

一种深度学习改进算法的图像处理技术研究

常春雷¹, 运凯¹, 王辉¹, 张辉², 徐赢³, 马崇瑞²

(1. 国网新疆电力有限公司信息通信公司, 乌鲁木齐 830002; 2. 北京中电普华信息技术有限公司;
3. 国网辽宁省电力有限公司大连供电公司)

摘要: 针对 YOLOv5 算法对目标物体及局部特征定位不准确的问题, 提出了一个基于 YOLOv5 的 YOLOv5-ABN 目标检测网络。首先为了得到准确的定位信息, 本文将卷积注意力机制中的空间注意力机制与坐标注意力机制相结合, 提出了一个融合空间特征信息与坐标信息的空间-坐标注意力机制, 在 YOLOv5 的 Backbone 网络中加入空间-坐标注意力模块, 在 Neck 检测层引入加权双向特征金字塔网络, 并对锚框参数进行优化, 提高了网络对样本模型及局部特征的检测精度及定位能力, 分别用改进前后的算法对自定义数据集和康奈尔数据集进行训练。实验结果表明, 改进后的算法性能要优于 YOLOv5 算法, 在识别精度没有降低的前提下, 对目标物体的定位更加准确, 鲁棒性更好。

关键词: YOLOv5 算法; 目标检测; 注意力机制; 加权双向金字塔网络

中图分类号: TQ341

文献标识码: A

Research on Deep Learning Improved Algorithm in Image Processing Technology

Chang Chunlei¹, Yun Kai¹, Wang Hui¹, Zhang Hui², Xu Ying³, Ma Chongrui²

(1. State Grid Xinjiang Information and Telecommunication Company, Wulumuqi 830002, China;
2. Beijing China-power Information Technology Co., Ltd.; 3. State Grid Dalian Electric Power Supply Company)

Abstract: In order to solve the problem that YOLOv5 algorithm is not accurate in locating the target object and local features, a YOLOv5-ABN target detection network based on YOLOv5 is proposed. Firstly, in order to get accurate positioning information, this paper combines spatial attention mechanism and coordinate attention module in convolution attention mechanism. A spatial and coordinated attention module is proposed, which integrates spatial feature information and coordinate information. The spatial and coordinated attention module is added to the Backbone network of YOLOv5. The weighted bidirectional feature pyramid network is introduced into the Neck detection layer, and the staggered box parameters are optimized, which improves the detection accuracy and positioning ability of the network for sample models and local features. The customized data set and Cornell data set are trained with the improved algorithms respectively. The experiment results show that the improved algorithm is better than YOLOv5 algorithm, and the target object is more accurate and more robust without reducing the recognition accuracy.

Keywords: YOLOv5 algorithm; target detection; attention mechanism; weighted bidirectional pyramid network

0 引言

随着经济的快速发展, 我国对工业化、自动化以及智能制造的技术水平提出了更高要求。智能制造旨在将工业化和自动化技术相互融合, 这也是未来智能制造业发展与变革的重要方向。机器视觉^[1-2]是移动机器人平台十分关键的技术, 代表机器人智能化、自动化及先进性的判定标准。检测物体及其 6D 位姿(3D 位置和方向)是许多机器人应用系统中的重要任务, 包括物体的拾取操作、工厂零件装配等。在复杂环境中, 机器视觉一般包括两个阶段: 第一阶段是目标检测, 在 RGB 图片中使用目标检测算法^[3]或者分割网络获取目标类别; 第二阶段是 6D 位姿估

计, 对检测出来的物体进行 6D 位姿估计。因此, 对目标检测算法进行改进是非常有必要的。

1 YOLOv5 改进算法设计

1.1 整体框架

YOLOv5 目标检测算法^[4]检测速度快、检测精度高, 在开源数据集上具有比较好的检测性能, 但是在对自定义数据集样本模型及局部特征准确识别和准确定位方面性能还有待加强。首先为了得到精确的定位信息, 本项目将空间注意力机制(SA)和坐标注意力机制(CA)结合, 提出了一个空间-坐标注意力机制(SCAA), 然后在 YOLOv5 主干网络中加入 SCAA 模块, 在 Neck 检测层中引入 BiF-

