

基于改进 YOLOv5 的电力设备检测算法*

郑婷婷 周浩 王秋忆
(云南大学信息学院 昆明 650500)

摘要: 针对电力设备背景复杂、小目标密集等特点导致无人机智能电力巡检精度低、效果不佳等问题,提出了一种改进 YOLOv5 的目标检测算法。首先在原模型上增加一层检测层,重新获取锚点框以便能更好地学习密集小目标的多级特征,提高模型应对复杂电力场景的能力;其次对模型的特征融合模块 PANet 结构进行改进,通过跳跃连接的方式融合不同尺度的特征,增强信息的传播与重用;最后结合协同注意力模块设计主干网络,以聚焦目标特征,增强复杂背景中密集目标区域的显著度。实验结果表明:所提算法的平均精度均值(IoU=0.5)达到 97.1%,比原网络检测性能提升了 5.6%,有效改善了复杂背景下小目标的错测、漏检现象。

关键词: YOLOv5 算法;电力设备检测;注意力机制;多尺度检测;特征融合

中图分类号: TP391 **文献标识码:** A **国家标准学科分类代码:** 510.99

Power equipment detection algorithm based on improved YOLOv5

Zheng Tingting Zhou Hao Wang Qiuyi
(School of Information, Yunnan University, Kunming 650500, China)

Abstract: Aiming at the problems of low accuracy and poor effect of UAV intelligent power inspection caused by the complex background and dense small targets of power equipment, an improved target detection algorithm of YOLOv5. Firstly, a detection layer is added to the original model to re-obtain the anchor frame so as to better learn the multi-level features of dense small targets and improve the ability of the model to deal with complex power scenarios. Secondly, the feature fusion module PANet structure of the model is improved, and the features of different scales are fused by jumping connection to enhance the dissemination and reuse of information. Finally, combined with the collaborative attention module, the backbone network is designed to focus on the target characteristics and enhance the visibility of dense target areas in complex backgrounds. The experimental results show that the average accuracy of the proposed algorithm (IoU=0.5) reaches 97.1%, which is 5.6% higher than the original network detection performance, and effectively improves the false detection and missed detection of small targets in complex background.

Keywords: YOLOv5 algorithm; power equipment detection; attention mechanism; multi-scale detection; feature fusion

0 引言

电力是支撑国民经济发展的基础性产业,电力设备作为其重要组成部分,其安全稳定运行对于整个电力系统至关重要。因此,如何便捷高效的对电力设备进行定时巡检,及时掌握设备的状态信息及周围环境状况,防止因设备故障造成的停电、火灾和其他潜在风险,是当前保障电网系统安全运作待解决的问题之一^[1]。传统的电力设备巡检方式主要靠人力,难以按时完成巡检任务且极不安全,随着无人机及智能巡检技术的快速发展,搭载目标检测算法的无人机巡检逐渐成为当下发展得新潮流,运检效率大大提升^[2]。基于深度学习的巡检影像数据自动分析与检测也受到人们

的广泛关注,具有很好的应用前景^[3]。

目前基于深度学习的目标检测算法主要分为两类。一类是先生成可能包含目标物体的候选区域,再进一步分类和校准的两阶段检测算法,代表有 RCNN^[4]、Fast-RCNN^[5]、Faster R-CNN^[6]等,大多数检测精度较好,但相应的时间成本和硬件消耗较大。另一类是候选框的产生和分类同时进行的一阶段检测算法,如 YOLO 系列^[7-10]、SSD^[11]、RetinaNet^[12]等,可实现端到端的训练,减少了模型复杂度并且能大幅提升图像的处理速度。随着目标检测算法日趋成熟,人们开始思考将其应用到电力设备巡检工作中。2018年,汤踊等^[13]使用 Faster R-NN 来对电力部件进行识别与分

收稿日期:2022-07-05

* 基金项目:云南省重大科技专项(202202AD080004)、国家自然科学基金(12263008)项目资助

类,其通过调整 CNN 卷积核的大小和对图像进行预处理操作扩充数据集,提高检测的有效性和可靠性。2020 年 Zheng 等^[14]提出了一种改进的变电站绝缘子图像 FSSD 检测模型,该模型设计了一种新的特征增强模块,提高了对变电站绝缘子红外图像特征的提取能力。同年 Liu 等^[15]提出了一种基于 YOLOv3 的外部电力绝缘设备自动位置识别与诊断方法,与 Faster R-CNN 相比该方法具有更好的实时性。

