

YOLO-SCD——多晶硅太阳能电池片缺陷的识别与分割

刘明辉 刘怀广 苗泽志

YOLO-SCD: identification and segmentation of defects in polycrystalline silicon solar cell wafers

LIU Minghui, LIU Huaiguang, MIAO Zezhi

引用本文:

刘明辉, 刘怀广, 苗泽志. YOLO-SCD——多晶硅太阳能电池片缺陷的识别与分割[J]. 应用光学, 2026, 47(1): 156–168. DOI: 10.5768/JAO202647.0102007

LIU Minghui, LIU Huaiguang, MIAO Zezhi. YOLO-SCD: identification and segmentation of defects in polycrystalline silicon solar cell wafers[J]. Journal of Applied Optics, 2026, 47(1): 156–168. DOI: 10.5768/JAO202647.0102007

在线阅读 View online: <https://doi.org/10.5768/JAO202647.0102007>

您可能感兴趣的其他文章

Articles you may be interested in

多特征融合的太阳能电池片缺陷检测

Defect detection of solar cells based on multi-feature fusion

应用光学. 2023, 44(3): 605–613 <https://doi.org/10.5768/JAO202344.0303005>

基于轻量化卷积神经网络的光伏电池片缺陷检测方法研究

Defects detection method of photovoltaic cells based on lightweight convolutional neural network

应用光学. 2022, 43(1): 87–94 <https://doi.org/10.5768/JAO202243.0103003>

基于卷积神经网络与特征融合的天气识别方法

Weather recognition method based on convolutional neural network and feature fusion

应用光学. 2023, 44(2): 323–329 <https://doi.org/10.5768/JAO202344.0202004>

基于改进YOLOv5模型的车载红外目标检测

Vehicle-mounted infrared target detection based on improved YOLOv5 model

应用光学. 2025, 46(5): 1044–1053 <https://doi.org/10.5768/JAO202546.0502004>

基于激光雷达图像全局特征融合的无人机巡检缺陷识别

Defect recognition for unmanned aerial vehicle inspection based on global feature fusion of LiDAR images

应用光学. 2025, 46(5): 1054–1062 <https://doi.org/10.5768/JAO202546.0502005>

基于改进YOLOv5s的曲面光学镜片缺陷检测

Defect detection of curved optical lenses based on improved YOLOv5s

应用光学. 2024, 45(4): 781–789 <https://doi.org/10.5768/JAO202445.0403003>



关注微信公众号, 获得更多资讯信息

文章编号: 1002-2082 (2026) 01-0156-13

引用格式: 刘明辉, 刘怀广, 苗泽志. YOLO-SCD——多晶硅太阳能电池片缺陷的识别与分割 [J]. 应用光学, 2026, 47(1): 156-168.

LIU Minghui, LIU Huaiguang, MIAO Zezhi. YOLO-SCD: identification and segmentation of defects in polycrystalline silicon solar cell wafers[J]. Journal of Applied Optics, 2026, 47(1): 156-168.



在线阅读

YOLO-SCD——多晶硅太阳能电池片 缺陷的识别与分割

刘明辉¹, 刘怀广^{1,2}, 苗泽志³

(1. 武汉科技大学 冶金装备及其控制教育部重点实验室, 湖北 武汉 430081; 2. 武汉科技大学 机器人与智能系统研究院, 湖北 武汉 430081; 3. 无锡鑫瑞能智能装备有限公司 技术部, 江苏 无锡 214204)

摘 要: 多晶硅太阳能电池片成本较低, 利于普及。但其内部绒丝较多, 在线检测困难, 通过研究深度网络通道运算特性, 在 YOLOv8 的基础上提出了改进的 YOLO-SCD (you only look once- solar cell defects) 网络。在骨干网络中, 以通道混洗操作代替密集卷积操作, 在轻量化网络的同时增强通道之间的信息交流; 通过引入注意力模块增强网络对关键特征的学习能力。在颈部网络中, 以多通道乘法代替层间的加法操作, 提出 C2f_star 模块, 增强网络拟合数据的能力; 提出部分卷积下采样模块, 以部分卷积操作来减少特征图冗余信息以达到轻量化网络的目的。YOLO-SCD 对多晶硅太阳能电池板缺陷的检测精度达到了 0.970, 分割精度达到了 0.962, 模型权重只有 5.7 MB, 并且帧速 (frame per second, FPS) 达到了 90.03。最后, 通过对比实验表明, YOLO-SCD 在具有高识别精度的同时更加适合移动端部署。

关键词: 太阳能电池片; 缺陷检测; 深度学习; 特征融合; YOLOv8

中图分类号: TN206

文献标志码: A

DOI: 10.5768/JAO202647.0102007

YOLO-SCD: identification and segmentation of defects in polycrystalline silicon solar cell wafers

LIU Minghui¹, LIU Huaiguang^{1,2}, MIAO Zezhi³

(1. Key Laboratory of Metallurgical Equipment and its Control (Ministry of Education), Wuhan University of Science and Technology, Wuhan 430081, China; 2. Institute of Robotics and Intelligent Systems, Wuhan University of Science and Technology, Wuhan 430081, China; 3. Technology Department, Wuxi Xinruineng Intelligent Equipment Co., LTD., Wuxi 214204, China)

Abstract: Polycrystalline silicon solar cell wafers have lower cost, which is favorable for popularization. However, their internal fluffy filaments are more, and PL (photoluminescence) on-line detection is difficult. By studying the channel operation characteristics of the deep network, an improved YOLO-SCD (you only look once- solar cell defects) network is proposed on the basis of YOLOv8. In the backbone network, channel blending operation is used instead of dense convolution operation to enhance the information exchange between channels while lightening the network; and the network's ability to learn key features is enhanced by the introduction of attention module. In the neck network, multi-channel multiplication is used instead of inter-layer addition operation, the C2f_star module is proposed to enhance the ability of the network to fit the data,

收稿日期: 2024-07-30; 修回日期: 2024-10-25

基金项目: 国家自然科学基金项目 (52272377)

第一作者: 刘明辉, E-mail: 2323533023@qq.com

通信作者: 刘怀广, E-mail: liuhuaiguang@wust.edu.cn

and the partial convolutional downsampling module is proposed to reduce the redundant information of the feature maps with partial convolutional operation to achieve the purpose of lightweighting the network. Finally, YOLO-SCD achieves a detection accuracy of 0.970 for polycrystalline solar panel defects, a segmentation accuracy of 0.962, a model weight of only 5.7MB, and an FPS(frame per second) of 90.03. Comparative experiments show that YOLO-SCD is more suitable for mobile deployments while having high recognition accuracy.

