

基于改进 YOLOv5 的安全帽小目标检测算法

刘一江¹ 樊福景¹ 王 通¹

LIU Yijiang FAN Fuming WANG Tong

摘 要

针对复杂施工场景下安全帽检测算法存在小目标漏检问题,提出一种基于 YOLOv5s 的安全帽改进算法。首先,为提升多尺度特征融合效果和小目标信息利用率,在 Efficient RepGFPN 基础上引入浅层分支及转置卷积替换 PAFPN,补充浅层小目标特征,减少上采样中小目标丢失的边缘信息。其次,采用 DBBNet 结构替换主干网络 C3 瓶颈层中的残差结构,通过多分支结构将小目标与周围信息相关联,增强主干网络的小目标提取能力,同时使用空洞卷积及通道注意力改进 SPP 结构,保留更多小目标信息,为 Efficient RepGFPN 网络提供更优质特征图。最后,在自建施工场景安全帽数据集与最新的目标检测算法进行比较。实验结果表明,改进 YOLOv5 算法召回率为 84.9%,平均精度达 90.1%;比原始 YOLOv5s 算法召回率提升 4.9%、平均精度提升 4.4%;对比最新的 YOLOv6s 3.0、YOLOv7-tiny、YOLOv8s 算法平均精度分别提升 2.4%、3.1% 和 1.8%,在施工场景下对小安全帽具有较强的检测性能。

关键词

YOLOv5; Efficient RepGFPN; DBB; 注意力机制; 安全帽检测; 小目标检测

doi: 10.3969/j.issn.1672-9528.2024.05.025

0 引言

在施工场景中,佩戴安全帽是保护施工人员免受事故伤害的重要保护措施之一,及时检测到工人是否佩戴安全帽,能减少安全事故的发生,对安全生产意义重大。近年来,随着视频监控系统的普及,基于深度学习的目标检测技术发展迅速,相比于通过肉眼检查工人是否按照规定佩戴安全帽,目标检测算法在保证精度的前提下,不仅高效快捷,还能解放人力,降低企业生产成本。

目前,基于深度学习的安全帽检测算法根据原理主要分为两类:一种是基于候选区的二阶段检测算法,另一种是基于回归的一阶段检测算法。由于二阶段检测算法需要去除冗余的候选框,检测速度较慢。而一阶段检测算法检测速度快,精度较高成为热门算法,YOLO 系列作为一阶段目标检测代表算法,在安全帽检测领域占有较高的应用比例,但直接应用 YOLO 算法进行安全帽检测过程中,发现复杂施工现场摄像头多采用远距离的拍摄方式覆盖整个施工场景,这导致安全帽在视频画面中所占像素尺寸较小,容易受背景信息干扰,存在漏检、误检的现象。针对此问题,研究人员通常在预处理、特征提取、特征融合和后处理等环节采用改进措施提高检测效果。文献[1]基于 YOLOv5 采用 Mosaic-9 数据增强,增加小安全帽样本多样性;文献[2]基于 YOLOv5 增加检测头,并在主干网络插入 SENet(squeeze and excitation net)注意力模块压缩全局空间信息,提高对小安全帽的关注度;文献[3]

基于 YOLOv5 在颈部网络增加特征图的跨尺度、跨层连接,提高了检测网络信息交换能力,增强了小安全帽检测效果;文献[4]基于 YOLOv4-tiny 引入标签平滑策略,提升小安全帽的检测精度。上述各阶段改进方法取得了一定的效果,但在特征融合部分使用跨层连接、增加检测头等方式促进深、浅层特征信息传递,其跨层信息融合率有待提升。针对此问题,RepGFPN 网络在跨层融合基础上,结合重参数化和高效聚合网络技术,提升特征融合效果,其改进措施在小麦穗、烟火、违禁品等检测^[5]中应用效果明显。为进一步降低施工现场小安全帽漏检率,本文在 RepGFPN 结构基础上,基于 YOLOv5 进行如下改进。

