

许立兴, 雷钧, 杨洋, 等. 基于 GSS-YOLOv5 的汽车半轴目标检测算法研究[J]. 智能计算机与应用, 2026, 16(4): 103-109.
DOI: 10.20169/j.issn.2095-2163.24053001

基于 GSS-YOLOv5 的汽车半轴目标检测算法研究

许立兴, 雷钧, 杨洋, 王伟鸣, 徐洪胜

(湖北汽车工业学院 电气与信息工程学院, 湖北 十堰 442002)

摘要: 在工业产品目标检测领域, 尽管 YOLOv5 和其他深度学习算法展现出了显著的性能, 但仍然存在一些挑战和改进空间, 包括模型参数庞大、边界框定位精度不佳等。本文设计了一种基于改进 YOLOv5 的汽车半轴检测模型。首先, 该模型将 YOLOv5 C3 结构与 Ghost 融合成 C3Ghost 结构, 旨在降低模型计算量 (GFLOPs) 的同时保持原 YOLOv5 的检测精度; 其次, 引入平滑交并比 (SIoU) 损失函数代替原有的 CIoU 损失函数, 以提升边界框生成精度和模型训练效率, 减少因定位不准确而导致的误检和漏检情况; 最后, 加入 SKNet 注意力模块, 使模型更加关注待检测区域, 抑制无用信息并降低复杂背景的干扰, 提高算法的检测准确率。实验数据表明, 该算法相较于其他主流算法 mAP@0.5 提升到了 0.971, 参数量和计算量上得到了较好的均衡, 是一个性能全面的模型。同时通过消融实验, 进一步验证了本文所做改进在汽车半轴目标检测上的有效性。

关键词: 汽车半轴; 目标检测; YOLOv5; 轻量化; SIoU; 注意力机制

中图分类号: TP391

文献标志码: A

文章编号: 2095-2163(2026)04-0103-07

Research on vehicle axle target detection algorithm based on GSS-YOLOv5

XU Lixing, LEI Jun, YANG Yang, WANG Weiming, XU Hongsheng

(School of Electrical and Information Engineering, Hubei University of Automotive Technology, Shiyan 442002, Hubei, China)

Abstract: In the field of industrial product object detection, although YOLOv5 and other deep learning algorithms have shown remarkable performance, there are still some challenges and room for improvement, including large model parameters and poor boundary frame positioning accuracy. In this paper, an automobile axle detection model based on improved YOLOv5 is designed. First, the model fuses the YOLOv5 C3 structure with Ghost into C3Ghost structure, aiming to reduce the model computation (GFLOPs) while maintaining the detection accuracy of the original YOLOv5. Secondly, the Smooth Intersection over Union (SIoU) loss function is introduced to replace the original CIoU loss function, so as to improve the accuracy of boundary frame generation and model training efficiency, and reduce the false detection and missing detection caused by inaccurate positioning. Finally, the SKNet attention module is added to make the model pay more attention to the region to be detected, suppress useless information and reduce the interference of complex background, and improve the detection accuracy of the algorithm. Experimental data show that, compared with other mainstream algorithms mAP@0.5, the proposed algorithm is improved to 0.971, and has a good balance between the number of parameters and the amount of computation. It is a comprehensive model. At the same time, the ablation experiment further verifies the effectiveness of the proposed improvement on vehicle axle target detection.