Key words: solar cell; defects detection; deep learning; feature fusion; YOLOv8

引言

太阳能电池片作为太阳能发电的关键部件,其质量决定着光伏发电效率,但是太阳能电池片的很多缺陷通过人眼难以识别,目前主要需通过光致发电(photoluminescence, PL)或电致发光(electroluminescent, EL)成像技术来进行在线检测^[1]。

目前,基于深度学习的图像检测算法在光伏电池片的在线检测中得到广泛应用, HUANG J等^[2]通过对 YOLOv5 添加改进后的协调注意力机制 LCA(leaky coordinate attention)和加权双向特征金字塔结构(bidirectional feature pyramid network, BiFPN)^[3],把耦合头替换为解耦头^[4],最终所提网络模型对太阳能电池板缺陷检测的均值平均精度(mean average precision, mAP)达到 95.5%。徐威等^[5]在 YOLOv7-tiny 中引入轻量化卷积模块^[6]和 GhostNetv2^[7]中的解耦全连接注意力机制(decoupled fully connected, DFC),在损失函数部分,将原本的 CIoU(complete intersection over union)^[8]替换为 EIou(expected IoU)^[9],最终所提出的网络模型在参数量和计算量方面分别相较于 YOLOv7-tiny 模型下降了 18.3% 和 16.7%,模型大小仅有 4.9 MB, mAP 相较于 YOLOv7-tiny 提升了 5.3%。

以上学者在太阳能电池片的缺陷检测方面取得了不错的成绩,但主要集中在背景简单的单晶硅,而多晶硅由于内部存在绒丝且分布差异性较大,掩盖了电池片本身的缺陷,其缺陷的识别更加困难,单晶硅和多晶硅太阳能电池片的对比图如图 1 所示。

在对多晶硅太阳能电池片的缺陷检测研究中, TANG W Q等^[10]以 Inceptionv4 和 Inception-ResNet-v2^[11]为主干网络对多晶硅太阳能电池片进行缺陷检测,通过对骨干网络添加和减少不同的残差模块,最终对多晶硅太阳能电池片缺陷的检测精度达到了 0.930。CAO Y K等^[12]将 GSConv(group-shuffle conv)、BiFPN 结构添加到 YOLOv8 中,在 PVEL-AD 多晶缺陷数据集上的 mAP 为

92.8%,模型权重减少了 16.7%。对多晶硅太阳能电池片缺陷检测的研究较少,且缺陷表达不明确,对其缺陷检测的精度和识别速度都还有研究空间。

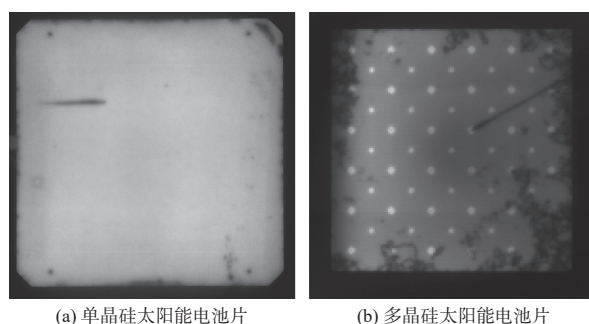


图 1 单晶硅和多晶硅太阳能电池片对比图

Fig. 1 Comparison of monocrystalline silicon and polycrystalline silicon solar cells

基于此,本文在模型权重最小的 YOLOv8 实例分割模型(YOLOv8-seg-n)的基础上提出了 YOLO-SCD 的实例分割算法,完成了对太阳能电池片缺陷更快、更高效的实例分割检测任务,旨在研究精度和速度之间的平衡。

1 YOLO-SCD 整体架构

YOLOv8 是 ultralytics 团队继 YOLOv5 之后推出的全新的 SOAT(state-of-the-art)模型,其在一次前向传播的过程中就能得到所有目标的检测结果, YOLOv8 在实际工程应用的实时检测中具有较低的延迟,且检测精度高,但是在目标尺寸变化较大的情况下会出现误检或者错检的情况, YOLOv8 需要较高的计算资源和储存空间,所以在嵌入式或者移动端部署时会受阻。

YOLOv8-seg 是基于 YOLACT(you only look at coefficients)^[13]的实例分割模型,相比 YOLOv8 目标检测模型输出头多了 Mask 分支,多晶硅太阳能电池片缺陷目标尺寸变化较大,用 YOLOv8-seg 可能会出现误检和漏检,且多晶硅太阳能电池片缺陷需要实时检测和分割,所以原始 YOLOv8-seg 可

能会不满足部署的内存大小要求。

综上所述,本文选取 YOLOv8-seg-n 为基准模型,提出了 YOLO-SCD, YOLO-SCD 通过轻量化网络的方法降低了 YOLOv8-seg-n 模型的参数量和计算量,减少了网络的推理时间,在轻量化网络的同时保证了识别和分割的精度。YOLO-SCD 网络结构图如图 2 所示,主要的改进点如下:1) 将骨干网络的最后两个卷积块换成 Shuffle_Block 模块,在

实现对骨干网络轻量化的同时增强通道之间的信息交流;2) 将 CBAM 注意力机制加入骨干网络,提升轻量化网络后的检测精度和分割精度;3) 将颈部网络的 C2f 模块替换为 C2f_star 模块,使颈部网络具有更好的非线性和更高的维度,增强网络拟合数据的能力;4) 将颈部网络的卷积模块换成部分通道下采样模块,在减少冗余特征图的同时使网络更轻量化。

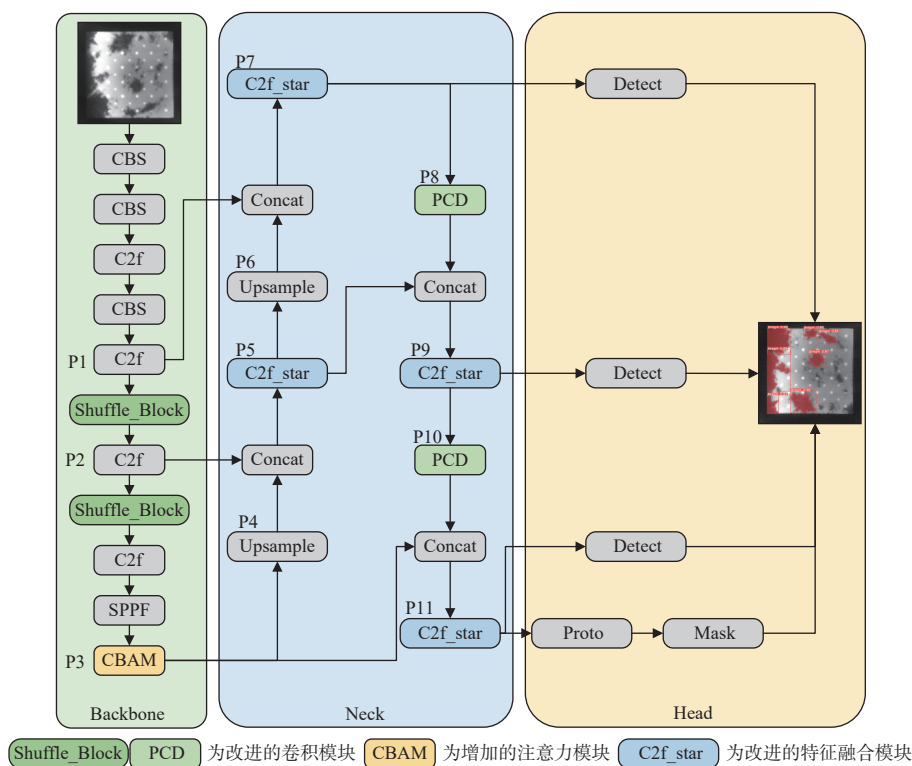


图2 YOLO-SCD 网络模型结构图

Fig. 2 Structure diagram of YOLO-SCD network model

1.1 骨干网络

YOLOv8-seg-n 的骨干网络主要是对输入进行特征提取,首先经过两个 CBS(Conv2d+BatchNormal2d+SiLU)模块对输入进行 4 倍下采样,输入图片尺寸从 640×640 像素变为 160×160 像素,通道数从 3 扩大为 32;然后经过一个 C2f 模块,C2f 模块中通过残差连接将浅层特征和深层特征进行融合,增强了骨干网络特征提取的能力;后面紧接着 3 组 CBS 模块和 C2f 模块进行特征提取;最后一层是 SPPF(spatial pyramid pooling fast)模块,SPPF 模块可以提取不同尺度的特征信息,进行多尺度特征融合,提高网络的目标识别精度。最终 YOLOv8-seg-n 骨干网络的输出特征图尺寸为 20×20 像素,通道数为 256。

