

引用:童源,费树岷. 使用 YOLOv8-OD 和 DeepSORT 的车辆跟踪算法[J]. 聊城大学学报(自然科学版),2026,39(1):24-31.
Citation:TONG Yuan,FEI Shumin. Vehicle tracking algorithms using YOLOv8-OD and DeepSORT[J]. Journal of Liaocheng University (Natural Science Edition),2026,39(1):24-31.
文章编号 1672-6634(2026)01-0024-08 DOI 10.19728/j.issn1672-6634.2025020008

使用 YOLOv8-OD 和 DeepSORT 的车辆跟踪算法

童源^{1,2},费树岷²

(1. 江苏旅游职业学院 信息工程学院,江苏 扬州 225000; 2. 东南大学 自动化学院,江苏 南京 210096;)

摘要 为解决传统多目标跟踪算法在检测和跟踪精度及鲁棒性方面存在的不足,提出一种基于 Tracking-By-Detection 模式的新方法,用于车流量检测。该方法运用 YOLOv8 目标检测算法实现了对车辆目标的快速定位与识别,并整合了一种改进的基于深度学习的 DeepSORT 多目标跟踪算法,从而确保了对车辆的精准实时跟踪和计数。实验结果显示,该方法在快速移动车辆的检测与复杂光照环境中表现出较高精度,平均精度达到 94.7%。这种端到端的方法在车辆视频的批处理应用中表现出良好的可行性和有效性。

关键词 YOLOv8;DeepSORT;深度学习;车辆跟踪

中图分类号 U495;TP391.41

文献标志码 A

开放科学(资源服务)标识码(OSID)



Vehicle tracking algorithms using YOLOv8-OD and DeepSORT

TONG Yuan^{1,2}, FEI Shumin²

(1. School of Information Engineering, Jiangsu College of Tourism, Yangzhou 225000, China;
2. School of Automation, Southeast University, Nanjing 210096, China)

Abstract To address the limitations of traditional multi-object tracking algorithms in terms of detection accuracy, tracking precision, and robustness, this paper proposes a novel method based on the Tracking-By-Detection paradigm for vehicle flow monitoring. The method employs the YOLOv8 object detection algorithm to achieve rapid localization and identification of vehicle targets, and integrates an improved deep learning-based DeepSORT multi-object tracking algorithm to ensure accurate and real-time tracking and counting of vehicles. Experimental results demonstrate that the proposed method achieves high detection accuracy when handling fast-moving vehicles, with an average precision of 94.7%. This end-to-end approach exhibits good feasibility and effectiveness in batch processing applications of vehicle video data.

Keywords YOLOv8; DeepSORT; deep learning; vehicle tracking

0 引言

随着城市化的迅速推进以及智能交通系统的不断发展,车辆流量检测已成为研究的核心研究议题。随

收稿日期:2025-02-21

基金项目:国家自然科学基金项目(61973105)资助

通信作者:费树岷,男,汉族,博士,教授,博士生导师,研究方向:非线性控制系统分析与综合、模式识别等,E-mail:lolana_a@hotmail.com。

着智慧城市建设的深入推进,车流量检测技术的创新与优化对提升城市管理效能具有显著促进作用^[1]。研究表明,当前交通规划领域主要采用空气管道检测与磁感应检测两类技术手段^[2],然而这些传统方法在环境适应性与自动化程度方面存在明显局限性,难以充分满足现代智慧城市的发展需求。随着视觉感知算法与智能硬件系统的协同发展,视频分析技术在交通流量监测领域实现了智能化突破,逐步取代传统人工统计模式并成为行业研究热点。该技术体系通常涵盖运动目标识别与多目标追踪等核心功能模块,通过动态感知与时空关联算法实现交通流参数的自动化提取^[3]。传统的车辆跟踪技术可分为判别型和生成型两大类,其中一些算法由于背景建模和光流法等因素,存在实时性不足和复杂度较高的问题^[4-5];而其他算法如卡尔曼滤波、粒子滤波和均值漂移,则容易受到目标移动速率的影响^[6-7]。

