

基于改进YOLOv8的工程场景目标检测方法

丁张杰¹, 黄剑^{2,3}, 谢林袁¹, 卞永明¹

(1. 同济大学机械与能源工程学院, 上海 201804; 2. 中交第二航务工程局有限公司, 湖北 武汉 430040;
3. 中交公路长大桥梁建设国家工程研究中心有限公司, 北京 100120)

摘要: 随着目标检测技术在无人化施工等场景中的广泛应用, 传统目标检测算法在复杂工程环境中面临识别精度低、计算量大、处理速度慢等挑战。基于这些挑战, 本文提出了一种基于改进YOLOv8的工程场景目标检测方法, 通过在YOLOv8中引入星乘模块改进C2f结构, 在保证检测精度的同时显著减少模型参数数量和计算量。以此为基础, 本文引入了轻量化共享与细节增强卷积检测头, 进一步提高检测头的细节捕获能力, 并显著降低计算负担。实验结果显示: 在基于Roboflow的工程场景数据集上, 与YOLOv8n相比, 改进后的模型在mAP@0.5和mAP@0.5:0.95上分别提高了0.3%和2.0%, 参数和计算量分别减少了36.7%和34.6%, 帧率(FPS)则相应提高了23.3%, 验证了提出的改进算法在轻量化和检测精度方面的优越性。

关键词: 目标检测; YOLOv8; 工程场景; 模型轻量化

中图分类号: TP 391.41 **文献标志码:** A **文章编号:** 1672-5581(2025)02-0366-05

Engineering scene target detection method based on improved YOLOv8

DING Zhangjie¹, HUANG Jian^{2,3}, XIE Linyuan¹, BIAN Yongming¹

(1. School of Mechanical Engineering, Tongji University, Shanghai 201804, China; 2. CCCC Second Harbor Engineering Company Ltd., Wuhan 430040, Hubei, China; 3. CCCC Highway Bridge National Engineering Research Centre Ltd., Beijing 100120, China)

Abstract: With the wide application of target detection technology in unmanned construction and other scenarios, traditional target detection algorithms face challenges such as low recognition accuracy, large computational volume, and slow processing speed in complex engineering environments. Based on these challenges, this paper proposes a target detection method based on improved YOLOv8 for engineering scenes, which improves the C2f structure by introducing the star block in YOLOv8, significantly reducing the number of model parameters and computational volume while ensuring the detection accuracy. Based on this, this paper introduces a lightweight shared and detail-enhanced convolutional detection head, which further improves the detail-capturing ability of the detection head and significantly reduces the computational burden. The experimental results show that compared with YOLOv8n on Roboflow-based engineering scene dataset, the mAP@0.5 and mAP@0.5:0.95 of the improved model improves by 0.3% and 2.0%, while the number of parameters and computational volume decreases by 36.7% and 34.6%, and improves the frames per second (FPS) by 23.3% accordingly, which verifies the superiority of the improved algorithm in terms of lightweight and detection accuracy.

基金项目: 国家重点研发计划资助项目(2023YFF0613200)

作者简介: 丁张杰(2000—), 男, 硕士生。E-mail: 2232651@tongji.edu.cn

通信作者: 卞永明(1966—), 男, 教授, 博士。E-mail: ymbianmail@163.com

Key words: target detection; YOLOv8; engineering scene; model lightweight

在工程机械无人化施工等场景中,目标检测技术应用日益广泛。目标检测算法不仅需在复杂环境中保持高精度,还必须满足实时性要求,例如无人挖掘机的周围车辆、障碍物及作业人员的识别。这要求算法在有限硬件资源下快速处理大量图像数据并及时响应,以确保高效、准确的识别^[1-2]。

针对视觉Transformer中频繁的元素加法和归一化操作导致内存访问开销大的问题,Liu等^[3]优化了视觉Transformer的推理效率,以提升模型的内存效率和性能。然而,EfficientViT依然计算量大,对边缘设备的要求较高。针对快速神经网络中浮点运算减少未必有效降低延迟的问题,Chen等^[4]提出了PConv,减少冗余计算和内存访问,从而更高效地提取特征。然而,FasterNet的参数量依然较大,对于常规工控机的计算负担仍然较重。