但是对多晶硅太阳能电池片的缺陷检测需要具有实时性和高效性,所以 YOLO-SCD 在骨干网络中引入了 Shuffle_Block 模块和 CBAM(convolutional block attention module)^[14] 注意力模块,Shuffle_Block 模块实现对骨干网络的轻量化,CBAM 来弥补轻量化网络后网络对多晶硅太阳能电池片缺陷的识别和分割精度下降问题。

1.1.1 Shuffle_Block 模块

为了提高 YOLO-SCD 对多晶硅太阳能电池片缺陷的检测速度,对骨干网络中部分的下采样模块进行了替换。图 3 为轻量化的 Shuffle_Block 下采样模块,输入特征图分为两个路线,分别通过一个 3×3 的 DWConv 进行下采样,两个分支的输出通道数都是 OutChannels 的一半,然后通过特征图拼

接使输出通道数为 OutChannels, 最后经过 Channels Shuffle 处理输出, 用 Shuffle_Block 代替骨干网络

最后两个普通卷积进行下采样, 可以减少网络的参数量和计算量, 同时增加网络的推理速度。

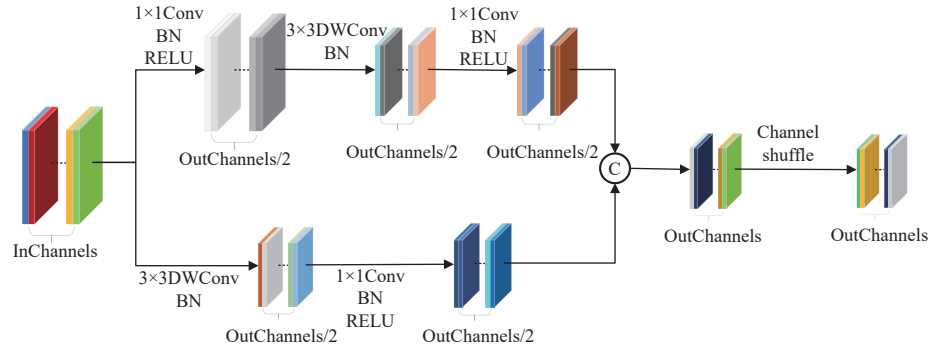


图3 Shuffle_Block 网络结构图

Fig. 3 Structure diagram of shuffle_Block network

Channels Shuffle 操作是 Shuffle_Block 的核心。在卷积神经网络中, 普通卷积会最大限度地保持特征图通道之间的信息交流, 但计算成本很高, 而分组卷积可以降低网络的计算成本, 但是各组之间没有通道通信, 会降低网络的特征提取能力。虽然使用 1×1 密集卷积可以解决这个问题, 但是会削弱分组卷积的轻量化效果。为了解决这个问题, Shufflenet^[15] 提出了 Channels Shuffle 的思想, 将输入特征图分为 G 组后经过分组卷积 (GConv), 将得到的特征图再分为 G 组, 然后经过 Channels Shuffle 操作, 这样就可以融合各组的特征信息, 增加网络提取特征的能力, 其原理如图 4 所示。

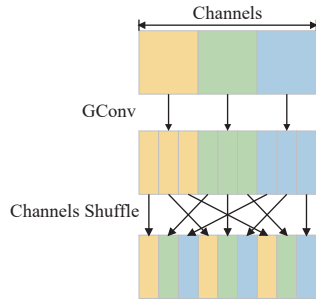


图4 Channels Shuffle 原理

Fig. 4 Principle of Channels Shuffle

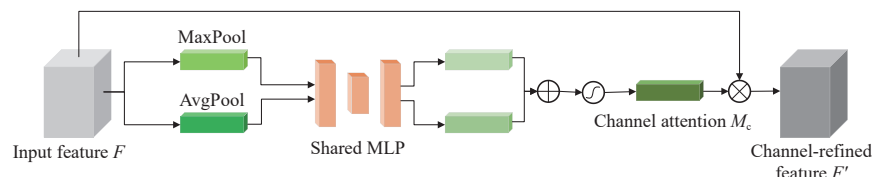


图5 CBAM 通道注意力模块

Fig. 5 CBAM channel attention module

如图 6 所示, 空间注意力模块将通道注意力模块的输出特征图沿通道维度通过最大池化和平均

池化得到两个 $1 \times H \times W$ 的特征图, 然后经过 Concat 操作对两个特征图进行拼接, 通过普通卷积后经过

1.1.2 CBAM 模块

在轻量化网络的同时会伴随着 YOLO-SCD 多晶硅太阳能电池片检测和分割精度的下降, 本文引用 CBAM 注意力机制来提升轻量化网络后的识别精度和分割精度, CBAM 是一种轻量化的注意力模块, 串联了通道注意力模块和空间注意力模块。

如图 5 所示, 在通道注意力模块中, 首先将输入特征图经过两个并行的最大池化层和平均池化层, 将输入特征图 $C \times H \times W$ 变为 $C \times 1 \times 1$, 通过 Shared MLP 模块将通道数压缩再扩到原通道数, 最后将两个特征图相加后通过 Sigmoid 激活函数将权重变为 $0 \sim 1$, 然后与原图相乘得到通道注意力模块的输出结果, 用公式表示为

$$F' = F \times (\sigma(\text{MLP}(\text{AvgPool}(F)) + \text{MLP}(\text{MaxPool}(F)))) \quad (1)$$

式中: F' 为通道注意力模块输出特征图; F 为输入特征图; σ 为 Sigmoid 激活函数; MLP 为 Conv-ReLU-Conv 结构; AvgPool 和 MaxPool 分别为全局平均池化层和全局最大池化层。

Sigmoid 得到空间注意力权重,最后将空间注意力权重乘输入特征图得到输出特征图,用公式表示为

$$F'' = F' \times (\sigma(f([\text{AvgPool}(F'); \text{MaxPool}(F')])))) \quad (2)$$

式中: F'' 为空间注意力模块输出特征图; σ 为 Sigmoid 激活函数; f 表示卷积运算; AvgPool 和 MaxPool

分别为全局平均池化层和全局最大池化层。

将 CBAM 插入到骨干网络中 SPPF 的下一层,能够帮助网络关注和学习关键特征,增强网络的泛化能力并提高网络检测和分割太阳能电池片缺陷的精度。CBAM 总体网络结构如图 7 所示。