目前,在目标识别领域应用深度学习的模型主要分为两大技术方向:一是依赖于候选区域的策略,另一种则是基于回归分析的思路。前者通过特征空间分析与候选框筛选实现目标定位,后者则直接在卷积特征图上进行像素级定位与分类预测。在目标检测算法研究领域,基于候选区域的方法如 R-CNN 和 Fast R-CNN^[8-9]存在计算效率较低的局限性。相比之下,采用回归思想的 YOLOv3 算法^[10]在保证检测精度的基础上显著提升了实时处理能力。在车辆跟踪技术方面, SORT 算法^[11]创新性地整合了卡尔曼滤波器与匈牙利算法,通过建立前后帧间的运动模型与数据关联机制,实现了高效的在线跟踪,然而其在处理长期遮挡和目标识别方面仍存在不足。DeepSORT 算法^[12]引入卷积神经网络进行特征提取与训练,采用更为精确的关联度量方法,显著提升了系统在目标丢失与遮挡情况下的稳定性,从而有效克服了传统跟踪算法中的遮挡难题^[13-14]。YOLOv8 是由 Ultralytics 推出的新一代端到端目标检测算法,其在结构设计与性能优化方面均实现了突破。YOLOv8 继承了 YOLO 系列以速度见长的优良传统,并在模型结构、特征提取能力、推理效率和泛化性能等方面进行全面革新。与前代模型如 YOLOv5 和 YOLOv7 相比, YOLOv8 在保持高检测精度的同时进一步提升了模型轻量化与推理速度,更加契合智慧交通等实时性要求高的应用场景。在网络结构设计方面, YOLOv8 秉承了“高效主干+深度可分离卷积+自适应检测头”的理念。其主干网络 Backbone 采用 CSPDarknet 的演化版本,引入了 C2f(Cross-Stage Partial with Fully Connected Fusion)模块,该模块通过分支连接与融合操作在保持模型结构轻量的基础上有效提升了特征表达能力。然而原始 C2f 模块在处理浅层特征时仍存在细节保留不足的问题,尤其在复杂背景或多目标遮挡情况下,容易出现目标边界模糊、尺度不敏感等识别瓶颈^[2,15]。

针对复杂交通场景下多目标持续追踪的技术需求,本研究构建了 YOLOv8-DeepSORT 融合框架以进行车流量检测研究。YOLOv8 算法难以充分保留浅层特征细节。基于此,我们在 C2f 模块的基础上引入了全维动态卷积 ODConv,并借鉴 DenseNet 网络的特点,提出了 C2f-OD 模块。C2f 模块通过连接每个瓶颈层提取的特征,但为保持计算效率和模型大小,瓶颈层的堆叠次数受限,削弱了 C2f 模块的特征提取能力。经过改进的 C2f-OD 模块通过在瓶颈层中用 ODConv 替代传统卷积,并结合残差结构对输入与输出特征图进行融合,从而更有效地提取出目标的尺度和轮廓等关键特征。将 DenseNet 网络的密集连接机制融入 C2f 模块,改进后的 C2f-OD 模块结构使每个瓶颈层能获得前一层的输出,充分保留目标边缘和纹理等浅层特征。提升了传统车流量检测算法的性能和鲁棒性。

1 改进的 YOLOv8 模型

作为 YOLO 架构的工程化演进版本, YOLOv8 通过动态标签分配策略和自适应特征提取机制,在目标检测领域展现了显著的工程实践价值。其网络采用模块化设计理念(如图 1 所示),主要由 4 个部分组成,分别是输入端、主干网络、颈部网络、输出端。在进行检测任务时,输入图像首先通过包含 CBS(Convolutional Block with Shortcut)、C2f、SPPF(Spatial Pyramid Pooling-Fast) 模块的主干网络进行特征提取,得到不同大小的特征图,然后将它们输入到颈部网络(由 Conv、C2f、Concat、Upsample 模块构成的 PA-FPN 架构)中进行特征融合。检测头采用无锚解耦头结构,在增强模型的泛化能力和鲁棒性的同时,有效地减少了参数数量和计算复杂度。

