

# 基于 DeepLabV3+改进的光伏板语义分割模型研究<sup>\*</sup>

王 银 孙海顺 谢 刚 赵志诚 谢新林

(太原科技大学电子信息工程学院 太原 030024)

**摘要:** 从光伏板的红外图像中分割提取出光伏板区域信息,可以极大提高光伏板故障检测的精度。而传统语义分割算法对光伏板的边界信息处理效果不佳,存在光伏板边界呈波浪状、互相黏连以及背景误分割等情况。针对此类情况,本文提出了一种基于改进 DeepLabV3+ 的光伏板语义分割算法模型,将主干网络更改为 MobileNetV2,引入 Canny 边缘检测算法输出新的浅层特征语义信息;设计 SE-ASPP 模块对特征通道进行重新校准,增强网络表达能力;增加浅层特征语义信息通道数,加强对浅层特征语义信息的关注。实验结果表明,改进后 DeepLabV3+ 算法模型的精准率、mIoU、召回率和 F1 分数分别达到 99.50%、99.21%、99.61% 和 99.55%,与原 DeepLabV3+ 模型相比,分别提高了 2.24%、1.58%、1.57% 和 1.72%,在实际分割任务中表现出色,具有更高的检测精度和可靠性。

**关键词:** 光伏板;语义分割;DeepLabV3+;MobileNetV2

中图分类号: TN29 文献标识码: A 国家标准学科分类代码: 510.4050

## Research on the improved semantic segmentation model of photovoltaic panels based on DeepLabV3+

Wang Yin Sun Haishun Xie Gang Zhao Zhicheng Xie Xinlin

(School of Electronic Information Engineering, Taiyuan University of Science and Technology, Taiyuan 030024, China)

**Abstract:** The segmentation and extraction of PV panel region information from infrared images of PV panels can greatly improve the accuracy of PV panel fault detection. However, the traditional semantic segmentation algorithm is not effective in processing the boundary information of PV panels, and there are cases that the boundary of PV panels is wave-like, sticking to each other, and the background is mis-segmented. To solve this situation, this paper proposes a semantic segmentation algorithm model for PV panels based on improved DeepLabV3+, which changes the backbone network to MobileNetV2, introduces the Canny edge detection algorithm to output new shallow feature semantic information, and designs the SE-ASPP module to re-calibrate the feature channels to enhance the network expression capability, and increase the number of channels of shallow feature semantic information to strengthen the attention to shallow feature semantic information. Experimental results show that the precision, mIoU, recall and F1 score of the improved DeepLabV3+ algorithm model reach 99.50%、99.21%、99.61% and 99.55%, respectively, which are 2.24%、1.58%、1.57% and 1.72% higher than the original DeepLabV3+ model, respectively. Improved DeepLabV3+ model performs well in real segmentation tasks and has higher detection accuracy and reliability.

**Keywords:** photovoltaic panels; semantic segmentation; DeepLabV3+; MobileNetV2

## 0 引言

化石能源是不可再生的,且化石燃料的燃烧已经引起了全球性的环境问题,例如全球变暖、污染问题。作为一种清洁、可再生的能源,光伏能源近年来在全球范围内得到了广泛应用和快速发展<sup>[1]</sup>。全球光伏发电装机容量在过去十

年中呈现爆发式增长<sup>[2]</sup>,这一趋势不仅推动了能源结构的转型,还在减少温室气体排放、应对气候变化等方面发挥了重要作用。

尽管光伏技术的应用前景广阔,但在实际运行过程中,光伏板常常会出现诸如二极管故障、碎裂和热斑等问题。这些故障不仅会降低光伏板的发电效率,还可能对设备造

成严重损害,进而影响整个光伏发电系统的稳定性和经济性。因此,快速准确实时地检测和诊断光伏板故障具有重要的实际意义。

传统的检测方法主要依靠人工检测和电气特性分析测量<sup>[3]</sup>,工作效率低下,故障检测的准确率也不能得到保证。近年来,深度学习在图像处理和故障检测领域展现出强大的应用潜力,特别是基于卷积神经网络 CNN 的深度学习方法已被广泛应用<sup>[4]</sup>,国内外研究人员提出了多种目标检测算法,一阶段算法有 SSD、YOLO 系列等,二阶段算法有 R-CNN、Fast R-CNN 和 Faster R-CNN 等。针对光伏板的二极管故障、碎裂和热斑等问题,王道累等<sup>[5]</sup>提出了 SSD 300 模型,具有快速的检测效率。Hong 等<sup>[6]</sup>同时采用 YOLOv5 和 ResNet 算法进行图像分割和故障检测,将光伏阵列的分离速度显著提高到 36Fps,算法框架适用性强,精度高。Cao 等<sup>[7]</sup>提出了基于 YOLOv5 改进的 YOLOv5s-GBC 算法,在许多评价指标上表现出了更好的性能,和原方法相比,准确率和推理速度分别提高了 2% 和 20.3%。伊欣同等<sup>[8]</sup>提出了基于 Faster R-CNN 的改进网络,用于检测光伏电池片内部缺陷,精度提升到 83.29%。这些算法都是直接对光伏板故障直接进行目标检测,虽然表现出了高效快速的性能,但在实际工作场景中,检测算法利用红外相机传回的红外照片来检测光伏板故障点,因红外成像的原因会对光伏板边缘部分以及背景误检,如:光伏板支撑架、金属板材屋顶、和其他红外成像与光伏板类似的物体。误检情况在分布式光伏场景中尤其严重。