Key words: automobile axle; target detection; YOLOv5; lightweight; SIoU; attention mechanism

0 引言

随着汽车与新能源产业的蓬勃发展, 汽车半轴作为汽车核心零部件, 其质量决定了汽车在行驶中的安全与使用寿命^[1]。为解决效率问题, 由人工上下料改为机器人自动化上下料, 须精准识别半轴并确定抓取轴杆相对于轴头的位置, 否则在半轴放入

淬火机床时会发生碰撞造成二次损害。

应用机器视觉技术进行物体识别, 能有效减少对人力的依赖, 避免了对操作人员的潜在伤害。基于深度学习的物体检测技术在目标检测中广泛应用, 目标检测的主流方法可分为 Two-Stage (两阶段) 和 One-Stage (单阶段) 两大类。Two-Stage 算法首先通过区域建议网络 (RPN) 生成候选区域, 为后

基金项目: 湖北省教育科学规划重点课题(2022GA049)。

作者简介: 许立兴(2001—), 男, 硕士研究生, 主要研究方向: 深度学习, 机器视觉。

通信作者: 雷钧(1975—), 男, 硕士, 教授, 主要研究方向: 工业互联网与智能制造。Email: hustlei@163.com。

收稿日期: 2024-05-30

Backbone 部分采用 Focus 模块实现快速下采样, 保证信息的完整性, 为后续的特征提取工作打下基础。C3 模块融合为 C3Ghost 结构, 用于增强模型学习能力, 并在保证检测精度不变的前提下, 实现了模型的简化。而 SPP 模块的引入, 则有效扩大了主干网络对特征接收范围, 增强了模型的尺度适应性, 使得网络的训练更加稳定, 进而提升预测准确性。Neck 模块对骨干网络提取的特性进行进一步的强化与整合。Head 是 YOLO v5s 模型中进行目标检测的核心部分, 经过骨干网络的特征在颈部网络压缩和融合后进行分类和预测。头部通常包括多个输出层, 对应不同分辨率的特征图。

1.1 融合 C3Ghost 结构

由于嵌入式设备的内存和计算资源有限, 部署卷积神经网络存在较大困难。本文引入了一种名为

Ghost^[7] 的模块, 该模块通过一系列简洁的线性变换操作, 生成额外的特征图, 这些特征图可以更深入地挖掘原始特征的信息。Ghost 模块可轻松集成到现有的卷积神经网络中, 以提升其性能。Ghost 模块相比于 MobileNet V3 有更高的识别性能, 并且在数据集上具有相似的计算成本。

在普通卷积神经网络中引入 Ghost 模块, 能够将卷积过程划分为两个阶段: 首先执行传统的卷积运算, 随后通过一系列简化的线性变换产生原特征图的 Ghost 特征图。这些 Ghost 特征图随后被合并, 形成最终的输出特征图。这种处理方式有效减少了特征图的冗余度, 并显著降低了网络中卷积层的参数数量和计算量, 从而大幅度减少模型的整体复杂性。Ghost 模块的示意图如图 2 所示。

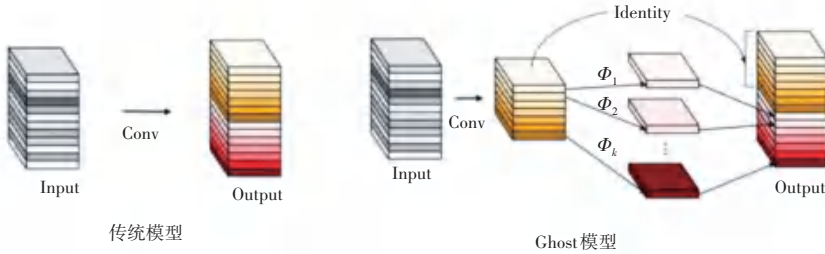


图 2 Ghost 模块示意图

Fig. 2 Ghost module schematic diagram

1.2 改进损失函数

损失函数能在一定程度上决定目标检测所能达到的效果。作为目标检测损失函数的重要组成部分, 边界框损失函数的合理定义能有效提升目标检测模型的性能。在早期的目标检测研究中, IoU 被用作定位损失, 当目标框与预测框没有交集时, 其会出现梯度消失的现象, 进而无法进一步优化模型。这启发了一系列以 IoU 为基础的改进损失设计, 其中包括 CIoU^[8]、DIoU^[9]。

