

本文引用格式: 武果,高守正,张德福,等.基于特征重组和注意力机制的 YOLOv5 光伏面板缺陷检测方法[J].自动化与信息工程,2024,45(5):47-53.

WU Guo, GAO Shouzheng, ZHANG Defu, et al. YOLOv5 photovoltaic panel defect detection method based on feature recombination and attention mechanism[J]. Automation & Information Engineering, 2024,45(5):47-53.

## 基于特征重组和注意力机制的 YOLOv5 光伏面板缺陷检测方法

武果<sup>1</sup> 高守正<sup>2</sup> 张德福<sup>3</sup> 肖国龙<sup>1</sup> 龚正平<sup>1</sup> 覃健<sup>1</sup>

(1.中科院广州电子技术有限公司, 广东 广州 510070

2.华南理工大学自动化科学与工程学院, 广东 广州 510641

3.广州理工学院智能制造与电气工程学院, 广东 广州 510800)

**摘要:** 针对光伏面板缺陷检测存在的小目标检测精度低、泛化能力差等问题, 提出基于特征重组和注意力机制的 YOLOv5 光伏面板缺陷检测方法。首先, 分别使用卷积注意力模块和压缩激励模块, 提取目标图像在通道和空间维度上的特征, 使网络更有效地学习目标特征; 然后, 使用感受野扩展模块进行特征融合, 确保网络能够结合大感受野特征和微小特征进行综合判断; 最后, 引入特征重组模块, 对不同尺寸的特征图进行特征重组, 解决了网络对小目标信息缺失和不敏感的问题。通过消融实验和对比实验, 验证了该方法的有效性。

**关键词:** 特征重组; 注意力机制; 光伏面板; 缺陷检测; YOLOv5; 神经网络

中图分类号: TP183

文献标志码: A

文章编号: 1674-2605(2024)05-0007-07

DOI: 10.3969/j.issn.1674-2605.2024.05.007

开放获取

## YOLOv5 Photovoltaic Panel Defect Detection Method Based on Feature Recombination and Attention Mechanism

WU Guo<sup>1</sup> GAO Shouzheng<sup>2</sup> ZHANG Defu<sup>3</sup>  
XIAO Guolong<sup>1</sup> GONG Zhengping<sup>1</sup> QIN Jian<sup>1</sup>

(1.Guangzhou Electronic Technology Co., Ltd., CAS, Guangzhou 510070, China

2.School of Automation Science and Engineering, South China University of Technology, Guangzhou 510641,

China 3.Guangzhou Institute of Technology, College of Intelligent Manufacturing and Electrical Engineering,  
Guangzhou 510800, China)

**Abstract:** Aiming at the problems of low small object detection accuracy and poor generalization ability in defect detection of photovoltaic panels, a YOLOv5 photovoltaic panel defect detection method based on feature recombination and attention mechanism is proposed. Firstly, the convolutional attention module and compression excitation module are used separately to extract the features of the target image in the channel and spatial dimensions, enabling the network to learn the target features more effectively; Then, the receptive field extension module is used for feature fusion to ensure that the network can combine large receptive field features and small features for comprehensive judgment; Finally, a feature recombination module was introduced to perform feature recombination on feature maps of different sizes, solving the problem of missing and insensitive information for small targets in the network. The effectiveness of this method was verified through ablation experiments and comparative experiments.

