

# 一种基于注意力机制目标检测的无人机的 光伏板巡检方法

姚瑛,顾冰,方正楠

(浙江泰仑电力集团有限责任公司,浙江 湖州 313000)

**摘要:**随着太阳能光伏板的广泛应用,光伏板的隐患及故障缺陷检测变得越发重要。为了提高光伏板的巡检效率和缺陷检测的准确性,设计了一种基于注意力机制目标检测算法,并结合无人机技术提出了一套完整的光伏板巡检方案。该方法利用无人机搭载的高分辨率相机进行图像采集,并通过注意力机制的目标检测和无人机技术相结合的方式进行光伏板缺陷的检测和识别。实验结果表明,所提方法在光伏板巡检方面具有较高的准确性和可靠性,平均精度均值(mAP)达到 83.2%,F1 得分为 84.5%,有效提升了光伏板巡检的质量和效果,具有很高的实用价值。

**关键词:**目标检测;无人机;注意力机制;光伏板巡检

**中图分类号:**TM28 **文献标识码:**A **DOI:**10.19457/j.1001-2095.dqcd25344

## A Photovoltaic Panel Inspection Method for UAV Based on Attention Mechanism Target Detection

YAO Ying, GU Bing, FANG Zhengnan

(Zhejiang TaLent Electric Group Company Limited, Huzhou 313000, Zhejiang, China)

**Abstract:** With the wide application of solar photovoltaic panels, the detection of hidden dangers and faults of photovoltaic panels has become more and more important. In order to improve the inspection efficiency of photovoltaic panels and the accuracy of defect detection, a target detection algorithm based on attention mechanism was designed and a complete inspection scheme of photovoltaic panels combined with UAV technology was proposed. The method uses the high-resolution camera mounted on the UAV for image acquisition, and detects and identifies the defects of photovoltaic panels through the combination of the object detection of the attention mechanism and the UAV technology. The experimental results show that the proposed method has high accuracy and reliability in photovoltaic panel inspection, mean average accuracy (mAP) reaches 83.2%, F1 score is 84.5%, effectively improve the quality and effect of photovoltaic panel inspection, and has high practical value.

**Key words:** object detection; unmanned aerial vehicle (UAV); attention mechanism; photovoltaic panel inspection

随着全球能源需求的不断增长和对环境可持续性的关注,光伏发电作为清洁能源的重要来源之一,正逐渐成为全球能源转型的关键技术<sup>[1]</sup>。光伏板作为光伏发电系统的核心组成部分,其性能的稳定性和可靠性对系统的效能和安全至关重要。然而,在长时间运行过程中,光伏板受气候变化、尘埃、湿度、温度波动、机械应力等外界因素影响,可能引出灰尘覆盖、损坏、连接器松动、电池老化等潜在问题,不仅会降低光伏板的寿命,还有可能影响局部区域的供电进而造成损

失。为了确保光伏发电系统的稳定性和高效性,对光伏板进行定期巡检十分必要<sup>[2]</sup>。

目前,光伏板巡检主要依赖于人工巡视和手动检测,但这些方式存在一些不足。首先,人工巡检需要耗费大量的人力和时间,在大规模的光伏电站中难以实施。其次,人工巡检容易受到主观因素的影响,检测结果可能存在误差和不确定性<sup>[3]</sup>。相比之下,无人机巡检具有速度快、效率高、精确性好、安全可靠和节省成本等特点,结合无人机和目标检测技术对光伏板巡检会提升巡

**基金项目:**浙江泰仑电力集团有限责任公司科技项目(HZJTKJ2022-01)

**作者简介:**姚瑛(1984—),男,本科,工程师,主要研究方向为电力自动化,Email:21123155@qq.com

检的效率和质量<sup>[4]</sup>。

目前流行的目标检测算法包括:多阶段的Faster-RCNN (faster regions with convolutional neural network)<sup>[5]</sup>系列目标检测算法和单阶段的SSD (single shot multibox detector)<sup>[6]</sup>和YOLO(you only look once)<sup>[7-8]</sup>目标检测算法。Faster-RCNN采用了RPN(region proposal network)的区域提议算法,能够检测和精确定位图像中的对象。虽然精度和稳定性较高,但计算资源需求较大,检测速度相对较慢,因此不适用于实时性要求较高的场景。SSD虽然是单阶段检测,但是其对每个像素点能生成多个锚框,导致在绝大部分情况下,相邻像素点所生成的锚框的重叠率相当高,从而产生大量重复计算。而YOLOv5(you only look once version 5)网络则是一种端到端的目标检测方法,善于快速且精确地定位图像中的物体。它的思想是尽量让锚框不重叠,因此检测速度远远快于SSD,适用于实时性要求较高的场景。然而,YOLOv5在检测精度方面略显不足,并在复杂场景下很难精准识别和定位小目标。