以上基于深度学习的检测方法虽然在电力设备检测精度和速度上略优于一般方法,但是在检测小目标和密集目标方面仍有不足之处。根据 MSCOCO 数据集定义,像素的绝对值小于 32×32 即可称之为小目标,由于其特征信息少、像素低等原因导致特征表达能力弱,在实际应用中存在漏检、误检及定位精确度不高的情况^[16-17]。本文具体针对电力塔、电线杆、销钉、绝缘子和电力工程车这 5 类目标进行检测,由于在无人机摄像头以及一些架设较高的图像捕获装置所拍摄的图像中,存在绝缘子、销钉、远距离电力塔

等较小密集目标,相应检测将会变得困难。针对以上存在问题,本文选择对设备要求不高,容易部署的 YOLOv5 算法基础上进行改进,通过增加一层小目标检测层,提高对小目标的特征提取能力,并在四尺度的基础上通过跳跃连接的方式进一步改进 neck 结构,开展多尺度特征融合,进一步利用浅层信息。然后结合协同注意力机制对网络模型架构进行调整,增强复杂背景中密集目标区域的显著度,有利于进一步提升网络性能,开展智能巡检工作。

1 YOLOv5 网络模型介绍

YOLOv5 是由 UltralyticsLLC 公司发布的一种单阶段目标检测算法,结构如图 1 所示。根据网络结构的深度和宽度不同可将其分为 YOLOv5s、YOLOv5m、YOLOv5l、YOLOv5x 四种模型,其中 YOLOv5s 为模型深度和宽度均最小的网络,更容易部署在无人机等移动硬件平台上且速度最快,因此本文选取 YOLOv5s 作为研究对象。

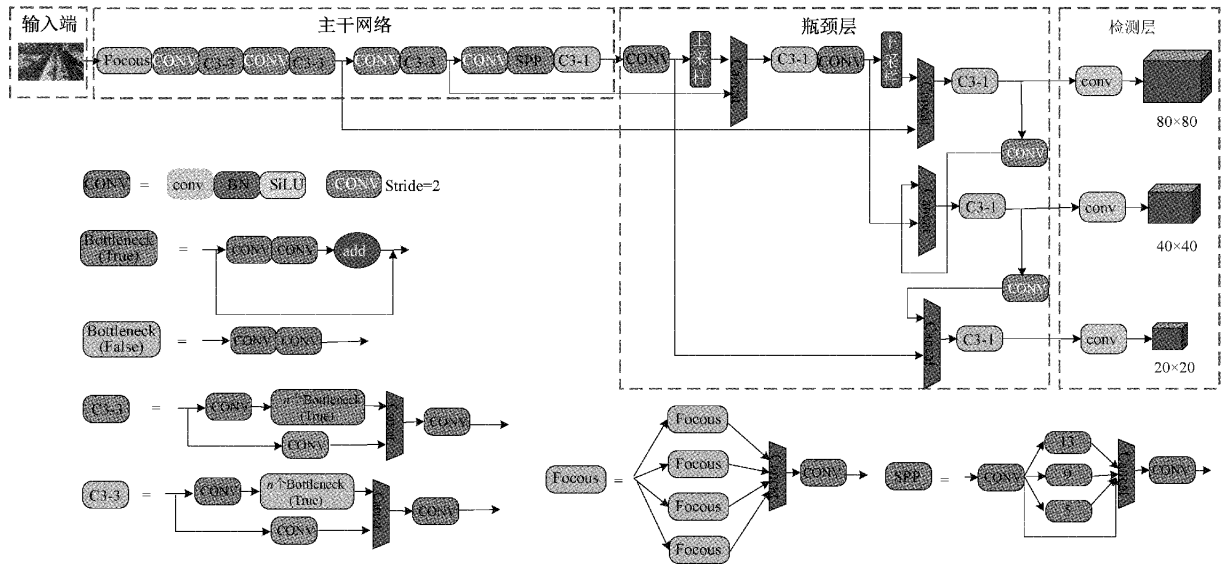


图 1 YOLOv5s 算法结构

YOLOv5 主要由输入端、主干网络、瓶颈层和检测层 4 部分构成。其中输入端针对输入图片进行预处理,包括 Mosaic 数据增强、图片尺寸处理以及自适应锚框计算,将输入图像统一为 $640 \times 640 \times 3$ 的大小;主干网络采用 Focus、C3、SPP^[18] 等网络结构从输入图像中提取特征信息;瓶颈层采用 FPN^[19] + PAN^[20] 双塔结构,将所提取的高层语义特征与底层位置特征进行融合,使模型获取更加丰富的特征信息;检测层使用 3 种不同尺度的特征图分别检测大、中、小不同尺寸的目标对象,生成边界框并预测类别,记录位置信息和分类信息。