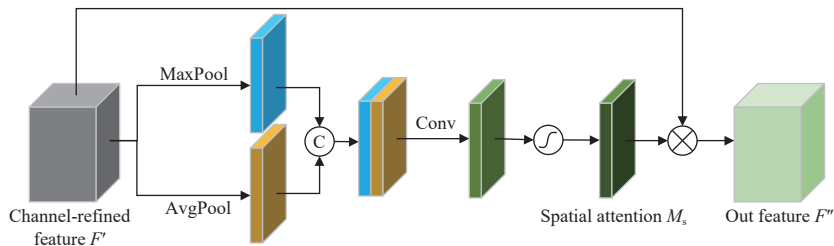


图 6 CBAM 空间注意力模块

Fig. 6 CBAM spatial attention module

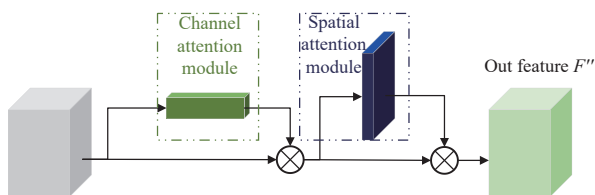


图 7 CBAM 总体网络结构图

Fig. 7 Structure diagram of CBAM general network

为了探究 CBAM 帮助网络学习关键特征的能力,验证 CBAM 对本文数据集的效果,在只对 YOLOv8-seg-n 添加 CBAM 注意力模块的情况下,对网络进行了输出特征热力图可视化实验,效果如图 8 所示。从图中可看出,加入 CBAM 模块后的网络比原网络更加关注缺陷信息,增强了网络提取关键特征的能力。

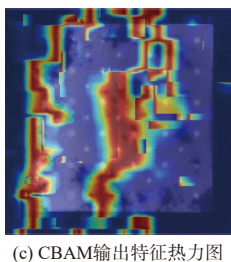
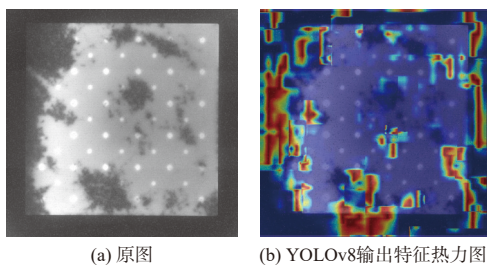


图 8 输出特征热力图

Fig. 8 Output characteristic heat map

1.2 颈部网络

YOLOv8-seg-n 的颈部网络采用 FPN^[16]+PANet^[17] 的结构进行多尺度特征融合,如图 2 所示,先将通过骨干网络输出的特征图进行上采样,把特征图尺寸变为 40×40 像素,然后将 P4 层输出的特征图和 P2 层输出的特征图进行拼接,经过一个 C2f 模块进行特征提取,接着将特征图上采样到 80×80 像素,再将 P6 层输出的特征图和 P1 层输出的特征图进行拼接,接着经过两次下采样将特征图尺寸再次降为 20×20 像素,在下采样的过程中,将 P8 层输出的特征图和 P5 层输出的特征图、P10 层和 P3 层输出的特征图进行拼接,每次拼接后都跟着一个 C2f 进行特征提取。

但是在 C2f 的结构中,通常会使用多个 Bottleneck 增强网络的特征提取能力,本文将 C2f 模块中的多个 Bottleneck 模块替换为 1 个 Star_Block 模块,形成 C2f_star 模块,在减少网络宽度的同时增加了网络的检测精度。同时,在下采样的过程中,冗余的特征图信息会增加网络的计算量,降低网络的推理速度,对此,本文提出的部分通道下采样模块,解决了冗余特征图的问题。

1.2.1 C2f_star 模块

在 C2f 中利用特征图矩阵元素相乘的方法进行特征融合。在特征聚合网络的设计中,最常用的是特征图矩阵元素相加的方法,而元素相加需要通过增加网络宽度和使用激活函数来将低维输入映射到高维非线性的特征空间中。最近特征图矩阵元素相乘的方法得到越来越多学者的研究,元素相乘可以把低维输入映射到超高维度的非线性空间中,虽然元素相乘看起来比元素相加更复

杂,但在移动端的延迟却更低^[18]。

在单层神经网络中,元素相乘可以写成 $(W_1^T X + B_1) \times (W_2^T X + B_2)$ 。为简化带偏置的神经元输出表达式,我们引入增广权重向量 W 和增广输入向量 X 。其中, $W = \begin{bmatrix} W' \\ B \end{bmatrix}$, $X = \begin{bmatrix} X' \\ 1 \end{bmatrix}$, W' 为原始权重向量, B 为偏置项, X' 为原始输入向量。则可以将元素相乘写成 $(W_1^T X) \times (W_2^T X)$ 。当输入通道为 $(d+1)$ 时,可以把元素相乘写为

$$w_1^T x \times w_2^T x = \left(\sum_{i=1}^{d+1} w_1^i x^i \right) \times \left(\sum_{j=1}^{d+1} w_2^j x^j \right) = \alpha_{(1,1)} x^1 x^1 + \dots + \alpha_{(5,6)} x^5 x^6 + \dots + \alpha_{(d+1,d+1)} x^{d+1} x^{d+1} \quad (3)$$

式中: i, j 为通道数的索引; α 为 x 的系数; w 为权重。

$$\alpha_{(i,j)} = \begin{cases} w_1^i w_2^j & i = j \\ w_1^i w_2^j + w_1^j w_2^i & i \neq j \end{cases} \quad (4)$$

从式(3)中可以看出,展开后的每一项都与 x 呈非线性,且显著放大了维度,使用元素相乘的方法可在不增加网络宽度的同时增加特征维度和非线性,能更好地提取特征信息。

图9是Star_Block的网络结构。在Star_Block的网络中,输入特征图首先经过DWConv进行特征提取,之后将特征图分为两条路径进行特征提取,在对这两条路径进行特征融合时,不再是传统的特征图元素相加的方式,而是特征图元素相乘的方式,接着再对特征图进行特征提取。用Star_Block替换C2f中的Bottleneck,形成C2f_star结构,把YOLOv8-seg-n中颈部网络的C2f更换成C2f_star(见图10),在保证检测精度和分割精度的同时减少了网络的宽度和参数量,并加快了网络的推理速度。