针对上述方法存在的不足,本文综合运用深度学习、计算机视觉算法和无人机技术,提出了一种高效、准确的光伏板巡检方法,通过设置精细化巡检的检测头来提高小目标的检测精度和覆盖范围,并结合注意力机制扩大网络感受野,增强特征表征能力和模型可解释性,实现光伏板全面性和定期性的缺陷检测和故障诊断。

## 1 无人机光伏板巡检

无人机巡检主要由3大部分组成:配备高清摄像机的无人机、机载AI计算终端和后端演示平台。巡检过程图如图1所示。



图1 无人机巡检过程图

Fig.1 UAV inspection process diagram

在光伏现场,无人机搭载着高分辨率相机,对光伏板进行图像采集。这些图像随后会被传输到机载AI计算终端,并使用深度学习算法进行快速而准确的缺陷识别和故障检测。因此,通过结合无人机技术、高分辨率相机和深度学习算法,就可以有效地对光伏板进行缺陷检测,并及时采取必要的措施进行处理。检测结果会通过4G网络传输到可视化平台,实现在线监测和远程观察。这种方式大大提高了巡检效率,而无需人工逐一检查。同时,由于高分辨率相机的使用,可以精确地捕捉到光伏板上的细微缺陷和故障,确保光伏系统的高效运行。

## 2 注意力机制目标检测算法

注意力机制目标检测算法的整个网络结构如图2所示,该结构包括骨干网络、颈网络和检测头三部分,分别用于共享特征提取、特征融合与目标检测。

通过提取共享特征不仅可以用较少的信息表示原始数据来减少存储和计算成本,而且能够捕获数据的核心信息而忽略那些不重要或是噪声造成的变化。此外,也有助于模型更好地泛化到未见过的数据,而不是依赖于特定样本的特异性。同时,为了缓解目标尺度变化的负面影响,本文在骨干网络分别提取了 $20 \times 20$ ,  $40 \times 40$ ,  $80 \times 80$ ,  $160 \times 160$ 四种尺度的特征,并将其输送到颈网络中,将不同空间尺度的边缘、纹理信息以及语义信息进行融合,达到充分提取共享特征的目的。

为了对提取的共享特征进行充分融合,本文使用注意力机制+FPN(feature pyramid network)+PAN(path aggregation network)相结合的方式作为颈网络的基本架构。

首先,通过引入注意力机制,模型可以专注于对特定任务更为关键的特征部分,从而提高其性能。其次,通过FPN和PAN结合可以从多个尺度捕获图像的信息,这对于检测不同大小的物体特别有效。最后,PAN增强了信息在不同尺度特征之间的传递,有助于细节和语义信息的交互。

为了精准识别无人机飞行过程中捕获的细小目标,设计了一个针对小目标的检测头模块。该模块分别对应了 $20 \times 20$ ,  $40 \times 40$ ,  $80 \times 80$ ,  $160 \times 160$ 四种感受野,保证对微小目标的检测,同时也有效缓解了剧烈的目标尺度变化带来的负面影响。



表1 骨干网络模块表

Tab.1 Table of backbone network modules

| 序号 | 模块    | 数量 | 宽×高×通道      | 卷积核    | 步长 |
|----|-------|----|-------------|--------|----|
| 0  | Focus | 1  | 320×320×64  | 6      | 2  |
| 1  | Conv  | 1  | 160×160×128 | 3      | 2  |
| 2  | C2f   | 3  | 160×160×128 |        |    |
| 3  | Conv  | 1  | 80×80×256   | 3      | 2  |
| 4  | C2f   | 6  | 80×80×256   |        |    |
| 5  | Conv  | 1  | 40×40×512   | 3      | 2  |
| 6  | C2f   | 3  | 40×40×512   |        |    |
| 7  | Conv  | 1  | 20×20×1 024 | 3      | 2  |
| 8  | SPP   | 1  | 20×20×1 024 | 5,9,13 | 1  |
| 9  | Trans | 3  | 20×20×1 024 |        |    |