**Keywords:** feature recombination; attention mechanism; photovoltaic panel; defect detection; YOLOv5; neural network

## 0 引言

光伏发电能提供清洁和可持续的能源。然而，光伏发电系统长期暴露在户外环境中，不仅受温度、雨水和风等气候变化的影响，还可能被人类活动或生物破坏，造成电力生产损失。同时，光伏面板的表面缺陷，如划痕、破损、污渍等，也降低了光伏发电系统的稳定性。因此，对光伏面板进行缺陷检测，有助于提高光伏发电系统的可靠性，并降低运行成本<sup>[1]</sup>。

目前，光伏面板缺陷检测的研究主要包括构建模型和优化算法两方面。文献[2]提出一种基于改进 Alexnet 模型的光伏电池片瑕疵分类方法，融合了多层次特征进行分类，提升了瑕疵分类的稳定性、实时性和正确率。文献[3]提出一种基于改进 U-Net 网络模型的光伏电池板缺陷检测方法，有效减少了模型参数量并提升了检测精度。文献[4]设计了一种包含 12 层卷积的残差网络，并与特征金字塔结合，以获取输入图像不同尺度的特征；将特征图分割池化后作为 Transformer 的输入，在不增加计算量的情况下，提升了缺陷识别的准确度。文献[5]针对动态开放场景设计了一种光伏缺陷检测数据归一化控制器，以提高网络的域适应能力，解决了光伏部件图像质量低和易失真的问题。文献[6]针对光伏面板缺陷类内差异大、类间差异小和背景特征复杂的特点，提出融合局部和全局特征的卷积视觉 Transformer 网络，有效提升了光伏面板缺陷检测的精度。文献[7]提出基于超分辨率网络和双池化融合的光伏组件缺陷检测算法，可有效应对无人机巡检光伏组件时效率和识别准确率低的问题。但光伏面板缺陷检测仍存在如下问题：1) 采集的图像缺少多样性，因光伏面板处于室外，受光照、雨水等因素的影响，其环境多变，而目前研究实验大多在单一场景下进行，缺少不同场景下的检测图像，导致网络泛化性能较低；2) 在实际部署中，摄像头捕捉的光伏面板目标大小不一，而距离较远的光伏面板目标较小，神经网络对小目标的检测精度较低。

针对上述问题，本文提出一种基于特征重组和注意力机制的 YOLOv5 光伏面板缺陷检测方法。通过

改进的 YOLOv5 网络结构，提高光伏面板的缺陷检测精度。

## 1 改进的 YOLOv5 网络

YOLOv5 网络主要分为 Backbone、Neck、Head 三部分。针对 YOLOv5 网络存在特征表达能力有限的问题，对其结构进行如下改进：1) 在 Backbone 部分引入注意力机制，使 YOLOv5 网络能够选择性地关注图像中更重要的信息；2) 在 Neck 部分引入感受野扩展模块（receptive field expansion module, RFEM），以提高网络的特征感知能力；3) 用特征重组模块（feature reorganization module, FRM）替换 Head 部分的 Detect 层，以增强网络的检测性能。改进的 YOLOv5 网络结构如图 1 所示。

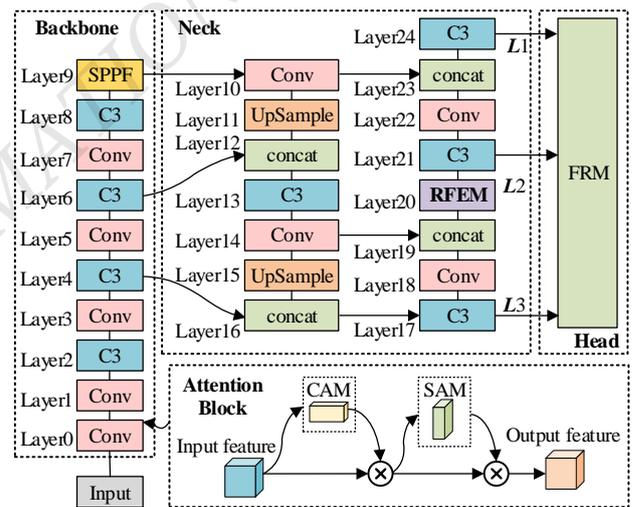


图 1 改进的 YOLOv5 网络结构

在改进的 YOLOv5 网络结构中，快速空间池化金字塔（space pooling pyramid fast, SPPF）是一种多尺度的特征融合技术<sup>[8]</sup>，可捕获图像中不同尺度和大小的信息，扩大网络的感知范围，提高网络理解图像全局和局部信息的能力。