PN 模块,并根据自定义数据集的需要优化锚框参数,提出了一个基于 YOLOv5 的 YOLOv5-ABN 目标检测算法,改进后的 YOLOv5-ABN 算法结构如图 1 所示。

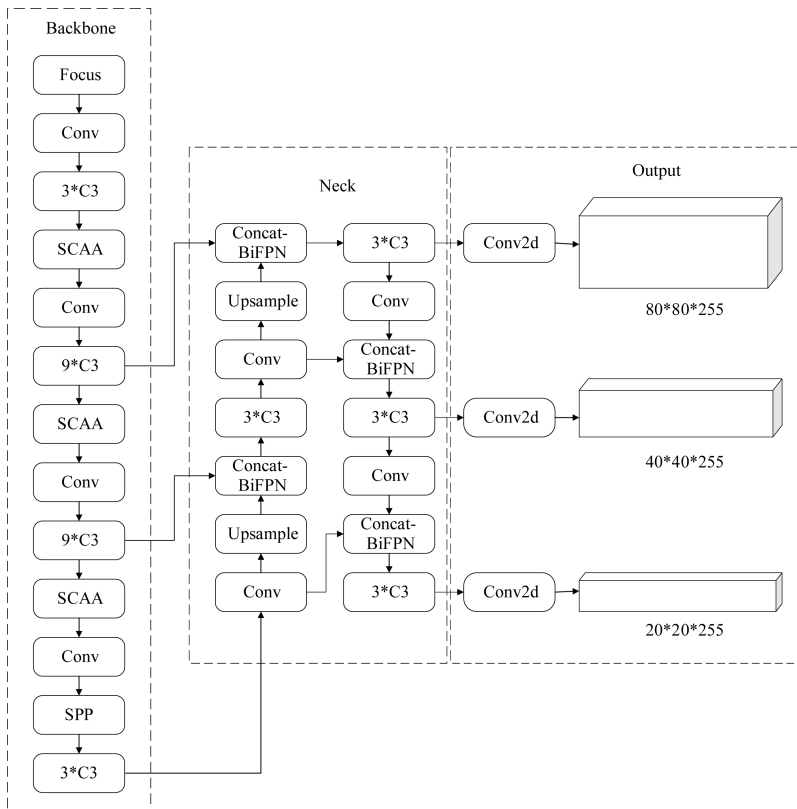


图 1 YOLOv5 改进算法整体框架

1.2 空间-坐标注意力机制

为了实现对目标物体及局部特征信息的准确定位,本项目受卷积注意力机制(CBAM)启发^[5],将空间注意力机制(SA)和坐标注意力机制(CA)结合,提出了一种空间-坐标注意力机制(SCAA),结构如图 2 所示。首先特征信息通过空间注意力机制后,提取两通道的最大值与平均值并进行加权融合,输出带有空间位置信息的特征信息;然后将带有空间位置信息的特征通过坐标注意力机制分别沿水平方向和垂直方向进行聚合加权,形成两个独立的特