1.1 ODConv 全维动态卷积机制

浅层特征是指神经网络中接近输入层的特征表示,包括目标的形状、纹理、位置等信息。因此,本文在 C2f 模块的基础上,引入全维动态卷积 ODConv,同时借鉴了 DenseNet 网络的特点,提出了 C2f-OD 模块。普通卷积卷积核参数固定不变,不具备动态属性,能够处理一些静态或比较规则的检测目标。如果采用普通卷积进行处理,则容易丢失关键特征,如果采用堆叠卷积的方法提高性能,又会导致增加大量计算成本。因此本文引入全维动态卷积 ODConv (Omni-dimensional Dynamic Convolution)。作为动态卷积网络的核心机制,ODConv 通过赋予卷积核自适应特性,实现了基于目标特征参数动态调节。该机制能够同步学习空间维度、输入输出通道以及卷积核数量等多个关键参数,最终得到全维度的卷积核注意力值。具体运算过程图 2 所示。

在 ODConv 的运算过程中,输入特征向量 x 首先经历全局平均池化处理,将其维度缩至 c_{in} ,随后该向量通过全连接层实现降维映射,并经由线性修正单元激活后产生四个并行分支。其中,首支路由输出维度为 $k \times k$ 的全连接层构成,经 Sigmoid 函数转换后输出归一化注意力 α_{si} ($i=1, \dots, n$)。另外三个分支采用相似的架构,分别通过输出维度为 $c_{in} \times 1, c_{out} \times 1, n \times 1$ 的全连接层,并配合 softmax 或 Sigmoid 函数最终生成注意力 $\alpha_{ci}, \alpha_{fi}, \alpha_{wi}$ (其中 $i=1, \dots, n$)。在卷积神经网络架构中,针对特定卷积核 W_i ($i=1, \dots, n$) 的优化机制主要包含四个互补的注意力模块,注意力 α_{si} 通过对空间位置赋予差异性注意力标量,实现对目标区域的精确定位;注意力 α_{ci} 则在输入通道维度上分配注意力标量,用于解析目标的结构特性并提取其表现特征;注意力 α_{fi} 在输出通道层面施加注意力标量,有效降低卷积运算过程中的特征损失;注意力 α_{wi} 则负责调节各卷积核的权重分配。这四类注意力机制通过渐进式地对卷积核 W_i 的四个维度施加差异化注意力标量,实现了对输入特征的多维度差异性建模,从而增强了上下文信息的捕获能力,显著提升了网络性能。

$$Y = \left(\sum_{i=1}^n \alpha_{wi} \otimes \alpha_{ci} \otimes \alpha_{fi} \otimes \alpha_{si} \otimes W_i \right) * X,$$

式中 $\otimes$ 表示沿内核空间不同维度的乘法操作, Y 为输出特征, X 为输出特征, W_i 为卷积核, $*$ 为卷积运算, α_{wi} 为沿卷积核数量维度 w 的注意力标量, α_{ci} 为沿卷积核输入通道维度 c 的注意力标量, α_{fi} 为沿卷积核输出通道维度 f 的注意力标量, α_{si} 为沿卷积核空间核尺寸 s 的注意力标量。

1.2 浅层特征增强模块 C2f-OD

C2f 模块将每个瓶颈层提取的特征进行连接,但为了保证计算效率和模型大小,C2f 模块中瓶颈层的堆叠次数受到了限制,这在一定程度上削弱了 C2f 模块提取特征的能力。本文的研究策略首先是对 C2f 模块中的瓶颈层进行革新,具体操作为以 ODConv 算子替代其中原有的标准卷积操作;进一步地借助残差连接,将输入与输出特征图进行有效融合,以此强化网络对目标物体的尺度以及轮廓等关键特征信息的提取能力。改进后的瓶颈层具体结构可参见图中详细展示。然后将