首先,为减少小目标向下传递特征损失、提高浅层特征利用率,采用转置卷积、增加浅层支路,对 Efficient RepGFPN 网络进行改进;其次,为将更高质量的特征图提供给 Efficient RepGFPN 网络,采用 DBB(diverse branch block)改进 C3 结构,减小冗余背景信息对小目标的影响,增强主干特征提取能力;最后,为保留更多高层语义中的小目标信息,采用空洞卷积和 CAM(channel attention module)通道注意力改进 SPP 结构,突出小目标特征。

1 改进 YOLOv5

改进 YOLOv5s 网络结构如图 1 所示,网络模型由四部分组成,分别是输入端(Input)、主干网络(Backbone)、颈部网络(Neck)和预测端(Prediction)。Input 端作数据预处理、Mosaic 数据增强;BackBone 网络由 CBS、C3-DBB 和改进 SPP 等模块组成,主要用于特征提取。Neck 网络层

1. 沈阳工业大学电气工程学院 辽宁沈阳 110870

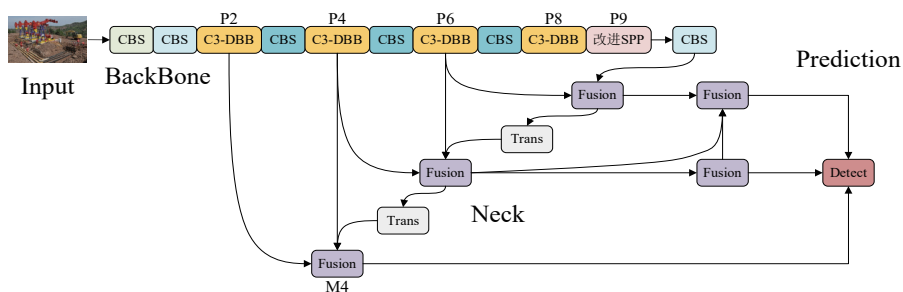


图1 改进 YOLOv5 网络结构图

是由转置卷积（Trans）、P2层支路和 Efficient RepGFPN 的 Fusion 模块组成，融合来自主干的多尺度特征。Prediction 端经非极大抑制处理消除大量冗余预测框，保留置信度最高的框预测类别，并返回目标位置的边框坐标。

1.1 改进 Efficient RepGFPN 网络

在安全帽检测中，监控广角覆盖整个施工场景，安全帽在整幅图像中占比较少，为增强 Efficient RepGFPN^[6] 特征融合网络对小目标的表征能力，将最临近上采样替换为转置卷积，缓解目标较小情况下，像素的数量相对较小，容易被背景同化的问题；同时引入 P2 层特征，将几何特征丰富的浅层信息补充到小目标支路中。本文提出了如图 2 所示的改进 Efficient RepGFPN 网络。

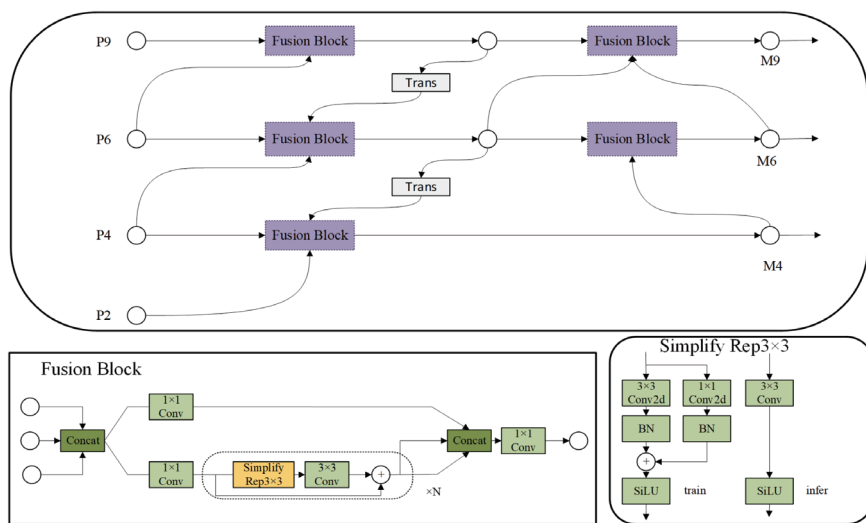


图2 改进 Efficient RepGFPN 网络

在改进 Efficient RepGFPN 网络中，以 P9 为起点自上而下传递语义信息，先经 Fusion Block 结构对有利于目标分类的抽象语义信息和来自下层的跨层特征进行融合，Fusion Block 将跨层特征与当前层特征在通道上进行拼接，在瓶颈层支路利用重参数化结构 Simplify Rep3×3 与高效层聚合网络思想，分析最短和最长的梯度路径，充分交换跨层信息，提升融合效果，其通用公式为：