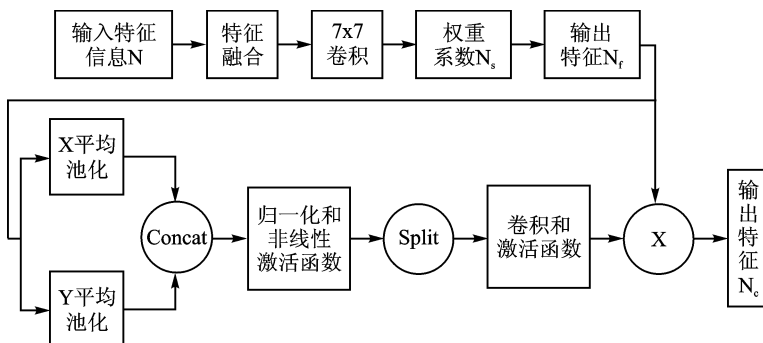


图 2 SCAA 网络结构

征图;最后存储带有方向编码信息的特征。这样对特征信息在空间维度以及坐标维度的双重注意将会使定位更加准确。

SCAA 模块具体的工作步骤如图 2 所示,首先假设输入的特征信息为 N ,在通道维度层面对输入特征进行压缩,然后对压缩以后的特征进行最大池化和平均池化操作,即分别在两个通道提取最大值和平均值,最后将两个通道的特征图进行融合,经过一个 7×7 的卷积层得到权重系数 N_s ,输出特征 N_f 是输入特征 N 与 N_s 的乘积,如式(1)和式(2)所示:

$$N_s = \partial(f^{7 \times 7}([\text{AvgPool}(N); \text{MaxPool}(N)])) \\ = \partial(f^{7 \times 7}([N_s^{\text{avg}}; N_s^{\text{max}}])) \quad (1)$$

$$N_f = N \times N_s \quad (2)$$

整个空间范围高为 H ,宽为 w ,在 $(H,1)$ 的垂直范围内,用池化核对垂直方向的坐标信息编码并保存;在 $(1,w)$ 的水平范围内,用池化核对水平方向的坐标信息进行编码并保存,如下所示:

$$D_c^h(h) = \frac{1}{w} \sum_{0 \leq i < w} N_f(h,i), D_c^w(w) = \frac{1}{H} \sum_{0 \leq j < H} N_f(j,w) \quad (3)$$

将具有精确编码信息的特征进行拼接,然后经过 1×1 卷积变换函数 F_1 和非线性激活函数 ∂ 得到编码空间信息的中间特征 f ,如下所示:

$$f = \partial(F_1([D^h, D^w])), f \in \mathbb{R}^{C \times (H+W)} \quad (4)$$

然后用 1×1 卷积 F_h 和 F_w 生成注意力权重 m^h 和 m^w ,如下所示:

$$m^h = \partial(F_h(f^h)), m^w = \partial(F_w(f^w)) \quad (5)$$

最终通过式(6)计算注意力模块的输出:

$$N_c(i,j) = N_f(i,j) \times m_c^h(i) \times m_c^w(j) \quad (6)$$

为了突出目标的特征信息、对标定的特征信息更精确地识别和定位,本文在主干网络 C3 模块后加入 SCAA 模块,在经过 C3 模块获得目标深层次的特征信息后,再经过 SCAA 模块,先得到空间维度的权重,然后对空间权重大的特征信息在水平方向和垂直方向进行编码,对目标模型特征更好地识别和定位,在 YOLOv5 算法的主干网络中加入 SCAA 模块,Backbone 网络由 10 层变成了 13 层,新的网络结构如表 1 所列,其中改进模块用黑体表示。

1.3 多尺度特征融合网络

YOLOv5 的 Neck 结构采用 FPN+PANet 相结合的方式^[6],根据 BiFPN 网络特征融合方式的优越性,本项目将借鉴 BiFPN 结构的特点对 YOLOv5 算法模型网络进行改进,改进后的模型



结构如图 3 所示。

表 1 改进后的 Backbone 网络结构

Module	Number	Filters	Arguments	Output
Focus	1	32	[3,32,3]	320×320
Conv	1	64	[32,64,3,2]	160×160
C3	3	64	[64,64,1]	160×160
SCAA	3	64	[64,64]	160×160
Conv	1	128	[64,128,3,2]	80×80
C3	9	128	[128,128,3]	80×80
SCAA	3	128	[128,128]	80×80
Conv	1	256	[128,256,3,2]	40×40
C3	9	256	[256,256,3]	40×40
SCAA	3	256	[256,256]	40×40
Conv	1	512	[256,512,3,2]	20×20
SPP	1	512	[512,512,[5,9,13]]	20×20
C3	3	512	[512,512,1,False]	20×20

