

基于YOLOv8的人车智能检测网络

蒋春蕾¹ 张劲松¹ 张瑞龙²

1 西昌学院, 四川 西昌 6150002; 2 深圳元创兴科技有限公司, 广东 深圳 518107

摘要: 针对道路交通事故频发问题, 本文提出了一种基于YOLOv8的人车检测网络, 旨在实现对周围车道中的车辆与行人进行快速识别, 及时为行驶车辆提供预警信息。本文在数据上进行了随机遮挡和混合增强, 在YOLOv8网络中使用了GIoU损失函数和多尺度训练进一步增强网络性能, YOLOv8在自定义交通场景数据集中, 车辆检测精度达到94.70%, 行人检测精度达到90.30%, 在复杂背景、部分遮挡场景, 以及夜间、雨天等复杂条件中的检测精度均达到了87%以上, 并且检测速度约为50帧/秒, 能够实现目标的实时检测。

关键词: YOLOv8; 车辆检测; 行人检测

Intelligent Human and Vehicle Detection Network Based on YOLOv8

Jiang, Chunlei¹ Zhang, Jingsong¹ Zhang, Ruilong²

1 Xichang University, Xichang, Sichuan, 615000, China

2 Shenzhen Yuan Zhenxing Technology Co., Ltd., Shenzhen, Guangdong, 518107, China

Abstract: In view of the frequent traffic accidents, this paper proposes a vehicle detection network based on YOLOv8, which aims to quickly identify vehicles and pedestrians in the surrounding lanes and provide warning information for drivers in time. In this paper, random occlusion and mixed augmentation are performed on the data, and the GIoU loss function and multi-scale training are used in the YOLOv8 network to further enhance network performance. In a custom traffic scene dataset, YOLOv8 achieves a vehicle detection accuracy of 94.70% and a pedestrian detection accuracy of 90.30%. Detection accuracy in complex backgrounds, partially occluded scenes, as well as complex conditions such as night and rainy days, all reach above 87%, and the detection speed is about 50 frames per second, enabling real-time detection of targets.

Keywords: YOLOv8; Vehicle detection; Pedestrian detection

DOI: 10.62639/sspis55.20240103

引言

近年来随着社会经济的飞速发展, 以及人民生活水平的不断提高。根据中国汽车工业协会的统计数据, 过去的十年间国内乘用车销量快速上涨, 截止2023年底, 国内乘用车销量达到了约2170万辆。随着汽车数量的不断增加^[1], 保障公共安全变得尤为重要, 特别是在车流巨大的市区中, 车辆如何有效避免交通事故是公共安全中亟待解决的问题之一。若车辆的车载摄像头能及时有效识别周围车道中的行人与车辆, 及时提供预警制动信息, 就能降低交通事故发生的概率。较为传统的检测方法一般通过方向梯度直方图^[2] (Histogram of Oriented Gradient, HOG)、局部二值模式 (Local Binary Pattern, LBP) 等方法对目标图像的特征进行提取, 再结合支持向量机 (Support Vector Machines, SVM)、逻辑回归 (Logistic Regression, LR) 等分类方法实现对待检测目标的分类。

近年来, 基于卷积神经网络^[3] (Convolutional Neural Network, CNN) 的目标检测算法不断出现, 检测的精度和速度有了巨大的提升。彭红星等人^[4]通过引入高斯混合聚类和坐标注意力方法, 提出了一种改进的YOLOv5高速公路隧道车辆和人员检测算法; 胡倩等人^[5]在YOLOv5的骨干网络中引入了Conv-SPD网络模块, 并设计了全新的双层渐进金字塔网络DAFPN, 实现了对行人的高效检测; 梁奕延等人^[6]选用轻量化EfficientnetV2作为YOLOv5的骨干网络, 并引入GAM注意力机制, 提升了算法的检测效率。

目前, 基于CNN的目标检测算法可分为一阶段检测算法和两阶段检测算法。一阶段算法以YOLO系列^[7]和SSD算法为主, 这种类型的算法初始阶段就提取输入图像的特征, 并实现目标分类和边界框回归。此类算法对图像的推理速度较快, 但是检测效果不佳。二阶段算法以R-CNN算法系列为主, 此类型算法在输入的图像上会先生成一个区域候选框, 再进行特征提取, 最后对

(稿件编号: IS-24-3-J001)