C3 模块采用跨阶段连接的方式，将不同阶段的特征进行拼接，使网络能够跨层级传递信息，有效避免了梯度消失的问题，同时增强了网络的表征能力。此外，C3 模块还具有特征重组功能。通过一层卷积支路和残差块支路对输入进行拼接，可生成更丰富和更

具有表征能力的特征,帮助网络更好地理解目标的语义信息,提高检测的准确性和稳定性。

Conv 模块由卷积层、BatchNorm 和 Sigmoid 激活函数构成。其中, BatchNorm 能加快网络的学习速度,并有效缓解梯度衰减或过冲的问题; Sigmoid 激活函数提供了非线性的变换功能,增强了网络的表征能力。

UpSample 和 concat 分别表示上采样和拼接操作。

### 1.1 注意力模块

为了使网络能够选择性地关注图像中更重要的信息,提升网络的表征能力和鲁棒性,本文引入卷积注意力模块 (convolutional block attention module, CBAM)<sup>[9]</sup>和压缩激励模块 (squeeze-and-excitation block, SE-Block)<sup>[10]</sup>。

#### 1.1.1 CBAM

CBAM 结合了通道注意力和空间注意力机制,旨在增强网络的特征表达能力,有助于网络在卷积操作过程中更好地分类。YOLOv5 网络通过卷积运算将浅层特征变为更深层的抽象特征,并将这些抽象特征输入到注意力模块,使网络聚焦于主要的抽象特征<sup>[9]</sup>。输入被送到通道注意力模块 (channel attention module, CAM),可获得通道注意力权重。先利用高速分叉路径和通道注意力路径提取关键特征,再把这些特征传递给空间注意力模块 (spatial attention module, SAM),以获得空间注意力权重。利用高速分叉路径和空间注意力路径来提取特征,并将它们应用于原始图像上,可生成含有焦点信息的关键特征图像。

在 CAM 中,首先,数据经过最大池化和平均值聚合函数处理后,被多层感知机 (multilayer perceptrons, MLP) 的输出覆盖,并通过元素级别的相乘方式计算新的特征信息;然后,将 S 型曲线作为非线性转换器,产生通道注意力特征的图形表示形式。通道注意力机制可表示为

$$M_c(F) = \sigma(MLP(AvgPool(F)) + MLP(MaxPool(F))) \\ = \sigma(W_1(W_0(F_{avg}^c)) + W_1(W_0(F_{max}^c))) \quad (1)$$

式中:  $F$  为输入特征,  $MLP(\cdot)$  为  $MLP$  层对输入

的作用,  $AvgPool(\cdot)$  和  $MaxPool(\cdot)$  分别为对输入进行平均池化和最大池化操作,  $F_{avg}^c$  和  $F_{max}^c$  分别为特征  $F$  经过平均池化和最大池化操作后的输出,  $W_0$  和  $W_1$  为组成  $MLP$  的两个全连接层,  $\sigma(\cdot)$  为 Sigmoid 函数对输入的激活作用。

SAM 用于捕捉特征图在空间维度上的重要性,重新分配特征图中不同空间位置的权重,使网络更关注图像的重要信息。

CBAM 关注基于图像的局部关系生成区域注意力的特性映射表征。与传统的流向式处理方式相比, CBAM 更注重捕捉有用的视觉元素位置,而非其流动方向或路径长度等属性。为此,首先,在通道维度进行平均池化和最大池化操作;然后,将它们生成的特征图拼接起来;最后,将拼接后的特征图输入到滤波器,得到具有显著性的空间注意力特征图。记空间注意力函数为  $M_s(F) \in R^{H \times W}$ ,则在通道维度上利用平均池化和最大池化操作产生的二维特征图为

$$F_{avg}^s \in R^{1 \times H \times W}, F_{max}^s \in R^{1 \times H \times W} \quad (2)$$