近年来,YOLO系列算法以其优越的实时性和检测精度,逐渐成为工程场景目标检测的主流选择^[5]。YOLOv8作为该系列的较新版本,进一步提升了检测性能。然而,尽管YOLOv8在速度和精度上取得了显著进步,其最轻量模型YOLOv8n仍然较为复杂,在某些计算资源和设备受限的无人工程机械中,仍有进一步优化的空间^[6-7]。

针对工程场景中常见的工程机械(如推土机、自动倾卸卡车、挖掘机、平路机、装载机、搅拌车、吊车、滚压机)、集装箱及工人等检测对象,本文提出了一种基于星乘模块和轻量化共享和细节增强卷积检测头的改进YOLOv8n。通过对YOLOv8n模型的轻量化改进,该方法在显著降低参数量和计算量的同时保证较高的检测精度,实现了不同视角和距离下的目标检测。基于Roboflow^[8]的工程场景数据集(共11类),对所提方法在精度、召回率、参数量、计算量等指标上进行了评估,验证了其有效性。

1 改进模块简介

1.1 星乘模块

StarNet由Ma等^[9]于2024年提出。StarNet中运用元素相乘来融合不同的子空间特征,称为“星形运算”。星形运算具有将输入映射到超高维、非

线性特征空间的能力。StarNet中的星乘模块(star block)正是利用星形运算的高效性,精度上超过了MobileNetv3^[10]、FasterNet^[4]等模型。本文使用Star Block对YOLOv8中的C2f进行改进,将加和操作改为星乘,主要出于轻量化的考虑,并且模型的精度也能得到保证。

星乘模块(Star Block)的结构如图1所示。其中,DW-Conv为深度卷积层,FC为全连接层。

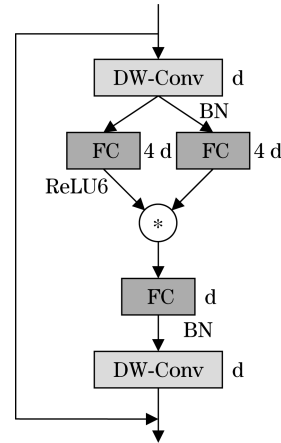


图1 星乘模块结构

Fig.1 Diagram of star block structure

星乘(*)的计算过程如下:

$$w_1^T x * w_2^T x = \left(\sum_{i=1}^{d+1} w_1^i x^i \right) * \left(\sum_{j=1}^{d+1} w_2^j x^j \right) = \sum_{i=1}^{d+1} \sum_{j=1}^{d+1} w_1^i w_2^j x^i x^j = \underbrace{\alpha_{(1,1)} x^1 x^1 + \dots + \alpha_{(4,5)} x^4 x^5 + \dots + \alpha_{(d+1,d+1)} x^{d+1} x^{d+1}}_{(d+2)(d+1)/2 \text{ items}} \quad (1)$$

式中: $w_1, w_2, x \in \mathbb{R}^{(d+1) \times 1}$; d 输入通道数; i, j 为通道的索引; α 为每项的系数:

$$\alpha_{(i,j)} = \begin{cases} w_1^i w_2^j, & i=j \\ w_1^i w_2^j + w_1^j w_2^i, & i \neq j \end{cases} \quad (2)$$

1.2 轻量化共享卷积检测头

分组归一化(group normalization, GN)是针对批归一化算法对批次大小依赖性强这一弱点而提出的改进算法,且GN在FCOS^[11]中证实可以有效提升检测头定位和分类的性能。

本文使用共享卷积来改进YOLOv8中的检测头,可显著减少参数量和计算量,便于在硬件资源受限的无人化设备上应用。在使用共享卷积的是同时,为了对应YOLOv8三个检测头所检测目标尺度不一致的问题,使用Scale层对特征进行缩放以