2 改进的 YOLOv5 检测算法

2.1 多尺度检测

YOLOv5 原模型利用 3 个检测层进行多尺度检测,当

输入图像为 640×640 时,分别进行 8 倍、16 倍、32 倍下采样,来检测图中大小在 8×8 、 16×16 、 32×32 以上的目标。由于在实际场景下,电力设备遍布在场地的远近高低各个层次、各个角落,航拍图像高空拍摄时映射出的设备尺寸大小也大不相同,所以当目标在原输入图像中像素小于 8×8 时,相关小目标会出现漏检、错检现象,检测将会变得困难。为了应对电力设备复杂的密集场景,提升小目标的识别能力,本文在原始 YOLOv5 网络 3 个检测层的基础上,增加了一个 160×160 的检测尺度,将 YOLOv5 原来的三尺度检测改进为四尺度检测。

改进后的模型如图 2 所示,在原网络的 17 层 C3 模块之后通过增加卷积层和上采样等操作,使得到的特征图继续扩大,并在 20 层经过拼接操作将骨干网络中第 2 层特征图与获取到的 160×160 的特征图进行张量拼接融合,同时

在21层增加一层小目标检测层。改进后的模型拥有四层检测层,增强了在密集目标下多尺度学习的能力,从而更好地学习电力设备小目标的多级特征,提升模型的检测性能。

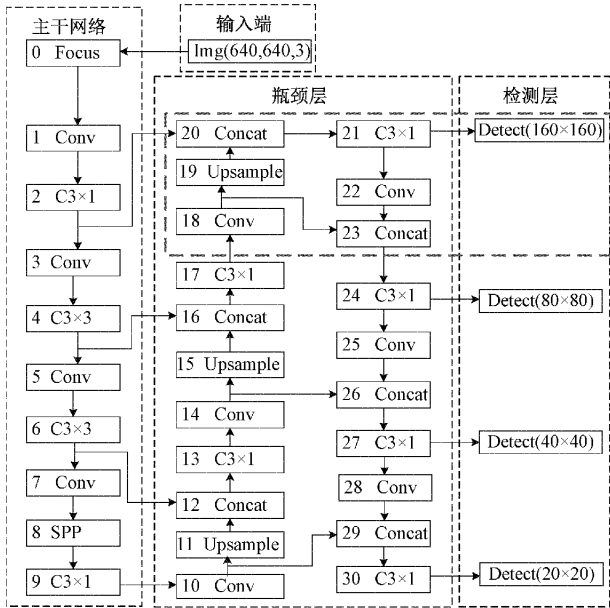


图2 改进的YOLOv5算法结构

2.2 特征融合

通过改进1中将YOLOv5的检测尺度加深为4层,充分利用了网络的特征提取能力,但随着每一层网络的加深都会造成一定的特征损失,影响检测效果,为了进一步利用浅层特征信息,本文对四尺度的YOLOv5进一步改进。

YOLOv5加强特征融合网络采用PANet结构,该结构通过双向融合深层网络的语义信息和浅层网络的位置信息来提升模型性能,但其采用直接拼接不同层级特征的方式,未考虑来自不同路径的特征对网络的重要程度,对有效信息没有特殊关注,本文对原PANet网络结构进行改进,对同一层的节点通过跳跃连接的方式按通道进行拼接,从而实现进一步特征融合。考虑到模型的运行效率,为了在不消耗更多成本的情况下获得更多的特征,仅对两个中间层F2、F3进行跨层加权融合,如图3所示。