### 2.2 颈网络

颈网络采用了 Attention+FPN+PAN 的结构,通过自顶向下的特征传播实现了多层次特征融合和上下文信息提取,并通过自底向上的连接缩短了低层与顶层特征之间的信息路径。每一层底部都嵌入了全局注意力机制模块(GAM)<sup>[11]</sup>,将融合的信息以放大到全局跨空间通道维度的相互作用,从而使得在不同层次和不同尺度捕获的重要目标对象特征深度融合。此外GAM作为轻量级模块嵌入到颈网络中,减少了冗余特征的数量,并优化了网络的计算效率。GAM模块网络结构图如图5所示。

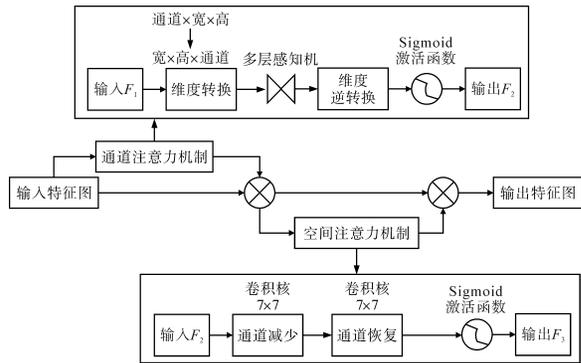


图5 GAM模块网络结构

Fig.5 GAM module network structure

在通道注意力子模块中,首先将通道、宽度、高度三个维度上的信息进行维度转换,然后通过一个两层的多层感知器用于放大通道和空间依赖性,再转换为原来的维度,进行Sigmoid激活后得到通道注意力的特征图。在空间注意力子模块中,为了关注空间信息,使用两个卷积层进行空间信息融合。首先通过卷积核为7的卷积缩减通道数量,缩小计算量,再经过一个卷积核为7的卷积操作,增加通道数量,保持通道数量的一致,最后经过Sigmoid得到空间注意力的特征图。

GAM模块的输入输出特征对应关系如下式所示:

$$F_2 = M_c(F_1) \otimes F_1 \quad (1)$$

$$F_3 = M_s(F_2) \otimes F_2 \quad (2)$$

式中: $F_1$ 为输入特征图; $F_2$ 为中间特征图; $F_3$ 为输出特征图; $M_c, M_s$ 分别为通道注意力和空间注意力特征图谱; $\otimes$ 表示对应元素相乘。

表2展示了颈网络各个模块的拼接结构。由此表可知,颈网络在序号10的Conv模块和序号12、序号17、序号22的Concat模块接入了骨干网络提取的不同尺度的共享特征信息,经过特征融合与注意力机制模块的特征筛选后,分别从序号23、序号27、序号31、序号35的Trans模块将融合的不同尺度的空间特征及语义信息输送到序号36的Conv模块中。