语义分割可以通过计算机解决图片中的每个像素的归属问题,能够将需要的语义信息以精确的描边方式进行提取,生成不带背景信息的掩码,为后续的视觉信息处理提供了极大的帮助。在光伏故障检测的实际工作中,利用语义分割算法将采集到的红外图像处理后,可以极大地避免光伏板故障检测的误检问题<sup>[9]</sup>。最先提出的语义分割网络是 FCN<sup>[10]</sup>,它把传统的全连接层转换成了卷积层,可以灵活地适应任意尺寸的输入图像,不受输入尺寸固定的限制,使得语义分割在实际应用中变得更加实用。继 FCN 之后,出现了许多改进的语义分割算法,例如 U-Net、PSPNet、HRNet 和 DeepLabV1~V3+ 系列。这些算法在 FCN 的基础上进行了优化和改进,进一步提高了分割精度和处理速度。U-Net<sup>[11]</sup>是一种主要用于医学图像分割的算法结构,通过对称的编码器-解码器结构,可以在不同尺度上捕捉到更多的上下文信息。汪方斌等<sup>[11-12]</sup>改进了 U-Net 网络模型并用于光伏电池板缺陷检测,检测精度达到了 97.05%。Zhao 等<sup>[13]</sup>提出的 PSPNet 通过引入金字塔池化模块,可以有效捕捉全局上下文信息,提高了分割结果的准确性和鲁棒性。Ren 等<sup>[14]</sup>对 HRnet 轻量化设计,在多个数据集中表现出色。DeepLab 系列中的 DeepLabV3+<sup>[15]</sup>结合了编码器和解码器的结构和空洞空间金字塔池化结构 (atrous spatial pyramid pooling, ASPP),可以灵活地处理

不同尺度的输入信息,使其成为语义分割的主流算法之一。黄聪等<sup>[16]</sup>对 DeepLabV3+ 进行改进,将 mIoU 提高到了 97.07%,表现出了优秀的图像分割效果。在光伏板的检修工作中,DeepLabV3+ 算法可以快速地将光伏板从背景中分割出来,去除背景干扰项,以供后续二极管故障、碎裂和热斑等问题的目标检测,但 DeepLabV3+ 对光伏板的边界信息处理不够精准,存在边界黏连、波浪状以及背景误分割等情况。

为解决以上问题,本文提出了一种基于改进 DeepLabV3+ 的算法,改进如下:将主干网络替换为 MobileNetV2;将 MobileNetV2 前 4 层引入 Canny 边缘检测算法输出浅层特征语义信息,利用 Canny 算法的边缘检测能力增强网络的特征提取能力;构建 SE-ASPP 模块,对特征通道进行重校准,增强提高了语义分割任务中对光伏板边界和细节判断的准确性;将浅层特征语义信息进入解码部分的单通道改为并行双通道,增强对浅层特征语义信息的表达能力。

## 1 网络结构

### 1.1 DeepLabV3+基础网络框架

DeepLabV3+ 是一种用于语义分割的深度学习模型,是 DeepLab 系列的最新版本,由 Google Brain 团队开发。DeepLabV3+ 引入了编码器-解码器结构,通过上采样和跳跃连接来改进分割结果的空间细节和准确性,使得网络能够更好地恢复细节信息;在编码器顶部使用空洞空间金字塔池化结构 (ASPP),通过 3 组不同的膨胀卷积率捕获多个尺度的特征信息,扩大了感受野;使用了特征融合策略,将不同层次的特征信息进行有效整合,提高了分割结果的一致性和连贯性<sup>[17]</sup>。DeepLabV3+ 算法模型结构如图 1 所示。

### 1.2 改进的 DeepLabV3+ 网络模型

语义分割任务中,浅层特征语义信息通常包含图像的细节信息,如边缘、纹理和颜色。这些特征对于捕捉对象的细节和形状至关重要,尤其是在样本边缘的分割<sup>[18]</sup>。所以本文在 DeepLabV3+ 网络结构的基础上特别增强了对浅层特征语义信息的表达能力,更适合于光伏板故障检测中语义分割任务,如图 2 所示:将 DeepLabV3+ 的主干网络更改为 MobileNetV2,取其第 4 层(即瓶颈层 3,输出深度为 32)输出的通道信息送入 Canny 边缘特征提取模块后,输出的特征信息为改进算法的浅层特征语义信息,其一方面通过并行的两个  $1 \times 1$  卷积通道直接进入解码部分,另一方面与深层特征语义信息特征融合后送入改进的 SE-ASPP 模块中,经  $1 \times 1$  卷积和 4 倍上采样后再进入解码部分后。两条不同通道的语义信息在解码器中先后经过特征融合、 $3 \times 3$  卷积和 4 倍上采样后得到经改进的 DeepLabV3+ 网络模型语义分割后的光伏板图像,此时背景信息被完全去除,只保留光伏板的像素信息。

