recall / %	Precision / %	mAP@0.5 / %	mAP@0.5:0.95 / %
YOLOv8	11.14	14.28	82.75	77.29	86.93	86.21	52.00
YOLOv8+FAM	12.95	15.22	82.50	77.66	89.44	87.92	53.30
YOLOv8+FBM	11.31	14.29	83.00	79.47	86.97	87.54	54.00
YOLOv8+FBM+FAM	13.04	15.23	82.00	78.09	88.25	88.56	53.50
YOLOv8+FAM+FBM	13.04	15.23	84.50	79.50	90.52	88.79	54.00

5.6 MADH 改进实验

为确定 MADH 中各注意力机制的有效性, 在保证模型其他结构参数相同的情况下, 在定位预测端加入 CA 模块, 在分类预测端分别加入 CCA 和 SE 模块进行对比, 结果如表 4 所示。分析可知: 在定位预测端引入

CA 模块可以提高模型对小目标的定位能力, 检测精度 mAP@0.5:0.95 相比 YOLOv8 提高了 2.7 个百分点; 在分类预测端引入 SE 和 CCA 机制, F1 相比 YOLOv8 分别提升 0.6% 和 2.4%, Recall 分别提升 3.47 和 4.77 个百分点; 仅加入 SE 模块会导致模型检测精度降低;

表 4 MADH改进实验结果
Table 4 MADH improvement experimental results

Model	Params /10 ⁶	Flops /10 ⁹	F1	Recall /%	Precision /%	mAP@0.5 /%	mAP@0.5:0.95 /%
YOLOv8	11.14	14.28	82.75	77.29	86.93	86.21	52.00
YOLOv8+CA	11.03	13.97	83.75	78.35	87.42	87.64	54.70
YOLOv8+SE	10.69	13.04	83.25	80.76	86.47	85.99	52.90
YOLOv8+CCA	10.71	13.07	84.75	82.06	87.87	87.44	52.50
YOLOv8+CA+CCA	10.59	12.76	84.00	81.79	86.60	88.35	54.20

仅加入 CCA 模块后,模型的检测精度 mAP@0.5 相比 YOLOv8 提高了 1.23 个百分点,证明了加入 CCA 模块有较好的效果;同时加入 CA 和 CCA 模块,提高了模型的分类和定位能力,同时提高了模型检测小目标的检测精度。

5.7 消融实验

5.7.1 模块消融实验

为了分析不同改进策略对模型检测性能的影响,设计了三组实验。在保证数据集和训练参数相同的前提下完成训练,实验结果如表 5 所示。分析可知:加入 CAM 使得模型可以自适应地调整不同层级特征图的融合比例,对齐局部特征,因此模型的分类指标均有所

提升,检测精度 mAP@0.5 和 mAP@0.5:0.95 相比 YOLOv8 分别提高了 2.58 和 2.00 个百分点;CAM 虽然增强了模型对小目标缺陷的检测性能,但是增加了参数量,降低了 FPS。MADH 使得解耦头更加专注于预测端的不同任务,提高了模型的分类和定位能力,检测精度 mAP@0.5 和 mAP@0.5:0.95 相比 YOLOv8 分别提高了 2.14 和 2.20 个百分点。将两种改进策略同时加入 YOLOv8 模型,以增加少量参数为代价,在提高分类指标、满足实时性要求的同时,使检测精度 mAP@0.5 和 mAP@0.5:0.95 相比 YOLOv8 分别提高了 3.69 和 2.70 个百分点,证明了改进策略的有效性。

表 5 消融实验结果
Table 5 Ablation experimental results

Model	Params /10 ⁶	Flops /10 ⁹	FPS /(frame/s)	F1	Recall /%	Precision /%	mAP@0.5 /%	mAP@0.5:0.95 /%
YOLOv8	11.14	14.28	81.25	82.75	77.29	86.93	86.21	52.00
YOLOv8+CAM	13.04	15.23	74.06	84.50	79.50	90.52	88.79	54.00
YOLOv8+MADH	10.59	12.76	78.96	84.00	81.79	86.60	88.35	54.20
YOLOv8-EL	13.13	15.43	73.30	84.25	79.90	90.12	89.90	54.70