$M_s(F)$  的计算公式为

$$M_s(F) = \sigma(f^{7 \times 7}(AvgPool(F); MaxPool(F))) \\ = \sigma(f^{7 \times 7}(F_{avg}^s; F_{max}^s)) \quad (3)$$

#### 1.1.2 SE-Block

SE-Block 由压缩和激励两个模块组成。首先,对特征进行挤压操作,通过在主干结构上添加特征映射来形成通道信息的表现形式,这种表现形式能够将通道特征反映的全局分布纳入考虑范围,使整个神经网络都能够利用整体视野获取的信息;然后,在激励阶段采用 Sigmoid 激活函数:

$$s = F_{ex}(z, W) = \sigma(g(z, W)) = \sigma(W_2 \delta(W_1 z)) \quad (4)$$

式中:  $\delta$  为 ReLU 函数,  $W_1 \in R^{\frac{c}{r} \times c}$ 、 $W_2 \in R^{\frac{c}{r} \times c}$ ,  $z$  为激励阶段的输入,  $W$  为可训练的参数。

为了增加模型的复杂性和灵活性,在下次变换前,对滤波器响应进行挤压和激励,重新校准响应,

可有效提升模型性能。

## 1.2 RFEM

随着网络层数的增加，小目标的局部语义可能会缺失，需要较大感受野的非局部上下文信息作为重要补充。为此，在 Neck 部分引入了 RFEM。RFEM 结构如图 2 所示。

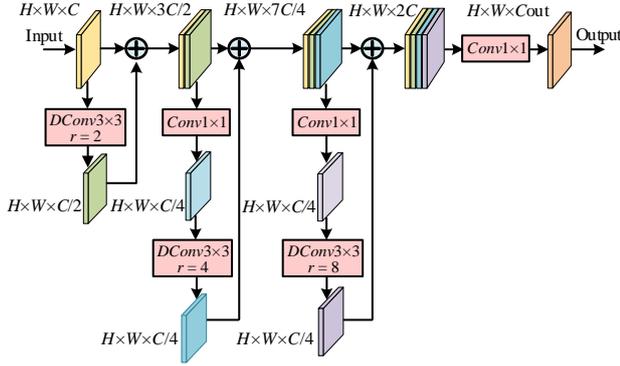


图 2 RFEM 结构

RFEM 计算过程可以用公式(5)~(8)表示：

$$R_1 = DConv_{3 \times 3}^{[1]}(F) \quad (5)$$

$$R_2 = DConv_{3 \times 3}^{[2]}(Conv_{1 \times 1}(R_1 \oplus F)) \quad (6)$$

$$R_3 = DConv_{3 \times 3}^{[3]}(Conv_{1 \times 1}(R_1 \oplus R_2 \oplus F)) \quad (7)$$

$$P = \begin{cases} R_1 \oplus R_2 \oplus R_3 \oplus F, & \text{if } C_{out} = 2C \\ Conv_{1 \times 1}(R_1 \oplus R_2 \oplus R_3 \oplus F), & \text{if } C_{out} \neq 2C \end{cases} \quad (8)$$

式中： $F$  为上一层获得的特征图； $R_i (i = 1, 2, 3)$  为中间变量； $DConv_{3 \times 3}^{[1]}$  为  $3 \times 3 \times C \times (C/2)$  的扩张卷积； $DConv_{3 \times 3}^{[2]}$ 、 $DConv_{3 \times 3}^{[3]}$  为  $3 \times 3 \times (C/4) \times (C/4)$  的扩张卷积，扩张速率分别为  $r = 4$ 、 $r = 8$ ； $Conv_{1 \times 1}$  为  $1 \times 1$  的标准卷积； $\oplus$  为通道拼接操作； $P$  为输出特征图；RFEM 结束处的  $1 \times 1$  标准卷积用于调整输出的特征映射通道； $C_{out}$  为 RFEM 的输出通道。

