

李彬彬, 丁纪峰. 基于 YOLOv5 改进的不良驾驶行为检测算法实现[J]. 智能计算机与应用, 2025, 15(5): 216-220. DOI: 10.20169/j. issn. 2095-2163. 250531

# 基于 YOLOv5 改进的不良驾驶行为检测算法实现

李彬彬, 丁纪峰

(大连民族大学 信息与通信工程学院, 辽宁 大连 116605)

**摘要:** 为解决已有不良驾驶行为检测算法中存在模型参数量过多、计算量较大、实时性差等不足, 并且很难部署到边缘设备等问题, 文章提出一种基于 YOLOv5 的不良驾驶行为检测改进算法 YOLOv5-MBi。在该模型中将 YOLOv5 主干网络替换为轻量级的 MobileNetV3 网络模块, 减少训练参数, 同时在 YOLOv5 中的 Neck 层结合 BiFPN 特征融合网络来提高其性能, 实现自上至下和自下至上的深浅层特征双向融合, 最后对改进的模型以及原网络进行相关的测试, 验证所修改方法的有效性以及实时性。实验结果表明, YOLOv5-MBi 算法在 State Farm 数据集上准确率达到 93.3%, 相比于原始的 YOLOv5s 算法, 参数量相对原有网络降低了 46.7%, 每秒传输帧数比原有网络提高了 53.3%。实验证明改进后的算法在保证较高检测准确率的同时, 模型参数量大幅下降, 能更好地满足检测实时性, 对交通安全方面具有重要的实际应用价值。

**关键词:** 不良驾驶行为检测; 轻量化; YOLOv5; MobileNetV3; BiFPN

中图分类号: TP391

文献标志码: A

文章编号: 2095-2163(2025)05-0216-06

## Implementation of bad driving behavior detection algorithm based on YOLOv5 improvement

LI Binbin, DING Jifeng

(School of Information and Communication Engineering, Dalian Minzu University, Dalian 116605, Liaoning, China)

**Abstract:** In order to solve the problems of excessive model parameters, high computational complexity, poor real-time performance, and difficulty in deploying to edge devices in existing bad driving behavior detection algorithms, the article proposes an improved algorithm for bad driving behavior detection based on YOLOv5, YOLOv5-MBi. In this model, the YOLOv5 backbone network is replaced with a lightweight MobileNetV3 network module to reduce training parameters. At the same time, the Neck layer in YOLOv5 is combined with the BiFPN feature fusion network to improve its performance, achieving bidirectional fusion of deep and shallow features from top to bottom and bottom to top. Finally, relevant tests are conducted on the improved model and the original network to verify the effectiveness and real-time performance of the modified method. The experimental results show that the YOLOv5-MBi algorithm has an accuracy of 93.3% on the State Farm dataset. Compared to the original YOLOv5s algorithm, the parameter count is reduced by 46.7% compared to the original network, and the number of frames transmitted per second is increased by 53.3% compared to the original network. Experiments have demonstrated that the improved algorithm can ensure high detection accuracy while significantly reducing the number of model parameters, which can better meet the real-time detection requirements and has important practical application value in traffic safety.

**Key words:** detection of bad driving behavior; Lightweight; YOLOv5; MobileNetV3; BiFPN

## 0 引言

随着时代的进步, 人们的生活水平不断提升, 汽车已成为一种普遍的出行方式。到 2022 年 3 月底, 国内的机动车总数已经超过 4.02 亿辆<sup>[1]</sup>, 这种飞速的增长无形中也对道路交通安全构成了严峻的

考验。

为了防止驾驶员的一些不良驾驶行为, 自从上世纪 70 年代至今, 很多科研人员都在研究不良驾驶行为检测技术, 因此不良驾驶行为检测一直以来都是国内外有关学者关注的重点。早期的检测方法大多是基于驾驶员生理特征和车辆运动特征进行的,