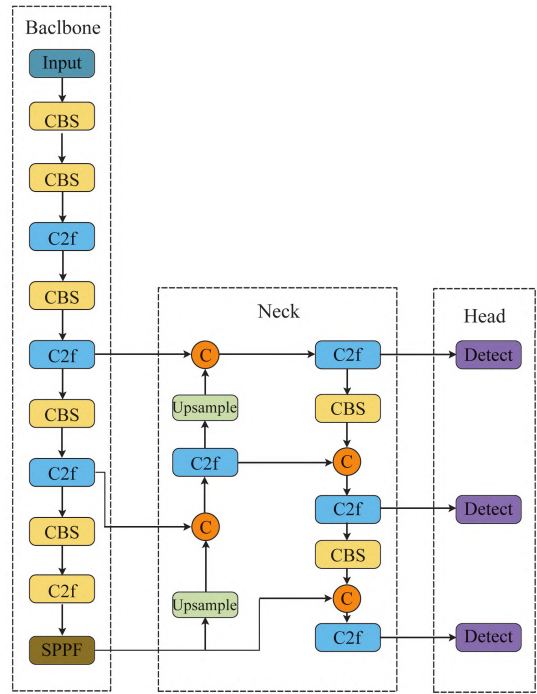


图 1 YOLOv8 架构

Fig. 1 Structural light measurement system model

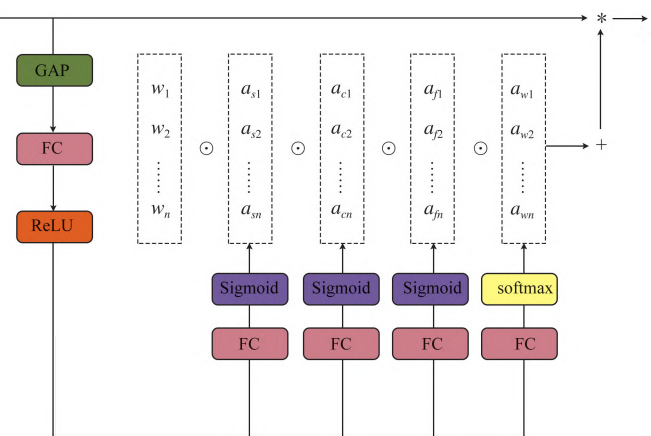


图 2 ODConv 结构

Fig. 2 ODConv structure

Fig. 3 C2f-OD structure

1.3 改进后的模型架构

[illegible]

Fig. 4 The improved structure

作为SORT跟踪算法的升级版，DeepSORT在复杂场景下的目标跟踪性能得到显著提升，特别是在处理目标遮挡和交叉等难点问题上表现出色。该算法主要采用卡尔曼滤波技术对目标运动轨迹进行预测，进而估算目标在后续帧中的可能位置，并运用匈牙利算法实现预测值与实际观测值之间的最优匹配。在此过程中，算法通过计算预测值与观测值之间的空间距离或特征相似度，建立目标间的关联关系，从而确保跟踪的准确性和连续性。在完成目标关联后，系统即可输出跟踪结果。然而由于该算法仅以图像中的检测框为处理单元，缺乏对目标外观特征的考量，在目标密集场景下易因遮挡与交叉而导致误判。为解决这一问题，DeepSORT在SORT框架的基础上引入了基于外观特征的相似度度量机制，即将目标从单一的轴向矩形几何描述扩展至包含内部像素信息与外观特征的综合表征，通过特征向量对目标进行更全面的描述，并以此增强匹配关联的准确性。此外，算法还采用了级联匹配策略，通过为不同跟踪目标分配差异化的关联优先级，有效增强了跟踪过程的一致性、稳定性。DeepSORT算法总体流程如下。

中国知网 <https://www.cnki.net>

(2) 自视频序列的第 2 帧起,采用 IOU 匹配方法关联当前帧检测到的目标边界框与前一帧经卡尔曼滤波预测后的边界框,并以此匹配结果为基础构建代价矩阵,进而实现目标跟踪。

(3) 将步骤(2)得到的代价矩阵输入匈牙利算法进行线性匹配,此过程将产生三种结果:1) 轨迹匹配失败。若该轨迹处于未确认状态,则直接将其删除;若为已确认状态,则只有当其连续匹配失败的帧数达到预设阈值(DeepSORT 算法默认为 30 帧)时,才予以删除该轨迹;2) 检测匹配失败。这表明有新目标出现,需对新检测到的目标边框初始化其轨迹信息;3) 成功匹配。说明前后两帧目标匹配成功,卡尔曼滤波会继续更新该目标轨迹信息。

(4) 不断重复步骤(2)和步骤(3),直到出现已确认的轨迹或者视频帧播放完毕。

(5) 将确认态轨迹预测框和目标检测框进行级联匹配。同样会得到 3 种结果:轨迹失配、检测失配和成功匹配。针对轨迹失配与检测失配的现象,本研究采用计算交并比(IOU)的方式,对先前处于未确认状态的轨迹预测框与当前匹配失败的检测框进行关联,并以此构建代价矩阵。当出现成功匹配的情况时,卡尔曼滤波会更新目标对应的轨迹信息。