当  $F$  和  $P$  的步幅为 8，输入图像大小为  $640 \times 640$  时， $F$  的分辨率为  $80 \times 80$ 。若  $F$  的感受野为 1，则  $R_1$ 、 $R_2$ 、 $R_3$ 、 $P$  的感受野分别扩展为 5、13、29、29。此时  $P$  的感受野几乎覆盖了  $F$  的一半，包含了足够的非局部信息。RFEM 的密集连接金字塔结构对原

始感受野和不同扩展感受野进行叠加和重用，缓解了网格化问题。

## 1.3 FRM

FRM 从网络的不同层收集特征，并通过特征提取、特征融合、特征重组 3 个步骤重新组合这些特征，以形成更丰富的特征表示。本文利用 FRM 替换 YOLOv5 网络 Head 部分的 Detect 层，以增强网络对输入不同尺寸特征图的代表能力。

### 1.3.1 特征提取

对从 Neck 中获得的特征图  $L_1$ 、 $L_2$ 、 $L_3$  分别进行上采样、不做处理、下采样操作，得到 3 个尺寸一致的特征图  $R_1$ 、 $R_2$ 、 $R_3$ 。

### 1.3.2 特征融合

将  $R_1$ 、 $R_2$ 、 $R_3$  输入到两条特征融合支路中进行特征提取。采用通道注意力机制，即先对输入图像进行最大池化和平均池化操作，再通过卷积运算和 ReLU 激活函数处理得到特征映射  $F_1^i, F_2^i, F_3^i, i \in \{1, 2\}$ ， $i$  表示上下两条支路。在特征融合的另一条支路上，输入图像通过 Softmax 函数获得特征图中比较活跃的特征点。

### 1.3.3 特征重组

特征重组操作可表示为

$$Y_j = F_j^i \otimes R_j, j \in 1, 2, 3 \quad (9)$$

通过特征重组操作融合了网络的深层和浅层特征，有效改善了模型对小目标的检测性能。将融合后的特征图输入到 Head 部分的输出层，用于预测对象的类别、边界框坐标和置信度。

## 2 实验

### 2.1 数据集构建与预处理

为验证本文缺陷检测方法的有效性，从 roboflow 上的“Broken Solar Panel Detection”、“clean-dirty solar Panel Computer Vision Project”、“Yeni Computer Vision Project”和“solar-panel1 Computer Vision Project”开源数据集中选出 4 890 幅光伏面板缺陷图像，制作为

实验数据集, 图像大小为  $640 \times 640$ 。这些图像被标注为正常 (normal, 951 幅)、覆盖 (cover, 1 320 幅)、碎裂 (crack, 1 743 幅)、尘土 (dust, 1 633 幅) 4 种类型, 并按 7 : 2 : 1 的比例划分为训练集 (3 423 幅)、验证集 (978 幅) 和测试集 (489 幅)。

在深度学习中, 通常数据量越大, 数据种类越多, 训练的模型越准确, 泛化性能越好。但现实中少有方法可以获得丰富多样的图像, 且本实验数据集还存在不同类型的样本数量不平衡的问题。为此, 本文采用自适应旋转、随机裁剪、直方图均衡化等方法进行数据增强, 以提高实验数据集的规模。

本实验优化算法采用小批量梯度下降方法, 初始学习率设置为 0.01, 最终循环的学习率设置为 0.001。YOLOv5 的损失函数由边界框损失 ( $box_{loss}$ )、分类损失 ( $cls_{loss}$ )、置信度损失 ( $obj_{loss}$ ) 三部分组成, 总损失函数计算公式为

$$Loss = box_{gain} \times box_{loss} + cls_{gain} \times cls_{loss} + obj_{loss} \times obj_{gain} \quad (10)$$