**作者简介:** 李彬彬(1999—), 男, 硕士研究生, 主要研究方向: 人工智能, 计算机视觉。

**通信作者:** 丁纪峰(1978—), 男, 副教授, 硕士生导师, 主要研究方向: 智能信息处理, 计算机视觉。Email: 535377131@qq.com。

收稿日期: 2023-10-16

但 these 方法容易受到外界条件以及设备成本等因素的影响, 导致实际应用价值较低。

近年来, 由于计算机硬件的不断进步, 深度学习已经成为许多领域的重要工具<sup>[2-4]</sup>。例如, Huang 等学者<sup>[5]</sup>通过应用多尺度特征提取、特征深度融合以及技术, 成功地研发出了一种新型的 CNN 模型, 可以有效地检测出分心驾驶的情况。尹智帅等学者<sup>[6]</sup>通过使用分类网络 Res Net 和 2D 姿态估计技术, 将其应用于分心驾驶行为的检测, 从而显著改善了检测的准确性, 但是 2D 姿态估计也存在只能构造有限模型的问题。Dey 等学者<sup>[7]</sup>首先利用机器学习算法从图像中检测并提取出可能引起驾驶分心的物体, 例如手机和水杯等, 并利用这些物体的相对位置来预测分心驾驶行为, 同时还目标检测算法引入了分心驾驶检测领域, 以提高检测效率和准确性。

以上基于深度学习的检测方法相对于前 2 种方

法在性能上有所提升, 但是检测算法的参数数量还是较多, 且检测速度较慢, 难以达到实时性, 因此本文提出了一种基于 YOLOv5s 网络模型改进的轻量级不良驾驶行为检测算法, 通过实验验证了 YOLOv5-MBi 模型的检测准确性与实时性。

### 1 YOLOv5 网络模型

YOLOv5 算法是一种检测速度、准确率都很高的不良驾驶行为检测算法<sup>[8]</sup>, 其中包括 YOLOv5n、YOLOv5s、YOLOv5m、YOLOv5l、YOLOv5x 五种不同深度和宽度的版本, 但各版本的网络整体架构都是一样的, 本文选择的网络模型为 YOLOv5s 6.0 版本。YOLOv5s 的网络结构由 4 个主要部分组成: 输入端、骨干网络、颈部网络和检测推理部分。其网络架构如图 1 所示。