(6) 将步骤(5)中代价矩阵结合匈牙利算法进行计算,接下的步骤与步骤(3)相同,重复此步骤直至视频帧结束。图 5 所示为 DeepSORT 算法总体流程图。

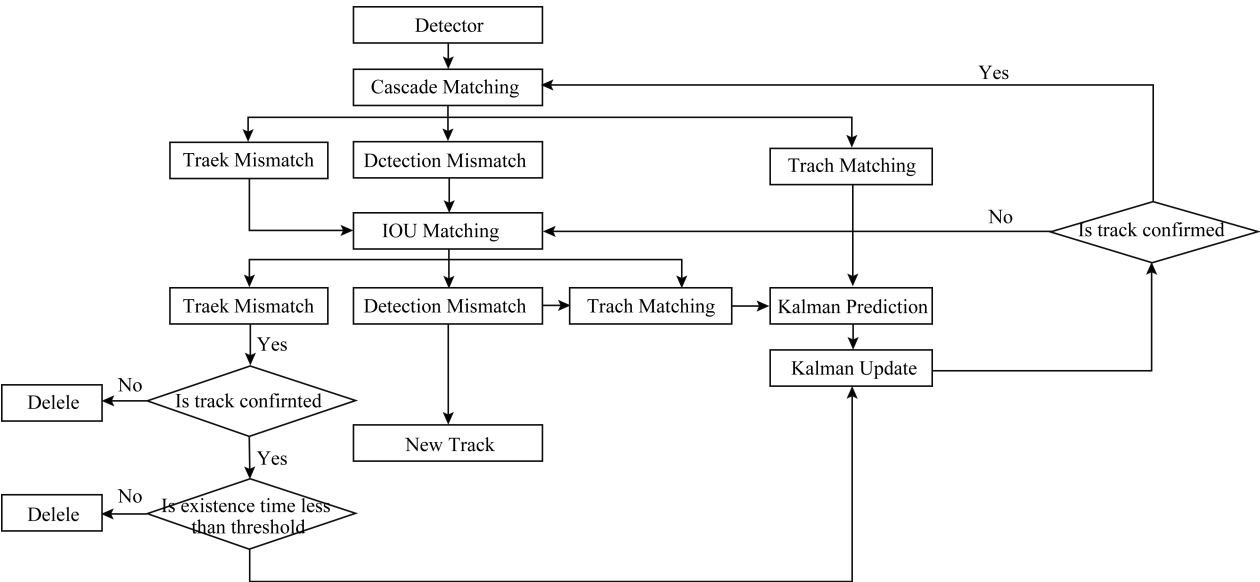


图 5 DeepSORT 算法整体流程
Fig. 5 Overall flow of DeepSORT algorithm

3 实验测试

3.1 实验平台与数据集

本文所用的实验平台与相关训练参数如表 1 和表 2 所示。其训练的可视化结果如图 6 所示。

表 1 实验平台与数据集

Tab. 1 Experimental platform and dataset

Hardware Parameters	Software Parameters	Dataset
GPU Model: NVIDIA GeForce RTX 4090	Python 3. 10	VeRi-776
CPU Model: Platinum 8352V	Pytorch 2. 1. 0	
Memory: 120 G	BCUDA 12. 1	
Graphics Memory: 24 GB		

表 2 训练参数

Tab. 2 Training parameters

Training Parameters	Training parameter values
Image size	1280×640×3
Optimizer algorithm	SGD
Number of iterations	200
Weight decay coefficient	0.0005
Batch size	16
Cyclical learning rate	0.01
Learning rate	0.001