式中:  $box_{gain}$ 、 $cls_{gain}$ 、 $obj_{gain}$  分别为边界框损失、分类损失、置信度损失的权重。本实验设置  $box_{gain} = 0.05$ ,  $cls_{gain} = 0.03$ ,  $obj_{gain} = 0.07$ 。

YOLOv5 的边界框损失函数用于衡量边界框的交并比 (intersection of union, IoU) :

$$IoU = \frac{A \cap B}{A \cup B} \quad (11)$$

YOLOv5 的边界框损失函数的计算公式为

$$L_{iou} = 1 - IoU + \frac{\rho^2(b, b^{gt})}{c^2} \quad (12)$$

式中:  $b$ 、 $b^{gt}$  分别为预测框和真实目标框的中心坐标,  $\rho^2$  为  $b$  与  $b^{gt}$  的欧氏距离,  $c$  为预测框和真实目标框的最小闭包区域的对角线距离。

YOLOv5 的分类损失函数是二元交叉熵损失:

$$L_{cls} = -y \log p - (1 - y) \log (1 - p) \quad (13)$$

式中:  $y$  为二元标签值,  $p$  为预测为正样本的

概率。

YOLOv5 的置信度损失包括有物体时的置信度损失和无物体时的置信度损失。其中, 有物体时的置信度损失用于评估含有目标边界框的预测置信度的精确程度, 如果模型提供的边界框与真实目标框交叠, 那么该损失会激励模型给出更高的置信度, 若模型提供的边界框并未覆盖到真实目标框, 则该损失对过分自信的预测做出处罚, 以提升模型识别目标的能力; 无物体时的置信度损失用于衡量模型预测边界框覆盖对象的程度, 可增强模型捕捉目标的效果。

## 2.2 评价指标

本实验采用的评价指标有准确率 (precision)、召回率 (recall)、均值平均精度 (mean average precision, mAP)。

准确率是指正确预测为正样本的样本数与所有预测为正样本的样本数之比, 计算公式为

$$P_{Precision} = \frac{TP}{TP + FP} \quad (14)$$

式中:  $TP$  为正确预测为正样本的样本数,  $FP$  为错误预测为正样本的样本数。

召回率是指正确预测为正样本的样本数占所有实际正样本的比例, 能够反映模型的漏检状况, 计算公式为

$$R_{Recall} = \frac{TP}{TP + FN} \quad (15)$$

式中:  $FN$  为错误预测为负样本的样本数。

mAP 用于计算多类别的平均精度, 而平均精度 (AP) 则表示不同类别的平均精度, 其关系为

$$AP = \sum_{k=0}^{k=n-1} [R(k) - R(k+1)] \times P(k) \quad (16)$$

$$mAP = \frac{1}{n} \sum_{k=1}^n AP_k \quad (17)$$

实验训练过程中, 将计算置信度为 0.5 时的均值平均精度表示为  $mAP@0.5$ , 置信度为 0.5~0.95 时的均值平均精度表示为  $mAP@0.5:0.95$ 。

### 2.3 实验过程

本实验采用小批量梯度下降的优化算法，以及上文所述的损失函数进行训练。训练集设置 `batchsize` 为 16，图像大小为  $640 \times 640$ ，每次训练 100 个 `epochs`。

在本实验中，对 YOLOv5s 网络的改进（在 Backbone 部分引入 CBAM 或 SE-Block；在 Neck 部分引入 RFEM；在引入 RFEM 的同时添加 CBAM 或 SE-Block）进行消融实验，实验结果如表 1 所示。

由表 1 可以看出：YOLOv5s 网络引入 CBAM 或 SE-Block 后，准确率和召回率有明显提升，说明模型的正类预测准确性、目标的漏检情况有所改善；引入 RFEM 后，*mAP* 和准确率都有所提升。综上所述，在消融实验中，YOLOv5s 网络引入 RFEM 和 CBAM 的效果最好。