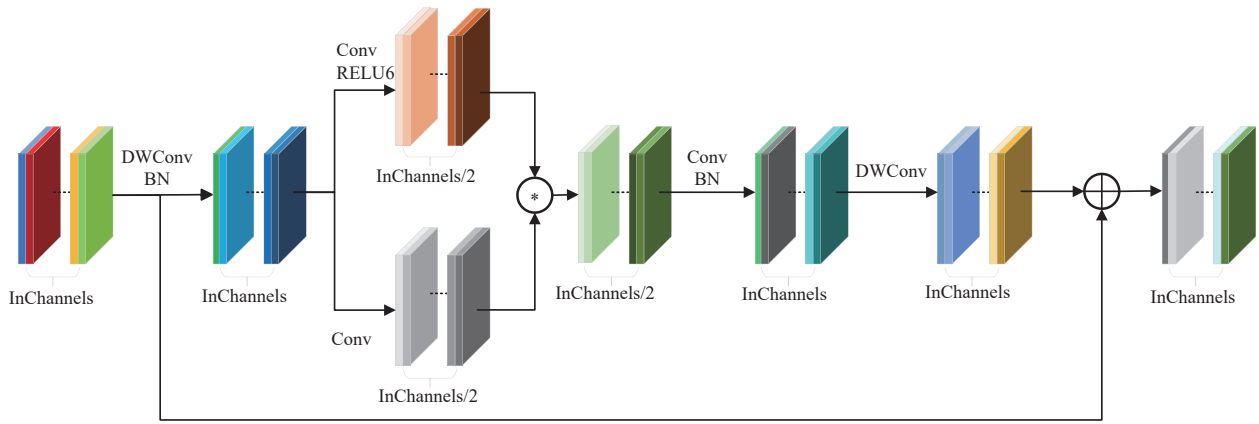


图9 Star_Block网络结构

Fig. 9 Architecture of Star_Block network

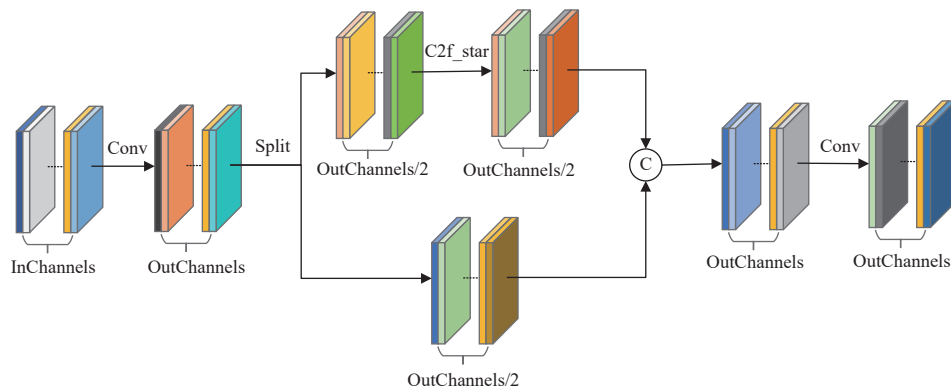


图10 C2f_star网络结构

Fig. 10 C2f_star network structure

1.2.2 部分通道下采样模块

通道数会随着下采样而增加,过多的通道数会

产生冗余的特征图,同时增加网络的参数量和计算量,从而使网络在推理过程中具有较高的延

迟。为了解决冗余特征图信息所带来的不足,借鉴了 FasterNet^[19] 中的 PConv(见 图 11), PConv 针对“过多的通道数会产生冗余的特征图”这个问题而设计,其设计核心就是把一部分通道的特征图进行卷积来得到新的特征图,另一部分通道的特征图恒等映射得到输出特征图,从而减少对通道的处理,减少网络的参数量和计算量。

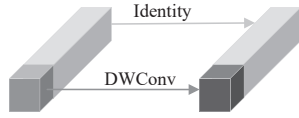


图 11 PConv 结构

Fig. 11 PConv structure

假设输入为 $h \times w \times c$ 的特征图, 普通卷积的 FLOPs(floating point operations per second)为

$$h \times w \times k^2 \times c^2 \quad (5)$$

式中: h 、 w 、 c 分别为特征图的高、宽、通道数; k 为卷积核大小。

而 PConv 的 FLOPs 为

$$h \times w \times k^2 \times c_p^2 \quad (6)$$

式中 c_p 为 PConv 中参与卷积的通道数。

当 $\frac{c_p}{c} = \frac{1}{4}$ 时, PConv 的 FLOPs 仅为普通卷积的 $\frac{1}{16}$, 明显减少了网络的计算量。

本文基于 PConv 设计了部分通道下采样模块 (partial channel downsampling, PCD)。如图 12 所示, 首先经过 PConv 进行特征提取, 但 PConv 中用的是 DWConv, DWConv 是特殊的分组卷积, 有多少个输入通道数就分为几组, 这样通道之间没有了信息交流, 仅用 PConv 会降低网络的精度, 所以 PConv 后通常会跟一个逐点卷积来进行通道信息之间的关联, 在残差连接后加入 SiLU 激活函数和归一化操作, 最后经过一个 3×3 的 DWConv 进行下采样和改变通道数。用 PCD 替换 Neck 中的 Conv 可以减少网络的计算量。

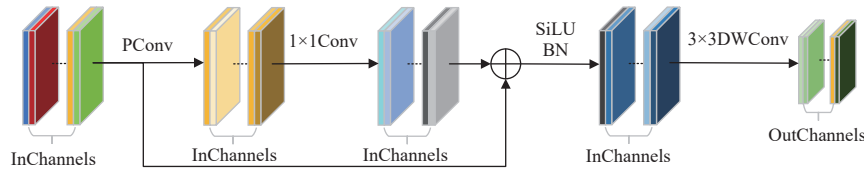


图 12 PCD 网络结构图

Fig. 12 Structure diagram of PCD network

1.3 头部网络

YOLOv8-seg-n 的头部网络对预测结果进行输出, 头部网络接收 P7、P9 和 P11 输出的特征图信息, 将特征图信息输入到解耦头进行预测。其结构如图 13 所示。

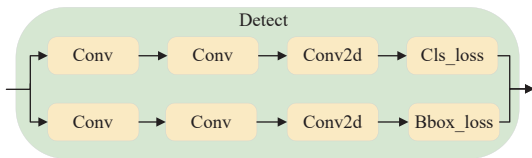


图 13 YOLOv8-seg 检测头结构图

Fig. 13 YOLOv8-seg detector head structure diagram

YOLOv8-seg-n 头部网络的 Loss 计算包括正负样本分配策略和 loss 计算两个部分, 正负样本分配策略采用的是 TOOD 的 TaskAlignedAssigner 动态分配策略, TaskAlignedAssigner 动态分配策略会根据分类和回归的加权分数来选择正样本, 首先计算真实框和预测框的匹配度。

$$\text{aligned_metric} = s^\alpha \times t^\beta \quad (7)$$

式中: s 为预测类别的分值; u 为预测框和真实框的 CIoU 值; α 和 β 为超参数; align_metric 就是真实框和预测框的匹配度。align_metric 的值越接近 1, 就越符合正样本的标准。

loss 计算也包括两个分支, 分别为分类 loss 计算和回归 loss 计算, 分类采用二元交叉熵损失 (binary cross-entropy loss, BCE Loss)。

$\text{loss}(0, t) =$

$$-\frac{1}{n} \sum_i (t[i] \times \log(o[i]) + (1-t[i]) \times \log(1-o[i])) \quad (8)$$

式中: $o[i]$ 为预测的概率; $t[i]$ 为真实标签值。BCE Loss 值越小, 分类越准确。

回归采用了 CloU Loss 和 DFL (distribution focal loss)^[20]。CloU Loss 是交并比 (intersection over union, IoU) 损失函数, 使锚框更加接近真实标签, DFL 是学习标签周围位置的损失, 可以使网络更快聚焦到目标位置或邻近区域。

$$\text{DFL}(S_i, S_{i+1}) = -((y_{i+1} - y) \log(S_i) + (y - y_i) \log(S_{i+1})) \quad (9)$$

式中: S 为输出分布; y 为浮点值; y_i 和 y_{i+1} 为 y 的左右整数值。

在训练中, CloU Loss 和 DFL 结合使用, 对预测框的每条边和整体预测框进行优化。Proto 模块通过上采样生成高质量的分割掩码, 通过与真实掩码比较来计算损失。