适应不同尺寸的目标,实现对不同视角和距离下的常见工程机械以及集装箱、安全帽和工人的检测。

轻量化共享卷积检测头(LSCD)的结构如图2所示,其中P3、P4、P5分别表示8倍、16倍和32倍下采样卷积层,Conv_GN 1×1为归一化层,Conv_GN 3×3为共享参数卷积层,Conv_Reg和Con_Cls为共享参数的识别和分类的卷积层,Scale为缩放层。

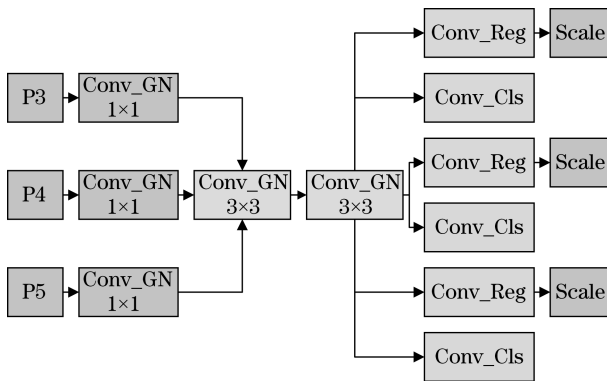


图2 LSCD结构

Fig.2 Diagram of LSCD structure

1.3 细节增强卷积

DEA-Net由Chen等^[12]于2024年提出,其设计了一个细节增强卷积(DEConv),DEConv将先验信息整合到普通卷积层,以增强表征和泛化能力。然后,通过使用重参数化技术,DEConv等效地转换为普通卷积,不需要额外的参数和计算成本。DEConv的结构如图3所示,其中,VC为原始卷积,CDC为中心差分卷积,ADC为角度差分卷积,VDC垂直差分卷积,HDC为水平差分卷积。

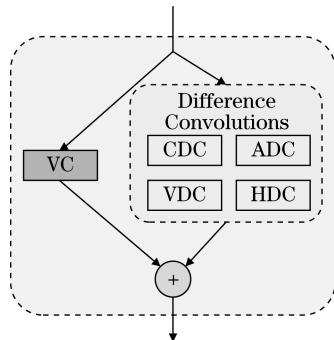


图3 DEConv结构

Fig.3 Diagram of DEConv structure

本文在轻量化共享卷积检测头(LSCD)的基础上加入细节增强卷积(DEConv),主要用于补偿因模型的轻量化造成的精度损失,提升模型的细节识

别能力。共享卷积(DEConv)具有5个并行卷积,利用卷积层的可加性,可将并行的卷积简化为单一的标准卷积,其计算过程如下:

$$F_{out} = \text{DEConv}(F_{in}) = \sum_{i=1}^5 F_{in} * K_i = F_{in} * \left(\sum_{i=1}^5 K_i \right) = F_{in} * K_{cvt} \quad (3)$$

式中:K_i=1:5分别为5个卷积核;K_{cvt}为转换后的卷积核。

2 改进YOLOv8n网络结构

YOLOv8算法是由Ultralytics于2023年发布的对象检测模型,主要应用于目标检测、实例分割、图像分类。其网络结构如图4所示,Backbone主干网络负责从输入图像中提取特征,Head头部网络负责检测,产生最终的检测结果,Neck颈部网络用于进行特征融合和增强^[13-14]。