表 1 消融实验

网络结构	准确率	召回率	mAP@0.5	mAP@0.5:0.95
YOLOv5s	0.759	0.784	0.774	0.408
YOLOv5s+CBAM	<b>0.793</b>	0.785	0.800	0.422
YOLOv5s+SE-Block	0.761	0.766	0.787	0.418
YOLOv5s+RFEM	0.784	0.789	0.800	<b>0.429</b>
YOLOv5s+RFEM+CBAM	0.784	0.793	<b>0.801</b>	0.418
YOLOv5s+RFEM+SE-Block	0.777	<b>0.807</b>	0.793	0.413

先利用 FRM 替换 YOLOv5s 网络 Head 部分的 Detect 层，再分别引入 CBAM、SE-Block 及 RFEM 进行对比实验，实验结果如表 2 所示。

表 2 FRM 替换 YOLOv5s 网络 Head 部分的 Detect 层后的对比实验

网络结构	检测层	准确率	召回率	mAP@0.5	mAP@0.5:0.95
YOLOv5s	Detect	0.759	0.784	0.774	0.408
YOLOv5s	FRM	0.821 ↑	0.723 ↓	0.811 ↑	0.443 ↑
YOLOv5s+CBAM	Detect	0.793	0.785	0.800	0.422
YOLOv5s+CBAM	FRM	0.811 ↑	0.739 ↓	0.795 ↓	0.423 ↑
YOLOv5s+SE	Detect	0.761	0.766	0.787	0.418
YOLOv5s+SE	FRM	0.841 ↑	0.713 ↓	0.794 ↑	0.411 ↓
YOLOv5s+RFEM	Detect	0.784	0.789	0.800	0.429
YOLOv5s+RFEM	FRM	0.845 ↑	0.767 ↓	0.821 ↑	0.453 ↑

由表 2 可以看出，用 FRM 替换 Head 部分的 Detect 层后，准确率和 *mAP* 均有明显提升，其中引入 RFEM 后提升效果最为明显，但召回率比替换前有所下降，这是因为特征重组后，目标识别能力虽然上升，但存

在关键信息漏检的情况。

利用 FRM+RFEM 模型对可见光图像光伏面板的表面缺陷进行检测，获得的图像目标锚框和类别如图 3 所示。

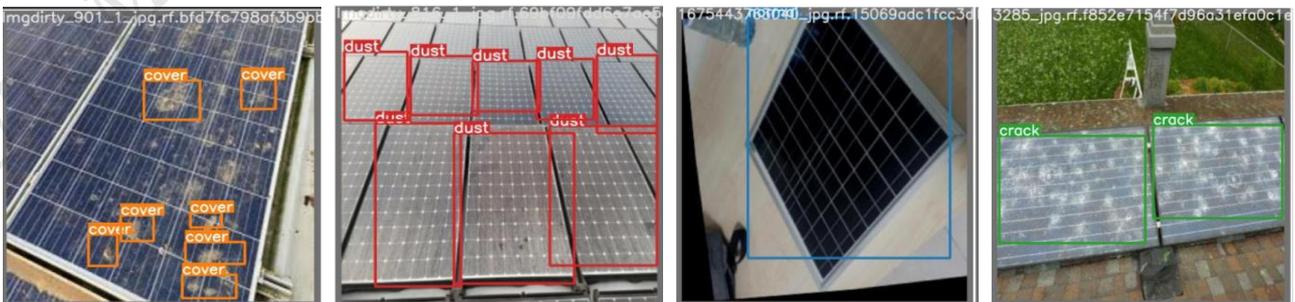


图 3 光伏面板表面缺陷检测效果图

由图 3 可知,本文提出的检测方法能够有效检测出光伏面板的覆盖(cover)、灰尘(dust)、破损(crock)等表面缺陷。

### 3 结论

本文基于 YOLOv5 网络提出了光伏面板缺陷检测方法。通过引入注意力机制,提升了网络对缺陷目标的聚焦和识别能力;引入 RFEM,减少了因样本本身或类内差异大等问题导致的训练效果不佳的情况;引入 FRM,提高了网络对目标位置的判断能力,从而有效地提升了模型检测的准确率和精度。