损失函数不仅能提高模型的收敛速度, 还能改善检测的精度。CIoU 损失函数考虑了 3 个重要的几何因素, 即中心点距离、重叠面积和宽高比。传统的 YOLOv5s 算法使用的 CIoU 在 IoU 的基础上引入预测框与真实框非重合区域的影响来进行边框回归, 能够反映真实框与预测框的远近距离, 但 CIoU 的最后一项仍未得到良好的定义, CIoU 损失的引入提高了长宽比的一致性, 但同时也导致了模型预测框与目标框的实际距离误差增大。CIoU 公式如下:

$$L_{CIoU} = 1 - \frac{A \cap B}{A \cup B} + \frac{\rho^2(A, B)}{c^2} \quad (1)$$

式中: c 为包住 A 框和 B 框的最小方框的对角线距离, $\rho(A, B)$ 为 A 框与 B 框中心点间的欧氏距离。

为解决上述问题, 本研究引入了 SIoU 损失函数, 其架构如图 3 所示, 包括 4 类损失: 角度、距离、形状和 IoU 损失, 这些组合形成了 SIoU 损失函数。具体可以表示为:

$$L_{SIoU} = 1 - X_{IoU} + \frac{\Lambda + \Omega}{2} \quad (2)$$

式中: Λ 为距离成本, 即两框中心点的欧氏距离, Ω 为形状成本。表达式如下:

$$\Lambda = \sum_{t=x,y} [1 - \exp(-\gamma \rho_t)] \quad (3)$$

$$\Omega = \sum_{t=w,h} [1 - \exp(-\omega_t)^\theta] \quad (4)$$

其中,

$$\rho_x = \left(\frac{b_{c_x}^{gt} - b_{c_x}}{c_x} \right)^2 \quad (5)$$

$$\rho_y = \left(\frac{b_{c_y}^{gt} - b_{c_y}}{c_y} \right)^2 \quad (6)$$

$$\gamma = 2 - \Delta \quad (7)$$

$$\omega_\omega = \frac{|\omega - \omega^{gt}|}{\max(\omega - \omega^{gt})} \quad (8)$$

$$\omega_h = \frac{|h - h^{gt}|}{\max(h - h^{gt})} \quad (9)$$

式中： c_y 和 c_x 为预测框和真实框最小外接矩形的高和宽， ω 、 h 、 ω^{gt} 、 h^{gt} 分别为预测框和真实框的宽和高， θ 的值控制对形状损失的关注程度， Δ 为角度成本。

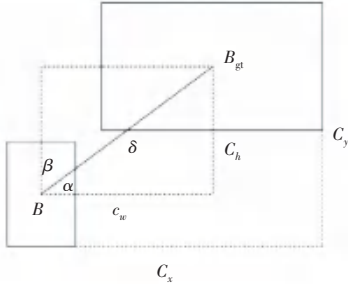


图 3 SIoU 损失示意图

Fig. 3 Schematic diagram of SIoU loss

1.3 引入注意力机制

SKNet^[10] 算法采用非线性多核信息整合策略，实现神经元自适应感受野大小的调整，其是一种基于通道域注意力的轻量级嵌入式模块。该模块旨在通过动态调整不同通道的卷积核大小，在保持模型性能的同时降低计算复杂度。SKNet 模块的核心思想是根据输入数据的特性自适应地选择合适的卷积核尺寸，从而在不同尺度上捕获特征。因此在特征提取后嵌入 SKNet，对特征进行重要性加权，再送入目标检测层，能够有效改善检测识别的性能^[11]。SKNet 网络结构如图 4 所示。