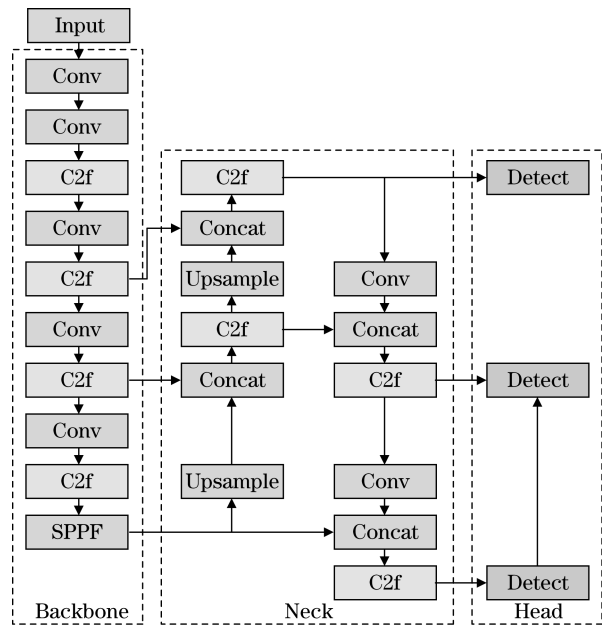


图4 YOLOv8n模型网络结构

Fig.4 The model network structure diagram of YOLOv8n

本文提出的YOLOv8n-C2f-Star-LSDECD模型的网络结构如图5所示,输入尺寸为640×640的RGB图像,经过卷积层进行下采样,C2f-Star利用星乘模块(star block)的高效性,在降低参数量和计算量的同时能提取更多特征。通过特征拼接将3个卷积层的输出输入到轻量化共享和细节增强卷积检测头(LSDECD),进行共享卷积和细节增强,

进一步降低参数量和计算量,最后通过 Scale 层的缩放(乘以缩放因子),实现不同视角、不同距离下的常见工程机械、集装箱、安全帽、工人等目标的检测。

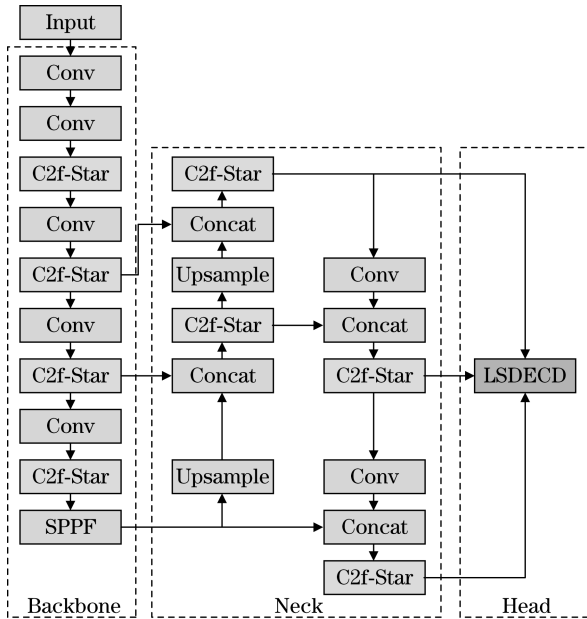


图5 YOLOv8n-C2f-Star-LSDECD 模型网络结构

Fig.5 The model network structure diagram of YOLOv8n-C2f-Star-LSDECD

3 实例验证

3.1 实验环境

本文基于 PyTorch 1.11.0 搭建深度学习框架,所有实验都基于表 1 软硬件环境进行。网络训练使用 SGD 优化器,图片输入尺寸为 640×640, batchsize 为 64,以 8 个工作线程加载数据,训练轮数为 300 轮。具体实验环境:操作系统 Ubuntu 20.04 3.8, Python 3.8, Cuda 11.3, GPU NVIDIA GeForce RTX 4090D, CPU Intel (R) Xeon (R) Platinum 8481C,内存 80 GB。

3.2 数据集

为验证本文提出方法的有效性,采用基于 Roboflow 的工程场景数据集进行验证,此数据集包含了 11 种工程场景中常见的对象如推土机、自动倾卸卡车、挖掘机、平路机、装载机、搅拌车、吊车、滚压机、集装箱、安全帽、工人,各标签数量如图 6 所示,共计 21 044 张图片。