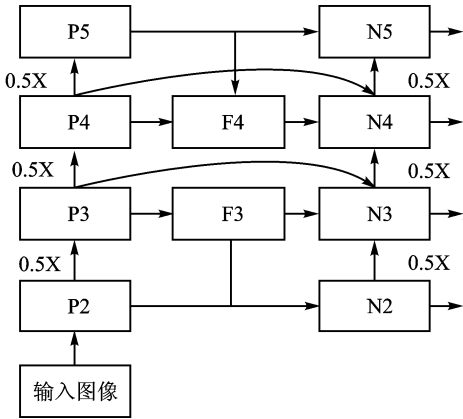


图 3 改进后的特征融合网络

下面将说明新型特征融合网络的工作方式,特征融合的方式如下所示:

$$\text{feature}=[f_1;f_2;f_3] \quad (7)$$

式中, f_1 、 f_2 、 f_3 表示双向特征融合网络的 3 个特征, $[:,:]$ 表示将两个特征按照通道维度数进行两两拼接,feature 表示特征按照通道维度拼接后生成的新特征。在图 3 中, P_i 表示前馈深度神经网络提取的多尺度融合特征。

在原始的融合网络中, N_3 、 N_4 、 N_5 分别代表小目标、中目标、大目标的多尺度特征检测,特征 N_3 并没有利用低层级特征 N_2 ,而是利用了经过上采样后的特征 F_3 ,考虑到需要对样本模型的局部小特征进行检测,为了提升局部特征检测的精确度,本文考虑充分利用低层级特征 P_2 ,特征 N_2 通过特征 P_2 和特征 F_3 融合拼接得到,将 YOLOv5 结构中的 CSP 算子记作函数 F ,如下所示:

$$N_2=F[P_2;F_3] \quad (8)$$

$N_3 \sim N_5$ 是通过新型的双向特征融合网络后生成的多

尺度融合特征,以特征 N_3 生成为例,特征 P_2 通过下采样生成特征 P_3 ,特征 P_4 与特征 P_5 按通道拼接融合得到特征 F_4 ,特征 F_4 通过上采样并与特征 P_3 按通道维度拼接生成特征 F_3 ,多尺度融合特征 N_3 则是特征 N_2 、特征 F_3 和特征 P_3 按照通道维度拼接后生成的,Downsample 是指通过卷积 2 倍下采样特征 N_3 。融合方式如下所示:

$$N_3=F[\text{Downsample}(N_2);F_3;P_3] \quad (9)$$

同理,特征 N_4 、 N_5 融合方式如式(10)和式(11)所示:

$$N_4=F[\text{Downsample}(N_3);F_4;P_4] \quad (10)$$

$$N_5=F[\text{Downsample}(N_4);P_5] \quad (11)$$

这样设计直接将底层特征信息与高级语义信息融合送入损失函数中,既包含底层的定位信息,也包含了高层语义信息,有助于检测局部小特征,性能更加优越,用改进后的新型双向特征融合网络对自定义模型数据集进行检测,相比于 YOLOv5,改进后的算法对样本模型检测与定位的准确率均有明显提高。

1.4 锚参数优化

YOLOv5 采用自适应锚定框 (Auto Learning Bounding Box Anchors) 对训练数据自动学习^[7],通过使用 Kmeans 聚类算法对数据集的目标框 (ground truth) 自动计算,以便获取合适的锚定框,所以需要对锚框参数进行优化,优化后的锚定框尺寸如表 2 所列。

表 2 锚定框尺度分配表

特征图	20×20	40×40	80×80
感受野	大	中	小
锚定框	(174,185)	(30,61)	(14,15)
	(190,218)	(62,45)	(18,17)
	(224,224)	(59,119)	(25,25)

2 实验环境及深度学习参数设置

2.1 实验环境

本实验所用的处理器为 Intel Core i5-10200H CPU @2.40 GHz、32 GB 运行内存;显卡为 NVIDIA GeForce RTX2080Ti GPU;操作系统为 Windows 10,64 位;所用的实验软件是 PyCharm,整个实验基于深度学习框架 PyTorch1.9.0;实验环境是 Python3.8;GPU 加速软件是 Cuda 11.1 和 cuDNN8.1.1;实验所用的深度相机是 Intel D435i 双目相机。