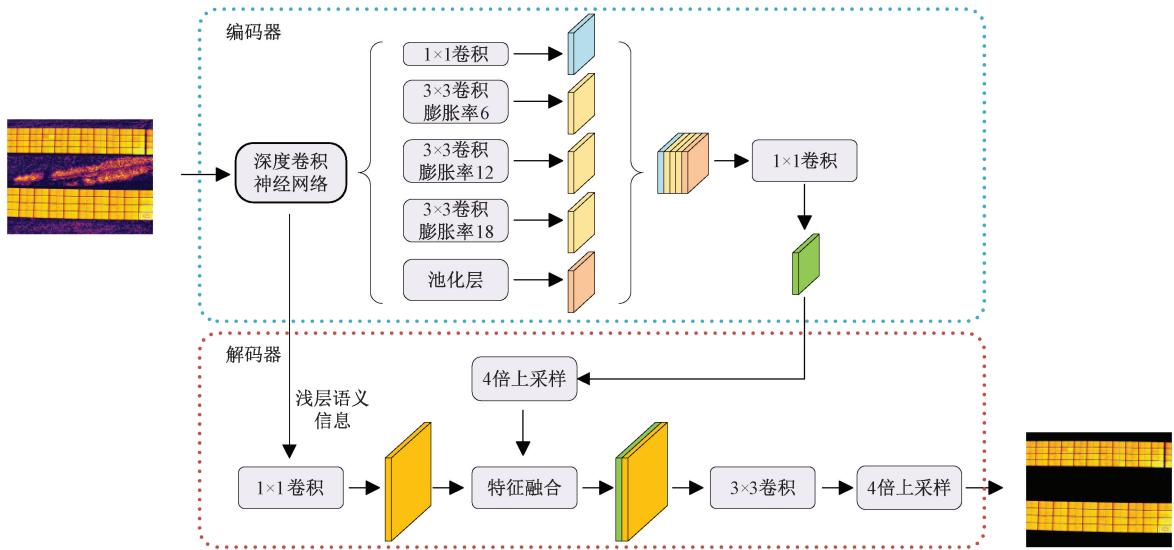


图 1 DeepLabV3+ 网络模型结构

Fig. 1 Structure of the DeepLabV3+ network model

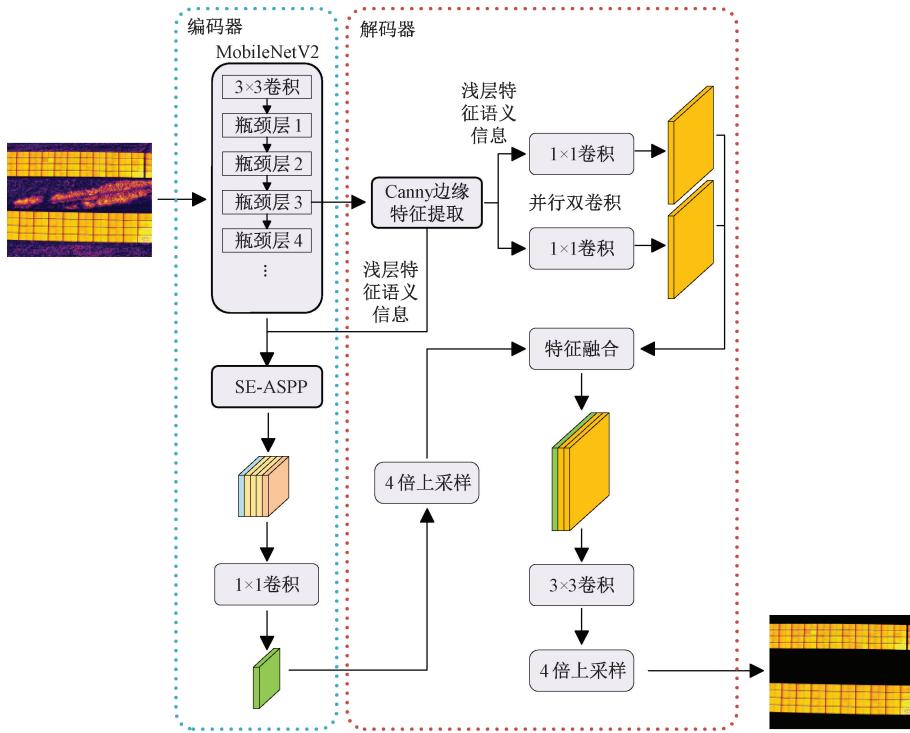


图 2 本文改进的 DeepLabV3+ 网络模型结构

Fig. 2 The structure of the improved DeepLabV3+ network model in this paper

### 1.3 主干网络 MobileNetV2

DeepLabV3+的主干网络是Xception,可以提取高层次语义信息,在复杂场景或需要较高精度的任务中有出色的表达能力,但是其运算模型参数量大、结构复杂,需要占用更多的计算资源和时间来训练,不适合轻量级的工作任务。在光伏故障检测的工作中,工作人员利用搭载红外相机的无人机来巡视检查光伏发电站,因此对网络算法模型要求轻量化更高。

MobileNetV2<sup>[19]</sup>是轻量级的卷积神经网络架构,采用深度可分离卷积,将传统卷积分成深度卷积和逐点卷积两个步骤,在结构中引入线性瓶颈和倒残差结构,在牺牲小部分精度的情况下,显著减小了训练模型的参数量。MobileNetV2被广泛用于图像处理的模型预训练中,将其替换为DeepLabV3+的主干网络可以准确高效地完成光伏板的语义分割任务。MobileNetV2网络结构如表1所示。