©The author(s) 2024. This is an open access article under the CC BY-NC-ND 4.0 License (<https://creativecommons.org/licenses/by-nc-nd/4.0/>)

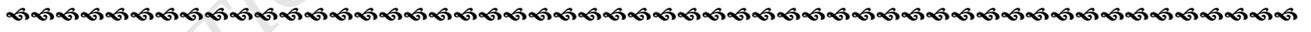
### 参考文献

[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.  
[2] 沈凌云,张洁,孙甲云,等.基于改进 Alexnet 模型的光伏电

池瑕疵分类方法[J].自动化与信息工程,2022,43(4):42-47.  
[3] 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.  
[4] 吕潇涵.基于改进 Vision Transformer 的光伏电池缺陷识别研究[J].计算技术与自动化,2023,42(4):33-40.  
[5] 陈海永,史世杰.基于归一化控制器的光伏图像无监督域适应缺陷检测[J].太阳能学报,2024,45(7):540-547.  
[6] 陶志勇,何燕,林森,等.局部和全局特征融合的太阳能电池片表面缺陷检测[J].光电工程,2024,51(1):93-107.  
[7] 艾上美,周剑峰,张必朝,等.基于改进 SSD 算法的光伏组件缺陷检测研究[J].智慧电力,2023,51(12):53-58.  
[8] HE K, ZHANG X, REN S, et al. Spatial pyramid pooling in deep convolutional networks for visual recognition[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2015, 37(9):1904-1916.  
[9] WOO S, PARK J, LEE J Y, et al. CBAM: Convolutional block attention module[C]. Proceedings of the European Conference on Computer Vision (ECCV), 2018:3-19.  
[10] HU J, SHEN L, SUN G. Squeeze-and-excitation networks[C]. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition(CVPR), 2018:7132-7141.

### 作者简介:

武果,男,1984年生,本科,工程师,主要研究方向:电子电路、工业自动化、光机电一体化。E-mail: 312830590@163.com  
高守正,男,2001年生,本科,主要研究方向:图像处理、目标检测方法。E-mail: 1213449232@qq.com  
张德福(通信作者),男,1967年生,本科,高级工程师,主要研究方向:机器人系统集成与应用、智能装备设计研发。E-mail: 13760841878@163.com  
肖国龙,男,1965年生,硕士研究生,高级工程师,主要研究方向:电路、信号信息系统。E-mail: jxdruil@giet.ac.cn  
龚正平,男,1972年生,本科,高级工程师,主要研究方向:工业电气自动化、仪器仪表技术。E-mail: zpjiong@giet.ac.cn  
覃健,男,1992年生,本科,主要研究方向:设备自动化、视觉检测、软件开发。E-mail: qinjian@giet.ac.cn



(上接第 46 页)

### 作者简介:

廉胤东,男,1992年生,博士研究生,工程师,主要研究方向:数字供应链、机器人调度、信息物理系统。E-mail: 601292628@qq.com  
苟彬,男,1996年生,博士研究生,工程师,主要研究方向:设备缺陷检测与仿真模拟。E-mail: wjg5267871@qq.com  
李超磊,男,1995年生,硕士研究生,主要研究方向:计算机视觉、人工智能。E-mail: 956901860@qq.com  
李泽明,男,1997年生,硕士研究生,主要研究方向:供应链管理、区块链。E-mail: 931717079@qq.com  
余锦伟,男,1995年生,博士研究生,工程师,主要研究方向:机器视觉、图像处理、人工智能。E-mail: yujw5@chinatelecom.cn  
曾俊海,男,1996年生,博士研究生,主要研究方向:智能优化算法及机器人调度。E-mail: jhaizeng@163.com