$$M_n = \text{Fusion}(\text{Concat}(M_{n-1} + M'_n + M_{n+1}')) \quad (1)$$

式中： M_n 表示融合后输出特征图， $n-1$ 、 n 、 $n+1$ 表示上一层、当前层及下一层， M' 表示融合前输入特征图，Concat 表示通道拼接，Fusion 表示特征融合。

融合后特征图采用转置卷积^[7]（Trans）学习最优的上采样网络参数，避免插值技术复制相邻像素，导致小目标边缘退化问题，降低向下传递过程中小目标信息损失；经过 P6 ~ M6 支路同理，最终传递至检测小目标的主要支路 P4 ~ M4 层。此时，小目标支路是经过多次融合后的结果，高级语义信息与浅层定位信息充分交互。为进一步提高小目标检测精度，获取更多的小目标特征，连接含有较多的边缘、轮廓、纹理等细节信息的 P2 层，对 P4 ~ M4 支路特征进行补充，使小目标检测支路具有强力的特征表示，有利于施工现场小安全帽的检测。

对于改进 Efficient RepGFPN 网络，主干输入特征图的好坏影响融合效果，C3 是主干特征提取的主要结构，其内部瓶颈层中残差块由普通卷积组成。然而普通卷积特征提取能力有限^[8]，P2、P4、P6、P8 层小目标信息容易淹没在背景信息中；

P9 层 SPPF 结构通过最大池化聚合感受野内的特征信息，会造成较为严重的特征损失^[9]，高级语义信息中小目标特征不突出，不利于改进 Efficient RepGFPN 网络语义信息向下传递。因此，需对 C3 及 SPPF 结构进行改进。

1.2 C3-DBB 结构

为提升 C3 的特征提取能力，可通过添加注意力机制、Res2Net 模块、DBB 结构等^[10-12] 方式改进 C3 模块，削弱背景噪声对目标特征的影响。其中，基于重参数化思想的多分支结构 DBB，在网络训练时通过多尺度特征增强卷积的提取能力，将局部范围内

小目标与周围信息相关联，推理时将多分支合并成一个卷积。该卷积保留各分支训练的结果，增强小目标特征，能够为改进 Efficient RepGFPN 网络提供背景噪声较少的特征图。因此，本文采用 DBB 和 SiLU 激活函数组成的 DBBNet 类残差结构替换 C3 瓶颈层中的残差结构，如图 3 所示。C3-DBB 结构瓶颈层内由 N 个 DBBNet 串联构成，在训练时 DBB 通过 1×1 、 3×3 卷积核和平均池化核，提取不同尺度和复杂多样的空间特征，增强特征提取能力；在 DBB 结构和残差后使用 SiLU 激活函数，为网络提供更好的非线性。