表 1 MobileNetV2 网络结构  
Table 1 MobileNetV2 network structure

序号	输入	操作	扩展因子	输出矩阵深度	Bottleneck 重复次数	步长
1	224×224×3	Conv2d	—	32	1	2
2	112×112×32	bottleneck	1	16	1	1
3	112×112×16	bottleneck	6	24	2	2
4	56×56×24	bottleneck	6	32	3	2
5	28×28×32	bottleneck	6	64	4	2
6	28×28×64	bottleneck	6	96	3	1
7	14×14×96	bottleneck	6	160	3	2
8	7×7×160	bottleneck	6	320	1	1
9	7×7×320	Conv2d 1×1	—	1 280	1	1
10	7×7×1 280	Avgpool 7×7	—	—	1	—
11	1×1×1 280	Conv2d 1×1	—	k	—	—

本文取 MobileNetV2 的第 4 层输出到 Canny 边缘检测算法后生成新的浅层特征语义信息, 在减少计算资源损耗的同时, 有助于减少信息丢失, 捕捉图像中更详细的边缘信息。

#### 1.4 Canny 边缘检测算法

Canny 边缘检测算法是一个优秀可靠的检测算法, 适用于检测图像边缘<sup>[20]</sup>, 算法流程图如图 3 所示。

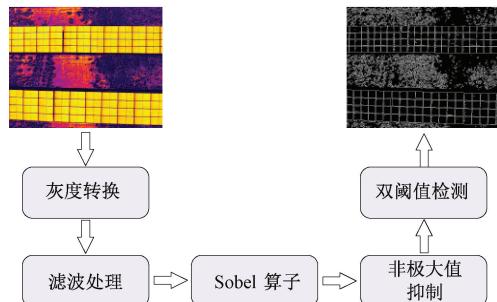


图 3 Canny 边缘检测算法流程图

Fig. 3 Flow chart of Canny's edge detection algorithm

Canny 边缘检测步骤通常先用将图像转为灰度图进行降噪滤波处理, 对图像中非边缘区域平滑处理, 突出边缘信息; 然后通过 Sobel 算子对图像中每个点的梯度的强度和方向, 计算公式如下:

$$G = \sqrt{G_x^2 + G_y^2} \quad (1)$$

$$\theta = \arctan\left(\frac{G_y}{G_x}\right) \quad (2)$$

式中:  $G_x$  为水平方向;  $G_y$  为垂直方向;  $G$  为每个点的梯度强度;  $\theta$  为梯度方向。

接着根据非极大值抑制的原则, 遍历每个像素点的梯度强度和方向, 去除梯度强度非最大值的点。最后, 使用双阈值(即低阈值和高阈值)来最终确定边缘像素。如果一个像素点的梯度强度高于高阈值, 则被认为是边缘像素; 如果低于低阈值, 则被抑制; 如果在两者之间, 则只有

当它与边缘像素相接时才被保留, 最后形成完整的边缘输出图像。

本文中, 经 Canny 算法计算输出后的语义信息为改进后的浅层特征语义信息, 与 SE-ASPP 输出的深层特征语义信息在解码器部分融合。并行处理浅层和深层的特征语义信息可以综合不同层次的图像信息, 有助于算法模型快速准确地定位和识别边缘, 提高光伏板边界图像信息的检测精度和鲁棒性。

#### 1.5 SE-ASPP 模块

挤压激励注意力机制(squeeze-and-excitation, SE)<sup>[21]</sup>是一种在深度学习和计算机视觉领域中用于提升卷积神经网络(CNN)性能的技术。它独特的“挤压-激励-重新加权”机制通过显式地建模特征通道之间的依赖关系, 增强了网络的表征能力, 其结构如图 4 所示。SE 注意力机制的挤压部分是通过全局平均池化挤压每个特征通道生成一个全局特征信息, 对应图 4 中的  $F_{sq}(\cdot)$ , 其表达公式如式(4)所示; 激励部分利用两个全连接层对网络先降维后升维, 并先后利用 ReLU 激活函数和 Sigmoid 激活函数生成通道注意力权重, 对应图 4 中  $F_{ex}(\cdot, W)$ , 其表达公式如式(5)所示; 重新加权则将通道注意力权重加权到原始特征通道, 提高了模型对重要特征的关注度, 对应图 4 中  $F_{scale}(\cdot, \cdot)$ , 其表达公式如式(6)所示。

图 4 中,  $F_{tr}$  是卷积算子, 特征图  $X$  经过  $F_{tr}$  操作变成特征图  $U$ ,  $F_{tr}$  的表达公式如式(3)所示。

$$u_c = V_c * X = \sum_{s=1}^{C'} V_c^s * X^s \quad (3)$$

$$z_c = F_{sq} u_c = \frac{1}{H \times W} \sum_{i=1}^H \sum_{j=1}^W u_{ci}, j \quad (4)$$

$$s = F_{ex}(z, W) = \sigma(g(z, W)) = \sigma(W_2 \delta(W_1 z)) \quad (5)$$

$$\tilde{X}_c = F_{scale}(u_c, s_c) = s_c u_c \quad (6)$$

式中:  $*$  为卷积操作,  $V_c$  为第  $c$  个滤波器参数,  $V_c^s$  是一个二维的卷积核,  $z_c$  为第  $c$  个特征通道的特征向量,  $s$  为特征