表2 颈网络模块表

Tab.2 Table of neck network modules

| 序号 | 入/出 | 模块       | 数量 | 宽×高×通道      | 卷积核 | 步长 |
|----|-----|----------|----|-------------|-----|----|
| 10 | 9   | Conv     | 1  | 20×20×512   | 1   | 1  |
| 11 |     | Upsample | 1  | 40×40×512   |     |    |
| 12 | 6   | Concat   | 1  | 40×40×1 024 |     |    |
| 13 |     | C2f      | 3  | 40×40×512   |     |    |
| 14 |     | GAM      | 1  | 40×40×512   |     |    |
| 15 |     | Conv     | 1  | 40×40×256   | 1   | 1  |
| 16 |     | Upsample | 1  | 80×80×256   |     |    |
| 17 | 4   | Concat   | 1  | 80×80×512   |     |    |
| 18 |     | C2f      | 3  | 80×80×256   |     |    |
| 19 |     | GAM      | 1  | 80×80×256   |     |    |
| 20 |     | Conv     | 1  | 80×80×128   | 1   | 1  |
| 21 |     | Upsample | 1  | 160×160×128 |     |    |
| 22 | 2   | Concat   | 1  | 160×160×256 |     |    |
| 23 | 36  | Trans    | 1  | 160×160×128 |     |    |
| 24 |     | GAM      | 1  | 160×160×128 |     |    |
| 25 |     | Conv     | 1  | 80×80×128   | 1   | 1  |
| 26 |     | Concat   | 1  | 80×80×256   |     |    |
| 27 | 36  | Trans    | 1  | 80×80×256   |     |    |
| 28 |     | GAM      | 1  | 80×80×256   |     |    |
| 29 |     | Conv     | 1  | 40×40×256   | 1   | 1  |
| 30 |     | Concat   | 1  | 40×40×512   |     |    |
| 31 | 36  | Trans    | 2  | 40×40×512   |     |    |
| 32 |     | GAM      | 1  | 40×40×512   |     |    |
| 33 |     | Conv     | 1  | 20×20×512   |     |    |
| 34 |     | Concat   | 1  | 20×20×1 024 |     |    |
| 35 | 36  | Trans    | 3  | 20×20×1 024 |     |    |

### 2.3 检测头网络

为了有效地检测无人机飞行中拍摄的微小目标,在检测头专门增加了一个检测小目标的预测头。该预测头与其他三个预测头相结合,形成

了一个分别对应 128×128, 256×256, 512×512 和 1 024×1 024 四种分辨率的结构,以缓解目标尺度变化带来的负面影响。尽管增加了部分计算开销和存储成本,但对无人机高空巡检微小物体精度方面带来了显著的提升。

鉴于分类任务注重图像的纹理特征,而回归定位任务则关注图像的边缘特征。为了更好地满足这两个任务的要求,在检测头设计了不同的分支进行运算,并对检测头部分进行了解耦设计,使得分类和回归任务能够独立地进行处理,针对不同类型的特征捕获更为精细的表征,从而提高网络的表征能力和分类性能。

表3展示了检测头网络各个模块的拼接结构。检测头网络的 128, 256, 512, 1 024 四种输出通道对应着 160×160, 80×80, 40×40 和 20×20 四种不同的检测尺度,可有效缓解目标尺度变化带来的负面影响。

表3 检测头网络模块表

Tab.3 Table of detection network modules

| 序号 | 模块   | 数量 | 输出通道                 | 卷积核 | 步长 |
|----|------|----|----------------------|-----|----|
| 36 | Conv | 1  | 128, 256, 512, 1 024 | 3   | 1  |
| 37 | Conv | 2  | 128, 256, 512, 1 024 | 1   | 1  |
| 38 | Conv | 2  | 128, 256, 512, 1 024 | 1   | 1  |

## 2.4 损失函数

本文设计的损失函数由分类损失、边框回归损失和目标置信度损失三部分组成。其中分类损失函数负责优化目标检测模型的分类能力,目标置信度损失函数负责优化模型对目标的识别能力,二者均采用交叉熵损失函数,即

$$L = -\frac{1}{N} \sum_{n=1}^N [y_n \times \log(x_n) + (1 - y_n) \times \log(1 - x_n)] \quad (3)$$

式中: $y_n$ 为样本的标签; $x_n$ 为样本预测为正的的概率。

边框回归损失函数负责优化检测模型的边框定位能力,衡量预测的边框与真实边框之间的差异,用于调整边框的位置和大小。本文采用 CIoU (complete intersection over union)<sup>[12]</sup>来计算, CIoU 可视化图如图7所示。

设预测框左上角、右下角坐标分别为 $(x_{p1}, y_{p1})$ 和 $(x_{p2}, y_{p2})$ ,标签框的左上角、右下角坐标分别为 $(x_{l1}, y_{l1})$ 和 $(x_{l2}, y_{l2})$ 。在计算两个外接矩形之间的距离时,采用对角线距离替代了交并比(intersection over union, IoU)中的欧式距离,同时还引入了