作者简介: 蒋春蕾 (1979-01), 女, 汉族, 籍贯: 四川布拖, 机械与电气工程学院, 职称: 副教授, 硕士学位, 研究方向: 电气工程。

张劲松 (1997-07), 男, 汉族, 籍贯: 四川德昌, 机械与电气工程学院, 职称: 助教, 硕士学历, 研究方向: 电力系统自动化、智能信息处理。

张瑞龙 (1992-10), 男, 汉族, 籍贯: 甘肃静宁, 深圳元创兴科技有限公司, 职称: 工程师, 本科学历, 研究方向: 智能机器人。

基金项目: 四川省人工智能实验室2022年项目: “基于深度神经网络的目标检测定位研究” (项目编号: 2022RY03)。

目标实现分类与边界框回归操作，此类算法对目标检测的精度较高，检测速度却相对较慢。
 本文针对车辆与行人较为密集的市区交通情况，构建一种准确、快速的车辆行人检测网络，实现对车辆周围行人与车辆的高效检测，及时对驾驶员进行预警，尽可能避免交通安全事故。

一、YOLOv8 人车检测网络设计

(一) YOLOv8 网络结构

YOLO 系列算法将输入图像分为 NN 个区域，每个区域的锚框对目标类别进行预测及预测框的坐标信息。主要对特征网络进行改进，先后使用了 GoogleNet、DarkNet19 和 DarkNet53 对图像特征进行提取，并将特征图金字塔网络^[8] (Feature Pyramid Networks, FPN) 引入到算法中，逐步提高了检测精度。YOLOv8^[9] 结合了先前 7 个版本的优点，在检测速度和精度上都有更好的表现，其主要由输入端、骨干网络 (Backbone)、Neck 网络、Head 网络组成，YOLOv8 网络结构图如图 1 所示。

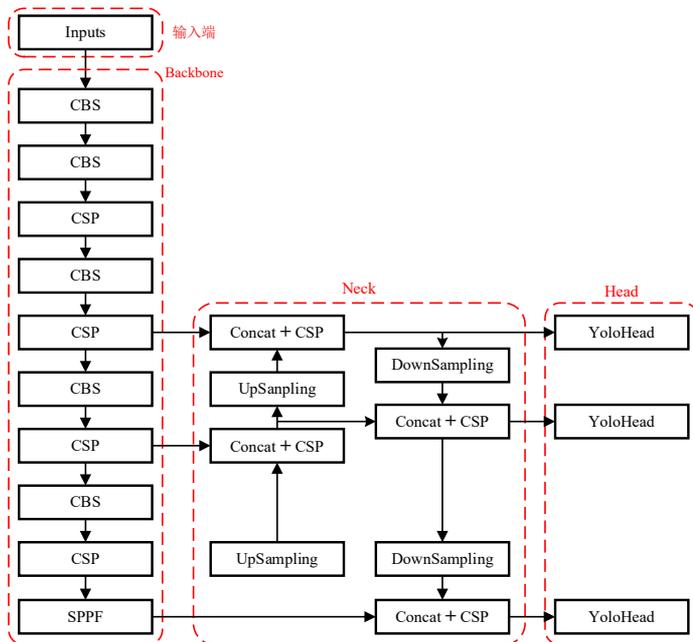


图 1 YOLOv8 网络结构图
Figure1

1. 输入端

主要作用是对输入网络的数据进行增强以及自适应锚框计算，主要通过平移翻转、裁剪拼接等操作，增加数据多样性。此外，YOLOv8 中，使用了 Anchor free^[10] 的锚框标定方法，对标定纵横比例较大的目标有明显优势。

2. 骨干网络

在输入端进行预处理后的图像被输入到 Backbone 中进行特征提取，生成特征图，获取图像中高维度的语义信息。Backbone 主要结构为 CSPDarkNet，如图 2 所示。

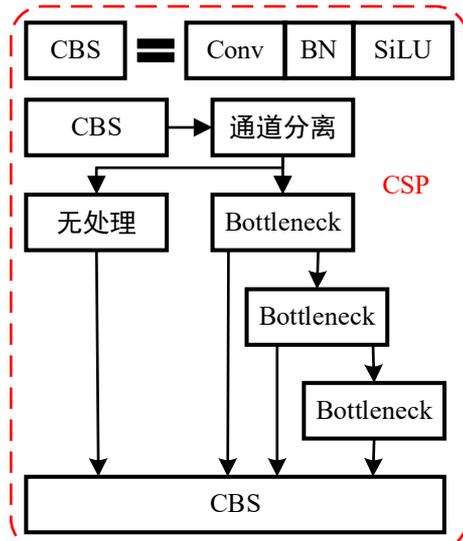


图 2 CSP 结构图
Figure2

在 Backbone 中，YOLOv8 还应用了空间金字塔快速池化结构 (Spatial Pyramid Pooling-Fast, SPPF)，如图 3 所示。SPPF 结构在空间金字塔池化 (Spatial Pyramid Pooling, SPP) 结构的首尾各增加了 1 个 CBS 模块，将并行的 Max Pooling 变为了串行与并行的方式，提高了池化效率。