2.2 数据集介绍

网络训练采用 Cornell 大学的机器人实验室开源数据集,一共 885 幅图像,都是单目标图像,另外再在数据集中添加 97 幅多目标的图像。原有图像尺寸是 640×480。

2.3 评价指标

为了验证算法改进的性能,本项目使用召回率 R(Recall)、精确度 P(Precision)、平均精度 AP(Average Precision)、平均精度均值 mAP@0.5(IOU=0.5)和 mAP@0.5:095 等评价指标,表 3 是算法的评估指标变量含义。

表 3 评估指标变量含义

评估指标	Positive(预测正确)	Negative(预测错误)
True(正样本)	TP	TN
False(反样本)	FP	FN

召回率含义为在实际为正的样本集中预测正确的概率,用式(12)表示,精确度含义为在预测正确的样本集中实际为正的样本,用式(13)表示,平均精度 AP 是 P-R 曲线中曲线与坐标轴围成的面积,mAP 是各类样本平均精度的均值,用式(14)和式(15)表示(mAP@0.5 是当 IOU=0.5 时的平均精度值):

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

$$\text{Precision} = \frac{TP}{TP + FP} \quad (13)$$

$$\text{AP} = \int_0^1 P(R) \quad (14)$$

$$\text{mAP} = \sum_{j=1}^c \text{AP}_j / C \quad (15)$$

2.4 模型参数设置

使用 Labellmg 标注工具进行人工手动标注位置及类别,生成的 XML 文件存储了样本数据的矩形位置信息和类别信息,将 xml 文件转换为 YOLO 测得的 txt 文件,按照 9:1 的比例分为训练集和测试集,具体设计参数如表 4 所列。

表 4 训练模型参数表

模型参数	数值设定
初始学习率	1e-4
批次大小	16
迭代次数	300
输入尺寸	416×416

3 实验结果与分析

通过 YOLOv5 算法和改进后的 YOLOv5 算法训练生成的两种检测模型分别对物体进行检测实验,对比两种算法生成的检测模型的识别精度,使用 YOLOv5 检测模型, mAP 为 92.4%,而使用改进后的算法, mAP 达到 96.6%,提升了 4.2%。另外在测试图像上,改进后的检测模型时间消耗为每帧 0.031 s,在单个 NVIDIA RTX 2080 T 显卡上约为每秒 32.26(FPS),满足检测的实时性要求。改进前 mAP 准确率和改进后的 mAP 准确率如图 4 和图 5 所示。