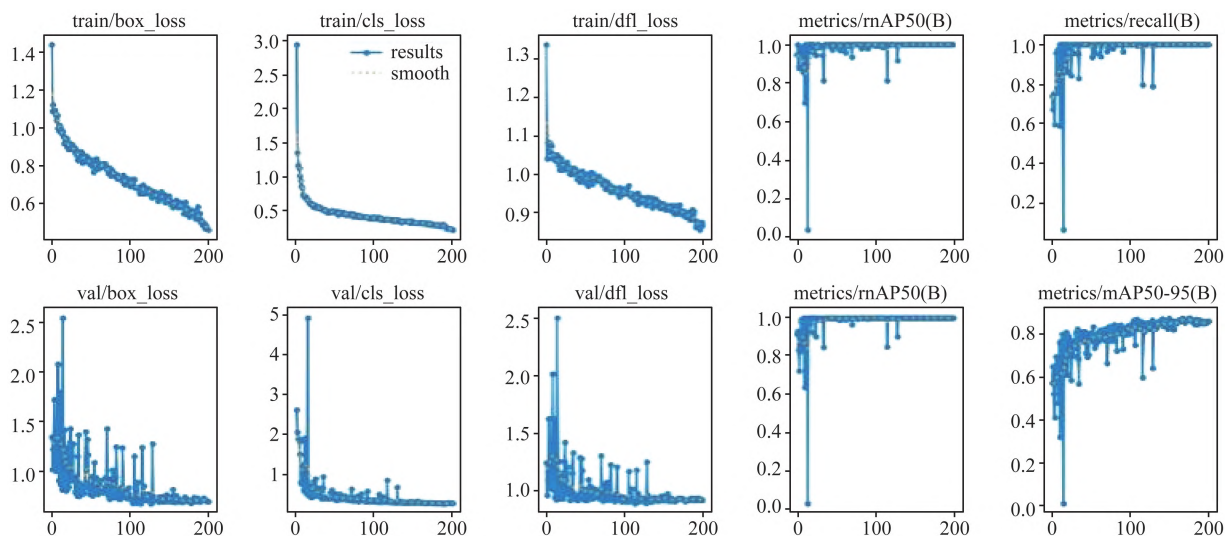


图 6 训练过程可视化

Fig. 6 Training process visualization

3.2 评价指标

本文使用如下的评价指标:

(1) 多目标跟踪准确度 M_{OTA} (Multiple Object Tracking Accuracy, MOTA): 正确跟踪目标的数量占有所有目标数量的比例。它综合考虑了多个方面的跟踪准确性。 M_{OTA} 计算公式为

$$M_{OTA} = 1 - \frac{\sum (N_{FN} + N_{FP} + N_{IDS})}{\sum N_{GT}},$$

式中 N_{FN} 为未能检测到的目标个数, N_{FP} 为被错误检测为目标的数量, N_{IDS} 为身份切换频次, 具体指目标身份被错误分配的次数, N_{GT} 为真实目标的总数量。

(2) 多目标跟踪精度 M_{OTP} (Multiple Object Tracking Precision, MOTP): 用于评估跟踪器在正确跟踪目标的情况对目标位置预测的准确性。 M_{OTP} 计算公式为

$$M_{OTP} = \frac{\sum_{t,i} d_{i,t}}{\sum_t c_t},$$

式中 i 表示当前检测的目标, t 表示当前时刻, $d_{i,t}$ 表示检测目标 i 与其匹配的真实标注边界框之间的平均距离度量, c_t 表示当前时刻匹配成功的个数。

3.3 车辆跟踪及计数

在获得目标检测结果后, 采用改进的 DeepSORT 算法对车辆进行持续追踪。该方法通过计算检测框中心点的运动特征, 结合目标外观信息进行数据关联, 建立车辆的运动轨迹。基于轨迹信息, 系统可实时统计通过检测区域的车辆数量, 以此实现交通流量的自动化监测。将改进后的 YOLOv8 算法与改进前的算法进

行对比实验,得到结果如表 3 和表 4 所示。

表 3 未改进 YOLOv8+DeepSORT 检测结果

Tab. 3 Detection results of YOLOv8 + DeepSORT without improvement

Serial Number	MOTA/%	IDS	Time Consumption/ms	MOTP/%
Video 1	68.3	35	33.6	86.2
Video 2	59.8	49	37.8	77.4
Video 3	66.4	31	35.7	82.7
Average	64.8	38	35.7	82.1

表 4 本文改进 YOLOv8+DeepSORT 检测结果

Tab. 4 This paper improves the detection results of YOLOv8 + DeepSORT

Serial Number	IDS	Time Consumption/ms	MOTA/%	MOTP/%
Video 1	28	34.1	69.6	87.8
Video 2	42	38.6	62.9	78.8
Video 3	26	36.5	67	83.5
Average	32	36.4	66.5	83.3