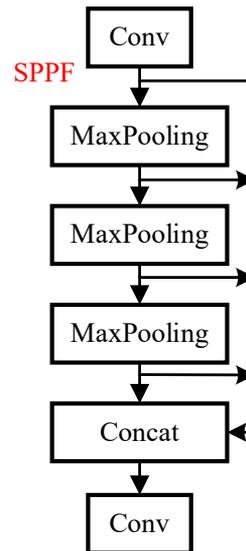


图 3 SPPF 结构图
Figure3

3. Neck 网络

Neck 网络中 FPN 网络对强 Backbone 有效特征层中不同尺度的特征信息进行特征融合，进一步增强了 Backbone 对图像的特征提取能力。

4. Head 网络

Head 网络将 Backbone 和 Neck 网络提取到的有效特征层视为一个特征点集合，将集合中的每个特征点作为先验点，根据该先验点判断生成的对应先验框是否与待检测目标对应，Head 网络是 YoloV8 的分类器与回归器。

(二) YOLOv8 网络优化

1. 损失函数优化

损失函数直接影响着模型的收敛速度和最终的检测精度。YOLOv8 初始采用交并比 (Intersection over Union, IoU) 损失函数, 该函数依靠计算预测框和真实框之间的重叠比例来衡量检测结果的准确性。该数值越接近 1, 表示预测框与真实框的重合度越高。

IoU 损失函数在处理小目标时存在不足, 特别是在预测框与真实框几乎没有重叠的情况下, IoU 值可能极小甚至为零, 难以借助优化 IoU 损失来有效调整小目标预测框, 且训练过程容易陷入局部最优。

为了解决这一问题, 本文引入基于广义交并比^[11] (Generalized Intersection over Union, GIoU) 的损失函数, GIoU 在 IoU 的基础上增加了对框间距离的度量, 不仅考虑了预测框与真实框的重叠情况, 还引入了包围框 (enclosing box), 能够完全包围预测框和真实框的最小矩形框。在计算 GIoU 时, 还计算预测框和真实框相对于包围框的面积比, 以此度量二者间的空间距离。GIoU 相比 IoU, 损失函数更关注于使得预测框和真实框的空间位置更加接近, 提高了小目标的检测精度。

2. 数据增强策略

本次研究为了增强模型对复杂场景的适应能力, 在传统数据增强策略的基础上引入了随机遮挡^[12]和混合增强策略^[13]。

随机遮挡 (Random Erasing) 是一种以随机选择图像的一部分并用特定颜色或噪声进行遮挡的方法, 不仅增加了训练数据的多样性, 还迫使模型学习到更加鲁棒的特征, 而不仅仅依赖于目标的全部信息进行检测。

混合增强 (Mixup) 策略的核心思想是在视觉特征空间中生成更多的中间状态样本, 使得模型能够更好地应对各种变换和噪声。将两张图像进行线性组合, 对其标签进行相应的线性组合, 从而生成包含多种视觉特征的样本, 使模型在学习过程中能够提取出更加普遍的特征, 提高其泛化能力。

3. 多尺度训练

目标检测任务中, 目标的大小变化是模型需要应对的重要难点, 尤其是在现实交通场景中, 车辆和行人可能出现在不同的距离和角度下, 导致目标在图像中的尺度变化较大。为了提高模型对不同尺度目标的适应能力, 本研究在 YOLOv8 的训练过程中采用了多尺度训练^[14]策略。

二、实验结果及分析

(一) 实验环境

本实验的硬件和软件配置如下表所示:

表 1 实验配置

Table 1

实验环境	具体配置
处理器 (CPU)	Intel Core i7-10700K, 8 核心 16 线程, 基础频率 3.8GHz

显卡 (GPU)	NVIDIA RTX 1080, 2560 个 CUDA 核心, 8GB GDDR5X 显存
内存 (RAM)	32GB DDR4, 频率 3200MHz
存储设备	1TB NVMe SSD, 高速读写能力
操作系统	Ubuntu 20.04 LTS
深度学习框架	PyTorch 1.7.1
CUDA/cuDNN	CUDA 11.1/cuDNN 8.0.4
Python 版本	Python 3.8