5.7.2 特征图可视化分析

注意力机制旨在帮助模型更加关注特征分布中的重要信息。类激活图利用特征图权重叠加的原理可以直观地展示模型对不同部分的关注程度。为验证

MADH 和 CAM 对较小目标缺陷的识别效果,使用类激活图对特征图进行可视化,结果如图 9 所示。该实验选取了两张具有小目标缺陷的 PV 电池缺陷图片,通过对比可以看出:加入 MADH 提升了模型对小目

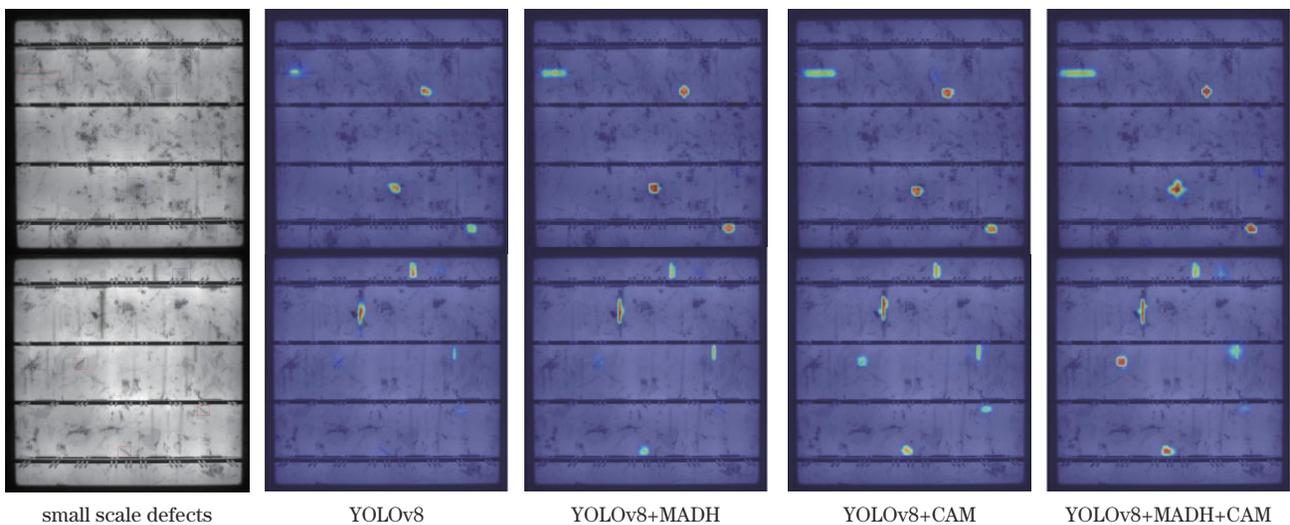


图 9 类激活图对比
Fig. 9 Comparison of class activation map

标的关注度,但是对于缺陷边界信息的预测能力较差;加入CAM提高了模型对缺陷目标的全局关注度,能更好地识别完整的缺陷特征;同时加入两个模块后热力图红色区域更加明显,说明进一步提高了模型的识别能力。

5.8 对比实验

5.8.1 不同算法对比实验

选取 Faster R-CNN、YOLOv5、YOLOX、YOLOv7 4 种主流目标检测模型的不同版本与所提模型进行对

比,模型的超参数和训练参数都设置为默认值,使用扩充后的数据集进行训练,结果如表 6 所示。改进模型相比不同版本的 YOLOv8 均有改善,且在小模型上提升效果更为明显,YOLOv8-s-EL 的检测精度 mAP@0.5 和 mAP@0.5:0.95 相比 YOLOv8-s 分别提高了 3.69 和 2.70 个百分点,且模型参数量仅为 13.13×10^6 ,兼顾了检测精度与实时性;YOLOv8-m-EL 的检测精度 mAP@0.5 和 mAP@0.5:0.95 分别达到了 90.58% 和 55.20%,是 11 种模型中的最优值,证明了所提改进模型的优势。