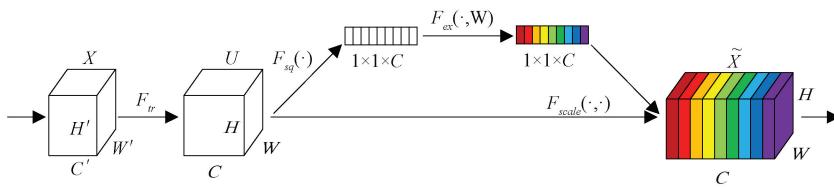


图 4 SE 注意力机制网络模型结构

Fig. 4 Structure of SE attention mechanism network model

图  $U$  中  $c$  个特征图的权重。

空洞空间卷积池化金字塔 (atrous spatial pyramid pooling, ASPP)<sup>[22]</sup> 是 DeepLabV3+ 中的关键组件, 使用 3 个并行的空洞卷积增加感受野, 在保持空间分辨率的同时, 有效地捕获不同尺度上的语义信息。ASPP 原始空洞卷积膨胀率 6、12、18 的组合有较大的感受野, 对局部特征提取的能力不足, 会容易丢失细微的空间信息, 尤其是边缘信息。

本文提出两者结合的 SE-ASPP 模块, 如图 5 所示。

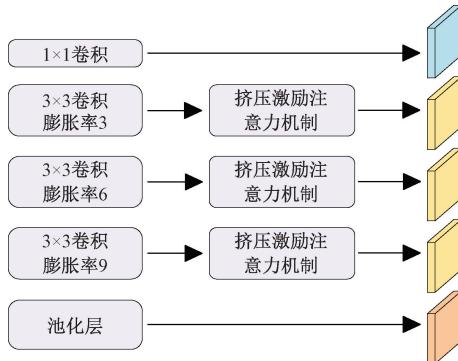


图 5 SE-ASPP 模块结构

Fig. 5 Structure of the SE-ASPP module

在 SE-ASPP 模块中, ASPP 的空洞卷积膨胀率组合选择为 3、6、9, 其中: 卷积膨胀率为 3 的带孔卷积适合捕获较小的细节和纹理特征, 如物体边缘的细微变化; 卷积膨胀率为 6 的带孔卷积能够扩展感受野, 涵盖更大的边缘和轮廓信息; 而卷积膨胀率为 9 的带孔卷积则进一步扩展感受野, 捕获更广泛的上下文背景信息, 有助于更准确地定位物体边缘。在 ASPP 每个空洞卷积的并行分支之后, 添加了一个 SE 模块, 在捕获多尺度语义信息的同时进一步增强每个尺度上的特征表示能力, 使每个分支学习到的特征能够更加准确和有区分度, 提升网络模型对复杂场景的感知能力和理解能力。这种结合方式可以有效提高语义分割任务中物体边界和细节的准确性, 有更多的非线性, 可以更好地拟合通道间的复杂关联, 同时保持了整体网络的高效性能。

## 1.6 并行双卷积通道

本文改进算法模型将编码器进入解码器的单  $1 \times 1$  卷积通道改为并行的双  $1 \times 1$  卷积通道, 创建了两个不同的特征空间变换。并行双卷积通道<sup>[23]</sup> 可以增加解码器接收

到的浅层特征语义信息的丰富度和多样性, 提高算法模型的特征表达能力和解码能力, 降低过拟合的风险, 改善整体的灵活性和泛化能力。

## 2 实验过程

### 2.1 实验数据集

本文的数据集由携带 HT20 红外热成像摄像头的大疆 M300RTK 无人机现场对光伏发电站巡回拍摄生成, 图像采集过程的场景主要为不同光照强度和不同天气下的山地地区, 图像分辨率为  $640 \times 512$ 。本文的研究过程中共采集到 1865 张图像, 筛选得到 1820 张含有光伏板的图像, 经旋转扩增至 3640 张图像, 共包含 11273 个光伏板样本数据。将数据集按 7:2:1 的比例随机地分为训练集、验证集和测试集, 对训练集和验证集的样本使用 Labelme 软件手动标注生成 JSON 文件, 再转换成包含光伏板样本掩码信息的 PNG 格式图像, 送入改进的 DeepLabV3+ 算法和各对照算法模型进行训练和测试。

### 2.2 实验环境

本文实验方法实现基于 Windows10 操作系统, 使用 Pytorch 深度学习框架, 版本为 2.1.2, 编程语言为 Python 3.8.18。硬件方面, CPU 为 13th Gen Intel(R) Core(TM) i9-13900K 3.00 GHz, GPU 为 NVIDIA GeForce RTX 4090。实验训练和验证时, 图像输入大小为  $512 \times 512$ , 初始学习率为 0.0001, 优化器选用 Adam, 每次训练时的图像批量 batch 大小为 16, 总共进行 100 轮迭代训练。