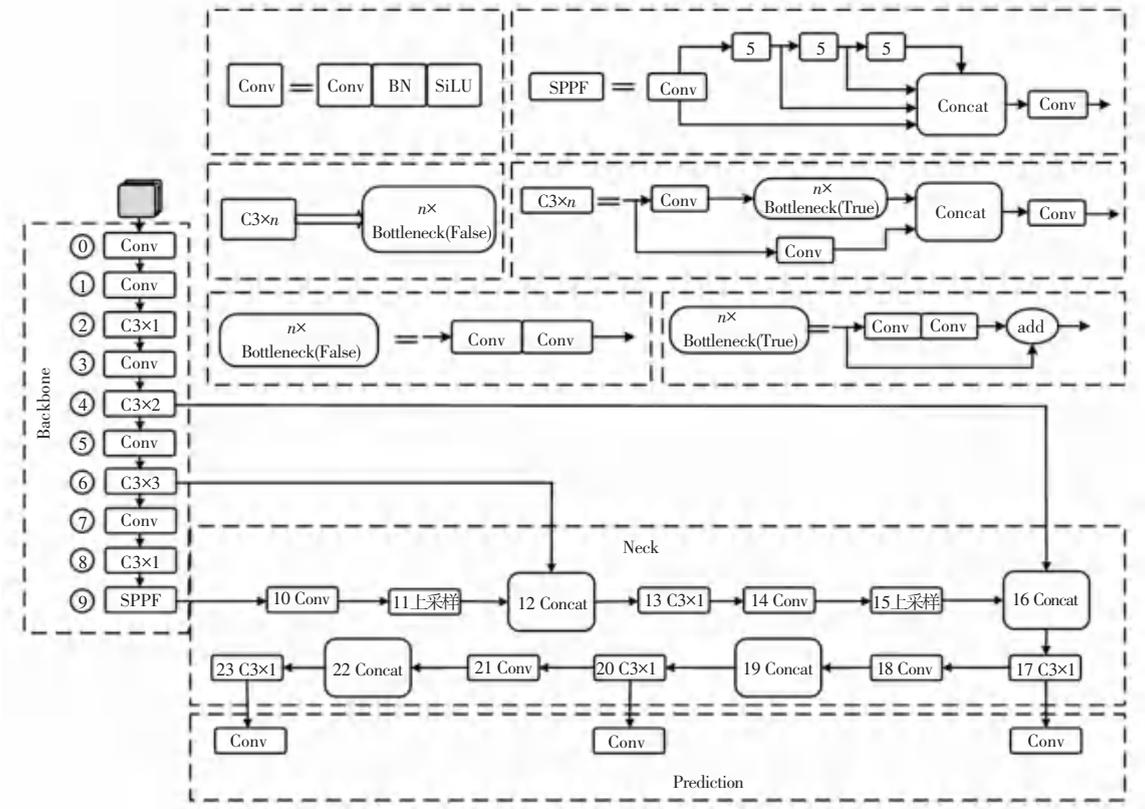


图 1 YOLOV5s 的网络架构图  
Fig. 1 Network architecture diagram of YOLOv5s

YOLOV5s6.0 的骨干网络主要分为切片单元 (Focus)、卷积模块 (Conv)、瓶颈层 (C3) 和快速空间金字塔池化 (SPPF)<sup>[9]</sup>。切片结构在高分辨率图像中获得像素点, 然后再构成低分辨率图像, 该模块主要用来加快计算速度。SPPF 模型采用了 5/5/5 的池化方法, 并且使用 Concat 融合来提高感受野。

YOLOV5s 在颈部网络中采用了 FPN+PAN 架构, 这一架构的设计比之前的版本更具有灵活性, 而且还新增了 sp2 架构, 从而大大提升了特征融合的效果。最后, 将损失函数 *GloU\_Loss* 应用于输出端, 可以有效地减少预测框和真实框之间的面积差异, 从而提升检测效率。