数据集标注格式统一为 YOLO 格式。其次,根据 7:1:2 的比例将数据集划分为训练集、验证集、

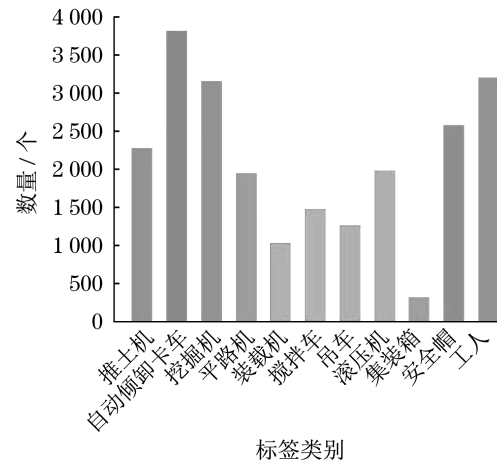


图6 数据集标签数量

Fig.6 Number and of labels in the dataset

测试集。

改进模型的 F1 分数-置信度曲线如图 7 所示。

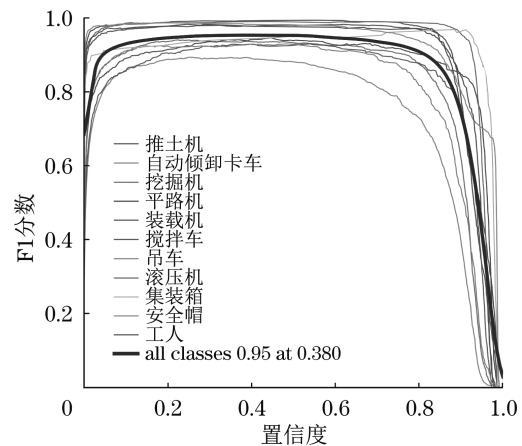


图7 改进模型的 F1 分数-置信度曲线

Fig.7 F1-confidence curve of the improved model

3.3 实验结果与分析

在实验中,以 YOLOv8n 作为基准模型,为了更好地评估模型的性能并与基准模型进行对比,本文使用目标检测中常用的 mAP@0.5、mAP@0.5:0.95、精度 P、召回率 R、每秒浮点运算次数、参数量、帧率(FPS)作为评价指标。其中,mAP@0.5、mAP@0.5:0.95、P 反应模型的精度、R 反应模型的检出率、每秒浮点运算次数反映模型的计算量,参数量反映模型的复杂度^[15]。

为验证所提算法的有效性,选择了当前主流的目标检测算法(如 FasterNet 和 StarNet)进行对比。在相同实验条件下,不同算法在工程场景混合数据集上的对比结果见表 1。改进模型与基准模型在相同实验条件下于混合数据集上的 mAP@0.5:0.95 对比见表 2。

表1 各模型主要评价指标对比

Tab.1 Comparison of the main evaluation indicators of the models

模型	mAP@0.5/%	mAP@0.5:0.95/%	P/%	R/%	每秒浮点运算次数/10 ⁹	参数量/10 ⁶	FPS
Baseline	97.5	86.2	96.4	94.1	8.1	3.0	202.4
FasterNet	98.2	86.8	97.1	94.7	10.7	4.2	182.2
StarNet	97.3	85.5	95.9	93.5	6.5	2.2	224.2
C2f-Star	97.3	87.9	96.7	94.0	6.9	2.5	234.5
C2f-Star-LSCD	97.7	88.5	96.0	94.4	5.3	1.9	223.4
C2f-Star-LSDECD	97.9	88.2	96.9	94.7	5.3	1.9	249.5

表2 各模型目标检测 mAP@0.5:0.95 对比

Tab.2 mAP@0.5:0.95 comparison of target detection of the models

模型	推土机/%	自动倾卸卡车/%	挖掘机/%	平路机/%	装载机/%	搅拌车/%	吊车/%	滚压机/%	集装箱/%	安全帽/%	工人/%
Baseline	90.0	72.3	79.6	91.8	78.1	90.3	76.9	96.5	87.2	92.9	90.2
FasterNet	89.2	73.8	80.1	90.5	79.6	89.1	76.8	96.4	87.0	92.9	89.0
StarNet	90.1	69.5	78.3	91.5	75.3	87.8	75.0	94.6	84.8	91.7	89.5
C2f-Star	91.4	73.5	82.9	92.6	81.2	91.3	81.5	97.9	88.9	93.9	92.4
C2f-Star-LSCD	91.2	72.1	82.4	92.7	80.9	92.1	80.3	97.8	85.8	93.6	91.8
C2f-Star-LSDECD	91.1	72.7	82.2	92.4	80.7	91.2	81.1	97.4	88.5	93.5	91.6