表 6 对比实验结果
Table 6 Comparison test results

Model	Params / 10^6	Flops / 10^9	F1	Recall / %	Precision / %	mAP@0.5 / %	mAP@0.5:0.95 / %
Faster R-CNN	28.31	253.5	58.83	68.85	75.07	75.80	42.30
YOLO5-s	7.03	15.82	68.75	56.63	90.52	82.80	45.40
YOLO5-m	21.07	25.22	77.50	72.07	84.28	85.42	50.60
YOLOX-s	8.90	26.84	83.00	82.11	84.75	86.43	49.40
YOLOX-m	28.50	39.44	83.75	84.37	83.43	88.05	52.80
YOLOv7-tiny	6.07	13.32	78.75	71.28	89.27	85.84	48.30
YOLOv7-m	37.76	59.53	83.00	76.55	83.71	87.39	50.30
YOLOv8-s	11.14	14.28	82.75	77.29	86.93	86.21	52.00
YOLOv8-m	25.86	39.44	84.60	83.86	85.56	88.77	53.70
YOLOv8-s-EL	13.13	15.43	84.25	79.90	90.12	89.90	54.70
YOLOv8-m-EL	29.71	41.51	86.00	87.49	89.41	90.58	55.20

5.8.2 检测结果可视化分析

为了更加直观地评价所提改进算法的性能,对比了 YOLOv8-s 与 YOLOv8-s-EL 在测试集图片上的检测效果,如图 10 所示。图 10(a)中存在大尺度线状隐裂缺陷,原始模型将完整的隐裂缺陷错误地识别成三部分,而改进后的模型能够较为完整地定位大尺度缺陷;图 10(b)中小尺度星状隐裂被栅线部分遮挡,导致

原始模型漏检,而改进后的模型具有更强的抑制干扰的能力,能够更好地区分背景因素,说明所提模型拥有更强的上下文感知能力和特征融合能力;图 10(c)、(d)的检测结果显示了改进模型对小尺度缺陷的识别能力,模型在提高置信度分数的基础上,能够减少漏检的情况,说明所提改进模型在 PV 电池 EL 缺陷图像检测任务中具有更好的性能。通过增强模型对上下文信息