4 结 语

本文将 CBAM 模块中的空间注意力与坐标注意力结

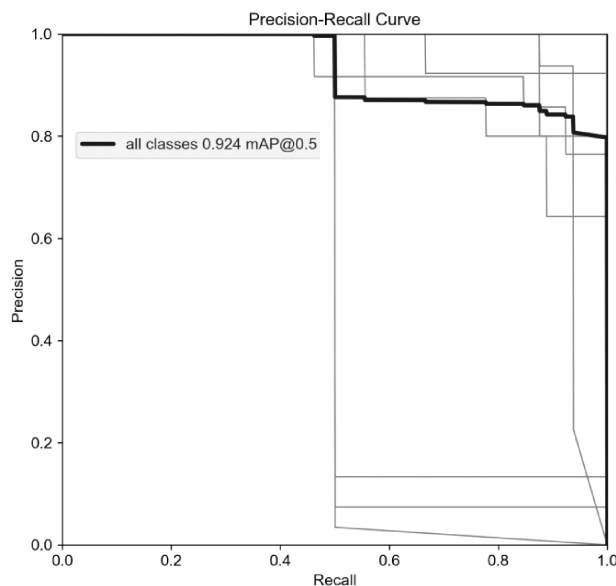


图 4 改进前 P-R 曲线

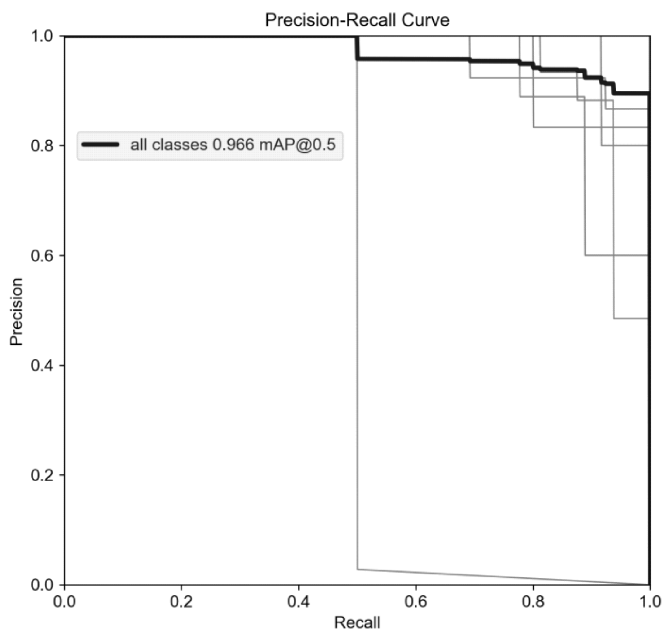


图 5 改进后 P-R 曲线

合,提出了一种空间-坐标注意力机制,在 YOLOv5 主网络中加入 SCAA 模块,在 Neck 检测层中引入 BiFPN 模块,并根据自定义数据集的需要对锚框参数进行优化,提出了一个基于 YOLOv5 的 YOLOv5-AB 目标检测算法,用改进后的 YOLOv5-ABN 网络对自定义数据集和康奈尔数据集进行训练,并对改进模块做了消融实验。与现有的目标检测算法在性能上进行对比,实验结果表明,改进后的算法无论是在自定义数据集中训练还是在公共数据集中训练,整体的算法性能都要优于 YOLOv5 算法,对目标物体的识别与定位效果更好,鲁棒性更强。

馈线上可以部署的分布式发电的最大量通常受到潜在的电压下降的限制,若沿着馈线的分布式能源资源输出突然下降,则会导致电压下降。此外,实现 TSO-DSO 无功功率管理为未来电力系统的主要挑战之一,因为连接到输电网络的无功功率资源被 DSO 中的分布式发电、负载和储能元件所取代,这需要更多的 TSO 和 DSO 之间的积极操作协调^[8]。因此本文选择了协调的 TSO-DSO 无功功率管理用例作为研究对象,因为它不仅是最有前途的支持因素,而且还是一个高度依赖通信的智能电网应用程序。

电压控制方案被开发成智能电网的一个用例,强调了安全通信的要求。此外,研究确定了适用于 IoT 启用的智能电网用例的 5G 安全规范特征,并讨论了它们实现安全要求的可行性。本文使用 STRIDE 威胁建模方法创建了智能电网用例的威胁模型,揭示了重要的安全挑战,其中一些可以通过 5G 网络来解决。表 1 列出了 5G 如何满足提出的安全要求并应对已识别的威胁。

表 1 利用 5G 应对威胁

安全要求	威胁	5G 贡献
认证方式	欺骗	防止虚假基站, eSIM 支持安全的设备身份验证
机密性	信息泄露	eSIM 密钥交换实现加密通信
可用性	拒绝服务	5G 拆分网络增强了单个基站的弹性, 物联网网络切片可以与其他切片隔离, 5G 支持回退到传统网络
保密性	信息泄露	eSIM 密钥交换实现加密通信
完整性	篡改	eSIM 密钥交换启用消息身份验证码和经过身份验证的加密

6 结 语

使用物联网设备承诺数字化转型到多个领域,包括当前的电力网络基础设施,然而在智能电网基础设施中,物联网安全仍然面临挑战,部分原因是考虑到成本和性能之间的权衡。本文确定了低压调节智能电网用例中的物联网通信要求,并开发了相关的威胁模型,研究了 5G 规范