由表1可知, YOLOv8n-C2f-Star-LSDECD的mAP@0.5和mAP@0.5:0.95相比基准模型分别提升了0.4%和2.0%, 计算量减少了34.6%, 参数量减少了36.7%, FPS提高了23.3%。这表明改进模型在表现出了良好的轻量化效果。

由表2可知, YOLOv8n-C2f-Star-LSDECD在本文数据集中对11类目标的mAP@0.5:0.95均有提升。例如, 推土机的检测精度提高了1.1%, 自动倾卸卡车提高了0.4%, 挖掘机提高了2.6%, 平路机提高了0.6%, 装载机提高了2.6%, 搅拌车提高了0.9%, 吊车提高了4.2%, 滚压机提高了0.9%, 集装箱提高了1.3%, 安全帽提高了0.6%, 工人提高了1.4%。由此可见, 改进模型在11类目标上的精度均有一定提升。

4 结论

本文提出了一种结合星乘模块(star block)和轻量化共享与信息增强卷积检测头的新型YOLOv8模型, 用于检测工程场景中的典型工程机械、集装箱、工人等目标。所提方法基于Roboflow数据集进行了全面评估。论文的主要贡献如下:

(1) 将星乘替换原模型中的C2f结构, 并引入轻量化共享与信息增强卷积检测头, 有效降低了模

型的参数量和计算量。

(2) 工程场景数据集的实验结果显示, YOLOv8n-C2f-Star-LSDECD模型相比原模型, mAP@0.5和mAP@0.5:0.95分别提升了0.4%和2.0%; 同时, 参数量减少了36.7%, 计算量减少了34.6%, FPS提升了23.3%, 模型在预测精度和轻量化方面表现优异。

(3) 改进的YOLOv8n在常见工程机械、集装箱、安全帽、工人等目标的检测上表现出色, 11类目标的检测精度均有提升。

因此, 本文提出的算法为工程场景中的目标检测轻量化提供了一种有效解决方案, 并有望在无人化工程机械的自主作业等场景中得到广泛应用。

参考文献:

- [1] 张利丰, 田莹. 改进YOLOv8的多尺度轻量化车辆目标检测算法[J]. 计算机工程与应用, 2024, 60(3): 129-137.
- [2] YOU Z, BIAN Y M, DING Z J, et al. A study on weighting factors in cost function of model predictive control algorithms [C]// International Conference on Advances in Construction Machinery and Vehicle Engineering. Singapore: Nature, 2023: 495-504.

(下转第376页)