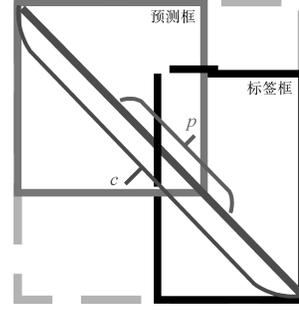


图7 CIoU可视化图

Fig.7 CIoU visualization

可调节的参数平衡边界框的宽高比,使得结果更加准确。CIoU的计算公式如下:

$$CIoU = IoU - \frac{p^2}{c^2} - \alpha v \quad (4)$$

其中

$$IoU = \frac{Intersection}{S_{PRE} + S_{LOC} - Intersection} \quad (5)$$

$$Intersection = [\min(x_{l2}, x_{p2}) - \max(x_{l1}, x_{p1})] \times [\min(y_{l2}, y_{p2}) - \max(y_{l1}, y_{p1})] \quad (6)$$

$$S_{PRE} = (x_{p2} - x_{p1})(y_{p2} - y_{p1}) \quad (7)$$

$$S_{LOC} = (x_{l2} - x_{l1})(y_{l2} - y_{l1}) \quad (8)$$

$$v = \frac{4}{\pi^2} \left( \arctan \frac{w_l}{h_l} - \arctan \frac{w_p}{h_p} \right)^2$$

$$= \frac{4}{\pi^2} \left( \arctan \frac{x_{l2} - x_{l1}}{y_{l2} - y_{l2}} - \arctan \frac{x_{p2} - x_{p1}}{y_{p2} - y_{p1}} \right)^2 \quad (9)$$

$$\alpha = \frac{v}{1 - IOU + v} \quad (10)$$

式中: $p$ 为标签框和预测框的中心点距离; $c$ 为标签框和预测框的最小包围矩形的对角线长度; $v$ 为标签框和预测框的宽高比相似度; $\alpha$ 为 $v$ 的影响因子,范围为0~1之间; $Intersection$ 为预测框和标签框交集的面积; $S_{PRE}$ 为预测框的面积; $S_{LOC}$ 为标签框的面积。

## 3 实验结果与分析

### 3.1 实验环境与参数

实验使用的数据集是由2 000张无人机采集的光伏板缺陷(灰尘和破损)图片组成,并按照6:2:2的比例划分训练集、验证集和测试集。这样划分不仅可以尽量保持数据分布的平衡,避免某一部分数据过少或过多,使得数据分布与整个数据集相似;而且还可以保证验证集和测试集的大小相对合适,能够较准确地评估模型的性能。其中,训练集用于训练模型,验证集用于调参和选择模型的最佳超参数,而测试集则用于最终评

估模型的泛化能力。具体而言,本文将数据集划分为1 200张图像作为训练集,400张图像作为验证集,400张图像作为测试集。首先在训练阶段用大部分数据进行模型的学习和参数调整,然后在验证集上进行模型的选择,最后基于测试集的验证来评估模型的泛化能力。

实验环境配置如表4所示。实验中,以初始值为0.01的学习率开始进行训练。通过学习率调整策略,每迭代300次,将学习率缩小为原来的0.9倍。每次输入的批大小(batch\_size)为16,并且采用随机梯度下降(SGD)<sup>[13]</sup>优化器优化网络权重。

表4 实验环境配置表

| 名称    | 配置                                  |
|-------|-------------------------------------|
| OS    | Ubuntu18.04                         |
| GPU   | NVIDIA GeForce RTX3090, 24GB Memory |
| CPU   | 32GB RAM                            |
| CUDA  | 11.3                                |
| CUDNN | 8.2.1                               |

### 3.2 实验评估指标

本文实验中采用了目标检测领域最为广泛使用的评价指标,包括精确率(Precision)、召回率(Recall),计算式如下:

$$Precision = \frac{TP}{TP + FP} = \frac{TP}{Predicted\ Positive} \quad (11)$$

$$Recall = \frac{TP}{TP + FN} = \frac{TP}{real\ Positive} \quad (12)$$

式中:TP为实际为缺陷(灰尘和破损),且被正确预测为缺陷的数量;FP为实际不是缺陷,但被错误预测为缺陷的数量;FN为实际为缺陷,但被错误预测为不是缺陷的数量。

然而,仅仅使用精确率和召回率来评价模型的性能是不准确的,还需要使用P-R曲线、mAP指标和F1得分等进一步评判模型的性能。

P-R曲线由横轴为Recall、纵轴为Precision绘制而成,其曲线下的面积定义为AP。AP值越高,则表示模型的性能越好,表示为

$$AP = \int_0^1 P(r) dr \quad (13)$$

$$mAP = \frac{1}{n} \sum_{i=1}^n AP_i \quad (14)$$

F1得分是将精确率和召回率的得分进行调和,从而达到衡量分类器质量的目的。F1得分取得最高值的点是达到精确度和召回率最佳平衡的关键点。

$$F1 = 2 \times \frac{Precision \times Recall}{Precision + Recall} \quad (15)$$

### 3.3 模型对比分析

本文在光伏板的灰尘和破损等缺陷检测中,采用原始的YOLOv5网络及其改进方法作为基线与本文提出的方法进行实验比较,并对结果进行分析和评估,如表5所示。

表5 YOLO算法与本文方法的性能对比表

Tab.5 Performance comparison table between YOLO algorithms and proposed method

| 算法     | 精确率          | 召回率          | mAP@0.5      | mAP@0.5~0.95 | F1           |
|--------|--------------|--------------|--------------|--------------|--------------|
| YOLOv5 | 0.859        | 0.768        | 0.809        | 0.419        | 0.811        |
| +检测头   | <b>0.899</b> | 0.774        | 0.821        | 0.435        | 0.831        |
| +Trans | 0.866        | 0.795        | 0.815        | 0.428        | 0.829        |
| +CBAM  | 0.877        | 0.789        | 0.817        | 0.430        | 0.830        |
| +GAM   | 0.878        | 0.792        | 0.823        | 0.437        | 0.832        |
| OURs   | 0.891        | <b>0.803</b> | <b>0.832</b> | <b>0.441</b> | <b>0.845</b> |

通过分析表5,可以得出结论:相较于原始的YOLOv5,新增加的检测头或者注意力机制模块都在不同程度上提升了性能。这说明本文所提出的改进模块能够有效地提高模型的性能和质量。与CBAM模块相比,GAM模块在精确率和召回率上分别提高了0.001和0.003,同时在mAP@0.5指标上提高了0.7%。尽管仅增加检测头在精确率方面取得了最高的结果(0.899),但在其他评价指标上,本文所提出的方法均表现出最优的性能。相较于其他方法的最优值,本文提出的改进方法在mAP@0.5, mAP@0.5~0.95和F1得分上分别提升了1.1%,0.9%和1.5%。

图8展示上述基线方法的损失随着训练步数下降的曲线图。整体损失按照分类损失、边框回归损失和目标置信度损失三者2:7:1的比例加权得到。从图中可以看出,整体上各类方法都随着训练步数的增加而减少损失,但本文所提出的方法具有最快的收敛速度,并且在同一阶段与其他方法相比有最低的误差损失,表现出了最佳的综合性能。

目前YOLOv5的检测水平已经与Faster-RCNN和SSD效果相当。基于此,本文新增加了Faster-RCNN和SSD的实验比较,进一步证明本文提出方法的速度与性能上的优势,性能表对比结果如表6所示。

由表6可知,在精度方面,Faster-RCNN的检测精度要好于SSD512,而本文提出的方法在mAP@0.5和mAP@0.5~0.95分别较Faster-RCNN

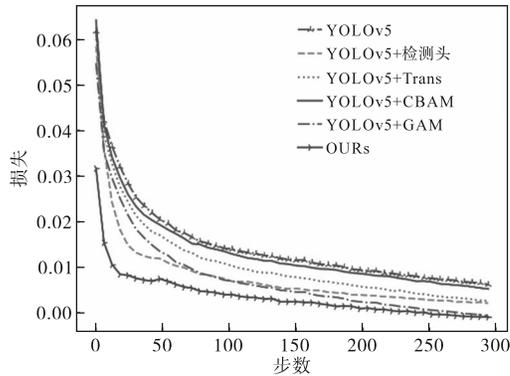


图8 算法的损失曲线图

Fig.8 The loss curves of the algorithm

表6 其他算法的性能表

Tab.6 Performance table of other algorithms

| 算法          | $mAP@0.5$    | $mAP@0.5\sim0.95$ | F1           | FPS |
|-------------|--------------|-------------------|--------------|-----|
| Faster-RCNN | 0.827        | 0.438             | 0.840        | 4   |
| SSD512      | 0.820        | 0.431             | 0.830        | 10  |
| OURs        | <b>0.832</b> | <b>0.441</b>      | <b>0.845</b> | 73  |