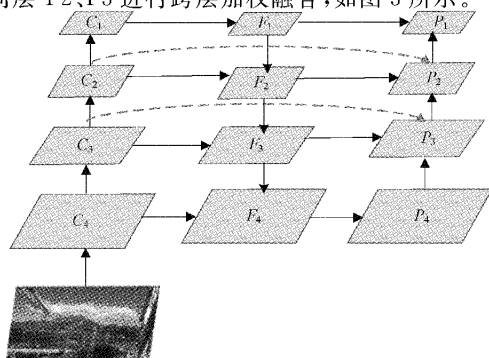


图3 改进后的PANet

具体的融合过程如下:

$$F_i = Conv(Concat(C_i, Resize(F_{i-1}))) \quad (1)$$

$$P_i = Conv\left(\frac{\omega_1 \times C_i + \omega_2 \times F_i + \omega_3 \times Resize(P_{i+1})}{\omega_1 + \omega_2 + \omega_3 + \epsilon}\right) \quad (2)$$

ω_1 是可学习的特征融合因子, ϵ 取 0.0001 保证数值的稳定, *Concat* 表示两个特征图在维度方向上拼接, *Resize* 操作用于调整特征图大小。改进后的模型,通过融合不同尺度的特征可以进一步将低层特征与高层特征高效融合,使预测网络具有强语义信息的同时对细节仍有较强的感知力,对不同尺寸的目标更敏感,从而降低电力设备漏检率与误检率,提高了模型的计算效率。

2.3 协同注意力机制

由于电力设备检测图片背景复杂、目标相对密集,使用YOLOv5模型算法在进行卷积采样时背景的迭代累积会形成大量的冗余信息,造成特征信息丢失,检测准确率不高。为了使模型更好的关注目标信息,在YOLOv5主干网络中引入协同注意力机制(coordinate attention, CA)^[21]模块,将其替换原网络第一个卷积层,通过强化重要信息的特征,抑制无关特征来提高模型在电力设备数据集上的特征提取能力和检测性能。

CA是将位置信息嵌入到通道注意力^[22]的注意力机制,不仅能捕获跨通道的信息,还能捕获方向感知和位置感知信息,使模型更加精准地定位和识别感兴趣的目标,具体实现过程包括坐标信息的嵌入和坐标注意力的生成,如图4所示。

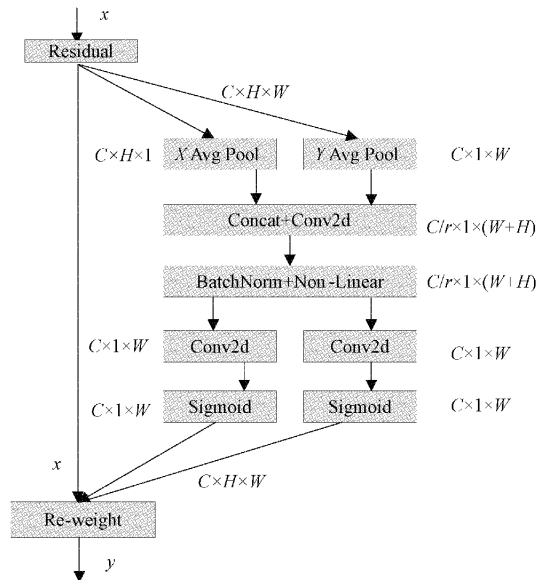


图4 协同注意力机制

1) 坐标信息的嵌入:为了使注意力模块能够捕捉具有精确位置信息的远程空间交互,首先对输入特征图 x 在水平和垂直方向上分别使用 $(H, 1)$ 和 $(1, W)$ 尺寸的池化核进行全局平均池化,来对各通道进行特征提取,如式(3)

和(4)所示。

$$z_c^h(h) = \frac{1}{W} \sum_{0 \leq j < W} x_c(h, j) \quad (3)$$

$$z_c^w(w) = \frac{1}{H} \sum_{0 \leq j < H} x_c(j, w) \quad (4)$$

W, H 分别为输入特征图的宽和高, $z_c^h(h)$ 和 $z_c^w(w)$ 表示特征图 x 第 c 个通道在高 h 和宽 w 处的输出。上述两种变换分别沿两个空间方向聚合特征,生成了一对具有方向感知的特征映射,能让注意力模块捕捉到沿一个空间方向的长期依赖关系,并保存沿另一个空间方向的精确位置信息,有助于网络更加准确地定位感兴趣的目标。