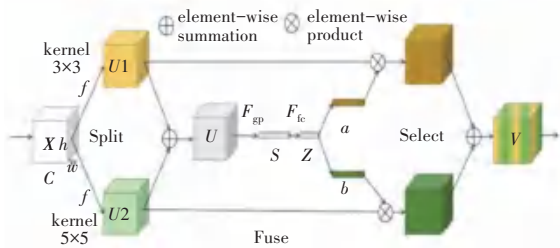


图 4 SKNet 网络结构图

Fig. 4 SKNet network structure diagram

SKNet 由 Split、Fuse 及 Select 三部分组成。Split 对特征图进行多分支分离卷积操作。Fuse 将多个分支的特征提取结果相加，并通过全局平均池化和全连接层进行降维处理。 F_{gp} 是全局平均池化操作， F_{fc} 是先降维再升维的 2 层全连接层。Select 对降维后的特征向量进行 Softmax 处理，得到每个分支的权重，并将权重应用于原始特征图，得到最终的输出特征图。使用 a 和 b 两个权重矩阵对 $U1$ 和 $U2$

进行加权操作，再进行相加得到输出 V 。

2 实验与结果分析

2.1 数据集

本文采用的数据集来自十堰某科技有限公司，共拍摄 800 余张汽车半轴图片，从中筛选出拍摄清晰及多角度不同类型汽车半轴的图片，共 655 张，按照 8 : 1 : 1 的比例进行划分，最终得到训练集 524 张、测试集和验证集各 131 张。通过 Labellmg 标注工具手动标注，共有两个标记类别，分别是 true (正确类型的汽车半轴) 和 false (错误类型的汽车半轴)，如图 5 所示。

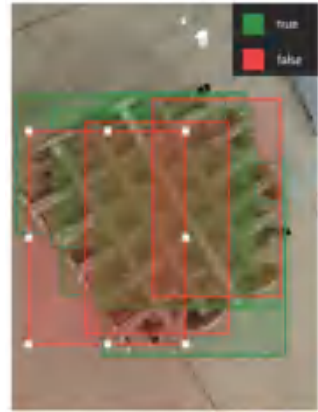


图 5 训练集类别标注示意图

Fig. 5 Schematic diagram of training set category annotation

2.2 实验环境与模型训练

实验配置的详细情况见表 1。实验总的训练次数为 150 次，迭代批量大小设置为 16，学习率为 0.01。^[12]

表 1 配置表

Table 1 Configuration table

名称	型号/版本
操作系统	Windows 10
CPU	Intel(R) Xeon(R) Platinum 8358 CPU
GPU	NVIDIA GeForce RTX 3090
CUDA	11.1.0
Pytorch	1.9.0
Python	3.8.10

2.3 评估指标

本文选取召回率(recall)、预测精度(precision)、平均精度均值(mean Average Precision, mAP)作为评估算法性能的指标，其中，mAP 为最主要的性能评价指标^[13]。各指标的计算公式如下：

$$R = \frac{T_p}{T_p + F_N} \quad (10)$$

$$P = \frac{T_p}{T_p + F_p} \quad (11)$$

$$mAP = \frac{1}{n} \sum_{i=0}^n AP_i \quad (12)$$

式中: R 表示被正确预测为缺陷的正样本占总正样本数的比例, P 表示正样本中被正确预测为缺陷的样本比例, mAP 表示在不同目标类别下的平均精准度, T_p 表示被正确预测为缺陷的正样本数量, F_N 表示被错误预测为正常的正样本数量, F_p 表示被错误预测为缺陷的负样本数量, n 表示目标检测任务中涉及的不同缺陷目标类别的数量。同时,选取 mAP 用来综合评估目标检测性能,一般情况下值越高,模型性能越好。

2.4 不同算法对比实验

为验证本文模型的有效性,在实验数据集和测试环境相同的情况下,将本文模型 GSS-YOLOv5 与 YOLOv6n、YOLOv4-tiny、YOLOv7-tiny、YOLOv8n、YOLOv6s 等主流模型进行对比实验,实验结果见表 2。

