

复杂环境下自动驾驶汽车视觉目标检测模型性能评估

卢嘉伟

(同济大学, 上海 201804)

【摘要】为提高光照变化、模糊、噪声和遮挡等复杂环境下视觉感知系统目标检测模型的性能,基于YOLO 11x模型,结合数据增强方法,分析了目标检测模型在不同复杂环境下的性能变化情况,结果显示,不同类型的环境干扰对模型性能的影响各有特点,其中高斯噪声和运动模糊的影响较为显著,亮度调整和随机遮挡的影响相对较小。

关键词: 目标检测 复杂环境 数据增强 自动驾驶

中图分类号: U467.3 **文献标志码:** A **DOI:** 10.20104/j.cnki.1674-6546.20240417

Performance Evaluation of Visual Object Detection Models of Autonomous Driving Vehicle in Complex Environments

Lu Jiawei

(Tongji University, Shanghai 201804)

【Abstract】In order to enhance performance of visual perception system object detection model in complex environments such as lighting variations, blur, noise, and occlusion, this study investigates the performance variations of object detection models under different complex scenarios using YOLO 11x and data augmentation techniques. The results indicate that various environmental disturbances have distinct impacts in model performance, with Gaussian noise and motion blur having more pronounced impacts, while brightness adjustments and random occlusions are relatively less disruptive.

Key words: Object detection, Complex environments, Data augmentation, Autonomous driving

【引用格式】卢嘉伟. 复杂环境下自动驾驶汽车视觉目标检测模型性能评估[J]. 汽车工程师, 2025(4): 17-28.

LU J W. Performance Evaluation of Visual Object Detection Models of Autonomous Driving Vehicle in Complex Environments[J]. Automotive Engineer, 2025(4): 17-28.

1 前言

视觉感知系统是自动驾驶技术的核心之一,但复杂环境(如光照变化、模糊、雾霾和噪声)对其功能安全构成了重大挑战。这些干扰可能导致目标检测性能下降,影响系统的可靠性^[1]。尽管视觉检测模型在标准环境下表现良好,但其在真实复杂场景中的鲁棒性仍缺乏系统性分析。而且现有研究多采用理想数据集对视觉检测模型进行评价,难以全面反映真实道路条件下复杂多变的运行环境。然而,真实道路环境往往包含更多不确定因素。为此,中国智能网联汽车产业创新联盟(China Industry Innovation Alliance for the Intelligent and Connected Vehicles, CAICV)等行业组织也针

对实车环境感知系统功能安全开展了相应研究,并发布了实车测试评估体系^[2],为后续深入研究提供了一定的参考。

近年来,深度学习技术的快速发展显著提升了视觉感知系统的目标检测性能,YOLO系列^[3]和更快速区域卷积神经网络(Faster Region-based Convolutional Neural Network, Faster R-CNN)^[4]等模型在标准数据集上的表现接近人类水平。然而,这些研究多聚焦于理想环境,未充分考虑复杂场景中的干扰因素。一些学者尝试通过数据增强(如高斯噪声、模糊处理和亮度调整)模拟复杂环境,探索模型鲁棒性^[5]。虽然相关研究表明,数据增强能在一定程度上提升适应性,但其方法往往难以全面反映真实道路中多样化干扰的叠加效应;此外,对误检率、漏

检率等安全性指标的分析有限,缺乏基于真实道路场景的大规模验证,难以准确评估模型的实际性能。针对这些不足,本文基于YOLO 11x模型,结合数据增强技术模拟典型复杂场景,并设计试验评估模型性能,通过分析关键指标,揭示模型在不同场景下的性能变化规律,明确模型的鲁棒性问题及优化方向,以期视觉感知模型优化与系统改进提供实践依据。

2 理论分析

2.1 数据增强方法与复杂环境模拟

数据增强技术可通过模拟自动驾驶场景中的复杂环境,为评估目标检测模型的鲁棒性提供有效工具^[6]。本文针对典型的光照变化、模糊、遮挡及噪声等挑战,设计了多种增强方法,包括亮度调整、高斯模糊、运动模糊、雾化、噪声添加和随机遮挡等,以模拟极端光照条件、动态模糊场景、低对比度天气以及部分目标被遮挡或受噪声干扰的实际情况,为模型性能测试奠定基础。

2.2 复杂环境对视觉目标检测性能的影响

目前,视觉目标检测模型主要依赖卷积神经网络(Convolutional Neural Network, CNN)提取的特征,这些特征经过分类器和定位回归模块生成目标类别与检测框^[7]。然而,复杂环境会对特征提取和模型输出造成显著干扰^[8]。

2.2.1 光照变化

假设图像中像素 (x,y) 处的值 $I(x,y)$ 在亮度增强的条件下被调整为 $I'(x,y)=k \cdot I(x,y)$,卷积操作依赖于像素梯度的变化,其梯度表达为:

$$\nabla I'(x,y)=k \cdot \nabla I(x,y) \quad (1)$$

当 $I(x,y)$ 接近饱和值时, $\nabla I'(x,y)$ 趋近于零,导致目标区域的边缘特征消失,使模型的特征图响应减弱,影响目标定位。

在低亮度条件下,像素值 $I(x,y)$ 接近零,对比度 C 可表示为:

$$C=(I_{\text{tgt}}-I_{\text{bgd}})/I_{\text{bgd}} \quad (2)$$

式中: I_{tgt} 为目标物的像素值, I_{bgd} 为背景的像素值。

当 $I_{\text{bgd}} \rightarrow 0$ 时,对比度 C 显著下降,使得目标的梯度特征难以区分,降低分类和定位性能。

2.2.2 图像模糊

摄像头对焦不准或车辆振动引起的图像模糊可以通过高斯模糊进行建模,其核心在于对图像进行平滑操作^[9]。具体来说,模糊后的像素值 $I'(x,y)$ 是

原始图像像素值 $I(u,v)$ 与高斯核函数 $G(x-u,y-v)$ 的二维卷积结果,表达为:

$$I'(x,y)=\iint_{-\infty}^{\infty} I(u,v) \cdot G(x-u,y-v) \text{d}u \text{d}v \quad (3)$$

其中,高斯核 $G(x,y)$ 的定义为:

$$G(x,y)=\frac{1}{2\pi\sigma^2} e^{-\frac{x^2+y^2}{2\sigma^2}} \quad (4)$$

式中: σ^2 为高斯核的方差,用于控制模糊的程度,方差越大,模糊效果越强。

高斯模糊会削弱图像中边缘特征的显著性,其梯度幅值随 σ^2 的增加而减小。图像的梯度近似表达为:

$$\nabla I'(x,y) \approx \iint_{-\infty}^{\infty} \nabla I(u,v) \cdot G(x-u,y-v) \text{d}u \text{d}v \quad (5)$$