### 2.3 评价指标

本文的实验评价主要指标为精准率 (precision,  $P$ )、平均交并比 (mean intersection over union, mIoU)、召回率 (recall,  $R$ ) 和 F1 分数。各计算公式分别如式(7)~(10) 所示。

$$P = \frac{TP}{TP + FP} \times 100\% \quad (7)$$

$$mIoU = \frac{1}{N} \sum_{i=1}^N \frac{TP}{TP + FP + FN} \times 100\% \quad (8)$$

$$R = \frac{TP}{TP + FN} \times 100\% \quad (9)$$

$$F1 = 2 \times \frac{P \times R}{P + R} \times 100\% \quad (10)$$

式中:  $TP$  表示正确预测为正类的像素数,  $FP$  表示错误预测为正类的像素数,  $FN$  表示错误预测为负类的像素数,  $N$

表示类别的总数,在本文实验中,因为只有背景和光伏板一个样本,所以公式中的  $N=2$ 。

### 3 实验结果分析

#### 3.1 对照模型实验结果分析

本文通过将改进后的 DeepLabV3+算法模型与 PSPNet、HRNet、U-Net、和原 DeepLabV3+算法模型进行同等实验环境和参数下的实验对比,评价指标参数对比如表 2 所示。

表 2 不同网络模型的实验结果对比

Table 2 Comparison of experimental results of different network models

模型	%			
	P	mIoU	R	F1
PSPNet	96.31	96.83	97.37	96.84
HRNet	97.01	97.29	97.62	97.31
U-Net	96.92	97.15	97.46	97.08
DeepLabV3+	97.26	97.63	98.04	97.83
本文改进模型	99.50	99.21	99.61	99.55

可以看出,原 DeepLabV3+模型的精准率为 97.26%,mIoU 为 97.63%,Recall 为 98.04%,F1 分数为 97.83%,均高于 PSPNet、HRNet 和 U-Net 的各项数据,说明相比于其他算法模型,原 DeepLabV3+模型在分割任务中对漏检敏感,表现优秀。而本文提出的 DeepLabV3+改进模型与原 DeepLabV3+相比,精准率、mIoU、Recall 和 F1 分数分别提高 2.24%、1.58%、1.57% 和 1.72%,达到了 99.50%、99.21%、99.61% 和 99.55%,这表明改进模型提升了模型的预测精度和边界判定的准确性,具有更加优秀

的分割性能。

但是从整体而言,各个算法模型的评价指标数值相差不是很大,为更明显地对比各个算法模型的实际分割效果,更全面直观地对改进模型评估,本文随机选取了 6 张数据集中不同背景情况和不同强度光照下的光伏板红外图像来展示分割效果,预测结果图中将光伏板标签对应的部分保留原图色彩,背景标签对应的像素值设置为 0,如图 6 所示,图中给出了原始图像、标签掩码图,和 PSPNet、HRNet、U-Net、原 DeepLabV3+、本文改进的 DeepLabV3+模型分别预测的效果图。可以看出,效果最差的是 PSPNet 模型,U-Net 和原 DeepLabV3+效果相对较好,本文改进的 DeepLabV3+模型效果则最佳。图像 1 中,左下角为光伏电站的金属材质建筑物,图像 2 中,中间部分为一辆汽车,对照模型的预测结果中,均对金属材质建筑物和汽车对应的像素部分产生误判,出现误分割的情况;在图像 3 中,对照模型预测结果都出现了边界附近像素的误分割;对图像 4 的分割任务,对照模型预测结果出现两块光伏板之间边界黏连的情况,在图中由绿色矩形框把分割缺陷部分放大标记;图像 5 和 6 中,对照模型预测结果中出现了光伏板边缘黑洞腐蚀、边界分割不完整以及内部黑洞腐蚀等情况,图像 5 中同样把分割缺陷部分放大标记。本文改进的 DeepLabV3+模型有效地避免了以上问题,对金属材质等非目标物体能够准确完整地排除;相隔距离较近的光伏板中间没有出现边界黏连情况;对光伏板的边界分割连贯清晰且精确,不会因背景中的因素导致误分割;没有光伏板内部黑洞现象。因此,本文提出的改进的 DeepLabV3+模型与其他对照分割算法模型相比,有着优秀的光伏板语义分割能力。