表 2 对比实验数据表

Table 2 Comparative experimental data table

模型	mAP@0.5	参数量/MB	计算量/G
YOLOv6n	0.915	5.3	11.0
YOLOv4-tiny	0.857	10.2	5.1
YOLOv7-tiny	0.902	7.9	13.3
YOLOv8n	0.926	4.6	8.2
YOLOv6s	0.962	17.8	44.7
GSS-YOLOv5	0.971	7.3	14.5

从表 2 中可以看出,YOLOv6s 在精度上表现最好,但参数量和计算量也最大,不利于在工业环境中部署;YOLOv4-tiny 在计算量上最低,但精度也相对较低,适合对实时性要求非常高且对精度要求不太严格的应用场景,而本文对半轴的识别要求精率较高;YOLOv7-tiny 和 YOLOv6n 在精度和计算量上均有不错的表现,但参数量稍大,可能需要根据具体应用场景进行权衡;YOLOv8n 在精度、参数量和计算量方面均表现出色,是一个性价比很高的模型,适合在资源受限的实时应用中使用^[14];本文提出的 GSS-YOLOv5 相较于其他算法,在精度和参数量之间取得了较好的平衡,计算量也适中,是一个性能全面的模型,适用于工业上对汽车半轴进行目标检测。

2.5 消融实验

为了验证各个模块的可行性和有效性,在相同的实验环境下设置消融实验。具体实验数据见表 3。实验数据表明,将 YOLOv5s 的 C3 模块与 Ghost 模块相结合后模型精度和 $mAP@0.5$ 略有下降^[15],召回率也下降。但值得注意的是,计算量显著减少,从 16.0 G 降至 10.4 G,这有助于在资源受限的设备上运行模型。在已添加 Ghost 模块的基础上,进一步添加 SIOU^[16] 损失函数后,模型的精度、召回率和 $mAP@0.5$ 均有所提高,尤其是召回率显著提升,从 0.862 提升至 0.932。同时,计算量也略有增加,从 10.4 G 升至 11.8 G。添加第 3 个模块后 $mAP@0.5$ 相比于原 YOLOv5s 有显著提高^[17],尤其是精度达到了 0.987。然而,召回率略有下降,但 $mAP@0.5$ 的提升表明模型在多个类别上的性能得到了整体提升。这对于实际的应用场景和需求来说至关重要,尽管添加更多的组件会增加计算量,但模型的性能得到了有效的提升。因此证明了各个模块在汽车半轴目标检测上的有效性。

表 3 消融实验对比表

Table 3 Comparison of ablation experiments

模型	precision	recall	mAP@0.5	计算量/G
YOLOv5s	0.960	0.886	0.947	16.0
YOLOv5s +Ghost	0.958	0.862	0.935	10.4
YOLOv5s +Ghost+SIOU	0.963	0.932	0.958	11.8
YOLOv5s+Ghost+SIOU+SKNet	0.987	0.920	0.971	14.5

2.6 训练结果

经过测试,改进后的 YOLOv5s 神经网络模型,精确率达到 98.7%,召回率达到 92%,置信度 0.5 的平均精度 ($mAP@0.5$) 指标超过 97%,训练结果如图 6 所示。

从图 6 中可以看出,随着训练次数的增加,精确度开始迅速上升,当达到 80 轮左右时,精确度整体上保持在 98% 的水平。这表明模型在区分正负样本方面具有良好的性能,能够有效减少误判的情况。与精确度不同,召回率随着训练次数的增加先急剧上升,达到 50 轮时召回率开始下降,但随着训练次数的增加,召回率继续上升,整体呈现上升趋势^[18]。这意味着,尽管类别数量增加带来了更大的挑战,但模型依然能够成功识别出更多的正样本,提高了整体的覆盖率。 $mAP@0.5$ 曲线在 40 轮处达到最高,随后一直保持在较高水平,这表明模型在多个类别上均具有良好的性能,证明了模型的稳定性和泛化能力。