对于高斯模糊图像,梯度幅值的下降表现为卷积特征图中边缘响应的减弱,直接影响目标检测框的生成和定位精度^[10]。模糊对检测性能的影响与目标的尺度直接相关。对于大目标,即使边缘特征经过平滑处理,特征仍然具有较高的显著性,因此对定位精度的影响较小。但对于小目标,由于边缘梯度信息本身较弱,模糊会导致小目标的细节特征完全丧失,使得检测框的生成概率显著下降,甚至导致漏检。

对于由车辆的快速移动或相机的相对位移引起的图像拖影现象,运动模糊可视为沿某一方向的线性卷积,其影响可表示为:

$$I'(x,y)=I(x,y) \otimes K(x,y) \quad (6)$$

式中: $K(x,y)$ 为一维运动模糊核,其长度反映了模糊强度,方向决定了模糊的方向分布。

在模糊方向上,目标的边缘信息被拉长,导致边界像素梯度沿模糊方向减弱。这种现象会使卷积神经网络难以精确定位目标边界^[11]。

2.2.3 雾霾

雾化场景可以用大气散射模型描述:

$$I'(x,y)=J(x,y) \cdot t(x,y) + A \cdot (1-t(x,y)) \quad (7)$$

式中: $J(x,y)$ 为无雾图像的真实像素值, $t(x,y)=e^{-\beta d(x,y)}$ 为透射率, $d(x,y)$ 为像素点到摄像头的深度距离, β 为散射系数, A 为大气光值。

雾化降低了透射率 $t(x,y)$,使目标与背景对比度下降,导致分类器对目标类别的区分能力减弱。此外,小目标的透射率变化更剧烈,使得小目标在特征提取阶段可能被忽略。

2.2.4 图像噪声

在车辆环境中,传感器信号干扰可能由信号传

输过程中的电磁干扰或传感器硬件故障引发。这种干扰可以用高斯噪声进行建模,其特点是像素值在均值为零、方差为 σ^2 的正态分布下随机波动。模糊后的像素值可以表达为:

$$I'(x,y)=I(x,y)+N(0,\sigma^2) \quad (8)$$

式中: $N(0,\sigma^2)$ 为服从均值为0、方差为 σ^2 的正态分布的随机变量,代表高斯噪声。

高斯噪声对图像的干扰表现为全局性随机像素波动,破坏目标区域的局部特征连续性,使得卷积层难以提取清晰的边缘和纹理特征^[12]。这种特征的丢失可能导致虚假激活值的增加,模型可能将噪声波动误认为目标特征,生成错误的检测框。同时,在特征不完整或被随机波动掩盖的情况下,目标的检测框可能偏离真实位置,导致定位精度下降。

外部环境干扰(如降雨、降雪)或传感器硬件故障(如接触不良)可能导致图像中出现局部极端像素值。这种现象可通过椒盐噪声建模,其像素值随机变为极值(黑点或白点):

$$I'(x,y)=\begin{cases} 0, & r < \frac{p}{2} \\ 255, & \frac{p}{2} \leq r < p \\ I(x,y), & r \geq p \end{cases} \quad (9)$$

式中: $r \in [0,1]$ 为均匀分布的随机变量,用于模拟椒盐噪声; p 为椒盐噪声的概率参数,表示噪点覆盖的比例。

椒盐噪声的干扰集中在目标边缘和细节特征上,可能引发边界信息的破坏^[13]。随机黑白噪点可能使目标的边界信息断裂,特别是在小目标或复杂背景条件下,对分类和定位的影响尤为显著。此外,当噪点覆盖目标关键特征时,模型可能完全忽略目标(漏检)或将背景误分类为目标(误检)。

2.2.5 随机遮挡

遮挡会导致目标区域部分特征缺失、检测框定位精度降低。对于卷积层,遮挡相当于对输入信号的非均匀性干扰。设被遮挡的特征区域为 R_v ,模型的特征提取响应值可以表示为:

$$F = \int_{R_v} I(x,y) \cdot W(x,y) dx dy \quad (10)$$

式中: $W(x,y)$ 为卷积核的权重参数,定义了模型如何提取特征。

遮挡导致 R_v 减小,使得目标特征的累积响应值

降低,模型的分类准确性显著下降。

基于上述推导,可以总结出复杂环境对目标检测性能的主要影响:

- a. 光照变化:过曝导致边缘特征消失,暗光削弱目标与背景的对比度,均显著降低定位和分类性能。
- b. 模糊干扰:模糊削弱了目标边缘和纹理特征,特别是运动模糊对小目标和快速目标的位置识别损失最为严重。
- c. 雾化干扰:透射率的降低削弱对比度,对小目标和远距离目标的检测尤为不利。
- d. 噪声干扰:高斯噪声破坏特征的连续性,椒盐噪声破坏边界特征,均增加了分类错误和漏检的概率。
- e. 随机遮挡:遮挡导致特征不完整,类别混淆显著增加,对小目标影响尤为明显。

3 试验设计

3.1 数据集与模型选用

3.1.1 数据集的选用

本文选用KITTI 2D目标检测数据集作为数据来源^[14]。数据集包含了多种真实驾驶场景,包括城市道路、高速公路和乡村道路,并且涵盖了多类别目标(如车辆、行人和自行车)的高精度标注,如图1所示。



图1 KITTI数据集及标注示例

3.1.2 模型的选择和训练

本文选择Ultralytics YOLO 11x作为目标检测模型,该模型是YOLO系列的最新版本,具备较强的特征提取能力、环境适应性和较高的计算效率。YOLO 11x模型采用改进的骨干网络和特征融合结构,其架构包含631层,总参数量为56 883 016个,浮点运算量达到 195.5×10^9 次/s。相较于前代模型,YOLO 11x在保持较高检测精度的同时显著降低了

计算成本,适合处理自动驾驶场景中的复杂目标检测任务。

训练过程基于YOLO 11x模型的官方预训练权重文件,该权重文件通过在COCO数据集上进行大规模训练得到,COCO数据集涵盖了80个类别的目标检测任务,具有丰富的多样性和代表性。预训练权重使模型具备了较强的特征提取能力和初始检测精度,为后续的针对性训练提供了基础。本文试验中使用KITTI数据集对模型进行进一步训练,目标类别包含轿车(Car)、厢式货车(Van)、行人(Pedestrian)、骑行者(Cyclist)、卡车(Truck)、其他(Misc)、有轨电车(Tram)和坐姿人类(Person_sitting)共8类。训练过程对模型的类别数量进行了调整,使其与KITTI数据集的目标定义相匹配。