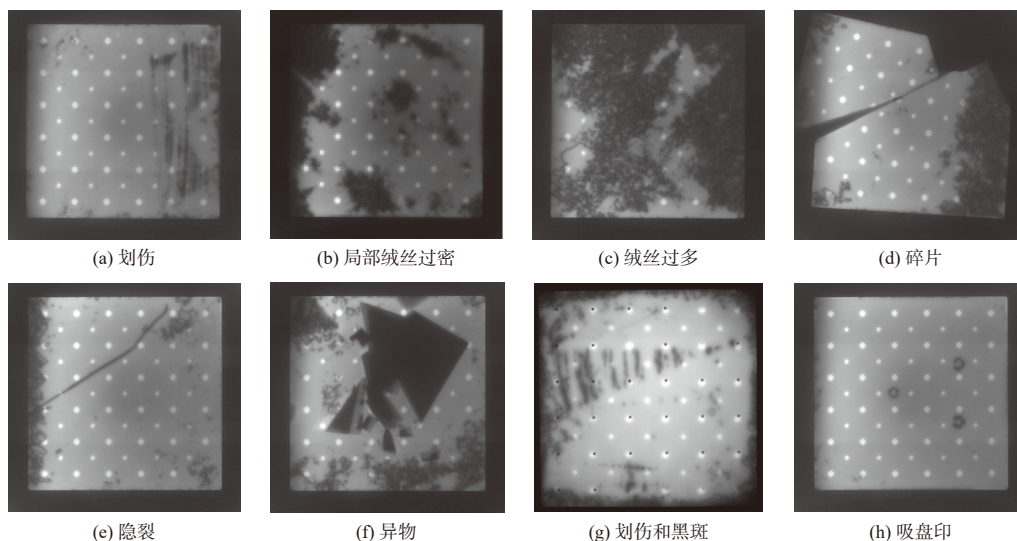


图 14 PL 缺陷数据集

Fig. 14 PL defective data set

由于多晶硅太阳能电池片 PL 缺陷数据集过少, 通过翻转、旋转、改变亮度、增加噪声的数据增强方式, 把数据集扩充为 1395 张, 并按 8 : 1 : 1 分为训练集、验证集和测试集。如图 15 所示, 使用 Labelme 标注软件对缺陷进行标注, 标注后的结果为 json 格式, 需转换成 YOLO 所需的 txt 格式来进行训练和验证。

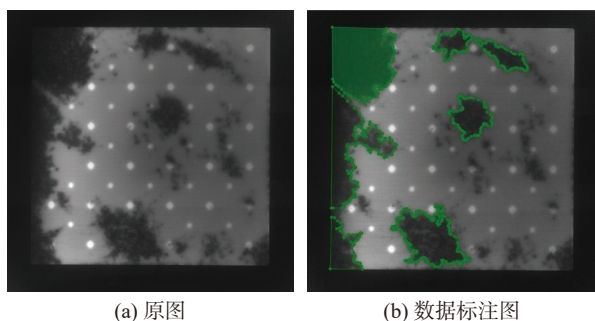


图 15 数据处理

Fig. 15 Data processing

2.2 实验环境

为了更加有效地突出本文的算法优势, 本文所有的模型都是在统一环境下进行训练和验证, 其

2 实验结果与分析

2.1 多晶硅太阳能电池片 PL 缺陷数据集

本文使用的数据集为自制多晶硅太阳能电池片的 PL 缺陷数据集, 数据集中多晶硅太阳能电池片 PL 缺陷共有 8 种, 分别是局部绒丝过密、划伤、绒丝过多、碎片、吸盘印、隐裂、异物和黑斑。如图 14 所示。

中 GPU 为 GeForce RTX4000(16 GB), CPU 为 6 核 E5-2680v4, PyTorch 版本为 2.0.1。训练参数如表 1 所示, 其中 YOLOv8-seg-n 采用 Mosaic 数据增强, 并在后 10 轮关闭数据增强。

表 1 训练环境参数

Table 1 Parameters of the training environment

Parameter type	Parameter value
Optimizer	SGD
Learning rate	0.01
Training epoch	100
Batch size	16
Image size/pixel	640 × 640

2.3 评价指标

为了在移动端部署模型, 模型的检测精度、分割精度、模型大小和推理速度都要综合考量, 检测精度用 $mAP(B)$ 来衡量, AP 是精度 (Precision, P) 和召回率 (Recall, R) 曲线所围成的面积, 计算公式为

$$P = \frac{N_{TP}}{N_{TP} + N_{FP}} \quad (10)$$

$$R = \frac{N_{TP}}{N_{TP} + N_{FN}} \quad (11)$$

$$AP = \int_0^1 P(R) dR \quad (12)$$

$$mAP(B) = \frac{1}{n} \sum_{i=1}^n AP_i \quad (13)$$

式中: N_{TP} 表示样本为正且被预测为正的样本数量; N_{FP} 为样本为负但被预测为正的样本数量; N_{FN} 表示样本为正, 但被预测为负的样本数量。

分割精度用 $mAP(M)$ 来衡量, 其计算公式和 $mAP(B)$ 相同, 模型大小用参数量 (Params)、计算量 (GFLOPs) 和模型权重大小来衡量, 推理速度用帧速 (frame per second, FPS) 来衡量, FPS 表示 1 s

中网络可以处理多少张图片, 而测量 FPS 会有波动, 所以本文的 FPS 是测量 10 次后取平均:

$$FPS = \frac{1}{10} \sum_{i=1}^{10} \frac{1000}{t_{pre} + t_{inf} + t_{post}} \quad (14)$$

式中: t_{pre} 为预处理时间; t_{inf} 为推理时间; t_{post} 为后处理时间。

2.4 PCD 替换位置实验

如图 16 所示, 为了探究 PCD 模块加在 YOLOv8 的哪个位置效果最好, 分别对 PCD 模块加在 YOLOv8-seg-n 的骨干网络 (实验 1)、颈部网络 (实验 2)、骨干网络+颈部网络 (实验 3) 进行实验。

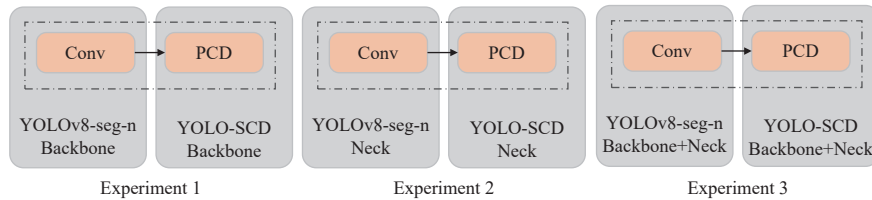


图 16 探究 PSD 模块位置实验

Fig. 16 Experiment for exploring position of PSD module

结果如表 2 所示, 实验表明, PCD 更适合颈部网络, 用 PCD 替换 YOLOv8-seg-n 颈部网络的 Conv, 检测精度达到了 0.976, 分割精度略微下降。虽然把 PCD 模块加在主干网络和把 YOLOv8-seg-n 的

Conv 都替换成 PCD 的 FPS, 相较于用 PCD 替换 YOLOv8-seg-n 颈部网络的 Conv 提升较多, 但检测精度和分割精度略微下降, 综合精度和速度考量, 把 PCD 加在 YOLOv8-seg-n 的颈部网络效果最好。

表 2 PCD 位置对比实验

Table 2 Comparison experiment of PCD position

Model	mAP(B)	mAP(M)	Params	GFLOPs	Size/MB	FPS
YOLOv8-seg-n	0.969	0.970	3.26M	12.0	6.8	85.20
Experiment 1	0.962	0.953	2.91M	11.4	6.1	90.90
Experiment 2	0.976	0.952	3.11M	11.9	6.5	88.14
Experiment 3	0.963	0.956	2.76M	11.4	5.8	92.03