(二) 实验数据

如下表 2 所示, 训练和验证所用的数据集包括 COCO 数据集和自定义交通场景数据集:

1. COCO 数据集: 含有 80 类常见物体, 总计约 20 万张标注图像。此数据集用于模型的基础训练, 以增强模型的泛化能力。

2. 自定义交通场景数据集: 选取实际交通环境中的不同类型车辆和行人作为数据集, 该数据集涵盖不同天气、时间段及交通密度下多种交通场景中的小型车、公交车、卡车等车辆图像和行人图像。

表 2 数据集

Table 2

数据集	图像数量	类别数量	使用目的
COCO	200,000	80	基础训练与泛化能力测试
自定义交通场景	15,000	2	交通环境下的模型验证与优化

(三) 结果与分析

本文实验采用平均精度 (Average Precision, AP)、平均精度均值 (mean Average Precision, mAP) 和推理时间 (Inference Time) 作为主要性能评价指标。

1. YOLOv8 在 COCO 数据集上的性能

下表 3 展示了 YOLOv8 与 YOLOv5 在 COCO 数据集上的性能对比。实验结果显示, YOLOv8 在检测精度和推理速度方面均有显著提升。

表 3 YOLOv8 与 YOLOv5 在 COCO 数据集上的性能对比

Table 3

模型	mAP (%)	推理时间 (ms/frame)	精度提升 (%)	速度提升 (%)
YOLOv5	47.3	20.5	-	-
YOLOv8	50.5	18.0	3.2	12.2

根据表 3, YOLOv8 在 COCO 数据集上的 mAP 为 50.50%, 相较于 YOLOv5 提升了 3.20%, YOLOv8 的推理时间为 18.00ms/frame, 比 YOLOv5 的 20.50ms/frame 快 12.20%。YOLOv8 能够在保持较高精度的同时, 实现更快的检测速度。

2. YOLOv8 在自定义交通场景数据集上的性能

下表 4 为 YOLOv8 在自定义交通场景数据集上对车辆和行人的检测性能。

表 4 YOLOv8 在自定义交通场景数据集上的检测性能

Table 4

检测对象	AP (%)	推理时间 (ms/frame)
车辆	94.7	20
行人	90.3	20

由表 4 可知, YOLOv8 针对不同类型车辆

的 AP 值达到了 94% 以上。在行人检测任务中, YOLOv8 针对行人的 AP 值为 90.3%, 具有较高的检测精度, 证明其在行人检测中具有可靠性。YOLOv8 针对不同车辆和行人进行检测的 mAP 值为 92.5%, 具有较高的检测精度, 同时检测速度约为 50 帧 / 秒, 能够满足实施检测的需求, 检测结果见图 1。

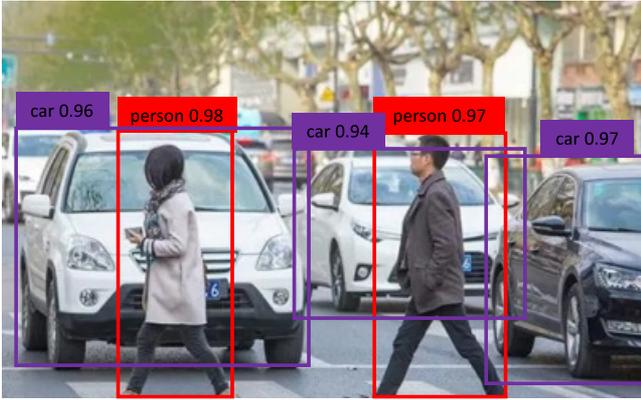


图 1 行人车辆检测

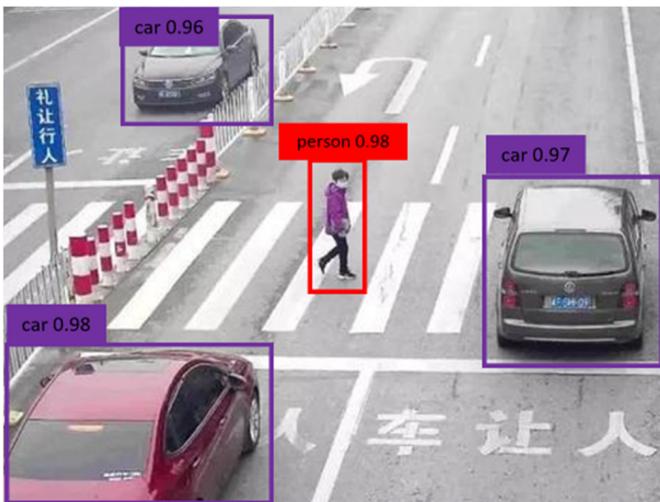
3. 行人检测效果分析

由于行人相较于车辆特征信息更为细微、不明显, 较难提取, 因此本文针对正常背景、复杂背景和部分遮挡场景中的行人进行了检测实验, 表 5 展示了 YOLOv8 在不同场景情况下的行人检测效果。