图2展示了模型在训练与验证过程中的边界框损失(Box Loss)、分类损失和分布式焦点损失(Distribution Focal Loss, DFL)的下降情况,以及交并比(Intersection over Union, IoU)阈值为0.5时的平均精度均值(mAP50)和交并比阈值为0.5~0.95时的平均精度均值(mAP50-95)的逐步提升。

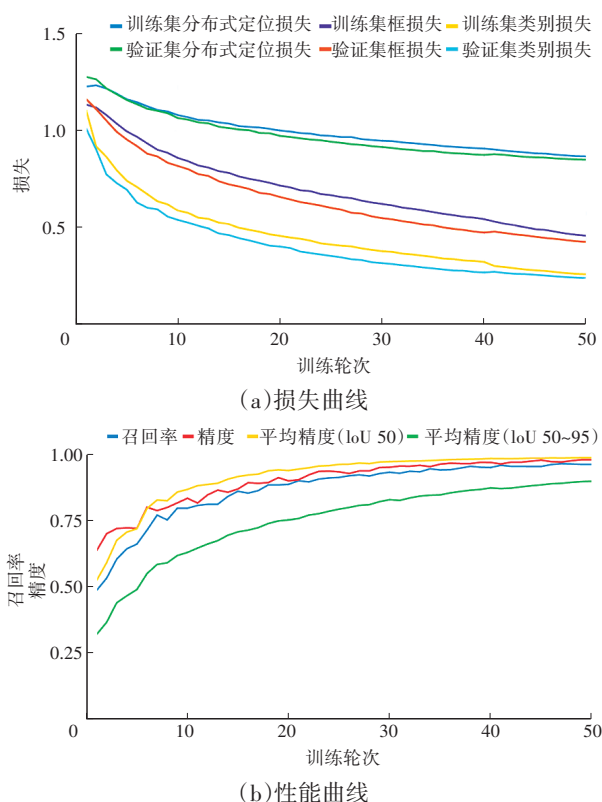


图2 训练过程中的损失与性能曲线

为了兼顾性能和硬件限制,输入图像尺寸设定为1 280像素×384像素,批量大小设置为8,并使用

AdamW优化器和自动混合精度(Automatic Mixed Precision, AMP)技术。

图3展示了模型在测试数据集上的精确率(Precision)-召回率(Recall)曲线,直观地反映了不同类别目标的检测性能,其中轿车(Car)和厢式货车(Van)等主要类别目标的精确率和召回率始终接近1,表现出极高的检测稳定性,而行人(Pedestrian)等目标在边界值处精确率略有下降,表明模型在处理稀疏目标时仍有一定改进空间。总体来看,总体的IoU50精度可以达到0.988,证明了模型在多类别任务中的高鲁棒性。

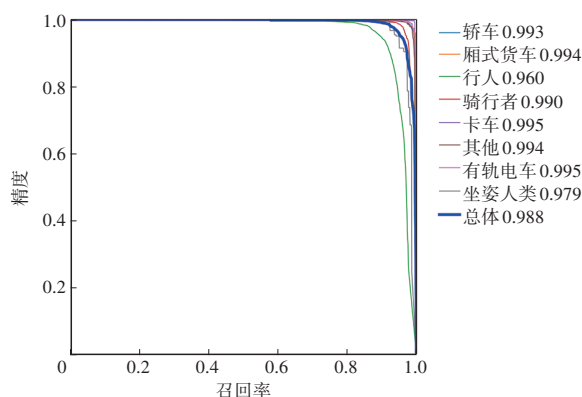


图3 预测结果精确率-召回率曲线

需要注意的是,上述性能表现是在训练和验证过程中使用的验证集上,以及IoU阈值为0.5的评估条件下得到的。在更严格的IoU阈值和更为复杂、多样化的原始数据集上进行独立测试时,模型对主要类别的检出率和定位精度可能会显著下降(如在原始数据中轿车的召回率为0.476 1)。因此,前文所描述的接近1的高精度、高召回率仅适用于训练/验证所使用的标准化环境和较宽松的评估标准,并不代表模型在所有场景下都能保持同样的性能。

为此,进一步分析模型在各类别上的分类性能,得到模型在测试数据集上的混淆矩阵如图4所示。轿车(Car)和行人(Pedestrian)等主要类别的对角线值较高,表明模型对这些类别的分类精确率和召回率非常高。此外,混淆矩阵还揭示了一些类别之间的误分类情况,例如卡车(Truck)/厢式货车(Van)和行人(Pedestrian)/骑行者(Cyclist)出现少量误分类,可能是由于目标的外观特征相似导致的。总体来看,模型对背景区域的误检测率较低,说明模型对背景干扰的抑制能力较强,体现了其分类鲁棒性。



图4 预测结构混淆矩阵

图5展示了F1分数(F1-Score)在不同置信度阈值下的变化趋势,进一步说明了模型的整体性能表现。对于轿车(Car)和厢式货车(Van)等大多数类别,F1分数在0.6~0.8的置信度区间内达到峰值,表明模型在该区间内很好地平衡了精确率和召回率。而对于行人(Pedestrian)等目标较小的类别,F1分数峰值略低,但总体依然表现出可靠的检测能力。从全类别的表现来看,在最佳置信度下,F1分数总体接近0.97,验证了模型在多种场景下的稳定性和检测准确性。

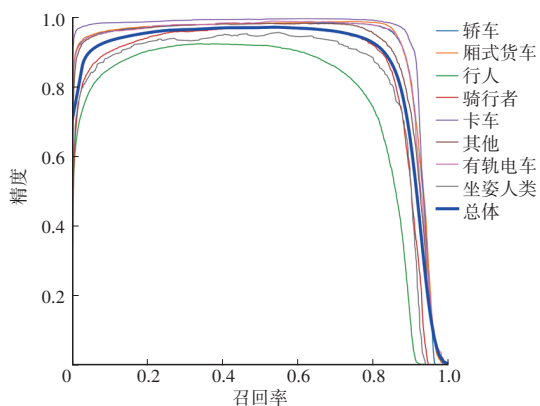


图5 预测结果F1分数与置信度曲线

需要指出的是,以上训练与验证阶段的性能指标是基于模型在训练过程中的验证集上测得的结果。该验证集的数据分布与模型训练数据较为接近,评估标准相对宽松,如IoU阈值设定为0.5(mAP50)、置信度与非极大值抑制(Non-Maximum Suppression, NMS)参数使用默认或较宽松的标准,从而在验证条件下展示出模型的较高指标和优异性能。这些结果并不能直接等同于在更复杂、严格条件下对原始数据进行评估时的