通过实验数据分析,改进型 YOLOv8 算法在目标跟踪性能上展现出显著优势,其中 MOTA 值较基准模型平均提升了 4.4 个百分点,这主要归因于 MOTA 评价体系中检测相关指标的优化,而检测精度的提升对多目标跟踪性能具有直接促进作用。此外,实验结果显示 MOTP 值平均提高了 2 个百分点,且 IDS 次数平均减少 6 次。

进一步验证 YOLOv8 结构中所引入的 ODConv 和 C2f-OD 模块对整体检测性能的影响,本文设计了基于结构组合的消融实验,实验结果如表 5 所示。从表中可以看出,在仅引入 ODConv 模块的实验组合中(YOLOv8+ODConv),MOTA 提升至 66.5%,MOTP 达到 83.0%,相比基础 YOLOv8 模型性能有明显改进,表明 ODConv 动态权重机制对提升特征提取能力具有积极作用;在仅使用 C2f-OD 模块时,MOTA 略提升至 65.7%,MOTP 提升至 82.5%,说明 C2f-OD 对浅层细节特征的保留和传播具备一定效果;当 ODConv 和 C2f-OD 模块联合使用时,MOTA 与 MOTP 分别提升至 69.6%和 87.8%,显著高于各单独模块,验证了两者协同融合对检测精度的促进作用。

表 5 消融实验

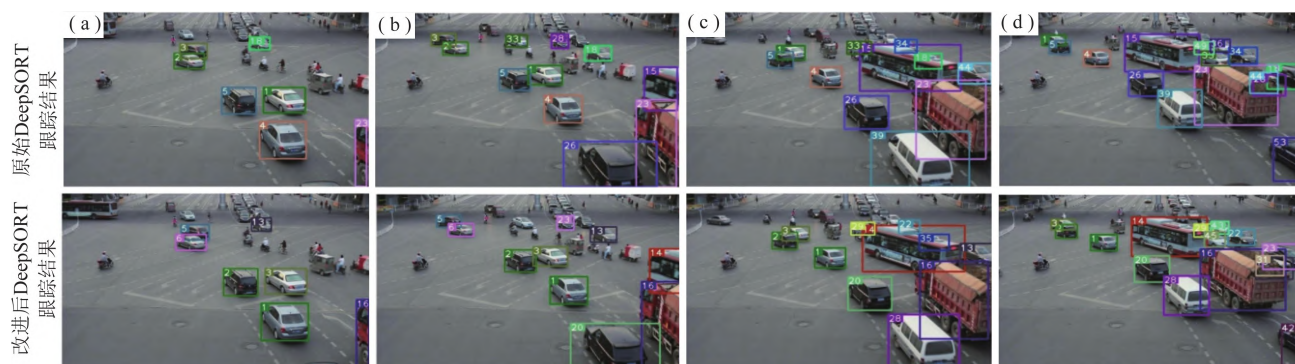
Tab. 5 Ablation Experiment

YOLOv8	ODConv Model	C2f-OD Model	MOTA/%	MOTP/%
✓	✓		66.5	83.0
✓		✓	65.7	82.5
✓	✓	✓	69.6	87.8

综上所述,C2f-OD 模块与 ODConv 卷积机制在目标尺寸、轮廓与细节特征方面具有互补性,最终共同构建了本文提出的 YOLOv8-OD 模型结构,有效提升了目标检测性能。

为评估算法在实际场景中的适用性,本研究选取交通流量密集的实际道路视频作为测试样本,分别采用原始算法和改进算法进行跟踪实验。实验结果以五帧为间隔进行采样,其可视化对比分析见图 7。

通过对 DeepSORT 算法的对比实验分析表明,在目标跟踪过程中,改进后的算法在 ID 一致性方面展现出显著优势。实验数据显示,原始 DeepSORT 算法在处理第 135 帧至第 215 帧的车辆跟踪时,目标 ID 由 18 切换为 44,而改进算法则保持了稳定的 ID 13;类似地,在第 175 帧至第 265 帧的跟踪过程中,原算法将目标 ID 从 28 变更为 18,而改进算法始终维持 ID 23 不变。这一实证结果表明,改进后的 DeepSORT 算法在目标跟踪的稳定性和准确性方面较原算法有显著提升。