2) 坐标注意力的生成:首先将坐标信息嵌入获得的两个方向的特征图通过 Concat 拼接在一起并使用 1×1 卷积变换函数 F_1 对其进行变换操作,如式(5)所示。

$$f = \delta(F_1([z^h, z^w])) \quad (5)$$

式中: $[\cdot, \cdot]$ 为沿空间维度的拼接操作, δ 为非线性激活函数 sigmoid, f 为具有水平和垂直方向空间信息的中间特征映射。接着,沿空间维度将 f 分解成两个张量 f^h 和 f^w , 并分别利用 1×1 的卷积变换 F_h 和 F_w 使其变换为具有相同通道数的张量,经过 Sigmoid 激活函数后得到特征图在水平和垂直方向上的注意力权重 g^h 和 g^w , 如式(6)和(7)所示。

$$g^h = \sigma(F_h(f^h)) \quad (6)$$

$$g^w = \sigma(F_w(f^w)) \quad (7)$$

最后将得到的注意力权重 g^h 和 g^w 与原始特征图 x 乘法加权计算,将得到在水平和垂直方向上带有注意力权重的特征图,如式(8)所示。

$$y_c(i, j) = x_c(i, j) \times g_c^h(i) \times g_c^w(j) \quad (8)$$

3 实验及结果分析

3.1 数据集

本研究实验数据由云南电网提供,通过设备采集到不同大小、不同光照、不同路况的电力场景图片,背景较为复杂,极大地增强了算法的鲁棒性。经过筛选与整理将得到的 6 688 张图片使用 labelimg 标注软件,采用 PASCAL VOC 格式进行标注,标注目标包含电力塔(tower)、电线杆(pole)、销钉(nail)、绝缘子(insulator)、电力工程车(car) 5 类,标注好的图像标注信息保存为 .xml 形式,包括目标物体的类别和目标框的坐标信息,样本示例如图 5 所示。

对数据集进行可视化分析结果如图 6 所示。其中图 6(a)表示将原图图片分辨率大小正则化以后,目标框的中心点在图中的位置坐标,图 6(b)表示目标框的长宽占图片长宽的比例。从图中可以看出,目标出现位置大致集中于图片的中间偏下位置,目标框大小不均,小目标数量较多。

3.2 实验环境

本文实验环境为 Windows10,64 位操作系 16 GB 运行

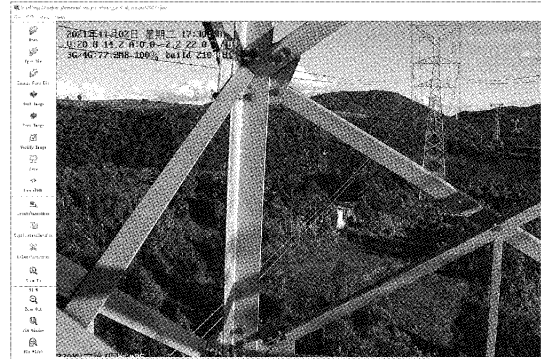


图 5 数据集标注示例

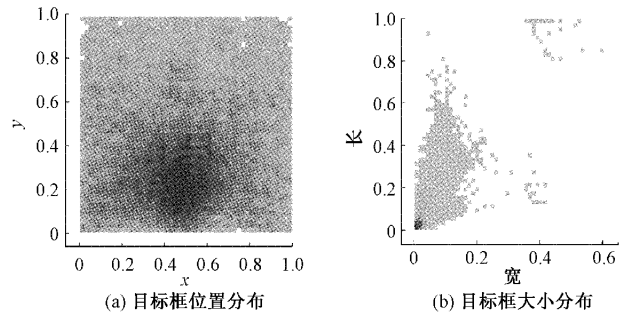


图 6 数据集分析可视化

内存, NVIDIA GeForce RTX 3060 显卡, Intel (R) Core (TM) i5-11600F/11600 处理器,基于深度学习框架 Pytorch 1.7.1, CUDA11.0 下实现模型的搭建及训练工作。

3.3 评价指标

为了验证改进 YOLOv5 算法的性能,本文采用准确率(Precision)、召回率(Recall)、平均精率均值(mean average precision)、模型大小等指标对模型性能进行评估。

其中,准确率表示二分类器预测出来的正例结果中有多少是真实正例,是针对模型预测结果而言;召回率表示测试集中有多少真实正例被二分类器挑选出来,是针对原样本而言的;mAP 指各类别平均精度(average precision, AP)的平均值,根据 IOU 阈值的不同可分为 mAP0.5 和 mAP0.5 : 0.95,值越高则表示性能越好。