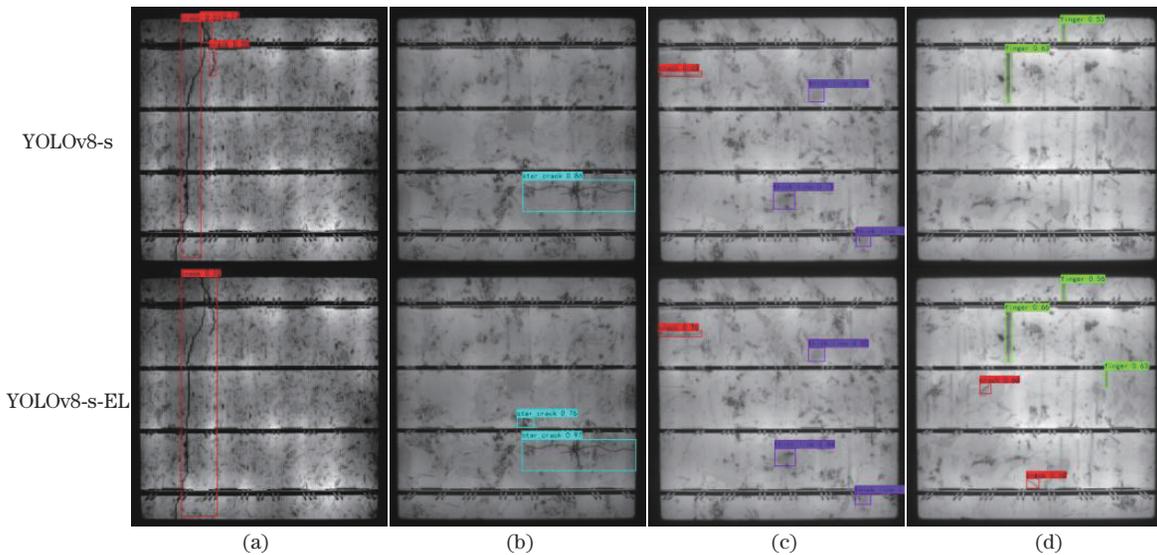


图 10 预测结果对比。(a) 大尺度线状隐裂;(b) 星状隐裂;(c) 粗线;(d) 断栅和小尺度线状隐裂

Fig. 10 Comparison of predicted results. (a) Large scale crack; (b) star crack; (c) thick line; (d) finger interruption and small scale crack

和特征融合能力的感知和理解,减轻背景因素的干扰,改进后的模型能够更准确地定位和识别缺陷,提高检测的准确性。

5.9 泛化性能验证

PASCAL VOC2007 通用数据集涵盖了 20 个常见的物体类别,共有 9963 张图像,24640 个实例标记,其中包括训练集 2501 张,验证集 2510 张和测试集 4952 张。实验使用的 PASCAL VOC2012 数据集共有

23080 张图像,54900 个实例标记,其中包括训练集 11540 张,测试集 11540 张。通用数据集相较于 EL 数据集有更多的图像和对象类别。使用 VOC2007 和 VOC2012 的训练集进行泛化实验,在 VOC2007 的测试集上进行测试。实验结果如表 7 所示,分析可知,与不同版本的模型相比,改进后模型的检测精度均有提升,YOLOv8-m-EL 相比于其他较大的主流网络仍有检测精度上的优势,证明了改进模型的泛化性能。

表 7 泛化性能实验结果

Table 7 Generalization performance experimental results

Model	F1	Recall /%	Precision /%	mAP@0.5 /%	mAP@0.5:0.95 /%
Faster R-CNN	78.25	80.68	82.01	79.70	51.30
YOLOv5-s	74.50	64.31	90.02	84.14	55.90
YOLOv5-m	83.45	77.74	90.28	86.38	63.60
YOLOX-s	81.10	80.22	82.29	84.83	58.00
YOLOX-m	83.70	84.13	83.37	88.08	63.90
YOLOv7-tiny	80.00	71.98	90.48	85.82	59.40
YOLOv7-m	87.80	86.21	89.93	90.02	68.10
YOLOv8-s	84.70	82.70	87.21	88.09	66.00
YOLOv8-m	87.50	87.33	87.83	90.72	70.20
YOLOv8-s-EL	84.45	80.74	88.99	90.11	67.30
YOLOv8-m-EL	88.15	86.45	90.18	92.88	72.30

6 结 论

提出了一种基于 YOLOv8 改进的 PV 电池 EL 图像缺陷检测方法——YOLOv8-EL。该方法在训练前使用 GauGAN 进行数据增强,在主干网络和特征提取网络之间加入 CAM,在检测头部分引入 CA 和 CCA 模块。实验结果表明,改进后模型的平均检测精度达到 89.90%,参数量为 13.13×10^6 ,有效提升了复杂背景下缺陷检测的精度,兼顾了模型部署的要求,具有一定的泛化性能。在未来的工作中拟开展如下研究:1) 探索生成更加真实的数据集;2) 对模型进行轻量化处理,在检测设备上进行部署;3) 针对不同种类的 PV 电池数据集进行微调,进一步提高模型的泛化性能。