注:(a) 第 135 帧;(b) 第 175 帧;(c) 第 215 帧;(d) 第 265 帧。

图 7 实验结果

Fig. 7 Experimental results

4 结论

为实现车辆的精确追踪与统计,本研究采用优化后的 YOLOv8 算法进行快速车辆检测,并结合 DeepSORT 多目标跟踪算法,实现了深度神经网络框架下的有效融合。改进的 YOLOv8 算法通过引入全维动态卷积(ODConv)以及借鉴 DenseNet 网络的优势,显著提升了对车辆目标的检测精度与鲁棒性。实验数据表明,在车辆快速移动、部分遮挡以及光照变化等复杂场景中,该方法能够有效增强跟踪稳定性,并大幅提高车流量检测的精度,平均精度可达 94.7%,MOTA 指标平均增长 4.4%,MOTP 指标平均增长 2%,IDS 出现次数平均减少 6 次。综上所述,将改进后的 YOLOv8 算法与 DeepSORT 算法协同应用,能有效优化车辆跟踪和计数性能,为智能交通系统中车流量检测提供了一种切实可行的解决方案。

参 考 文 献

- [1] 郭真子. 基于视频图像的城市交通流量统计算法研究[D]. 西安: 西安科技大学, 2019.
- [2] 李坤. 基于磁阻传感器的车流量检测系统的设计[J]. 测控技术, 2019, 38(1):114-116.
- [3] 刘鑫, 金珏宏. 四帧间差分与光流法结合的目标检测及追踪[J]. 光电工程, 2018, 45(8):35-42.
- [4] 程淑红, 马继勇, 张仕军, 等. 改进的混合高斯与 YOLOv2 融合烟雾检测算法[J]. 计量学报, 2019, 40(5): 798-803.
- [5] 程淑红, 高许, 程树春, 等. 基于计算机视觉的运动车辆检测[J]. 计量学报, 2017, 38(3):288-291.
- [6] 王忠民, 段娜, 范琳. 融合 YOLO 检测与均值漂移的目标跟踪算法[J]. 计算机工程与应用, 2019, 55(10):186-192.
- [7] 李雪林, 倪玉冰, 张健, 等. 欺骗攻击下 Markov 跳变系统的不匹配量化自适应安全控制[J]. 南京师大学报(自然科学版), 2024, 47(3):104-111.
- [8] WANG D, TANG J L, ZHU W J, et al. Dairy goat detection based on faster R-CNN from surveillance video[J]. Computers and Electronics in Agriculture, 2018,154:443-449.
- [9] REDMON J, FARHADI A. YOLOv3: an incremental improvement[EB/OL]. arXiv:1804.02767, 2018.
- [10] BEWLEY A, GE Z Y, OTT L, et al. Simple online and realtime tracking[C]//2016 IEEE International Conference on Image Processing (ICIP), Phoenix, AZ, USA, 2016:3464-3468.
- [11] WOJKE N, BEWLEY A, PAULUS D. Simple online and realtime tracking with a deep association metric[C]//2017 IEEE International Conference on Image Processing (ICIP), Beijing, China, 2017:3645-3649.
- [12] ZHANG X, HAO X Y, LIU S L, et al. Multi-target tracking of surveillance video with differential YOLO and DeepSORT[C]//11th Eleventh International Conference on Digital Image Processing, Guangzhou, China, 2019:174.
- [13] BOUACHIR W, GOUIAA R, LI B, et al. Intelligent video surveillance for real-time detection of suicide attempts[J]. Pattern Recognition Letters, 2018,110:1-7.
- [14] 程瑶, 赵雷, 成珊, 等. 基于机器视觉的车距检测系统设计[J]. 计量学报, 2020, 41(1): 11-15.
- [15] 葛雯, 史正伟. 改进 YOLOV3 算法在行人识别中的应用[J]. 计算机工程与应用, 2019, 55(20):128-133.
- [16] NANCY J L, CHABBERT M, DEKETELAERE N. The world's fragile skin[J]. Angelaki, 2021, 26(3/4): 12-16.

(责任编辑:赵海军)