表现。

3.2 数据增强处理

3.2.1 数据增强的目标

自动驾驶场景中的视觉目标检测模型常面临复杂环境的挑战,如光照变化、天气干扰和硬件噪声等。这些因素可能显著影响模型性能,暴露出其在复杂场景中的局限性。

本为利用数据增强方法模拟典型环境干扰,以评估这些干扰对视觉目标检测模型的影响。

3.2.2 数据增强方法及实现

为模拟自动驾驶环境中视觉系统可能遇到的多种干扰,本文设计了7种数据增强方法对KITTI数据集进行处理,以有效提升模型的鲁棒性和泛化能力:

a. 亮度调整(Brightness Adjustment):对图像的像素值进行线性缩放,通过模拟光照条件的变化增强模型在光线极暗或过曝环境下的适应能力,亮度因子调整范围为0.3~2.0,并考虑了光线极端情况下的细节保留和视觉信息的还原。

b. 高斯模糊(Gaussian Blur):通过高斯滤波器对图像进行平滑处理,以模拟由于摄像头对焦不准或车辆振动导致的图像模糊,核大小范围设置为3×3至15×15,模糊程度随核大小的增加而增强。

c. 运动模糊(Motion Blur):通过构造一维运动核实现不同方向和强度的模糊,以模拟因车辆高速运动而产生的动态模糊,运动核大小在15~30的范围内随机设定,随机生成方向参数以适配多种场景。

d. 雾化(Fog):通过在图像上叠加不同浓度的雾层,降低图像对比度并增加亮度,模拟雾天环境对视觉系统的影响,雾化强度参数的范围为0.3~0.7,能够有效覆盖从轻雾到浓雾的多种天气条件。

e. 高斯噪声(Gaussian Noise):通过添加随机分布的像素值波动模拟传感器噪声对图像的干扰,噪声均值固定为零,标准差在5~20的范围内随机调整,以反映不同强度的传感器噪声。

f. 椒盐噪声(Salt and Pepper Noise):通过在图像中随机添加黑白噪声点,模拟传感器故障或图像数据采集中的外部干扰,噪声比例在0.02~0.07的范围内随机变化,确保增强图像中噪声的密度和分布能够覆盖多种实际场景。

g. 随机遮挡(Random Occlusions):通过在图像中随机添加若干矩形遮挡区域模拟物体被遮挡或镜头附着异物的情况,遮挡面积占图像的比例范围为5%~30%,遮挡区域的位置、大小及数量均为随机设定,以提高模型对遮挡场景的鲁棒性。

图6所示为原始图像和采用7种数据增强方法处理后图像的对比。通过对光学干扰、环境干扰和硬件问题的全方位覆盖,这些增强方法为模型的训练提供了更加丰富且具有挑战性的样本,为验证目标检测模型在复杂场景下的性能提供了数据支持。此外,由于所有增强方法被均匀应用于数据集,增强数据呈现出全面均匀分布,确保模型对各种干扰场景均具备良好的鲁棒性和泛化能力。



图6 数据增强效果示例

3.3 试验步骤

3.3.1 数据准备

试验的输入数据集包括原始KITTI数据集以及通过7种增强方法生成的增强数据集,所有增强版本的数据集均与原始数据集在样本数量上保持一致。每个数据集独立存储并编号,为后续推理阶段提供明确的输入来源。数据集的构建旨在通过全面覆盖复杂场景,确保试验结果的一致性。

3.3.2 推理与预测

在推理阶段,本文使用预训练模型对8组数据集进行逐一推理,生成每张图像的目标检测结果,包括检测框的位置、类别和置信度。具体推理流程为:

a. 图像预处理:为了与模型训练阶段的输入保持一致,所有推理阶段的输入图像被调整为1280像素×384像素的分辨率,以确保推理结果的可比性

和模型性能的有效验证。

b. 推理执行:预训练模型逐一处理输入图像,并输出每张图像的检测结果。模型首先根据输入图像生成原始检测框,随后应用NMS算法消除冗余检测框,使用IoU阈值(设置为0.7)保留最高置信度的检测框。此外,为了提高处理效率,推理阶段采用流式处理方式逐张图像处理,有效节省内存消耗。

c. 结果存储:推理生成的检测结果以2种方式存储。首先将检测结果存储为文本文件,包括每张图片中目标物的检测框信息,坐标以归一化XYWH(中心点坐标和宽度、高度)格式表示,同时附加目标类别和置信度信息。然后生成每张图像的可视化结果,形式为带有检测框和类别标注的图像,同时叠加真值框,用于验证检测结果与真值之间的匹配情况。

3.3.3 核心指标定义与计算

试验中,首先从真值文件与预测文件中分别读取每张图像的目标检测结果,包括类别与边界框坐标,逐图计算真正例(True Positives, TP)、假正例(False Positives, FP)和假负例(False Negative, FN)数量,并累积至全局指标。同时,逐类统计指标,为每个类别的检测表现提供精确评估结果。为评估模型在不同增强数据集上的性能表现,选用精确率 P 、召回率 R 和F1分数 F_1 作为核心指标:

$$P = N_{TP} / (N_{TP} + N_{FP}) \quad (11)$$

$$R = N_{TP} / (N_{TP} + N_{FN}) \quad (12)$$

$$F_1 = 2PR / (P + R) \quad (13)$$

式中: N_{TP} 、 N_{FP} 、 N_{FN} 分别为真正例、假正例、假负例的数量。

F1分数的取值范围为[0,1]。当精确率和召回率均较高时,其值趋近于1;当二者不平衡时,F1分数会显著降低,揭示模型在不同场景中的短板。

为细化分析模型的表现,本文将上述核心指标扩展至不同类别目标的层面。

4 结果分析

使用训练完成的模型对原始数据集以及增强数据集进行重新检测和评估。需要强调的是,此处评估条件与前文在训练/验证阶段的指标计算方法有所差异:首先,评估数据集可能包含更高难度的

真实场景,而非训练过程使用的验证集;其次,采用更严格的评估标准,如更高的IoU阈值0.8来匹配真实框与预测框,这意味着许多在IoU阈值为0.5标准下被视为正确的预测框,在更严格的IoU阈值标准下可能不再符合匹配条件,从而导致召回率和F1分数等指标明显下降;此外,小目标在真实且复杂的场景中更难被准确识别,进一步降低了整体检测性能指标。

4.1 定量结果

4.1.1 总体关键指标对比分析

原始数据和增强数据集预测结果如表1所示。

表1 原始数据和增强数据的检测结果关键指标

数据集类别	精确率	召回率	F1分数
原始图像	0.763 0	0.399 5	0.524 4
亮度调整	0.761 2	0.382 8	0.509 4
雾化	0.784 2	0.174 3	0.285 3
高斯模糊	0.744 7	0.272 5	0.399 0
高斯噪声	0.713 4	0.035 2	0.067 1
运动模糊	0.701 0	0.111 5	0.192 4
随机遮挡	0.670 6	0.357 6	0.466 4
椒盐噪声	0.704 2	0.166 2	0.268 9