## 2 改进后的 YOLOv5-MBi 网络模型

本文将 MobileNetV3 网络替换 YOLOv5 中的骨干网络,使检测模型更加轻量化,提高模型检测的速

度。同时在 YOLOv5 中的 Neck 层结合 BiFPN 特征融合网络来提高 YOLOv5 算法检测准确度。YOLOv5-MBi 的网络结构如图 2 所示。

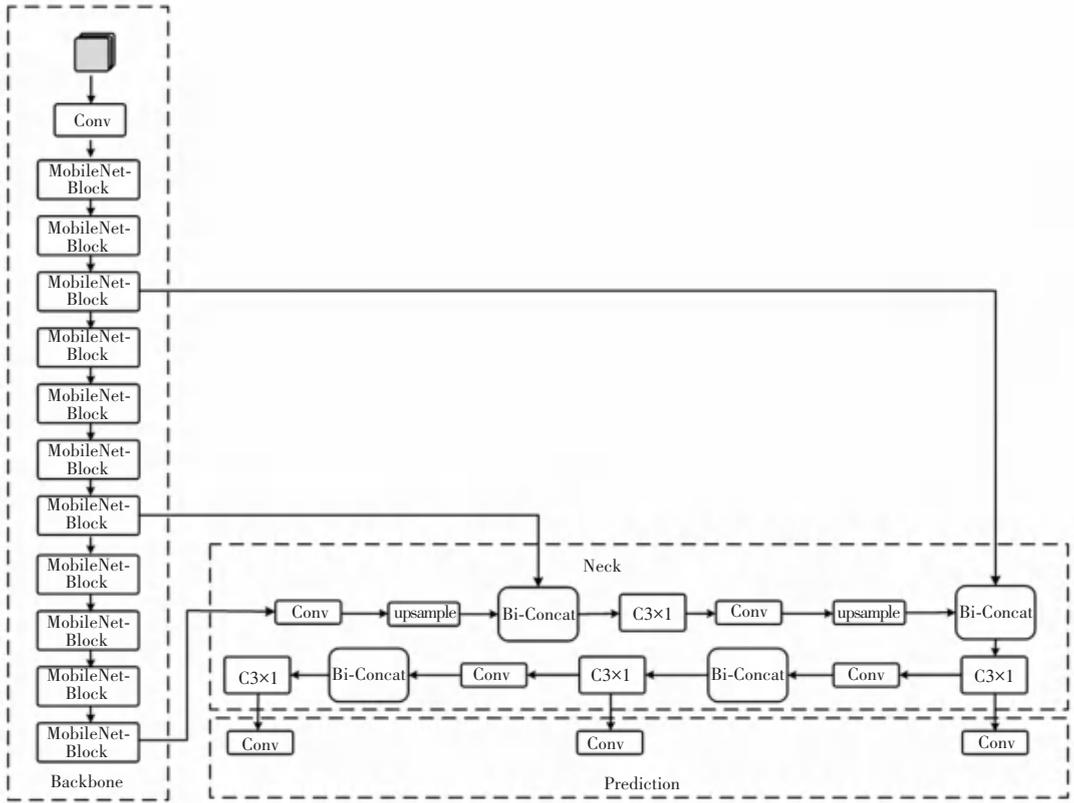


图2 YOLOv5-MBi 的网络结构图

Fig. 2 YOLOv5-MBi network structure diagram

### 2.1 轻量化网络模块

轻量化网络模块 (MobileNetV3) 的基本结构与传统卷积神经网络一致,包含卷积层、激活函数、批量归一化层、池化层和分类器, MobileNetV3 网络保留了 MobileNetV1 网络中的深度可分离卷积,同时也保留了 MobileNetV2 网络中的线性瓶颈的逆残差

结构,这样可以在避免信息损失的同时减少了计算量。并且 MobileNetV3 网络模型中添加了轻量级的 SE 注意力模块,将原有的 Swish 函数替换为 Hswish 函数,使模型的检测精度与效率得到了一定的提升。其中, MobileNetV3 基本模块如图 3 所示。

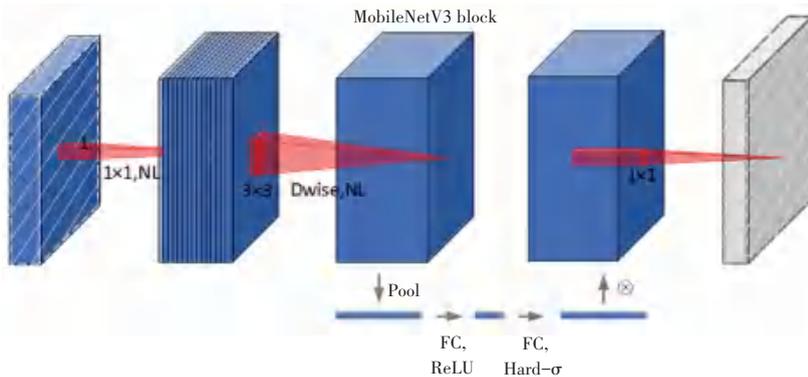


图3 MobileNetV3 基本模块

Fig. 3 Basic module of MobileNetV3

首先, MobileNetV3 基本模块采用  $1 \times 1$  卷积核进行特征提取, 在特征提取完成后采用  $3 \times 3$  的卷积核进行深度可分离卷积, 再经过 SE 注意力模块优化特征图数据, 在提高有价值特征的同时减少无用特征, 最终通过  $1 \times 1$  的卷积核, 实现特征的降维处理。这样既减少了网络运算量, 还提高了网络的计算效率。