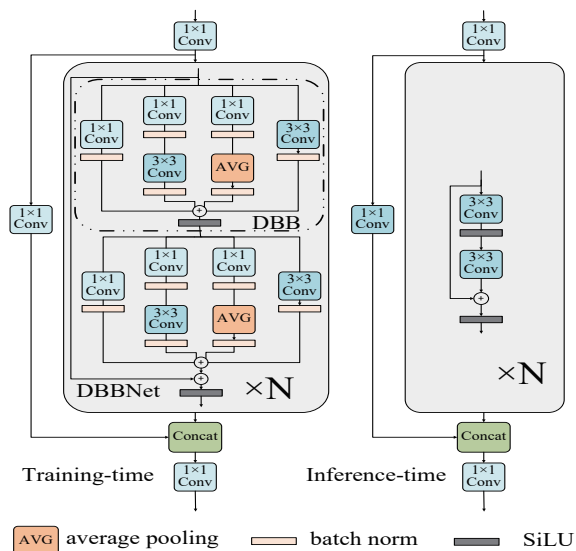


图3 C3-DBB 结构

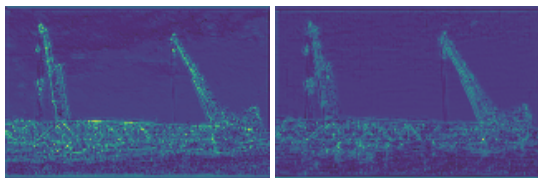
训练结束后，将各支路学习到的卷积权重，利用卷积的齐次性和可加性转化为统一卷积进行重组，其公式为：

$$M = M' \otimes F' + \text{REP}(b') \quad (2)$$

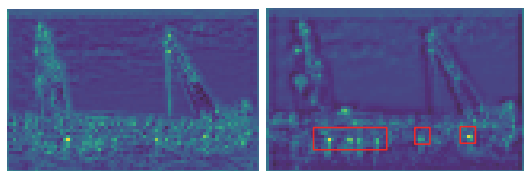
式中： M' 为输入特征图， M 为输出特征图， F' 为重构卷积权重， $\otimes$ 为卷积算子， b' 为偏置， $\text{REP}(b')$ 为重构后卷积偏置。通过图4改进前（左）和改进后（右）特征图可知，C3瓶颈层中卷积特征提取能力有限，经C3处理的特征图背景噪声较多，小安全帽特征不明显；而经过C3-DBB瓶颈层中重组后的卷积提取的特征图，融合不同感受野的多尺度信息，拥有多分支的特征提取特性，P2、P4、P6、P8层背景噪声相对较少，安全帽分布特征较为明显。



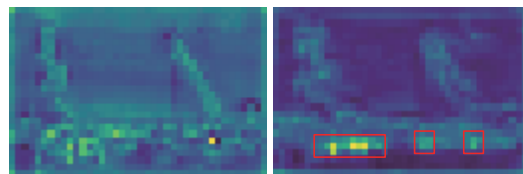
(a) 原图



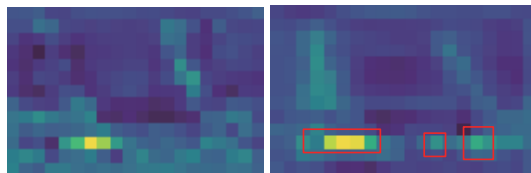
(b) P2 层



(c) P4 层



(d) P6 层



(e) P8 层

图4 改进前后 C3 输出特征图对比

1.3 改进 SPP 结构

对于 SPP 最大池化造成较多特征损失问题，使用空洞卷积替换最大池化减少特征信息丢失已取得较好效果^[13]。在此基础上，为进一步提升小目标召回率，有学者使用 MECA（more efficient channel attention）注意力^[14]、CA（channel attention）注意力^[15]对单独 1×1 卷积支路特征图进行学习，再将学习后的权重加在空洞卷积之后的特征图上，但通过空洞卷积的特征图，信息的保留程度发生变化，权重存在不适配的可能。针对上述问题，在使用空洞卷积降低特征损失的基础上，利用 CAM^[16]通道注意力对不同空洞卷积通道上拼接的特征图权重进行自适应调整，使其通道重要程度一一对应，改进后的结构如图5所示。