2.5 C2f_star 替换位置实验

如图 17 所示, 为了探究 C2f_star 添加到 YOLOv8-seg-n 的哪个位置效果最好, 也分别添加

在 YOLOv8-seg-n 的骨干网络 (实验 4)、颈部网络 (实验 5)、骨干网络+颈部网络 (实验 6) 进行实验。

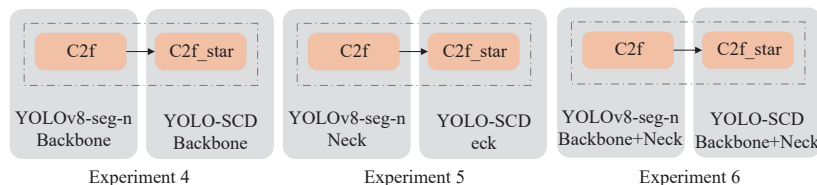


图 17 探究 C2f_star 位置实验

Fig. 17 Experiment for exploring position of C2f_star

结果如表3所示,在C2f_star加到特征融合网络中,C2f_star的元素相乘使模型具有更好的非线性,并且使特征权重达到了更高的维度,使模型具有更好的边界决策,在轻量化网络的同时提升了网络的检测精度。

而把C2f_star加在骨干网络中网络的识别精度和分割精度分别下降了0.02和0.021,说明元素相乘在特征提取网络优势不大,综合精度、模型大小和速度考量,将C2f_star加在YOLOv8-seg-n的颈部网络效果最好。

表3 C2f_star位置对比实验

Table 3 Comparison experiment of C2f_star position

Model	mAP(B)	mAP(M)	Params	GFLOPs	Size/MB	FPS
YOLOv8-seg-n	0.969	0.970	3.26 M	12.0	6.8	85.20
Experiment 4	0.949	0.949	3.10 M	11.7	6.5	89.63
Experiment 5	0.971	0.963	3.06 M	11.6	6.4	89.50
Experiment 6	0.947	0.945	2.90 M	11.3	6.1	90.74

2.6 对比试验

首先将本文提出的YOLO-SCD和YOLOv5-seg-n、YOLOv8-seg-n进行对比实验,YOLOv5-seg-n和YOLOv8-seg-n都是ultralytics团队所开发的单阶段实例分割网络模型,且参数量和计算量相对于YOLOv7和YOLOv9具有明显的轻量化优势(见表4)。YOLOv5-seg-n、YOLOv8-seg-n和本

文所提出的YOLO-SCD对多晶硅太阳能电池片PL缺陷数据集的mAP(B)和mAP(M)对比图见图18和图19。从图中可看出,YOLOv8-seg-n的mAP(B)和mAP(M)一直优于YOLOv5-seg-n,前40轮次更为明显,而本文所提出的YOLO-SCD比YOLOv8-seg-n略胜一筹,更加适用于PL缺陷数据集。

表4 对比试验

Table 4 Comparative tests

Model	mAP50(B)	mAP50(M)	Params	GFLOPs	Size/MB	FPS
YOLOv8-seg-n	0.969	0.970	3.26 M	12.0	6.8	85.20
YOLOv7-seg	0.978	0.964	37.88 M	142.0	76.4	50.00
YOLOv9-seg	0.985	0.967	27.37 M	144.3	55.7	46.00
YOLO-SCD	0.970	0.962	2.67 M	11.3	5.7	90.03

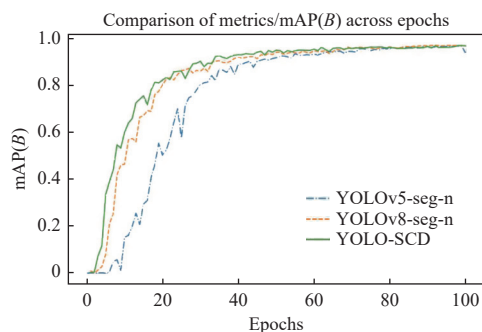


图18 YOLOv5-seg-n、YOLOv8-seg-n和YOLO-SCD mAP(B)对比图

Fig. 18 Comparison of YOLOv5-seg-n, YOLOv8-seg-n and YOLO-SCD mAP(B)

为了更好地评估本文模型的性能,在实验环境一致、默认超参数保持一致的情况下,用YOLOv8-seg-n、YOLOv7-seg、YOLOv9-seg来进行对比实

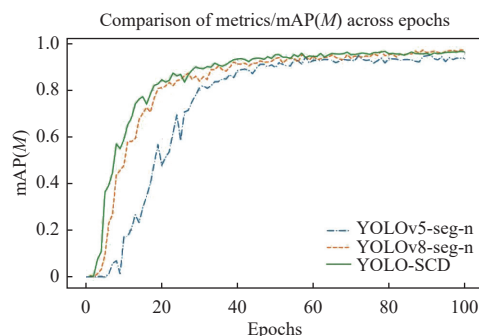


图19 YOLOv5-seg-n、YOLOv8-seg-n和YOLO-SCD mAP(M)对比图

Fig. 19 Comparison of YOLOv5-seg-n, YOLOv8-seg-n, and YOLO-SCD mAP(M)

验,实验结果如表4所示,从表中可看出YOLO-SCD与YOLOv8-seg-n的mAP(B)和mAP(M)相差无几,但由于PCD、Shuffle_Block、C2f_star轻量化

模块的加入, YOLO-SCD 的参数量和计算量分别比 YOLOv8-seg-n 减少了 18% 和 6%, FPS 增加了 4.83, 达到了速度与精度之间的平衡, 虽然 YOLOv7-seg、YOLOv9-seg 的 $mAP(B)$ 和 $mAP(M)$ 相对来说比 YOLO-SCD 较高, 但是 YOLOv7-seg 的参数量和计算量相比 YOLO-SCD 增加了 13 倍和 11.6 倍, YOLOv9-seg 的参数量和计算量相对于 YOLO-SCD 增加了 9.2 倍和 11.8 倍, 最后 YOLOv7-seg 和 YOLOv9-seg 的权重大小达到了 76.4 MB 和 55.7 MB。部署工作讲究精度和速度之间的平衡, YOLOv7-seg 和 YOLOv9-seg 在追求精度的同时忽略了模型大小和推理速度, YOLO-SCD 的 FPS 相对于其他模型最高, 且检测精度和分割精度分别达到了 0.97 和 0.962, 综合考量精度、参数量、计算量和推理速度, 本文所提出的 YOLO-SCD 相比于其他模型更适合部署。

2.7 消融实验

为了探究各个模块对网络的提升效果, 对各个模块进行了消融实验, 结果如表 5 所示。从表中

可看出, 在基准网络 YOLOv8-seg-n 中分别加入 Shuffle_Block 模块、PCD 模块、C2f_star 模块都降低了网络的参数量, 并且提高了网络的 FPS; 只加 Shuffle_Block 时, 网络的 $mAP(B)$ 达到了 0.973, 说明 Channels Shuffle 操作增加了通道之间的信息交流, 能使网络更好地提取缺陷特征信息; 而 PCD 剔除冗余通道信息也使网络的 $mAP(B)$ 增加, 当只加 PCD 模块、C2f_star 模块时网络的检测精度和分割精度达到了最大, 但是此时对基准网络的轻量化并不明显; 当在主干网络中加入 Shuffle_Block 模块后进一步达到更好的轻量化效果, 但此时检测精度和分割精度略微下降, 在主干网络后加上 CBAM 注意力机制后弥补了轻量化带来的精度下降问题。图 20 展示了 YOLO-SCD 检测和实例分割的效果, 从图中可以看出检测效果和分割精度都很精确, 检测的置信度也达到了很高的水平, 并没有因为对网络的轻量化而出现漏检、误检的问题, 且改进后的网络模型大小只有 5.7 MB, FPS 也达到了 90.03, 为之后在移动端的部署工作提供了可能。