### 2.2 双向加权特征金字塔网络

在 YOLOv5 网络模型中, Neck 层使用的是 FPN+PAN 结构, FPN 结构如图 4(a) 所示, FPN 结

构采用自顶向下横向连接的方式, 将高层特征传递到低层, 并进行特征融合。PANet 引入了自底向上的连接方式, 从而使底层特征信息更容易传递到高层。FPN 与 PAN 结合, 对来自不同骨干层的不同检测层进行参数聚合。虽然 FPN+PAN 结构提高了网络的特征融合能力, 但由于 PAN 结构输入的都是 FPN 结构输出的特征信息, 这就导致骨干特征提取网络的部分原始特征信息存在部分丢失现象。为了解决这一问题, 采用了双向加权特征金字塔网络 (BiFPN), 其结构如图 4(c) 所示。

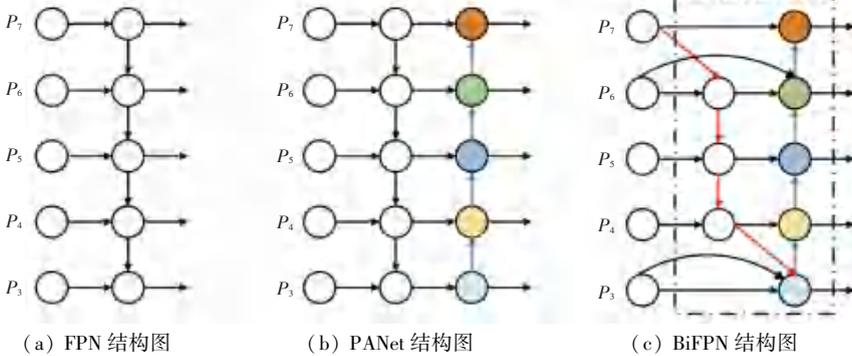


图 4 FPN、PANet、BiFPN 结构图

Fig. 4 Structure diagram of FPN, PANet, BiFPN

BiFPN 采用双向融合思想重新构建自上而下与自下而上的路径, 对不同尺度的特征信息进行了融合, 并且在同一尺度的特征图间建立了双向连接, 解决了部分特征信息丢失问题, 可以较好地对不同大小特征图进行特征信息融合。由 BiFPN 的结构图可知, 由于只有一个输入的节点对特征融合的贡献度较低, 所以为了简化网络, BiFPN 将其删去; 在原输入节点与输出节点间添加一条边, 目的是为了在不消耗成本的基础上进行更多的特征融合; 在模块中整合自上向下与自下而上的路径, 来达到更高层次的特征融合。

### 3 实验与结果分析

为了验证 YOLOv5-MBi 算法的检测准确性和

实时性, 以及与原 YOLOv5 之间的区别, 本次实验在 State Farm 公开数据集上进行算法的训练与测试。

#### 3.1 不良驾驶行为数据集

驾驶汽车时, 驾驶员可能会表现出多种行为, 其中一些行为可能会对其行车安全构成潜在威胁。在本文中, 主要研究以下驾驶行为: 正常驾驶、打电话、查看手机、喝水等情况。这些驾驶行为都导致不同程度的事故发生。本文采用的数据集为 Kaggle 机器学习竞赛平台公布的驾驶行为数据集 State Farm。State Farm 的数据集中包含 10 种不同类型的驾驶行为, 数据集中共标注了 22 424 张照片<sup>[10]</sup>。数据集中的部分图像如图 5 所示。



图 5 驾驶行为的示例图像

Fig. 5 Example images of driving behavior

State Farm 数据集中的样本类别数量分布相对均衡,因此无需特别关注类别分配是否不平衡问题。表1为 State Farm 数据集各图像类别对应数量。

表1 State Farm 数据集中各类别数据对应数量

Table 1 Corresponding quantity of data for each category in the State Farm dataset