表 5 YOLOv8 在自定义交通场景中行人检测的精度表现 Table 5

场景类型	AP (%)
正常背景	92.5
部分遮挡	90.3
复杂背景	89.7

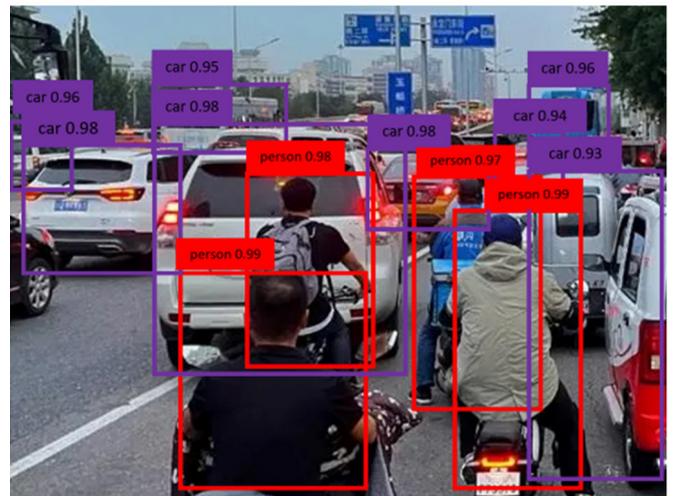
实验表明, 即使在复杂背景和部分遮挡的条件下, YOLOv8 的检测精度依然能够维持在 89% 以上, 能够较为准确的检测多种场景中的行人, 不同背景下的行人和车辆检测效果图见图 2。



a 正常



b 部分遮挡



c 复杂

图 2 不同背景检测结果

4. 复杂场景下的检测表现

表 7 展示了 YOLOv8 在夜间、雨天和拥堵路段等复杂场景中的 YOLOV 检测性能, 其中 AP1 表示车辆检测精度、AP2 表示行人检测精度。

表 7 YOLOv8 在复杂交通场景中的检测性能

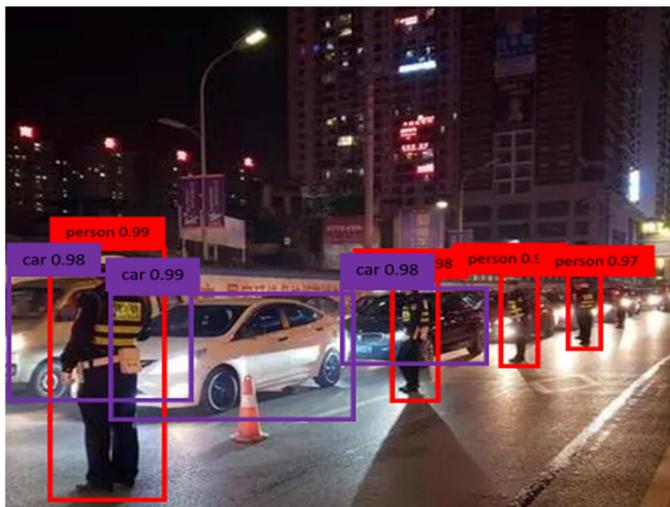
Table 7

场景类型	AP1 (%)	AP2 (%)	mAP (%)
夜间	91.5	88.7	90.10
雨天	89.2	86.5	87.85
拥堵	92.0	89.1	90.55