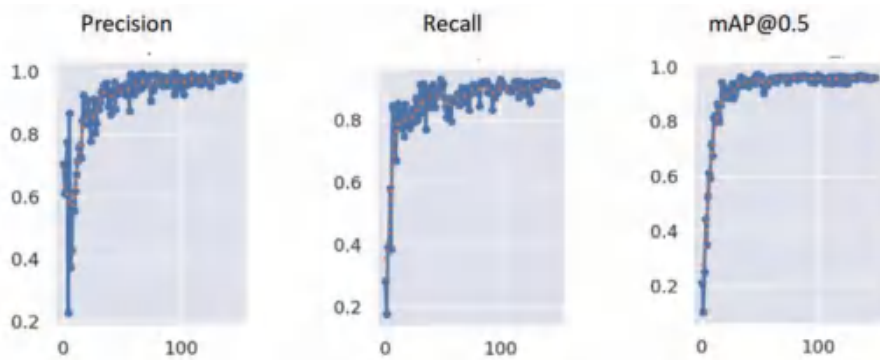


图6 改进后神经网络训练结果图

Fig. 6 Training results of the improved neural network

改进前后漏检情况对比如图7所示,改进后的算法为检测到的半轴提供了更高的置信度分数,这表明算法对这些目标的识别更加准确和可靠。这种提高可以降低漏检的可能性,从而提高整体的目标检测性能。从这张对比图中可以看到,改进后的YOLOv5s算法在目标检测任务中取得了显著的进步,特别是在减少漏检和提高置信度分数方面^[19]。这些改进有望使算法在各种实际应用中更加可靠和有效。



(a) 改进前 (b) 改进后

图7 漏检情况对比图

Fig. 7 Comparison of missed cases

边界框准确度对比如图8所示。对比改进前后的边界框,可以看到改进后的算法生成的边界框更加贴合目标物体的边缘,这意味着算法的准确度得到了提高。



(a) 改进前 (b) 改进后

图8 边界框准确度对比图

Fig. 8 Comparison of boundary box accuracy

3 结束语

针对汽车半轴检测容易出现漏检、边界框定位不精准、模型参数规模大、可移植性不高等问题^[20],本文提出了一种基于改进YOLOv5s的汽车半轴检测算法,通过YOLOv5s C3结构与Ghost结合成C3Ghost结构,在网络中利用更少的计算资源输出更多有效的特征信息,提升了网络的计算速度,降低参数数量和GFLOPs的同时,保持原YOLOv5s检测精度不下降,经实验证明精度甚至有所提高。同时使用SiLU损失函数替换原网络中的CIoU损失函数,提高边界框生成的准确性和模型收敛速度,减少因定位不准确导致的误检和漏检。改进后的算法与现有模型对比,计算精度有所提升,具有更好的鲁棒性和可移植性,且满足检测需求。最后,添加SKNet注意力机制模块,对特征通道进行加权,使特征图更关注半轴所在区域的特征通道,有效改善了检测识别性能。考虑到目前发展趋势,未来的研究重点应该放在轻量化设计,在各项性能不降低的情况下,使计算速度和精度达到平衡,对模型进行压缩,以满足部署工业领域等嵌入式设备的需求,实现轻量化模型应用于工业设备对产品的检测任务^[21]。

参考文献

- [1] 工业和信息化部装备工业发展中心,浙江吉利控股集团有限公司. 中国汽车产业与技术发展报告[R]. 北京:电子工业出版社, 2021:407.
- [2] 武越,丁航奇,何昊,等. 多任务优化算法及应用研究综述[J]. 计算机应用, 2024, 44(5):1338-1347.
- [3] 杨利春,田彬,党建武. 基于深度学习的图像拼接算法研究综述[J]. 计算机应用研究, 2024, 41(7):1930-1939. DOI:10.19734/j. issn. 1001-3695. 2023. 10. 0528.
- [4] 姜越夫,王青,吕绪山. 改进YOLOv5s的纱管目标检测方法[J]. 机械与电子, 2024, 42(2):29-34.
- [5] 牛柯,彭斌,杨小亮. 基于改进YOLOv5s的轻量化智能钻机管柱检测方法[J/OL]. (2024-05-07)[2024-05-16]. DOI:10.13700/j. bh. 1001-5965. 2024. 0088.