精确率的整体变化从原始数据的0.763下降到增强数据集(如高斯噪声场景)的0.713 4,理论上,精确率的下降主要与背景区域生成的误检框增加相关。例如,高斯噪声引起的全局像素随机波动会干扰模型的特征提取过程,使模型对背景区域的误分类显著增加。然而,随机遮挡和高斯模糊对精确率的影响相对较小,这表明模型在一定程度上能够容忍局部遮挡和轻度模糊带来的误检风险。这一现象与理论分析一致:随机遮挡导致的特征缺失通常局限于局部区域,而高斯模糊对大目标的主要特征影响较弱,因此误检率较低。

召回率的下降在高斯噪声、运动模糊和雾化场景中尤为显著,召回率的急剧下降主要源于漏检框数量的大幅增加,其背后的机制包括:高斯噪声的随机像素波动掩盖了目标的局部特征,使得检测框与真实框的IoU降低,难以满足匹配条件($IoU < 0.5$);运动模糊导致目标边缘的显著性削弱,卷积网络无法准确提取目标形状信息,从而增加漏检情况;雾化降低了图像的对比特度,尤其是对

小目标的影响最为严重,进一步增加了漏检框的数量。

F1分数在增强数据集中的下降趋势也较为明显,这是精确率和召回率同时下降的结果,其中召回率的下降对小目标检测的影响尤为显著。相比之下,亮度调整和随机遮挡对F1分数的影响较小,这与理论分析结果一致:亮度调整的干扰较为温和,模型仍能捕捉到大部分有效特征;随机遮挡主要影响局部区域,对整体检测性能的冲击有限。

4.1.2 按类别关键指标对比分析

通过分解不同目标类别的关键指标,可以更深入地理解增强场景对模型性能的细化影响,并将其与理论分析建立关联。

作为大基数目标类别,轿车目标的检测结果精确率在不同增强场景下相对稳定,但其召回率和F1分数在高斯噪声和运动模糊场景中显著下降。理论上,这种下降主要是由于高斯噪声对目标边缘和纹理特征的掩盖,使得模型无法提取稳定的特征;而运动模糊则通过边界拉伸效应增大了检测偏差,使得模型在这些场景下的漏检率大幅增加。

对于行人类别目标,由于目标体积较小且边缘特征本身较弱,其精确率和召回率在所有增强场景中均出现大幅下降,尤其是在高斯噪声和运动模糊场景下。结合第2章的理论分析,小目标对梯度变化的敏感性更高,高斯噪声对像素级波动的干扰会破坏小目标的局部特征,使其在特征提取阶段被忽略,同时,运动模糊削弱了边缘梯度特征,从而导致严重的漏检现象。

其他类别(如厢式货车和卡车)在增强场景下的表现介于大目标与小目标之间,其精确率相对较高,但召回率显著下降。这表明模型在中型目标类别上的捕捉能力同样受到干扰限制,特别是在动态模糊或低对比度场景下。

表2展示了轿车和行人目标的关键检测指标对比结果。总体来看,轿车目标的检测性能表现相对稳定,这与理论分析结果一致:大体积目标的显著特征使其对环境干扰的鲁棒性较强。同时,行人目标的性能显著下降,特别是在高噪声和模糊场景中,暴露了模型在复杂环境下的短板。这进一步验证了理论分析中对小目标在梯度和交并比敏感性上的预测。

表2 轿车和行人目标的检测结果关键指标

目标类别	数据集类别	精确率	召回率	F1 分数
轿车	原始图像	0.825 6	0.476 1	0.603 9
	亮度调整	0.823 6	0.459 5	0.589 9
	雾化	0.836 7	0.223 2	0.352 4
	高斯模糊	0.820 8	0.331 7	0.472 5
	高斯噪声	0.762 7	0.046 4	0.087 5
	运动模糊	0.725 9	0.148 5	0.246 5
	随机遮挡	0.762 6	0.431 4	0.551 1
	椒盐噪声	0.767 0	0.213 4	0.333 9
行人	原始图像	0.344 5	0.124 1	0.182 5
	亮度调整	0.322 3	0.108 8	0.162 6
	雾化	0.238 1	0.026 7	0.048 1
	高斯模糊	0.260 8	0.079 6	0.121 9
	高斯噪声	0.357 1	0.006 7	0.013 1
	运动模糊	0.237 3	0.009 4	0.018 0
	随机遮挡	0.312 7	0.102 3	0.154 2
	椒盐噪声	0.289 3	0.031 2	0.056 3

4.1.3 F1 分数趋势分析

通过不同数据增强场景检测结果的 F1 分数变化趋势,可以清晰地观察模型在复杂场景下的性能变化特征,尤其是在不同目标类别之间的差异性表现。

整体 F1 分数在不同数据增强场景下呈现明显波动。原始数据集的 F1 分数为 0.524 4,表现出较为稳定的检测性能。但在高斯噪声(0.067 1)和运动模糊(0.192 4)等极端增强场景中,F1 分数显著下降,显示模型在这些场景下的鲁棒性较差。然而,对于亮度调整(0.509 4)和随机遮挡(0.466 4)等增强场景,模型的性能接近原始数据集水平,表明模型在处理这些相对温和的干扰时具有一定适应能力。

轿车(Car)类别作为大目标类别,其 F1 分数在原始数据集中表现最佳(0.603 9),在亮度调整(0.589 9)和随机遮挡(0.551 1)场景中,F1 分数相对保持稳定。然而,高斯噪声和运动模糊场景下的 F1 分数分别骤降至 0.087 5 和 0.246 5,表明模型在这些场景中对大目标的定位能力显著受损。行人(Pedestrian)类别由于目标较小,表现出明显的受干扰特性。在原始数据集中 F1 分数为 0.182 5,远低于轿车(Car)类别。经数据增强后,其性能进一步下降,尤其是在高斯噪声和运动模糊场景中,F1 分数仅为 0.013 1 和 0.018 0。这种现象反映了模型在小

目标检测中的显著短板,特别是在高噪声和动态干扰条件下。

小体积目标类别总体 F1 分数在所有增强场景中表现不佳。例如,坐姿人类(Person_sitting)类别在高斯噪声和雾化场景下 F1 分数甚至为 0,表明模型对这些类别在恶劣环境中的检测完全失败。而随机遮挡和亮度调整对这些类别的影响相对较小,F1 分数有所提升,但仍显不足。

图 7 所示为不同场景下各类目标检测结果的 F1 分数。多类别性能波动的特点可总结为:高斯噪声和运动模糊是所有类别性能下降的主要原因,导致目标检测的可靠性严重削弱;随机遮挡和亮度调整对目标类别的影响较小,表现出模型对这些增强场景的较高适应能力;大体积目标类别对增强干扰表现出一定的鲁棒性,而小体积目标类别则表现出显著的脆弱性,特别是在高噪声或动态场景下。