3.4 结果与分析

为了进一步分析每个改进点对 YOLOv5s 算法的影响,进行了消融实验,如表 1 所示。该实验通过分别引入小目标检测层、多尺度特征融合和 CA 注意力机制模块进行了对比,每组实验使用相同的超参数以及训练技巧。

实验结果显示使用原始 YOLOv5 算法对电力设备进行目标检测时,最终的 mAP0.5 和 mAP0.5 : 0.95 分别达到 91.5%,61.9%,引入小目标检测层后分别提升了 4.3% 和 4.2%;在 4 个检测层基础上进一步特征融合更多的体现在模型的召回率上,mAP 分别提升了 0.3%;当原模型将三点改进同时引入时效果最好,模型大小微升的情况

表1 消融实验结果

网络模型	模型大小/M	准确率/%	召回率/%	mAP0.5/%	mAP0.5:0.95/%
YOLOv5s	13.8	95.2	87.4	91.5	61.9
YOLOv5s+检测层	14.6	96.3	91.3	95.8	66.1
YOLOv5s+检测层+特征融合	15.0	95.5	92.3	96.1	66.4
YOLOv5s+CA+检测层+特征融合	15.6	96.0	93.4	97.1	69.2

下,精度、召回率、mAP0.5和mAP0.5:0.95分别达到96.0%、93.4%、97.1%和69.2%,大幅度地提升网络性能。为了更加直观地呈现本算法的有效性,选取部分检测图像改进前后结果对比如图7(a)和(b)所示,从图中可看出,本算法对原模型存在的误检、漏检现象有所改善,且对电力设备目标的识别精度有明显提升。

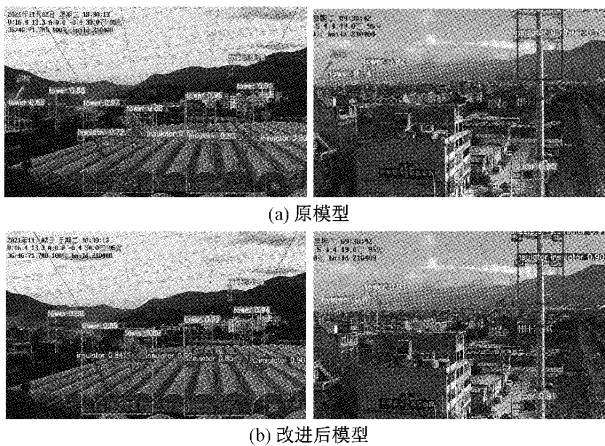


图7 实验结果对比图

为了进一步验证本方法的优越性,将其与目前性能优异的4种目标检测算法Faster R-CNN、SSD、YOLOv3、YOLOv4进行了对比,采用模型大小和mAP0.5两项指标进行评估,如表2所示。从实验结果可看出SSD和YOLOv3在检测效果上明显不占优势,Faster R-CNN和YOLOv4检测结果分别达到83.9%和92.2%,但其模型大小远超YOLOv5,不利于部署。本方法在原模型基础上模型大小有所增加,但仍远小于目前的主流方法,且mAP0.5达到了97.1%,比原模型和Yolov4分别提升了5.6%和4.9%,整体性能占有一定的优势。

表2 不同检测算法性能指标对比

网络模型	主干网络	模型大小/M	mAP0.5/%
Faster R-CNN	ResNet50	315.1	83.9
SSD	VGG16	191.0	77.0
YOLOv3	DarkNet53	235.0	80.3
YOLOv4	CSP DarkNet53	244.2	92.2
YOLOv5s	CSP DarkNet53	13.8	91.5
Our	改进网络	15.3	97.1

4 结 论

针对目前无人机自动巡检电力设备难以实现高精度检测的问题,本文提出一种面向电力设备的改进型目标检测算法。以YOLOv5s检测模型作为研究对象,通过在主干网络重新添加一个检测尺度,使小目标特征更有效的表达,并在此基础上对特征融合模块进行改进,进一步提升网络多尺度学习能力,再通过融入协同注意力机制,强化网络对目标的表征能力,缩减无用特征对模型的干扰和影响。实验结果表明:相较于原始YOLOv5,电力设备目标平均检测精度提升了5.6%,达到了97.1%,有效改善了电力设备目标检测的误检、漏检现象,对实现无人机智能巡检工作具有现实意义。然而,文中所提算法仍存在进步空间,网络模型中使用大量卷积,增加了模型复杂度,后续将通过模型剪枝等压缩方法,进一步降低模型参数量方便在自动化平台部署,更好的去辅助电力设备巡检工作人员。