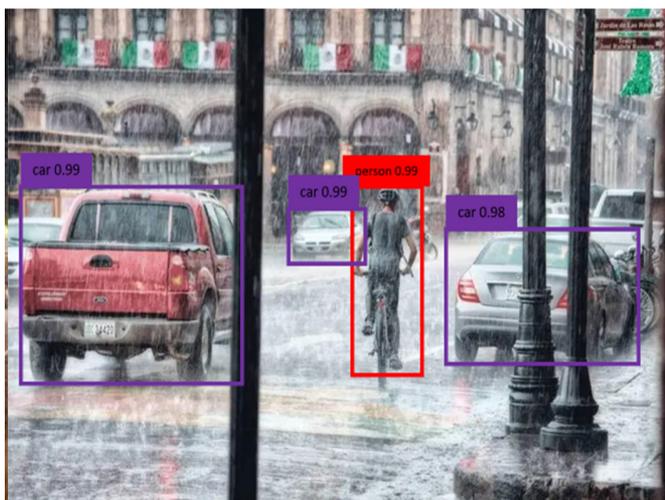
实验结果表明, YOLOv8 在对夜间、雨天、拥堵等复杂场景中的车辆和行人进行检测时, 对图像中的特征信息提取有所缺失, 导致检测精度有所下降, 但是车辆和行人检测的 AP 值仍然维持在 86% 以上, mAP 值仍然保持在 87% 以上, 能够完成检测任务, 具有较强的鲁棒性, 不同场景下的行人和车辆检测效果图见图 3。

三、结论

本文基于 YOLOv8 构建了车辆和行人检测网络, 并通过实验验证了其在复杂交通环境中的性能。相较于 YOLOv5, YOLOv8 在 COCO 数据集上的检测精度提升了 3.2%, 推理速度提高了 12.2%, 具有较高的效率和精度。本文在自定义交通场景数据集中, 使用随机遮挡和混合增强的方法对原始数据进行增强, 同时在 YOLOv8 的网络结构中使用 GIoU 损失函数和多尺度训练, 增强了网络对图像特征信息的提取能力。实验表明, YOLOv8 对车辆和行人检测的 AP 值分别达到了 94.7% 和 90.3%, 并在复杂背景、部分遮挡场景, 以及夜间、雨天等复杂条件下依然保持了较高的鲁棒性。因此本文设计的基于 YOLOv8 的人车检测网络具有一定应用价值, 能够为智能交通系统的安全运行提供一定技术支持。



a 夜间



b 雨天



c 拥堵

图 3 不同场景检测结果

参考文献:

- [1] 方茂东. 从比亚迪销量看新能源汽车趋势 [J]. 世界汽车, 2024, (07):52-55.
- [2] 宁爽, 宋辉. 帧间方向梯度直方图特征关联的行人检测方法 [J/OL]. 电子测量与仪器学报, 1-7 [2024-09-03].
- [3] 贺宏博. 基于深度学习的自动驾驶车辆目标检测与跟踪技术研究 [J]. 专用汽车, 2024, (07):61-65.
- [4] 彭红星, 袁畅, 柯威曳, 等. 基于改进 YOLOv5 的高速公路隧道车辆和人员检测 [J]. 科学技术与工程, 2024, 24(06):2453-2461.
- [5] 胡倩, 皮建勇, 胡伟超, 等. 基于改进 YOLOv5 的密集行人检测算法研究 [J/OL]. 计算机工程, 1-12 [2024-05-27].
- [6] 梁奕延, 陈昕, 郑明祥, 等. 基于 YOLOv5 的车辆目标检测算法轻量化改进 [J]. 汽车与新动力, 2024, 7(02):8-14.
- [7] 徐彦威, 李军, 董元方, 等. YOLO 系列目标检测算法综述 [J/OL]. 计算机科学与探索, 1-19 [2024-09-03].
- [8] 何志鹏. 基于全局渲染的特征金字塔网络目标检测算法 [J]. 兰州工业学院学报, 2024, 31(02):24-28.
- [9] 黄昆, 齐肇建, 王娟敏, 等. 基于改进 YOLOv8 的密集行人检测模型 [J/OL]. 计算机工程, 1-11 [2024-09-03].
- [10] 陈恒星, 刘一鸣. Anchor-free 目标检测算法综述 [J]. 机电工程技术, 2024, 53(08):7-12.
- [11] 朱伟, 段先华, 程婧怡. 基于 YOLOv3-Tiny 改进的船舶目标检测研究 [J]. 计算机应用与软件, 2024, 41(06):169-174+229.
- [12] 任坤, 李峥琪, 桂源泽, 等. 低分辨率随机遮挡人脸图像的超分辨率修复 [J]. 电子与信息学报, 2024, 46(08):3343-3352.
- [13] 陈孝国. 基于两阶段混合策略实现的低照度图像增强算法的研究 [D]. 南昌: 南昌大学, 2022.
- [14] 程传祥, 金飞, 林雨准, 等. 应用多尺度融合策略和改进 YOLOV5 的道路病害无人机检测 [J]. 地球信息科学学报, 2024, 26(08):1991-2007.