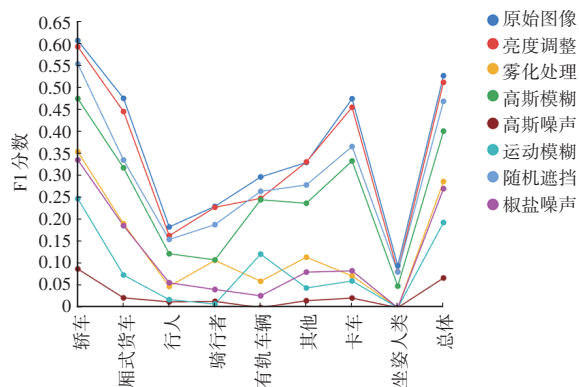


图 7 不同场景下各类目标检测结果 F1 分数

4.2 检测结果可视化分析

为了直观分析数据增强对目标检测模型性能的具体影响,本文结合检测结果的可视化展示,探讨不同增强场景对检测位置和分类的干扰效应,并分析其成因。

在原始数据集中,模型对“轿车”等主要类别表现出较高的检测精度和一致性。然而,在增强数据集中,复杂环境干扰显著影响了模型的检测性能,具体体现在目标定位偏差、类别误判和置信度下降等方面:

- a. 光照变化:在亮度调整场景中,模型的表现因光照变化而波动较大,强光条件(过曝)导致部分目标细节被淹没,使边缘梯度近似为零,从而导致漏检,如图 8a 所示;暗光条件下,目标对比度下降,背景和目标的特征变得难以区分,这直接降低了模

型的检测精度,如图8b所示。

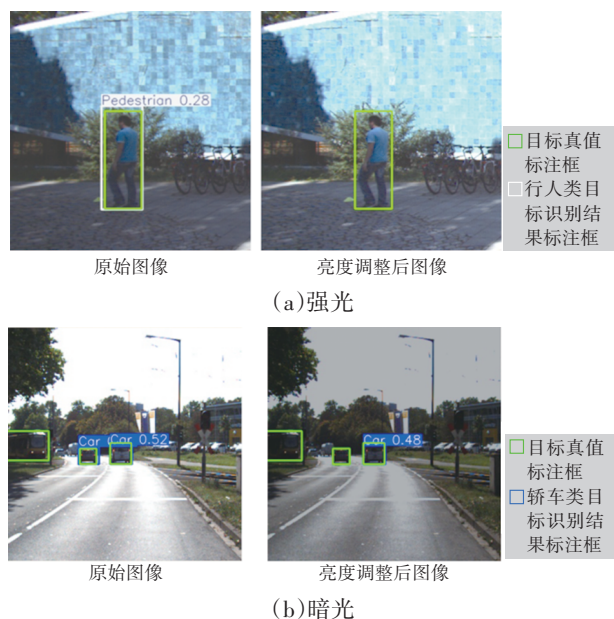


图8 亮度调整导致行人目标漏检

b. 模糊干扰:高斯模糊场景中,模糊削弱了图像的边缘信息,使目标定位和类别识别变得困难。理论上,高斯模糊降低了梯度幅值,尤其影响小目标的边缘特征。这在汽车类别的预测框边界错位以及行人和骑行者类别的显著漏检中表现得尤为明显,如图9所示。此外,运动模糊加剧了边界的拉伸和方向性丢失,使得检测框的位置偏差急剧增大,导致漏检或误检率显著上升。这种情况对动态小目标的影响最为严重。

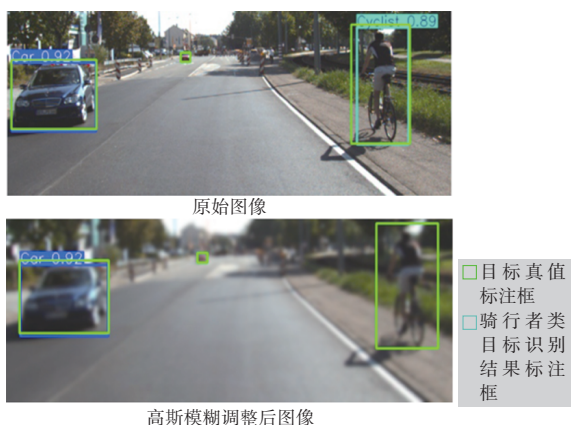


图9 高斯模糊调整导致骑行者目标漏检

c. 随机遮挡:随机遮挡场景中,部分目标区域因被遮挡导致特征缺失,这对模型的分类精度和定位能力提出了更高的挑战。理论上,遮挡会减少模型特征提取层的响应值,导致类别混淆和定位偏差。例如,部分背景在被随机遮挡后被误识别为厢式货车目标,如图10所示。



图10 随机遮挡导致厢式货车目标误检

通过上述可视化分析结果可以看出,不同数据增强场景对模型的干扰效应因目标类别和环境场景的不同而呈现显著差异。其中光照变化导致目标特征对比度减弱或边缘信息丢失;模糊(尤其是运动模糊)削弱了边缘梯度特征并导致位置检测偏差;遮挡显著减少了特征响应,导致分类错误和漏检的增加,特别是对小目标的影响尤为显著。

这些现象与理论分析结果一致,进一步验证了数据增强方法在暴露模型性能短板和识别潜在问题方面的有效性。

4.3 检测置信度分析

图11所示为目标检测结果在不同增强效果下的核密度估计(Kernel Density Estimation, KDE)曲线。根据总体置信度的均值、标准差和偏度,原始(Raw)数据集下置信度均值为0.769、标准差为0.202、偏度为-1.46,显示出置信度分布集中于高置信度区间([0.8,1.0])的特点。然而,在高斯噪声(平均值为0.706、标准差为0.221、偏度为-0.91)和雾化场景(平均值为0.747、标准差为0.211、偏度为-1.23)中,置信度均值显著下降,分布更加分散,偏度趋于平缓。这表明环境干扰削弱了模型的预测稳定性,在高斯噪声场景下表现尤为明显。