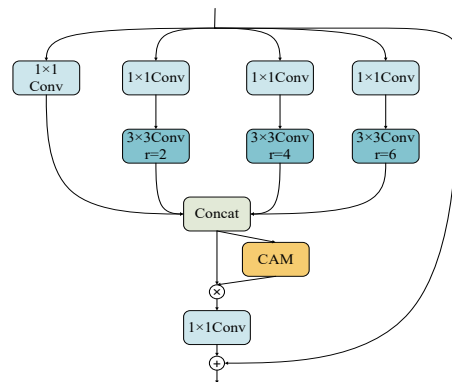


图5 改进 SPP 结构

在改进 SPP 结构中，在原有 SPP 四条支路基础上，首先，使用 1×1 卷积调整通道数，降低计算量；其次，为减少最大池化原感受野的特征损失，采用卷积核为 3×3 ，空洞率为 2、4、6 的空洞卷积提取局部特征及全局特征，汇聚更多的小目标语义信息；然后，由于不同空洞卷积提取特征时关注点不一样，为使有效信息在通道上得到加强，需将拼接后的特征图采用 CAM 通道注意力在不同通道感受野中进行自适应加权，动态聚合多尺度上下文信息，直接突出有利于小目标的特征层，提高小目标检测效果；最后，为防止网络层数增加梯度下降可能导致精度降低，在最外侧引入残差边。

2 实验结果与分析

2.1 实验设置

2.1.1 数据集制作

自建施工场景安全帽数据集，一部分来自网络爬虫，另一部分选取 SHWD 数据集中含中、远距离小目标的图片作为补充，扩充数据量，增强模型的泛化能力。经过筛选、清洗和标注，共得到数据集 7768 张，分为训练集 6275 张，测试 1493 张，数据集目标类型分为两类：佩戴安全帽（hat）和未佩戴安全帽（no hat）。

2.1.2 实验环境

本次实验采用的硬件环境 GPU 为 NVIDIA 3090 24 G，CPU 为 AMD EPYC 7543，软件环境为 Linux 操作系统，选用 Pytorch 1.9.0 为深度学习框架。采用 YOLOv5s 进行实验，输入图像尺寸大小为 640×640，批次为 64，Epochs 为 250，使用 SGD 优化器进行优化计算。

2.2 实验结果与分析

2.2.1 实验结果

将改进 YOLOv5 算法与原始 YOLOv5s 算法使用相同的训练策略分别训练 250 轮后，召回率、平均精确率对比曲线如图 6 所示，横坐标为训练轮次，纵坐标为评价指标。对于漏检问题，评价指标主要以召回率（R）、平均精度（mAP）为主。由对比曲线可知，两种模型在前 50 轮都迅速收敛，在 100 轮之后逐渐趋于稳定，训练效果良好，召回率、平均精度提升明显，验证了改进算法的可行性。

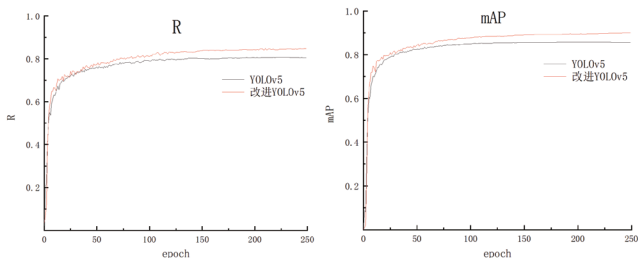


图 6 改进前后 R 和 mAP 对比曲线

2.2.2 消融实验

为全面分析改进 YOLOv5s 模型中各改进模块对算法的作用，以未改动的 YOLOv5 为基础模型，设计消融实验，将改进的各模块添加到原 YOLOv5 模型中，与原模型作对比，√表示增加改进模块，具体如表 1 所示。

表 1 消融实验

模型	RepG-FPN	改进 RepG-FPN	C3-DBB	改进 SPP	P	R	mAP
1					0.885	0.8	0.857
2	√				0.885	0.816	0.869
3		√			0.883	0.827	0.876
4		√	√		0.886	0.84	0.889
5		√	√	√	0.889	0.849	0.901