参 考 文 献

- [1] 张治, 邹鹏辉, 刘志锋, 等. 发射极掺杂工艺对产业化 IBC 太阳能电池性能的影响[J]. 太阳能学报, 2022, 43(3): 158-162.
Zhang Z, Zou P H, Liu Z F, et al. Influence of emitter doping process on performance of industrial IBC solar cell [J]. Acta Energetica Solaris Sinica, 2022, 43(3): 158-162.
- [2] Hoiaas I, Grujic K, Imenes A G, et al. Inspection and condition monitoring of large-scale photovoltaic power plants: a review of imaging technologies[J]. Renewable and Sustainable Energy Reviews, 2022, 161: 112353.
- [3] Liu M Y, Huang X, Yu J H, et al. Generative adversarial networks for image and video synthesis: algorithms and applications[J]. Proceedings of the IEEE,

2021, 109(5): 839-862.

- [4] 王云艳, 周志刚, 罗帅. 基于数据增强的光伏电池片缺陷检测[J]. 电子测量与仪器学报, 2021, 35(1): 26-32.
Wang Y Y, Zhou Z G, Luo S. Defect detection for solar photovoltaic cells based on data augmentation[J]. Journal of Electronic Measurement and Instrument, 2021, 35(1): 26-32.
- [5] Fan J G, Yuan X F, Miao Z M, et al. Full attention Wasserstein GAN with gradient normalization for fault diagnosis under imbalanced data[J]. IEEE Transactions on Instrumentation and Measurement, 2022, 71: 1-16.
- [6] Zheng M, Li T, Zhu R, et al. Conditional Wasserstein generative adversarial network-gradient penalty-based approach to alleviating imbalanced data classification[J]. Information Sciences, 2020, 512: 1009-1023.
- [7] Wei E, Ren S. Drowning detection based on YOLOv8 improved by GP-GAN augmentation[EB/OL]. (2023-03-02)[2023-04-11]. <https://openreview.net>.
- [8] Terven J, Cordova-Esparza D. A comprehensive review of YOLO: from YOLOv1 and beyond[EB/OL]. (2023-04-02)[2023-06-09]. <https://arxiv.org/abs/2304.00501>.
- [9] 李翔, 何森, 罗海波. 一种面向遮挡行人检测的改进 YOLOv3 算法[J]. 光学学报, 2022, 42(14): 1415003.
Li X, He M, Luo H B. Occluded pedestrian detection algorithm based on improved YOLOv3[J]. Acta Optica Sinica, 2022, 42(14): 1415003.
- [10] 唐政, 张会林, 马立新. 密集主干网络算法对光伏电池的缺陷检测[J]. 激光与光电子学进展, 2023, 60(14): 1412003.
Tang Z, Zhang H L, Ma L X. Defect detection for solar