- [6] 姚聪颖,申静,汤浩威,等. 基于级联网络的输电线路螺栓故障检测[J]. 电子设计工程, 2024, 32(7):26-31. DOI:10.14022/j.issn1674-6236.2024.07.006.
- [7] 刘菲,钟延芬,邱佳伟. 基于改进YOLOv5s的轻量化交通标志识别检测算法[J]. 激光与光电子学进展, 2024, 61(24):2412004. DOI:10.3788/LOP240672.
- [8] LU J, HUANG T, ZHANG Q, et al. A lightweight vehicle detection network fusing feature pyramid and channel attention[J]. Internet of Things, 2024, 26:101166.
- [9] LI Y, LI J, NA T Y H, et al. Detection of attack behaviour of pig based on deep learning[J]. Systems Science & Control Engineering, 2023, 11(1):1101616.
- [10] 周秦汉,贾杰,陈昊,等. 基于SKNet改进YOLOv5s的无人机对道路小目标的检测[J]. 南昌航空大学学报(自然科学版), 2023, 37(4):39-45.
- [11] ZHANG Y, MA M, WANG Z, et al. POD-YOLO object detection model based on bi-directional dynamic cross-level pyramid network[J]. Engineering Letters, 2024, 32(5):1-10.
- [12] SENTÍS A M, WANG K, CAO Z, et al. GrapeMOTS: UAV vineyard dataset with MOTs grape bunch annotations recorded from multiple perspectives for enhanced object detection and tracking[J]. Data in Brief, 2024, 54:110432. DOI:10.1016/j.dib.2024.110432.
- [13] CUI L, JIAO W. Remote sensing image location based on improved YOLOv7 target detection[J]. Pattern Analysis and Applications, 2024, 27(2):50.
- [14] XU W, CUI C, JI Y, et al. YOLOv8-MPEB small target detection algorithm based on UAV images[J]. Heliyon, 2024, 10(8):18. DOI:10.1016/j.heliyon.2024.e29501.
- [15] YANG C, CAO Y, LU X. Towards better small object detection in UAV scenes: Aggregating more object-oriented information[J]. Pattern Recognition Letters, 2024, 182:24-30. DOI:10.1016/j.patrec.2024.04.002.
- [16] LUO Z, TIAN Y. Improved infrared road object detection algorithm based on attention mechanism in YOLOv8[J]. IAENG International Journal of Computer Science, 2024, 51(6):673-680.
- [17] 吴瑞林,葛泉波,刘华平. 基于YOLOX的类增量印刷电路板缺陷检测方法[J]. 智能系统学报, 2024, 19(4):1061-1070.
- [18] NANDHAKUMAR T, SWETHA S, THRISHA T, et al. Driver distraction and drowsiness detection based on object detection using deep learning algorithm[J]. SSRN Electronic Journal, 2024. DOI:10.2139/ssrn.4851819.
- [19] TAN R J, GAO Y, RAGHURAMAN R, et al. Application of deep learning algorithms in classification and localization of implant cutout for the postoperative hip[J]. Skeletal Radiology, 2025, 54(1):67-75. DOI:10.1007/s00256-024-04692-6.
- [20] ALAZEB A, CHUGHTAI R B, MUDAWI A N, et al. Remote intelligent perception system for multi-object detection[J]. Frontiers in Neurorobotics, 2024, 18:1398703.
- [21] SALEH A A R, AREQI A F, KONYAR Z M, et al. Advancing tire safety: Explainable artificial intelligence-powered foreign object defect detection with xception networks and Grad-CAM interpretation[J]. Applied Sciences, 2024, 14(10):4267.