由表 2 可知，模型 1（YOLOv5s）的召回率和平均精度分为 0.8 和 0.857；模型 2 在替换 Efficient RepGFPN 网络后，

召回率提升 1.6%、平均精度提升 1.2%；模型 3 对模型 2 的 Efficient RepGFPN 网络进行 2 处改进，召回率提升 1.1%、平均精度提升 0.7%；模型 4 在模型 3 中加入 C3-DBB 结构，召回率提升 1.3%、平均精度提升 1.3%；模型 5（改进 YOLOv5）在模型 4 中加入改进 SPP 结构，召回率提升 0.9%、平均精度提升 1.2%；添加各改进模块评价指标均有提升。同时，通过现场图片进行检测对比，模型 1（左）与模型 5（右）检测结果如图 7 所示，图（a）中安全帽距离较远，模型 1 存在左侧 3 个，中间 1 个目标漏检；图（b）模型 1 将 2 个管口和 1 个佩戴安全帽错检为未佩戴安全帽，并漏检 1 个井台上目标；两个场景基本找出漏检的安全帽，错检问题亦有改善，实验效果良好。



(a) 野外施工场景



(b) 油田施工场景

图 7 检测结果对比

2.2.3 性能对比实验

为更好地展示改进 YOLOv5s 算法的性能，将改进 YOLOv5s 改进算法与最新发表的 YOLO 系列算法进行对比，在相同的实验条件下，进行模型的训练与测试，检测结果如表 2 所示。实验结果表明，改进 YOLOv5s 改进模型对比 YOLOv6s 3.0 召回率提升 2%，平均精度提升 2.4%；对比 YOLOv7-tiny 召回率提升 2.5%，平均精度提升 3.1%；对比 YOLOv8s 召回率提升 2.1%，平均精度提升 1.8%。各评价指标均有提升，体现了改进 YOLOv5s 算法的先进性。

表 2 性能对比实验

模型	P	R	mAP
YOLOv6s 3.0	0.884	0.829	0.877
YOLOv7-tiny	0.882	0.824	0.87
YOLOv8s	0.888	0.828	0.883
改进 YOLO	0.889	0.849	0.901

为进一步展示改进 YOLOv5 的优越性，图 8 为复杂施工场景中，YOLOv6s、YOLOv7-tiny、YOLOv8s 和改进 YOLOv5s 实验主观视觉对比图。从图 8 可以看出，在施工场景广角摄像头下，算法 YOLOv6s 3.0、YOLOv7-tiny、YOLOv8s 存在漏检、错检现象，而改进 YOLOv5 对上述目标全部正确检出。由此可见，改进 YOLOv5 算法对比最新的检测算法在复杂施工场景下的小目标有更好的检测能力。

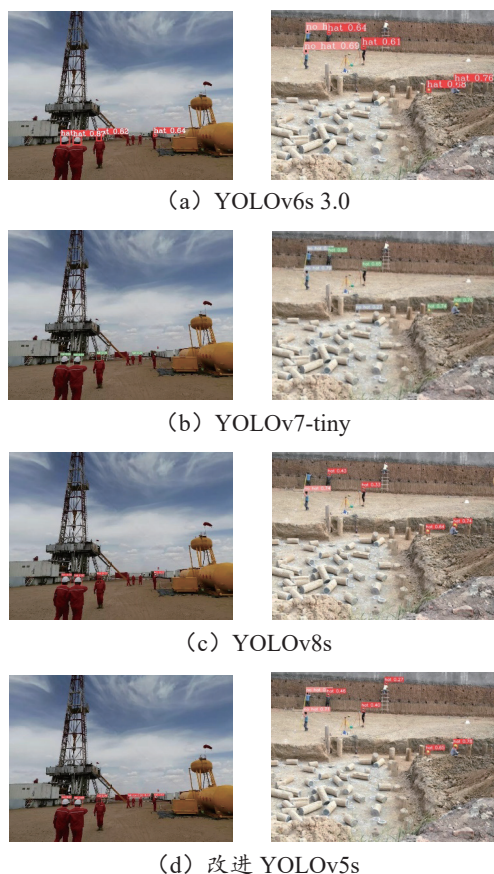


图8 性能实验对比图

3 结论

针对施工场景下, YOLOv5s 安全帽算法存在小目标漏检的问题, 采用改进 Efficient RepGFPN 替换 PAFPN, 增加浅层、高层信息交互, 增强特征图融合效果; 为提高改进 Efficient RepGFPN 各节点输入特征图质量, 使用 DBB 模块与 SiLU 激活函数改进 C3 结构; 使用空洞卷积及通道注意力机制改进 SPP 结构。通过自建数据集上的实验证明, 改进 YOLOv5s 算法比 YOLOv5s 算法及最新的 YOLOv6s 3.0、YOLOv7-tiny、YOLOv8s 等算法召回率、平均精度均有提升, 展示了改进算法优越的性能。