- cells using dense backbone network algorithm[J]. *Laser & Optoelectronics Progress*, 2023, 60(14): 1412003.
- [11] 张寅, 朱桂熠, 施天俊, 等. 基于特征融合与注意力的遥感图像小目标检测[J]. *光学学报*, 2022, 42(24): 2415001. Zhang Y, Zhu G Y, Shi T J, et al. Small object detection in remote sensing images based on feature fusion and attention[J]. *Acta Optica Sinica*, 2022, 42(24): 2415001.
- [12] 程松, 杨洪刚, 徐学谦, 等. 基于 YOLOv5 的改进轻量型 X 射线铝合金焊缝缺陷检测算法[J]. *中国激光*, 2022, 49(21): 2104005. Cheng S, Yang H G, Xu X Q, et al. Improved lightweight X-ray aluminum alloy weld defects detection algorithm based on YOLOv5[J]. *Chinese Journal of Lasers*, 2022, 49(21): 2104005.
- [13] Li L L, Wang Z F, Zhang T T. GBH-YOLOv5: ghost convolution with BottleneckCSP and tiny target prediction head incorporating YOLOv5 for PV panel defect detection[J]. *Electronics*, 2023, 12(3): 561-576.
- [14] Xu S W, Qian H M, Shen W Y, et al. Defect detection for PV modules based on the improved YOLOv5s[C]//2022 China Automation Congress (CAC), November 25-27, 2022, Xiamen, China. New York: IEEE Press, 2023: 1431-1436.
- [15] 吕志轩, 魏霞, 马志钢. 改进 YOLOX 的轻量级安全帽检测方法[J]. *计算机工程与应用*, 2023, 59(1): 61-71. Lü Z X, Wei X, Ma Z G. Improved YOLOX lightweight helmet detection method[J]. *Computer Engineering and Applications*, 2023, 59(1): 61-71.
- [16] Li Z, Song X, Chen S, et al. Data, language and graph-based reasoning methods for identification of human malicious behaviors in nuclear security[EB/OL]. (2023-02-15)[2023-04-06]. <http://dx.doi.org/10.2139/ssrn.4411800>.
- [17] Liu D F, Liang J, Geng T, et al. Tripartite feature enhanced pyramid network for dense prediction[J]. *IEEE Transactions on Image Processing*, 2023, 32: 2678-2692.
- [18] 亢洁, 王勃, 刘文波, 等. 融合 CAT-BiFPN 与注意力机制的航拍绝缘子多缺陷检测网络[J]. *高电压技术*, 2023, 49(08): 3361-3376. Kang J, Wang Q, Liu W B, et al. Aerial insulator multi-defect detection network by fusing CAT-BiFPN with attention mechanism[J]. *High Voltage Technology*, 2023, 49(08): 3361-3376.
- [19] Dai X Y, Chen Y P, Xiao B, et al. Dynamic head: unifying object detection heads with attentions[C]//2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), June 20-25, 2021, Nashville, TN, USA. New York: IEEE Press, 2021: 7369-7378.
- [20] Lou H T, Duan X H, Guo J M, et al. DC-YOLOv8: small-size object detection algorithm based on camera sensor[J]. *Electronics*, 2023, 12(10): 2323.
- [21] Ruiz-Ponce P, Ortiz-Perez D, Garcia-Rodriguez J, et al. POSEIDON: a data augmentation tool for small object detection datasets in maritime environments[J]. *Sensors*, 2023, 23(7): 3691.
- [22] Phan Q B, Nguyen T. A novel approach for PV cell fault detection using YOLOv8 and particle swarm optimization [EB/OL]. (2023-04-24)[2023-04-26]. <https://doi.org/10.36227/techrxiv.22680484.v1>.
- [23] Park T, Liu M Y, Wang T C, et al. Semantic image synthesis with spatially-adaptive normalization[C]//2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), June 15-20, 2019, Long Beach, CA, USA. New York: IEEE Press, 2020: 2332-2341.
- [24] Zhang H Y, Wang Y, Dayoub F, et al. VarifocalNet: an IoU-aware dense object detector[C]//2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), June 20-25, 2021, Nashville, TN, USA. New York: IEEE Press, 2021: 8510-8519.
- [25] Li X, Wang W H, Wu L J, et al. Generalized focal loss: learning qualified and distributed bounding boxes for dense object detection[EB/OL]. (2020-06-08)[2023-03-05]. <https://arxiv.org/abs/2006.04388>.
- [26] Huang S H, Lu Z C, Cheng R, et al. FaPN: feature-aligned pyramid network for dense image prediction[C]//2021 IEEE/CVF International Conference on Computer Vision (ICCV), October 10-17, 2021, Montreal, QC, Canada. New York: IEEE Press, 2022: 844-853.
- [27] Hou Q B, Zhou D Q, Feng J S. Coordinate attention for efficient mobile network design[C]//2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), June 20-25, 2021, Nashville, TN, USA. New York: IEEE Press, 2021: 13708-13717.
- [28] Hu J, Shen L, Sun G. Squeeze-and-excitation networks [C]//2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition, June 18-23, 2018, Salt Lake City, UT, USA. New York: IEEE Press, 2018: 7132-7141.