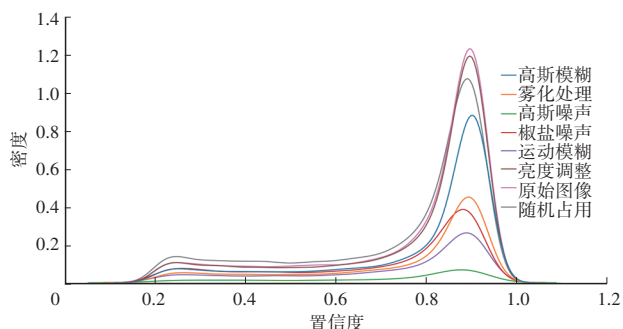
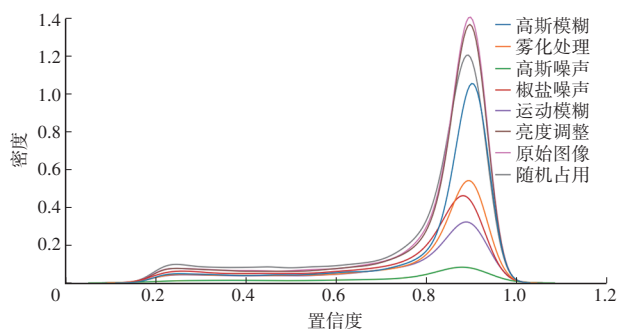


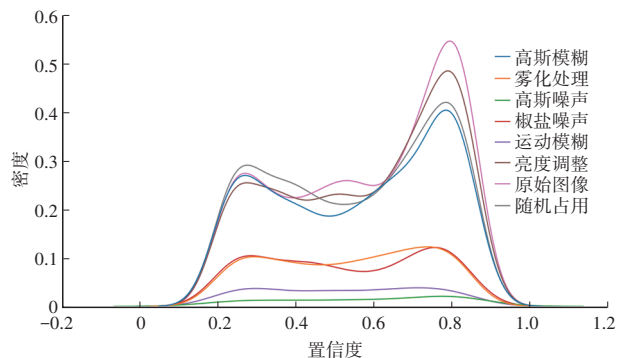
图11 检测结果总体置信度分布KDE曲线

图12所示为轿车类和行人目标检测结果KDE曲线。轿车类目标的检测置信度分布在大多数增强场景中保持相对稳定,其均值约为0.77,在随机遮挡和亮度调整场景中变化较小。然而,高斯噪声和运动模糊场景显著降低了置信

度均值,且标准差和偏度均有显著变化。这表明,尽管模型对大目标较为鲁棒,但复杂环境干扰对特定场景仍有较大影响。行人类目标的检测置信度均值显著低于轿车,在原始数据集中为0.58,但在高斯噪声和运动模糊场景中分别下降至0.46和0.44。同时,偏度趋于正值,显示出置信度分布更多集中在低置信度区间([0.0,0.5])。这一结果验证了小目标在高噪声和模糊环境下的不确定性更高,难以稳定检测。



(a)轿车目标



(b)行人目标

图12 轿车和行人目标置信度分布KDE曲线

图13所示为不同增强效果下检测结果置信度分布比例,呈现以下特点:

a. 高置信度区间([0.8,1.0]):原始(Raw)数据集中约66.8%的目标置信度位于高置信度区间。在高斯噪声场景中,该比例显著下降至52.3%。运动模糊和雾化场景中,比例分别下降至54.3%和61.0%,验证了复杂环境对模型置信度的削弱作用。

b. 中置信度区间((0.5,0.8)):高斯噪声场景中比例增至26.1%,运动模糊场景中也增加至25.8%,这与描述中的“26%和25%”一致,说明模型在这些场景下对目标的判断更加不确定。

c. 低置信度区间([0.0,0.5]):在高斯噪声场景中,低置信度目标比例达到21.7%,在运动模糊场景中为19.8%,与描述基本吻合。低置信度比例的增加进一步说明目标特征被噪声和模糊掩盖,导致漏

检和误检问题。

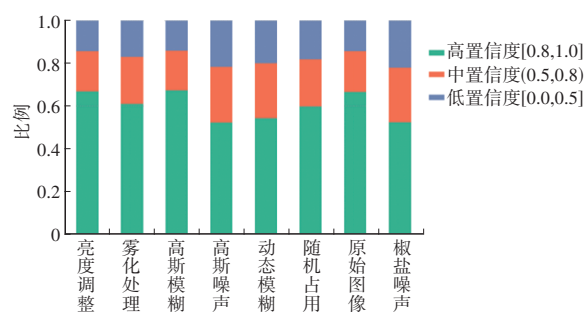


图13 不同增强效果下检测结果置信度分布比例

从统计结果和分布变化来看,不同增强场景对置信度分布的影响各异,噪声和模糊对模型的稳定性冲击最大,在小目标检测中表现尤为明显。这些结果与理论分析结果一致,进一步揭示了复杂环境中模型鲁棒性的问题。

5 对比与讨论

5.1 实车环境评价体系对比

本文基于标准数据集与单因素数据增强,针对不同干扰对视觉检测性能的影响开展了深入分析,但难以完全复刻真实道路的多重复杂性。《智能汽车感知系统预期功能安全评价体系与改进措施研究报告》^[2](简称“报告”)提出了一套面向实车的功能安全评价体系,覆盖多种天气、光照、道路类型等多维场景,并引入相关功能安全指标,用于评估感知系统在真实环境中的安全性能和鲁棒性。

表3给出了本文与实车环境评价之间的核心差异。可以看出,本文更注重拆分单一干扰的细化分析,研究不同环境因素的具体影响机理,而报告更关注多因素叠加时的综合表现,能够更全面地衡量系统在真实道路环境下的安全需求。两种评价方式各有侧重,也可形成互补。

5.2 实车环境数据对照分析

为进一步了解单因素数据增强与多因素叠加的实车场景在检测性能上的区别,本文结合报告中部分测试数据进行对照分析。报告指出,在夜间低照度或浓雾天气下,车辆检测精度和行人漏检率均显著恶化,降雨或路面湿滑等情况还会导致系统响应延迟与漏检率的同步升高。这与本文在“暗光、雾、高斯噪声、运动模糊”等干扰下发现的性能下降规律相吻合,表明环境干扰确实会显著削弱感知系统的准确度和稳定性。

表3 评价指标对比

指标维度	本文 (标准数据集评估)	报告 (实车环境测试评价体系)
基础指标类型	精确率、召回率、F1分数等图像级检测指标	实车场景下的误检率、漏检率随环境条件变化的统计、响应时间、最小可探测距离、感知稳定性等功能安全指标
环境场景	单一数据集(KITTI)下的标准场景与对象分布	多场景综合:不同天气(晴、雨、雪、雾)、光照条件(白天、夜晚、逆光)、道路类型(城市、高速、郊区)及交通密度
扰动因素建模	数据增强模拟单一因素(光照变化、模糊、噪声、遮挡)	实车环境中多因素叠加(天气变化、车速变化、传感器污染、路面湿滑、目标轨迹不规则)真实复杂工况的复合影响
测试流程与验证	以离线数据集评估为主,统一分辨率、交并比阈值和验证标准	从仿真测试、封闭场地测试到开放道路实车测试的渐进流程;强调在真实道路和随机场景下连续、多轮次测评,并记录响应延迟、决策稳定性
功能安全维度	无显式功能安全要求,仅关注检测精度与召回情况	强调功能安全要求:如在紧急条件下能否及时探测目标、可保证安全制动距离、在复杂动态场景中保持稳定的感知输出,减少潜在安全风险
优势与局限	优势:可定量剖析不同干扰的独立影响;试验条件易控制 局限:难以还原多因素交互;无法评估真实道路环境下传感器的可靠性	优势:综合真实场景多重干扰,贴近实际安全要求 局限:难以拆分具体干扰的影响;测试成本高、过程更复杂