表 5 消融实验
Table 5 Ablation experiments

Shuffle_Block	PCD	C2f_star	CBAM	$mAP(B)$	$mAP(M)$	Params	FPS
—	—	—	—	0.969	0.970	3.26 M	85.20
√	—	—	—	0.973	0.957	2.96 M	89.50
—	√	—	—	0.976	0.963	3.11 M	88.14
—	—	√	—	0.971	0.963	3.06 M	86.47
—	—	—	√	0.967	0.960	3.33 M	84.20
—	√	√	—	0.979	97.10	2.91 M	87.80
√	√	√	—	0.966	0.953	2.61 M	91.09
√	√	√	√	0.970	0.962	2.67 M	90.03

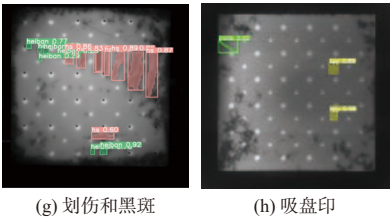
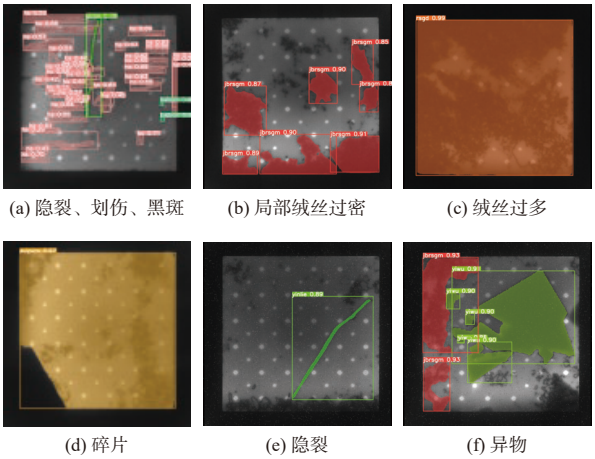


图 20 实例分割结果
Fig. 20 Example segmentation results

3 结论

针对多晶硅太阳能电池片缺陷的识别与分割,

本文基于 YOLOv8-seg-n 提出轻量化 YOLO-SCD 网络, 不仅对网络的参数量和计算量进行了轻量化, 还关注网络的推理速度, 通过在主干网络加入 Shuffle_Block 模块, 在轻量化网络的同时增强了特征图通道之间的信息交流, 通过增加 CBAM 注意力模块增加网络对关键特征的学习能力。通过在颈部网络加入 PCD 模块减少特征图冗余信息, 在颈部网络加入 C2f_star 使网络具有更好的非线性和更高的维度。最终, 本文所提出的网络检测精度达到了 0.970, 分割精度达到了 0.962, 模型权重只有 5.7 MB, 并且 FPS 达到了 90.03 的, 为之后的部署工作打下了基础。

参考文献:

- [1] 刘玉淇, 吴一全. 基于机器视觉的太阳能电池片缺陷检测算法综述[J]. *光学精密工程*, 2024, 32(6): 868-900.
LIU Yuqi, WU Yiquan. A review of defect detection algorithms for solar cells based on machine vision[J]. *Optics and Precision Engineering*, 2024, 32(6): 868-900.
- [2] HUANG J, ZENG K Y, ZHANG Z J, et al. Solar panel defect detection design based on YOLOv5 algorithm[J]. *Heliyon*, 2023, 9(8): e18826-e18826.
- [3] TAN M X, PANG R M, LE Q V. Efficientdet: scalable and efficient object detection[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. New York: IEEE, 2020: 10781-10790.
- [4] GE Z, LIU S T, WANG F, et al. YOLOx: exceeding YOLO series in 2021[EB/OL]. [2024-07-10]. <https://arxiv.org/pdf/2107.08430v2>.
- [5] 徐威, 李为相, 方志, 等. 基于改进 YOLOv7-tiny 的光伏电池缺陷检测算法[J]. *计算机工程与应用*, 2024, 15: 336-343.
XU Wei, LI Weixiang, FANG Zhi, et al. Photovoltaic cell defect detection algorithm based on improved YOLOv7-tiny[J]. *Computer Engineering and Applications*, 2024, 15: 336-343.
- [6] LI H I, LI J, WEI H B, et al. Slim-neck by gsconv: a better design paradigm of detector architectures for autonomous vehicles[EB/OL]. (2022-06-06) [2024-7-19]. <https://arxiv.org/abs/2206.02424.pdf>.
- [7] TANG Y, HAN K, GUO J, et al. Ghostnetv2: enhance cheap operation with long-range attention[J]. *Advances in Neural Information Processing Systems*, 2022, 35: 9969-9982.
- [8] ZHENG Z, WANG P, LIU W, et al. Distance-iou loss: faster and better learning for bounding box regression[J]. *Proceedings of the AAAI Conference on Artificial Intelligence*, 2020, 34(7): 12993-13000.
- [9] ZHANG Y F, REN W, ZHANG Z, et al. Focal and efficient iou loss for accurate bounding box regression[J]. *Neurocomputing*, 2022, 506: 146-157.
- [10] TANG W Q, YANG Q, HU X C, et al. Convolution neural network based polycrystalline silicon photovoltaic cell linear defect diagnosis using electroluminescence images[J]. *Expert Systems with Applications*, 2022, 202: 117087.
- [11] CHRISTIAN S, SERGRY I, VINCENT V, et al. Inception-v4, inception-resnet and the impact of residual connections on learning[EB/OL]. (2016-08-23) [2024-7-19]. <https://arxiv.org/pdf/1602.07261.pdf>.
- [12] CAO Y K, PANG D D, ZHAO Q C, et al. Improved YOLOv8-gd deep learning model for defect detection in electroluminescence images of solar photovoltaic modules[J]. *Engineering Applications of Artificial Intelligence*, 2024, 131: 107866.
- [13] BOLYA D, ZHOU C, XIAO F Y, et al. Yolact: real-time instance segmentation[C]//Proceedings of the IEEE/CVF International Conference on Computer Vision. New York: IEEE, 2019: 9157-9166.
- [14] WOO S, PARK J, LEE J Y, et al. CBAM: convolutional block attention module[C]//Proceedings of the European Conference on Computer Vision (ECCV). [S. l.]: [s. n.], 2018: 3-19.
- [15] ZHANG X, ZHOU X, LIN M, et al. Shufflenet: an extremely efficient convolutional neural network for mobile devices[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. New York: IEEE, 2018: 6848-6856.
- [16] LIN T Y, DOLLAR P, GIRSHICK R, et al. Feature pyramid networks for object detection[C]//2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Honolulu, HI, USA: IEEE, 2017: 936-944.
- [17] LIU S, QI L, QIN H F, et al. Path aggregation network for instance segmentation[C]//2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Salt Lake City, UT, USA: IEEE, 2018: 8759-8768.
- [18] MA X, DAI X, BAI Y, et al. Rewrite the stars[EB/OL]. (2024-03-29) [2024-07-19]. <https://arxiv.org/pdf/2403.19967.pdf>.
- [19] CHEN J, KAO S, HE H, et al. Run, don't walk: chasing

higher flops for faster neural networks[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. New York: IEEE, 2023: 12021-12031.

[20] LI X, WANG W, WU L, et al. Generalized focal loss: learning qualified and distributed bounding boxes for dense object detection[J]. Advances in Neural Information Processing Systems, 2020, 33: 21002-21012.