并讨论了其在选择的智能电网用例中的安全应用。虽然 5G 网络不能解决所有物联网安全问题,但它们提供了一种可靠的替代通信渠道,具有内置的安全优势,可用于智能电网基础设施。

就科学方面的影响而言,本文总结了使用 5G 进行物联网网络的智能电网安全的现状,记录了一个可用于评估未来安全机制的特定用例,确定了进一步工作的几个领域,尤其是需要对规模部署的 5G 物联网进行实验室研究。对于从业人员,本文提供了威胁模型(DSO 可以利用它来更新它们的变电站通信策略),关于 5G 在安全方面为 DSO 提供的信息,以及需要考虑的限制。

参考文献

- [1] 肖安南,张蔚翔,赵德盛,等. 基于物联网的智能电网监控系统构建方法研究[J]. 信息技术,2020,44(12):86-90,95.
- [2] 彭姣,刘明硕,杨力平面向智能电网的大数据降维管理方案[J]. 计算技术与自动化,2019,38(4):139-143.
- [3] 何立民. 从智能电网、物联网到泛在电力物联网[J]. 单片机与嵌入式系统应用,2022,22(4):3-5,10.
- [4] 瞿水华,郭剑杰,许乐飞. 5G 技术在智能电网中的实践和验证[J]. 供用电,2021,38(5):2-9,34.
- [5] Graditi G, Ciavarella R, Di Somma M, et al. Critical assessment of potential barriers for integrating the emerging technologies in the pan-european network considering the future eu grid scenarios[C]//2019 1st International Conference on Energy Transition in the Mediterranean Area (SyNERGY MED). IEEE,2019:1-6.
- [6] 王旭蕊,孟显,季名扬,等. 序列号同步下配电网 5G 通信的传输[J]. 单片机与嵌入式系统应用,2022,22(10):25-28.
- [7] D'ADAMO C, ABBEY C, BAITCH A, et al. Development and operation of active distribution networks[J]. IRIS,2011(1).
- [8] SUN H, GUO Q, QI J, et al. Review of challenges and research opportunities for voltage control in smart grids[J]. IEEE Transactions on Power Systems,2019,34(4):2790-2801.

陈旗展(高级工程师),主要研究方向为电力系统、人工智能。通信作者:陈旗展,pulouyue621021@163.com。

(责任编辑:薛士然 收稿日期:2023-05-26)

参考文献

- [1] 蒋超,张豪章,恩泽惠,等. 基于改进 YOLOv5s 的行人车辆目标检测算法[J]. 扬州大学学报:自然科学版,2022,25(6):45-49.
- [2] 赵元章,耿生玲. 基于改进 YOLOv5 的人脸遮挡物目标检测算法[J]. 信息通信,2021,34(11):32-35.
- [3] 张德春,李海涛,李勋,等. 基于 CBAM 和 BiFPN 改进 YoloV5 的渔船目标检测[J]. 渔业现代化,2022(3):49.
- [4] 孙丰刚,王云露,兰鹏,等. 基于改进 YOLOv5s 和迁移学习的苹果果实病害识别方法[J]. 农业工程学报,2022(11):38.
- [5] 吕晓玲,杨胜月,张明路,等. 改进 YOLOv5 网络的鱼眼图像

目标检测算法[J]. 计算机工程与应用,2023,59(6):10.

- [6] 向浩,田地. 微型计算机与分析仪器的通信技术研究[J]. 现代科学仪器,2020(6):51-55.
- [7] Xia X, Zhai J. Study on Small Target Detection of Construction Site Helmet Based on YOLOv5[C]//Chinese Intelligent Systems Conference. Springer, Singapore, 2022.

常春雷(高级工程师),主要研究方向为电力信息化、云计算技术。通信作者:常春雷,goufu43396953941@126.com。

(责任编辑:薛士然 收稿日期:2023-05-29)