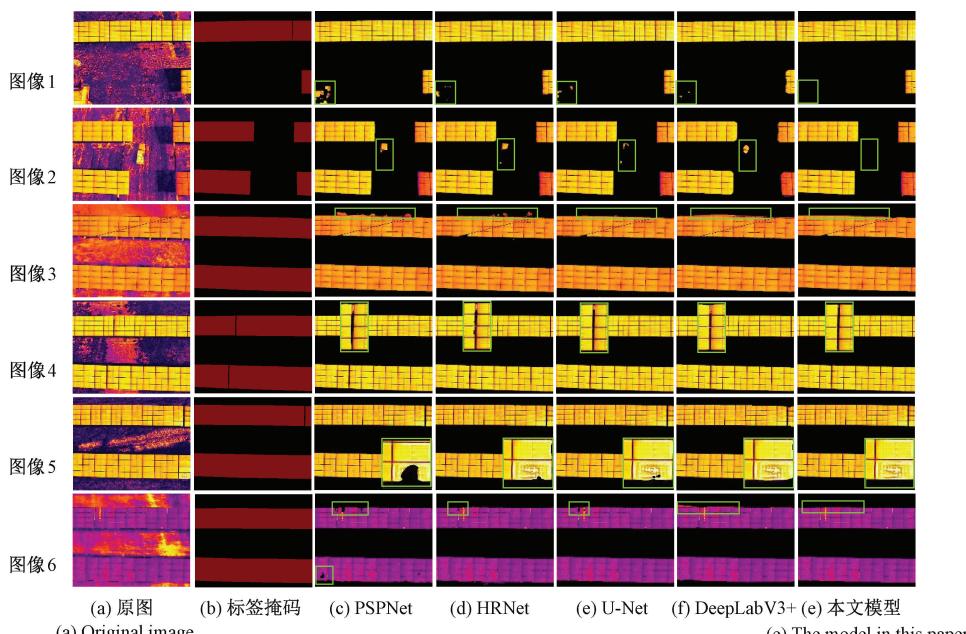


图 6 不同对照模型预测结果对比图

Fig. 6 Comparison of prediction results between different control models

### 3.2 消融实验结果分析

为验证本文改进 DeepLabV3+ 模型中主干网络 MobileNetV2 以及增加的 Canny 边缘检测、SE-ASPP 模块

和浅层特征语义信息并行双通道模块的有效性,以原 DeepLabV3+ 模型为基础,通过不同的模块组合,进行消融实验,实验结果如表 3 所示。

表 3 消融实验结果对比

Table 3 Comparison of ablation experimental results

实验组别	MobileNetV2	Canny 边缘检测	SE-ASPP 模块	浅层特征语义 信息双通道模块	P	mIoU	R	F1	%
1					97.26	97.63	98.04	97.83	
2	√				98.10	98.95	99.07	98.94	
3	√	√			98.54	99.01	99.52	99.01	
4	√	√	√		98.84	99.03	99.52	99.02	
5	√	√	√	√	99.50	99.21	99.61	99.55	

实验结果显示,主干网络 MobileNetV2 的替换,使精准率、mIoU、Recall 和 F1 分数分别提升了 0.84%、1.32%、1.03% 和 1.11%,起到了较大的改进效果,表明 MobileNetV2 有着优秀的分割能力。Canny 边缘检测模块的增加,各项评价指标分别提高 0.44%、0.06%、0.45% 和 0.07%,说明 Canny 模块的边缘特征信息提取有助于提高模型对光伏板边界的识别能力。SE-ASPP 模块在召回率提升方面表现不佳,但精准率提升了 0.30%,提升了模型对光伏板信息的正确识别能力。浅层特征语义信息双通道模块的增加,使精准率、mIoU、Recall 和 F1 分数各自达到了最佳值,分别提升了 0.66%、0.18%、0.09% 和 0.53%,提高了算法模型对各个尺度目标的识别与分割能力,对边缘信息的处理效果也达到最佳状态。表 3 证明,本文提出的 3 个模块对算法模型的图像语义分割能力有一定的提升作用,3 个模块同时使用时效果达到最佳状态。

## 4 结 论

针对光伏故障检测任务中背景其他因素容易干扰检测效果的问题,本文提出了一种基于改进 DeepLabV3+ 的光伏板语义分割算法模型,旨在提高光伏板故障检测的准确性和效率。本文在 DeepLabV3+ 模型的基础上,把主干网络替换为具有更轻量级的结构和更高的计算效率 MobileNetV2,将 MobileNetV2 前四层输出引入 Canny 边缘检测算法,构建新的浅层特征语义信息;增加浅层特征语义信息进入解码器的通道数;构建 SE-ASPP 模块,对特征通道进行重新校准。改进的 DeepLabV3+ 模型与对照的语义分割模型: PSPNet、HRNet、U-Net 和原 DeepLabV3+ 相比,有着优秀的光伏板语义分割能力,在评价指标 mIoU 和 Recall 上有所提高,实际的图像分割效果出色。此外,通过消融实验验证了各改进模块的有效性。因此,本文提出的基于 DeepLabV3+ 改进的算法模型在光伏板语义分割任务中可以有效地避免传统语义分割算法模型会出现的样本边界黏连、波浪状和背景误分割的

情况,对提高故障检测的精度和可靠性方面具有显著优势,可减少误检和漏检问题。未来的研究可进一步优化模型结构,提高实时监测的效率和准确率,为下一步的光伏板故障检测提供准确的图像输入。