- 试验[J]. 系统仿真学报, 2022, 34(11): 2485-2496.
- [3] 成凯, 徐晓龙, 韩毓文, 等. 履带式湿地系列推土机行走机构的仿真[J]. 中国工程机械学报, 2013, 11(4): 331-336.
- [4] 赵家丰, 汪伟, 孙中兴, 等. 均质土壤承压下陷模型改进及验证[J]. 农业工程学报, 2016, 32(21): 60-66.
- [5] 孙术发, 高靖萱, 王敬凯, 等. 适应湿地作业的履带式运输车设计与研究[J]. 森林工程, 2022, 38(3): 77-86.
- [6] 贾鑫, 谢锐, 丁小兵, 等. 山地履带车辆软坡路面稳态转向模型建立及验证[J]. 中国农业科技导报, 2022, 24(11): 97-111.
- [7] AL-MILLI S, SENEVIRATNE L D, ALTHOEFER K. Track-terrain modelling and traversability prediction for tracked vehicles on soft terrain [J]. Journal of Terramechanics, 2010, 47(3): 151-160.
- [8] VYAZNIKOV M V, VYAZNIKOV A M. Increasing the mobility of tracked vehicles during curvilinear motion with partial skidding [J]. IOP Conference Series: Materials Science and Engineering, 2020, 971(5): 1-7.
- [9] 董超, 成凯, 高学亮, 等. 离心力影响条件下的履带车滑移转向性能分析[J]. 振动、测试与诊断, 2017, 37(1): 76-83, 200.
- [10] BEKKER M G. Theory of land locomotion [M]. Michigan: The University of Michigan Press, 1956.
- [11] 曹明, 龙祖熙, 王永维, 等. 塑相水田土壤参数对履带式拖拉机行驶性能的影响[J]. 浙江大学学报(农业与生命科学版), 2022, 48(1): 125-134.
- [12] LU E, MA Z, LI Y, et al. Adaptive backstepping control of tracked robot running trajectory based on real-time slip parameter estimation [J]. International Journal of Agricultural and Biological Engineering, 2020, 13(4): 178-187.
- [13] RAHAMAN M L, BERNAL E, SPIRYAGIN M, et al. An investigation into the effect of slip rate on the traction coefficient behaviour with a laboratory replication of a locomotive wheel rolling/sliding along a railway track [J]. Tribology International, 2023, 187: 1-7.
- [14] WANG H, LU E, ZHAO X, et al. Vibration and image texture data fusion-based terrain classification using WKNN for tracked robots [J]. World Electric Vehicle Journal, 2023, 14(8): 1-14.

(上接第 370 页)

- [3] LIU X, PENG H, ZHENG N, et al. Efficientvit: memory efficient vision transformer with cascaded group attention [C]// Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE, 2023: 14420-14430.
- [4] CHEN J, KAO S, HE H, et al. Run, don't walk: chasing higher FLOPS for faster neural networks [C]// Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE, 2023: 12021-12031.
- [5] 邵延华, 张铎, 楚红雨, 等. 基于深度学习的YOLO目标检测综述[J]. 电子与信息学报, 2022, 44(10): 3697-3708.
- [6] 雷源毅, 朱文球, 廖欢. 复杂场景下的改进YOLOv8n安全帽佩戴检测算法[J]. 软件工程, 2023, 26(12): 46-51.
- [7] 高荣芝, 徐翔, 尹君, 等. 基于机器视觉的混凝土泵车支腿识别和开度检测方法[J]. 中国工程机械学报, 2023, 21(4): 323-327.
- [8] BRUCAL S G E, JESUS L C M, PERUDA S R, et al. Development of tomato leaf disease detection using YoloV8 model via RoboFlow 2.0 [C]// 2023 IEEE 12th Global Conference on Consumer Electronics (GCCE). Piscataway: IEEE, 2023: 692-694.
- [9] MA X, DAI X, BAI Y, et al. Rewrite the Stars [C]// Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE, 2024: 5694-5703.
- [10] HOWARD A, SANDLER M, CHU G, et al. Searching for MobileNetV3 [C]// Proceedings of the IEEE/CVF International Conference on Computer Vision. Piscataway: IEEE, 2019: 1314-1324.
- [11] TIAN Z, SHEN C, CHEN H, et al. FCOS: A simple and strong anchor-free object detector [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2020, 44(4): 1922-1933.
- [12] CHEN Z, HE Z, LU Z M. DEA-Net: single image dehazing based on detail-enhanced convolution and content-guided attention [J]. IEEE Transactions on Image Processing, 2024, 33: 1002-1015.
- [13] 韩强. 面向小目标检测的改进YOLOv8算法研究[D]. 长春: 吉林大学, 2023: 19-30.
- [14] 李松, 史涛, 井方科. 改进YOLOv8的道路损伤检测算法[J]. 计算机工程与应用, 2023, 59(23): 165-174.
- [15] 程换新, 乔庆元, 骆晓玲, 等. 基于改进YOLOv8的无人机航拍图像目标检测算法[J]. 无线电工程, 2024, 54(4): 871-881.