参考文献:

- [1] TAI W, WANG Z, LI W, et al. DAAM-YOLOV5: a helmet detection algorithm combined with dynamic anchor box and attention mechanism[J]. Electronics, 2023, 12(9): 2094-2110.
- [2] ZHANG Y, QIU Y, BAI H. FEFD-YOLOV5: a helmet detection algorithm combined with feature enhancement and feature denoising[J]. Electronics, 2023, 12(13): 2902-2902.
- [3] MAO Y, YUAN J, ZHU Y, et al. Surface defect detection of smartphone glass based on deep learning[J]. The international journal of advanced manufacturing technology, 2023, 127(11/12): 5817-5829.
- [4] ZHANG B, SUN C, FANG S, et al. Workshop safety helmet

wearing detection model based on SCM-YOLO[J]. Sensors, 2022, 22(17): 6702-6702.

- [5] LIU L, LI P. An improved YOLOv5-based algorithm for small wheat spikes detection[J]. Signal, image and video processing, 2023, 17(8): 4485-4493.
- [6] XU X, JIANG Y, CHEN W, et al. Damo-yolo: a report on real-time object detection design[EB/OL]. (2022-11-23)[2024-03-01]. <https://doi.org/10.48550/arXiv.2211.15444>.
- [7] KIM M, JEONG J, KIM S. ECAP-YOLO: efficient channel attention pyramid YOLO for small object detection in aerial image[J]. Remote sensing, 2021, 13(23): 4851.
- [8] QING Y, LIU W, FENG L, et al. Improved YOLO network for free-angle remote sensing target detection[J]. Remote sensing, 2021, 13(11): 2171.
- [9] 李雅雯, 孙浩然, 胡跃明, 等. 基于注意力机制与多尺度特征融合的电极缺陷 YOLO 检测算法[J]. 控制与决策, 2023, 38(9): 2578-2586.
- [10] WEN G, LI S, LIU F, et al. YOLOv5s-CA: a modified yolov5s network with coordinate attention for underwater target detection[J]. Sensors, 2023, 23(7): 3367-3381.
- [11] LUO X, WU Y, WANG F. Target detection method of UAV aerial imagery based on improved YOLOv5[J]. Remote sensing, 2022, 14(19): 5063.
- [12] 刘阔, 李喆, 李曜丞, 等. 基于重参数化 YOLOv5 的输电线路缺陷边缘智能检测方法[J/OL]. 高电压技术: 1-11. [2024-03-01]. <https://doi.org/10.13336/j.1003-6520.hve.2022086>.
- [13] WANG L, NI Q, CHEN C, et al. Lightweight target detection algorithm based on improved YOLOv4[J]. IET image processing, 2022, 16(14): 3805-3813.
- [14] LI G, ZHAO S, ZHOU M, et al. YOLO-RFF: an industrial defect detection method based on expanded field of feeling and feature fusion[J]. Electronics, 2022, 11(24): 4211.
- [15] FU X, LI A, MENG Z, et al. A dynamic detection method for phenotyping pods in a soybean population based on an improved YOLO-v5 network[J]. Agronomy, 2022, 12(12): 3209.
- [16] WOO S, PARK J, LEE J Y, et al. CBAM: convolutional block attention module[C]//Computer Vision-ECCV 2018, p. VII. Berlin: Springer, 2018: 3-19.

【作者简介】

刘一江(1994—), 男, 黑龙江七台河人, 硕士研究生, 研究方向: 深度学习、图像处理技术。

樊福景(1999—), 男, 山东菏泽人, 硕士研究生, 研究方向: 工况诊断。

王通(1976—), 男, 辽宁盘锦人, 博士, 副教授, 研究方向: 复杂工业过程控制、检测及故障诊断、软测量等。

(收稿日期: 2024-03-11)