参考文献

- [1] 臧国强,刘晓莉,徐颖菲,等.深度学习在电力设备缺陷识别中的应用进展[J].电气技术,2022,23(6):1-7.
- [2] 彭继慎,孙礼鑫,王凯,等.基于模型压缩的ED-YOLO电力巡检无人机避障目标检测算法[J].仪器仪表学报,2021,42(10):161-170.
- [3] 李鑫,刘帅男,杨楨,等.基于改进Cascade R-CNN的输电线路多目标检测[J].电子测量与仪器学报,2021,35(10):24-32.
- [4] GIRSHICK R, DONAHUE J, DARRELL T, et al. Rich feature hierarchies for accurate object detection and semantic segmentation[C]. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2014: 580-587.
- [5] GIRSHICK R. Fast R-CNN[C]. Proceedings of the IEEE International Conference on Computer Vision, 2015: 1440-1448.
- [6] REN S Q, HE K M, 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, 39(6): 1137-1149.
- [7] REDMON J, DIVVALA S, GIRSHICK R, et al.

- You only look once; Unified, real-time object detection[C]. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2016; 779-788.
- [8] REDMON J, FARHADI A. YOLO9000: Better, faster, stronger [C]. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2017; 7263-7271.
- [9] REDMON J, FARHADI A. Yolov3: An incremental improvement [J]. ArXiv Preprint, 2018, ArXiv: 1804.02767.
- [10] BOCHKOVSKIY A, WANG C Y, LIAO H Y M. Yolov4: Optimal speed and accuracy of object detection [J]. ArXiv Preprint, 2020, ArXiv: 2004.10934.
- [11] LIU W, ANGUELOV D, ERHAN D, et al. SSD: Single shot multibox detector [C]. European Conference on Computer Vision. Springer, Cham, 2016; 21-37.
- [12] LIN T Y, GOYAL P, GIRSHICK R, et al. Focal loss for dense object detection [C]. Proceedings of the IEEE International Conference on Computer Vision, 2017; 2980-2988.
- [13] 汤踊, 韩军, 魏文力, 等. 深度学习在输电线路中部件识别与缺陷检测的研究[J]. 电子测量技术, 2018, 41(6): 60-65.
- [14] ZHENG H B, SUN Y H, LIU X H, et al. Infrared image detection of substation insulators using an improved fusion single shot multibox detector [J]. IEEE Transactions on Power Delivery, 2020, 36(6): 3351-3359.
- [15] LIU Y P, JI X X, PEI S T, et al. Research on automatic location and recognition of insulators in substation based on YOLOv3 [J]. High Voltage, 2020, 5(1): 62-68.
- [16] 石欣, 卢灏, 秦鹏杰, 等. 一种远距离行人小目标检测方法[J]. 仪器仪表学报, 2022, 43(5): 136-146.
- [17] 解尧婷, 张丕状. 基于改进的 YOLOv4 输电线路小目标检测[J]. 国外电子测量技术, 2021, 40(2): 47-51.
- [18] HE K M, ZHANG X Y, REN S Q, 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.
- [19] LIN T Y, DOLLAR P, GIRSHICK R, et al. Feature pyramid networks for object detection [C]. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2017; 2117-2125.
- [20] LIU S, QI L, QIN H F, et al. Path aggregation network for instance segmentation[C]. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2018; 8759-8768.
- [21] HOU Q B, ZHOU D Q, FENG J S. Coordinate attention for efficient mobile network design [C]. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2021; 13713-13722.
- [22] TANG X L, ZHONG B, PENG J P, et al. Multi-scale channel importance sorting and spatial attention mechanism for retinal vessels segmentation [J]. Applied Soft Computing, 2020, 93(1): 106353.

作者简介

郑婷婷, 硕士研究生, 主要研究方向为目标检测与识别。

E-mail: 958052264@qq.com

周浩(通信作者), 硕士研究生导师, 主要研究方向为数字图像处理, 计算机视觉和智能视频监控。

E-mail: zhouhao@ynu.edu.cn

王秋忆, 硕士研究生, 主要研究方向为目标检测与识别。