## 参考文献

- [1] WAN L, ZHAO L, XU W, et al. Dust deposition on the photovoltaic panel: A comprehensive survey on mechanisms, effects, mathematical modeling, cleaning methods, and monitoring systems[J]. Solar Energy, 2024, 268: 112300.
- [2] GÓMEZ M, 徐国畅, 李彦, 等. 引领未来: 中国光伏产业发展路线图的挑战(英文)[J]. Science Bulletin, 2023, 68(21): 2491-2494.
- [3] GÓMEZ M, XU G CH, LI Y, et al. Navigating the future: China's photovoltaic roadmap challenges. [J]. Science Bulletin, 2023, 68(21): 2491-2494.
- [4] CHEN L C, PAPANDREOU G, SCHROFF F, et al. Rethinking atrous convolution for semantic image segmentation[J]. ArXiv preprint arXiv: 1706. 05587, 2017.
- [5] MINAEE S, BOYKOV Y, PORIKLI F, et al. Image segmentation using deep learning: A survey[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2021, 44(7): 3523-3542.
- [6] 王道累, 李明山, 姚勇, 等. 改进 SSD 的光伏组件热斑缺陷检测方法[J]. 太阳能学报, 2023, 44(4): 420-425.
- [7] WANG D L, LI M SH, YAO Y, et al. Method of hotspot detection of photovoltaic panels modules on improved SSD[J]. Acta Energiae Solaris Sinica, 2023, 44(4): 420-425.
- [8] HONG F, SONG J, MENG H, et al. A novel framework on intelligent detection for module defects of PV plant combining the visible and infrared images[J]. Solar Energy, 2022, 236: 406-416.

[7] CAO Y, PANG D, YAN Y, et al. A photovoltaic surface defect detection method for building based on deep learning [J]. Journal of Building Engineering, 2023, 70: 106375.

[8] 伊欣同,单亚峰. 基于改进 Faster R-CNN 的光伏电池内部缺陷检测[J]. 电子测量与仪器学报, 2021, 35(1): 40-47.

YI X T, SHAN Y F. Photovoltaic cell internal defect detection based on improved Faster R-CNN [J]. Journal of Electronic Measurement and Instrumentation, 2021, 35(1): 40-47.

[9] LEI Y, WANG X Y, AN AI M, et al. Deeplab-YOLO: A method for detecting hot-spot defects in infrared image PV panels by combining segmentation and detection[J]. Journal of Real-Time Image Processing, 2024, 21(2): 52.

[10] LONG J, SHELHAMER E, DARRELL T. Fully convolutional networks for semantic segmentation[C]. IEEE Conference on Computer Vision and Pattern Recognition, 2015: 3431-3440.

[11] 汪方斌,李文豪. 基于改进轻量化 U-Net 模型的光伏电池 EL 图像缺陷检测[J]. 电子测量技术, 2024, 47(5): 102-111.

WANG F B, LI W H. Defect detection of photovoltaic cell EL images based on improved lightweight U-Net model [J]. Electronic Measurement Technology, 2024, 47(5): 102-111.

[12] RONNEBERGER O, FISCHER P, BROX T. U-net: Convolutional networks for biomedical image segmentation [C]. Medical Image Computing and Computer-Assisted Intervention-MICCAI 2015: 18th International Conference, 2015: 234-241.

[13] ZHAO H, SHI J, QI X, et al. Pyramid scene parsing network[C]. IEEE Conference on Computer Vision and Pattern Recognition, 2017: 2881-2890.

[14] REN Q, LU Z, WU H, et al. Hr-net: A landmark based high realistic face reenactment network [J]. IEEE Transactions on Circuits and Systems for Video Technology, 2023, 33(11): 6347-6359.

[15] CHEN L, ZHU Y, PAPANDREOU G, et al. Encoder-decoder with atrous separable convolution for semantic image segmentation[C]. European Conference on Computer Vision(ECCV), 2018: 801-818.

[16] 黄聪,杨珺,刘毅,等. 基于改进 DeeplabV3+的遥感图像分割算法 [J]. 电子测量技术, 2022, 45 (21): 148-155.

HUANG C, YANG J, LIU Y, et al. Remote sensing image segmentation algorithm based on improved DeeplabV3+ [J]. Electronic Measurement Technology, 2022, 45(21): 148-155.

[17] WANG Z, WANG J, YANG K, et al. Semantic segmentation of high-resolution remote sensing images based on a class feature attention mechanism fused with Deeplabv3+ [J]. Computers & Geosciences, 2022, 158: 104969.

[18] ZHANG S, REN G, ZENG X, et al. Efficient cross-information fusion decoder for semantic segmentation[J]. Computer Vision and Image Understanding, 2024, 240: 103918.

[19] GULZAR Y. Fruit image classification model based on MobileNetV2 with deep transfer learning technique [J]. Sustainability, 2023, 15(3): 1906.

[20] MONICKA S G, MANIMEGALAI D, KARTHIKEYAN M. Detection of microcracks in silicon solar cells using Otsu-Canny edge detection algorithm[J]. Renewable Energy Focus, 2022, 43: 183-190.

[21] ASKER M E. Hyperspectral image classification method based on squeeze-and-excitation networks, depthwise separable convolution and multibranch feature fusion[J]. Earth Science Informatics, 2023, 16(2): 1427-1448.

[22] WANG Q, LIU H, WEN T, et al. Research on semantic segmentation algorithm for multiscale feature images based on improved DeepLab v3+[C]. Journal of Physics: Conference Series. IOP Publishing, 2024, 2747(1): 012038.

[23] YANG Y, CHEN Q, LIU Q. A dual-channel network for cross-domain one-shot semantic segmentation via adversarial learning [J]. Knowledge-Based Systems, 2023, 275: 110698.

## 作者简介

王银,博士,副教授,主要研究方向为计算机视觉、智能控制等。

E-mail:2902459153@qq.com

孙海顺(通信作者),硕士研究生,主要研究方向为计算机视觉、图像处理。

E-mail:sunsunshs@163.com