提升了0.6%和0.7%,这进一步证明本文方法的检测能力。另外在检测速度方面,在本文采用的RTX3090的GPU显卡上,本文提出的基于多头注意力机制的目标检测算法相较于Faster-RCNN和SSD算法在FPS(frame per second)参数上分别提升了18倍和7倍,在检测速度上遥遥领先。

图9和图10分别为本文提出的注意力机制光伏板巡检方法的P-R曲线图和缺陷检测效果图。图中的检测灰尘的AP值为0.829,破损的AP值为0.834。图中的数据曲线充分证明了本文提出的方法在光伏板缺陷检测方面具有高准确性和实用性,可有效应用于基于无人机的光伏板巡检任务。

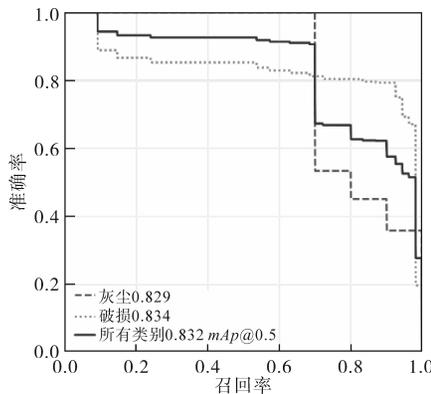


图9 P-R曲线图

Fig.9 P-R curves

为了验证本文提出的基于多头注意力机制的光伏板巡检算法的鲁棒性和泛化能力,本文除了在无人机采集的光伏板缺陷(灰尘和破损)的

样本验证了其性能之外,在开放的coco2017数据集上进行进一步验证。

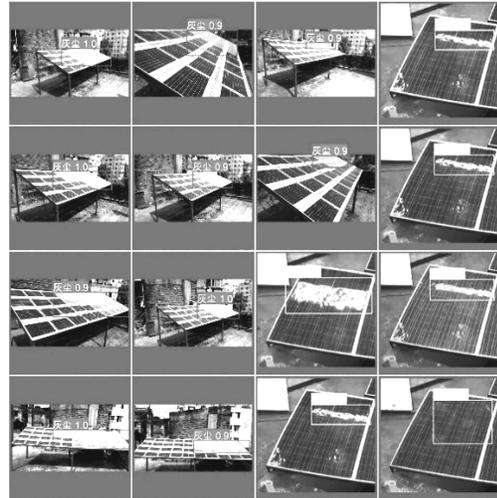


图10 光伏板灰尘和破损检测效果图

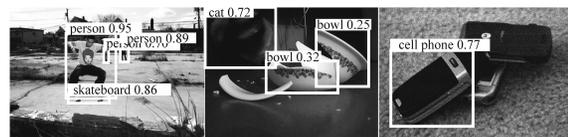
Fig.10 Dust and damage detection effect of photovoltaic panel

针对数据集的80类目标,本文统计了 $mAP@0.5$ 和 $mAP@0.5\sim0.95$ 两个指标来进行比较分析,基准方法为YOLOv5m。从表7可以看出,本文的注意力机制的目标检测方法在 $mAP@0.5$ 和 $mAP@0.5\sim0.95$ 分别提升了3.3%和2.9%,虽然在公开的coco2017数据集中检测的类别较多, $mAP$ 提升的幅度较于本文光伏板表面缺陷检测的结果有所下降,但依然有效地证明了本文提出方法的泛化性以及检测效果的提升。此外,图11展示了本文提出的注意力机制的目标检测方法与YOLOv5M在coco2017公开数据集的检测效果对比图,除了检测精度的明显提升之外,在小目标和漏检率方面,本文提出的方法也有明显的优势。