标签序号	图像类别	数量
c0	Safe-driving	2 489
c1	right hand typing	2 267
c2	call with the right hand	2 317
c3	left hand typing	2 346
c4	call with the left hand	2 326
c5	turn the radio	2 312
c6	drinking	2 325
c7	taking the things behind	2 002
c8	finishing hair and make up	1 911
c9	talking to other passengers	2 129

### 3.2 实验环境

为验证各模块的作用与相互之间的影响,针对 MobileNetV3 与 BiFPN 模块进行实验,并在 State Farm 分心驾驶行为数据集上进行训练和测试。在本实验中,将学习率折减系数设置为 0.2,初始学习率设置为 0.01,动量系数设置为 0.937, batch size 设置为 16, epochs 设置为 200 轮。实验平台配置见表 2。

表2 实验平台配置

Table 2 Experimental platform configuration

名称	配置
操作系统	Windows 10 专业版
CPU	2.80 GHz Intel(R) Core(TM) i7-7700HQ
GPU	Nvidia GeForce GTX 1050 Ti(4 G 显存)
深度学习框架	Pytorch1.9.0
加速环境	CUDA11.1.1
编程语言版本	Python 3.8

### 3.3 评价指标

为了能够准确评估改进网络模型的检测效果,本次实验中使用准确率( $P$ )、召回率( $R$ )、平均精度( $mAP$ )、模型大小、帧率( $FPS$ )作为评价指标,其中  $P$ 、 $R$ 、 $mAP$  的计算公式具体如下:

$$P = \frac{TP}{TP + FP} \quad (1)$$

$$R = \frac{TP}{TP + FN} \quad (2)$$

$$mAP = \frac{\sum_{i=1}^N AP_i}{N} \quad (3)$$

其中,  $TP$ 、 $FP$ 、 $FN$  分别表示正类正样本、正类负样本以及负类正样本;  $AP$  表示单个类别的平均准确率;  $N$  表示总类别的数量;  $mAP$  表示对准确率和召回率的综合评价指标,  $mAP@0.5$  表示  $IoU$  为 0.5 时的  $mAP$ 。FPS 指的是画面每秒传输帧数。FPS 值反映了模型的检测效率, FPS 值越高,则网络模型的检测效率越好。

### 3.4 实验结果

本实验以准确率 ( $Precision$ )、召回率 ( $Recall$ )、 $mAP@0.5$ 、模型大小、每秒传输帧数  $FPS$  为检测指标。本实验共训练了 4 个网络模型,分别为原始的 YOLOv5s、YOLOv5s-MobileNetV3、YOLOv5s-BiFPN 和 YOLOv5s-MBi,在 State Farm 数据集上进行训练后,得到了在数据集上的检测效果如图 6 所示,最后得到的性能指标见表 3。



图6 检测结果

Fig. 6 Detection results

表3 性能指标对比表

Table 3 Performance index comparison table

网络模型	$P/\%$	$R/\%$	$mAP@0.5/\%$	模型大小 /M	$FPS$
YOLOv5s	93.4	93.1	93.4	13.7	15
YOLOv5s-MobileNetV3	91.8	92.6	91.9	7.1	25
YOLOv5s-BiFPN	95.1	95.2	95.6	14.1	12
YOLOv5s-MBi	93.3	92.3	93.2	7.3	23