然而,实车场景的多因素叠加(如车辆速度、道路湿滑、传感器脏污、光学畸变等)不仅降低了漏检率、加剧了响应延迟,也在一定程度上难以拆分每一种干扰对测试结果的影响所占的权重。相比之下,本文通过单因素增强,能够在可控条件下定量评估不同干扰的具体影响,可为后续算法鲁棒性改进提供针对性的方向。

综上所述,实车测试与本研究的单因素评估本质上针对不同的角度:前者可验证整体系统在真实复杂环境中的安全可靠,后者可帮助研究者细分干扰因素、精确分析模型在特定环境下的性能瓶颈。要真正提高自动驾驶视觉感知的功能安全水准,仍需将两者紧密结合,既在真实道路中开展多轮次、多天气、多交通流的测试,又要在实验室环境下针对性地深化模型在特定干扰场景下

的鲁棒性改进。

6 结束语

本文通过在标准数据集上应用多种数据增强手段(亮度调整、模糊、噪声、遮挡等),系统分析了复杂环境对自动驾驶视觉目标检测模型(YOLO 11x)的影响。结果表明,不同类型的环境干扰均会显著降低检测精度和稳定性,特别是对小目标类别造成更为严重的识别挑战。其中,高斯噪声和运动模糊导致的性能下降最为明显,而亮度调整和随机遮挡的影响则相对温和。

然而,本文仍存在若干局限:仅使用KITTI数据集与YOLO 11x模型,难以涵盖更大规模或更具多样性的场景与算法;单因素模拟虽便于量化分析各种干扰对检测性能的影响,却难以全面复刻真实道路中多因素叠加的复杂环境;对功能安全指标(如响应时间、最小可探测距离等)的考量尚不充分,无法全面评估模型在实车条件下的功能安全表现。

结合与实车环境测试数据的对比可见,真实道路环境中往往涉及车辆速度、天气突变、路面湿滑等多重干扰因素,远超单因素数据增强的范畴。未来工作将进一步拓展数据来源与传感器种类,引入多因素叠加及更多功能安全指标,开展真实道路条件下的综合测试与优化研究。通过多传感器融合、自适应阈值策略以及完整的功能安全评估流程,有望持续提升自动驾驶感知系统的鲁棒性与稳定性,从而更好地满足智能驾驶对预期功能安全的高标准要求。

参考文献

- [1] 胡明娣, 吴怡, 宋尧, 等. 单幅图像去雨数据集和深度学习算法的联合评估与展望[J]. 中国图象图形学报, 2022, 27(5): 1359-1391.
HU M D, WU Y, SONG Y, et al. The Integrated Evaluation and Review of Single Image Rain Removal Based Datasets and Deep Learning Methods[J]. Journal of Image and Graphics, 2022, 27(5): 1359-1391.
- [2] 中国智能网联汽车产业创新联盟智能网联汽车预期功能安全工作组. 智能汽车感知系统预期功能安全评价体系与改进措施研究报告-发布版[R/OL]. 北京: 中国智能网联汽车产业创新联盟, 2022.
CAICV SOTIF Technical Alliance. Research Report on the Evaluation System and Improvement Measures of the Perception System Safety of Intelligent Vehicles (Release Version) [R/OL]. Beijing: China Industry Innovation

- Alliance for the Intelligent and Connected Vehicles (CAICV), 2022.
- [3] REDMON J, FARHADI A. YOLOv3: An Incremental Improvement[EB/OL]. (2018-04-08)[2023-12-30]. <https://arxiv.org/abs/1804.02767>.
- [4] REN S Q, HE K M, GIRSHICK R, et al. Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2017, 39(6): 1137-1149.
- [5] HOWARD A, ZHU M, CHEN B, et al. MobileNets: Efficient Convolutional Neural Networks for Mobile Vision Applications[EB/OL]. (2017-04-17)[2023-12-30]. <https://arxiv.org/abs/1704.04861>.
- [6] WU Y L, XIANG Y X, TONG E D, et al. Improving the Robustness of Pedestrian Detection in Autonomous Driving with Generative Data Augmentation[J]. IEEE Network, 2024, 38(3): 63-69.
- [7] LI X T, WANG S Z, ZHAO Y, et al. Hierarchical Scene Coordinate Classification and Regression for Visual Localization[C]// Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Seattle, WA, USA: IEEE, 2020.
- [8] 孙金萍, 丁恩杰, 鲍蓉, 等. 多特征融合的长时间目标跟踪算法[J]. 南京大学学报(自然科学), 2021, 57(2): 217-226. SUN J P, DING E J, BAO R, et al. Long-Term Object Tracking Algorithm Based on Multi-Feature Fusion[J]. Journal of Nanjing University (Natural Science), 2021, 57(2): 217-226.
- [9] CHEN R J, ZHENG Z J, YU Y X, et al. Fast Restoration for Out-of-Focus Blurred Images of QR Code with Edge Prior Information via Image Sensing[J]. IEEE Sensors Journal, 2021, 21(16): 18222-18236.
- [10] FELZENSZWALB P F, GIRSHICK R B, MCALLESTER D, et al. Object Detection with Discriminatively Trained Part-Based Models[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2010, 32(9): 1627-1645.
- [11] PAN L Y, DAI Y C, LIU M M, et al. Joint Stereo Video Deblurring, Scene Flow Estimation and Moving Object Segmentation[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2019, 29: 1748-1761.
- [12] ZHOU Y, REN C, ZHANG S G, et al. A Second-Order Method for Removing Mixed Noise from Remote Sensing Images[J]. Sensors, 2023, 23(17).
- [13] KUMAR A, KUMAR S, KAR A. Salt and Pepper Denoising Filters for Digital Images: A Technical Review[J]. Serbian Journal of Electrical Engineering, 2024.
- [14] GEIGER A, LENZ P, URTASUN R. Are We Ready for Autonomous Driving? The KITTI Vision Benchmark Suite [C]// Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Providence, RI, USA: IEEE, 2012: 3354-3361.

(责任编辑 斛 畔)

修改稿收到日期为2024年12月18日。