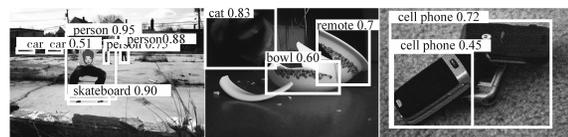
表7 coco2017数据模型性能表

Tab.7 coco2017 data model performance table

| 模型      | $mAP@0.5$ | $mAP@0.5\sim0.95$ |
|---------|-----------|-------------------|
| YOLOv5M | 64.1      | 45.4              |
| OURs    | 66.2      | 46.7              |



(a)YOLOv5M



(b)OURs

图11 coco2017数据集模型检测效果图

Fig.11 The model checking effect diagram of the coco2017 dataset

## 4 结论

本文提出了一种基于注意力机制的目标检测算法,适用于无人机太阳能光伏板灰尘及缺陷检测。该算法可部署在端侧机载 AI 计算终端,实时应用于无人机检测任务,具有极高的实用价值。本文方法设计了4个检测头,扩大了检测目标的感受野,并在特征提取与融合环节引入全局注意力网络和多头自注意力机制,有效结合了浅层空间特征与深层语义特征。实验对比分析表明,本文方法在无人机光伏板巡检中灰尘与破损的 *mAP* 值达到 83.2%, F1 得分为 84.5%, 优于 YOLOv5 及其他改进的检测方法,证实了本文提出方法具有更高的特征表征能力和准确性。未来研究工作计划考虑红外图像与可见光图像的多光谱融合分析,以进一步提高检测性能。

### 参考文献

- [1] 张晓飞. 储能型准 Z 源光伏并网逆变系统控制策略研究[D]. 西安:陕西科技大学, 2020.  
ZHANG Xiaofei. Research on control strategy of energy-stored quasi-Z-source photovoltaic grid-connected inverter system[D]. Xi'an: Shaanxi University of Science & Technology, 2020.
- [2] 李露露. 基于深度学习的光伏板巡检图像故障检测研究[D]. 北京:华北电力大学, 2020.  
LI Lulu. Research on fault detection of photovoltaic panel inspection image based on deep learning[D]. Beijing: North China Electric Power University, 2020.
- [3] 徐章禄. 低倍率聚光光伏系统的性能研究[D]. 杭州:浙江大学, 2011.  
XU Zhanglu. Investigation of the performance of a low concentration photovoltaic system[D]. Hangzhou: Zhejiang University, 2011.
- [4] 王浩, 闫号, 叶海瑞, 等. 基于无人机的光伏电站智能巡检[J]. 红外技术, 2022, 44(5): 537-542.  
WANG Hao, YAN Hao, YE Hairui, et al. Intelligent patrol inspection of photovoltaic power station based on UAVs[J]. Infrared Technology, 2022, 44(5): 537-542.
- [5] REN S, HE K, 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, 2017(6): 1137-1149.
- [6] LIU W, ANGUELOV D, ERHAN D, et al. SSD: single shot multibox detector[J]. Computer Vision-ECCV 2016, 2016: 21-37.
- [7] REDMON J, DIVVALA S, GIRSHICK R, et al. You only look once: unified, real-time object detection[C]//IEEE Conference on Computer Vision and Pattern Recognition, 2016: 779-788.
- [8] REDMON J, FARHADI A. YOLOv3: an incremental improvement[J]. ArXiv Preprint, 2018: 1804.02767.
- [9] WANG C Y, BOCHKOVSKIY A, LIAO H Y M. YOLOv7: trainable bag-of-freebies sets new state-of-the-art for real-time object detectors[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2023: 7464-7475.
- [10] VASWANI A, SHAZEER N, PARMAR N, et al. Attention is all you need[C]//Advances in Neural Information Processing Systems, 2017.
- [11] HU J, SHEN L, SUN G. Squeeze-and-excitation networks[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2018: 7132-7141.
- [12] ZENG W, HUANG J, WEN S, et al. A masked-face detection algorithm based on M-EIOU loss and improved ConvNeXt[J]. Expert Systems with Applications, 2023, 225: 120037.
- [13] WOODWORTH B, PATEL K K, STICH S, et al. Is local SGD better than minibatch SGD[C]//International Conference on Machine Learning, PMLR, 2020: 10334-10343.

收稿日期: 2023-09-01

修改稿日期: 2023-11-13