从表3中可以看出, YOLOv5s-BiFPN模型与原始的YOLOv5s模型相比, 准确率、召回率与*mAP*分别提升了1.7%、2.1%和2.2%, 模型大小增加了2.9%; YOLOv5-MobileNetV3对比YOLOv5s, 准确率降低了1.6%, 召回率降低了0.5%, *mAP*降低了1.5%, 但是模型大小降低了48.1%, *FPS*有了大幅提升, 提高了66.7%。YOLOv5-MBi模型相比于原始网络, 虽然准确率降低了0.1%, 但模型大小降低了46.7%, *FPS*提高了53.3%。YOLOv5-MobileNetV3模型相比于原有YOLOv5s模型*FPS*的大幅提高, 说明MobileNetV3 block模块具有减少模型运算量、提高模型检测速度的作用; YOLOv5-BiFPN网络相比于原有YOLOv5s模型检测准确度有所提高, 但同时带来了参数量变多和运算时间降低的问题。在原有YOLOv5s模型的基础上结合了MobileNetV3模块和BiFPN模块的YOLOv5-MBi的模型参数量已经大幅减少, 模型也更加轻量化, 大大提高了驾驶员不良驾驶行为检测的实时性。

## 4 结束语

由于目前不良驾驶行为检测算法的参数量较多<sup>[11]</sup>, 检测速度较慢, 难以达到实时性的要求。为了弥补YOLOv5算法的不足, 本次研究在其基础上进行了2处改进: 首先, 将YOLOv5s的骨干网络替换为更加轻量化的MobileNetV3网络, 以此来降低训练参数, 从而大大加快检测的效率。其次, 将YOLOv5中的Neck层结合BiFPN特征融合网络来提高其性能, 从而有效地解决不良驾驶检测中的精度和速度失衡的问题。经过实验对比, 由表3可知, YOLOv5-MBi算法相比于原始的YOLOv5s算法, 能在保证较高检测准确率的同时, 参数量相对原有网

络降低了46.7%, 每秒传输帧数比原有网络提高了53.3%。体现了改进后的YOLOv5-MBi算法更加轻量化, 能更好地满足检测实时性, 对交通安全方面具有重要的实际应用价值。

## 参考文献

- [1] 邱彬, 王芳. 2023年中国汽车行业发展趋势展望[J]. 汽车工业研究, 2023(1): 2-9.
- [2] 杨晓峰, 邓红霞, 李海芳. 基于计算机视觉的驾驶人低头行为检测[J]. 计算机科学, 2016, 43(6): 210-213.
- [3] 王宏雁, 赵明明, BEURIER G, 等. 汽车驾驶人姿态监测系统研究综述[J]. 中国公路学报, 2019, 32(2): 1-18.
- [4] LETH N, ZHENG Yutong, ZHU Chenchen, et al. Multiplescale faster RCNN approach to driver's cell-phone usage and hands on steering wheel detection [C]//2016 IEEE Conference Computer Vision and Pattern Recognition Workshops (CVPR). Piscataway, NJ: IEEE, 2016: 46-53.
- [5] HUANG Chen, WANG Xiaochen, CAO Jiannong, et al. HCF: A hybrid CNN framework for behavior detection of distracted drivers [J]. IEEE Access, 2020, 99: 109335-109349.
- [6] 尹智帅, 钟恕, 聂琳真, 等. 基于人体姿态估计的分心驾驶行为检测[J]. 中国公路学报, 2022, 35(6): 312-323.
- [7] DEY A K, GOEL B, CHELLAPPAN S. Context driven detection of distracted driving using images from in-car cameras [J]. Internet of Things, 2021, 14(6): 100380-100391.
- [8] WANG X Y. Design of traffic signal intelligent control device based on Yolov5 object detection algorithm [J]. Journal of Highway and Transportation Technology, 2022, 38(1): 142-148.
- [9] ZHAO Rui, LIU Hui, LIU Peilin, et al. A helmet detection algorithm based on improved YOLOv5s [J]. Journal of Beijing University of Aeronautics and Astronautics, 2023, 49(8): 2050-2061.
- [10] 郑磊. 基于深度学习的驾驶行为识别与预警系统的设计与实现 [D]. 重庆: 西南大学, 2021.
- [11] ZOU Peng, YANG Kaijun, LIANG Chen. Improving real-time detection of lightweight irregular driving behavior in YOLOv5 [J]. Computer Engineering and Applications, 2023, 59(13): 186-193.