

基于改进 YOLO 算法的地铁车厢 客流检测方法

郭 宁^{1,2}, 胡小晨^{1,2}, 董德存¹

(1. 同济大学 道路与交通工程教育部重点实验室, 上海 201804; 2. 同济大学 上海市轨道交通结构耐久与系统安全重点实验室, 上海 201804)

摘要: 在高峰时段, 地铁车厢的客流量剧增, 人群密集和目标遮挡等复杂场景导致传统方法难以准确识别每一个乘客, 容易导致漏检或误检。针对此问题, 提出了基于改进 YOLO 算法的地铁车厢客流检测方法。分析 YOLOv8 模型结构后, 将 ASF-YOLO 中的 TFE 模块加入 YOLOv8n 中, 结合时空模型, 考虑到站时乘客流动大和行驶时乘客活动少的特点, 及车门和车厢内乘客的不同流动特性, 将多帧检测结果相融合, 实现了对地铁车厢内乘客流量的精准检测。经实验比较: 原始 YOLOv8n 模型平均精度为 57.0%, 改进后的模型为 69.1%, 多帧融合处理后为 76.6%。基于该模型所得到的客流信息可用于乘客出行引导、应急救援、铁路运营管控等场景。

关键词: 地铁车厢; YOLOv8 模型; 客流检测; 时空模型

中图分类号: U293; TN919 **文献标志码:** A **文章编号:** 1671-5497(2025)04-1258-08

DOI: 10.13229/j.cnki.jdxbgxb.20240756

Passenger flow detection method of subway car based on improved YOLO algorithm

GUO Ning^{1,2}, HU Xiao-chen^{1,2}, DONG De-cun¹

(1. Key Laboratory of Road and Traffic Engineering, Ministry of Education, Tongji University, Shanghai 201804, China; 2. Shanghai Key Laboratory of Rail Infrastructure Durability and System Safety, Tongji University, Shanghai 201804, China)

Abstract: During peak hours, the passenger flow of subway carriages increases sharply. In complex scenarios such as dense crowds and target occlusion, it is difficult to accurately identify each passenger, which can easily lead to missed or false detections. To this end, a subway carriage passenger flow detection method based on the improved YOLO algorithm is proposed. After analyzing the YOLOv8 model structure, the TFE module from ASF-YOLO was added to YOLOv8n. Combined with the spatiotemporal model, the characteristics of high passenger flow at stations and low passenger activity during driving, as well as the different flow characteristics of passengers in train doors and carriages, were

收稿日期: 2024-06-12.

基金项目: 中国-东盟综合交通国际联合实验室建设项目(桂科 AD20297125).

作者简介: 郭宁(1994-), 男, 博士研究生. 研究方向: 基于图像处理和深度学习的铁路交通智能运维.

E-mail: 13273031888@163.com

通信作者: 董德存(1958-), 男, 博士, 教授. 研究方向: 交通信息控制及系统安全. E-mail: 1710910@tongji.edu.cn

considered. The multi frame detection results were fused to achieve accurate detection of passenger flow in subway carriages. Through experimental comparison, the average accuracy of the original YOLOv8n model is 57.0%, the improved model is 69.1%, and after multi frame fusion processing, it is 76.6%. The passenger flow information obtained based on this model supports multiple aspects such as passenger travel guidance, emergency rescue support, and railway operation control.

Key words: metro car; YOLOv8 model; passenger flow detection; spatial-temporal model

0 引言

随着城市化进程的持续推进,城市居民数量持续增加,这也对城市现有管理模式造成了一定的负担以及挑战,尤其是对公共交通系统提出了更高的要求。地铁由于特殊的运营模式,在时效性、载流量以及抗气候干扰等方面表现突出,不仅成为当下民众出行的主要选择之一,更是现代城市基建不可或缺的组成部分。城市人口聚集导致乘客流量增加,对地铁系统的运输服务能力提出了更高的要求,但地铁建设投资巨大,难以单纯依靠新增线路缓解运营压力。因此,如何在有限的运力基础上实现科学、精细化管理,以缓解供需不平衡问题,变得越来越重要。而这样的地铁管理模式需要大量准确的交通信息作为支撑,例如地铁营运过程中的乘客数量。通过对乘客数量的感知,可以为后续多项工作提供重要支撑:对车厢内乘客数量及密度实时监测并发布,能够帮助乘客更好地了解等车信息,确定所需要乘坐车次,从而优化出行计划,降低车厢拥挤程度,提高乘坐舒适度;同时,地铁运营管理部门可以根据这些信息,对地铁运力进行动态调度和对站台乘客流量进行引导,并可以有效预判以及防范危机事件,避免踩踏等情况的发生。

以检测范围为例,现有针对客流的研究多围绕车站范围内以及进出站区域的安全防范、闸机区域的客流管理,而针对地铁车厢内部乘客流量检测的研究较少。从检测手段来看,地铁列车和站台的乘客密度可以通过多种方法进行估算或者直接检测。例如,通过轴重数据估算列车载客量,使用手机信号数据推测列车乘客拥挤程度^[1]或交通枢纽的客流量^[2],还可以采用主动红外^[3]、被动红外^[4]、压力^[5]、视频等检测技术。然而,轴重估算的乘客密度不准确,手机信号数据存在信号干扰等问题,红外检测技术在人流密集区域很难对多个同时通过的目标进行计数。目前,地铁视频监控系统已广泛

部署,且基于图像处理的目标检测技术近些年来发展迅速,可根据现有的配套基础,对地铁客流数据进行精确研判。文献[6]基于运动目标光流特征,结合支持向量机进行目标识别。文献[7]基于BP神经网络构建了地铁车厢下车客流量的短时分析模型。文献[8]通过融合纹理数据和神经网络算法,对地铁站内人群密度进行预测。文献[9]利用YOLOv5对人员的口罩佩戴情况进行实时监测。文献[10]引入卷积神经网络提取特征信息,并引入支持向量机进行分类,判断教室中的人群信息。文献[11]采用Faster R-CNN实现头部的分析。文献[12,13]利用人脸监测开展人群分析。文献[14]结合Mask R-CNN实现了车厢内的客流量分析。文献[15]将YOLOv5和注意力机制相融合进行车厢内乘客检测。可以看到,现有研究大多适用于站点区域的乘客检测,而针对车厢内乘客的研究较少。针对车厢内乘客的探究,由于环境狭小、遮挡过多、人体形变等因素,检测效果不佳。此外,多数方法并没有考虑列车到站和行驶时的客流差异,以及前后多帧视频信息的融合。

在YOLO等检测技术的更新迭代^[16,17]下,检测精度不断提高,因此本文建立了以YOLOv8为基本的优化模型,结合自主搭建的车厢乘客信息时空系统,对检测失误导致的多帧乘客数目波动进行融合分析,从而提升检测的准确性。

1 YOLOv8模型结构分析

YOLOv8是2023年Ultralytics发布的YOLO系列模型,它的核心架构涵盖了主干特征提取网络(Backbone)、特征强化提取网络(Neck)、检测头(Head)3个部分。YOLOv8网络^[18]结构图如图1所示。

该模型为典型的SOTA模型,其主要基于之前的YOLO系统进行改进,同时增加了新型的应用,以此保证其性能水平以及灵活度。模型提出了新型的框架,其中包括P5 640以及P6 1280分

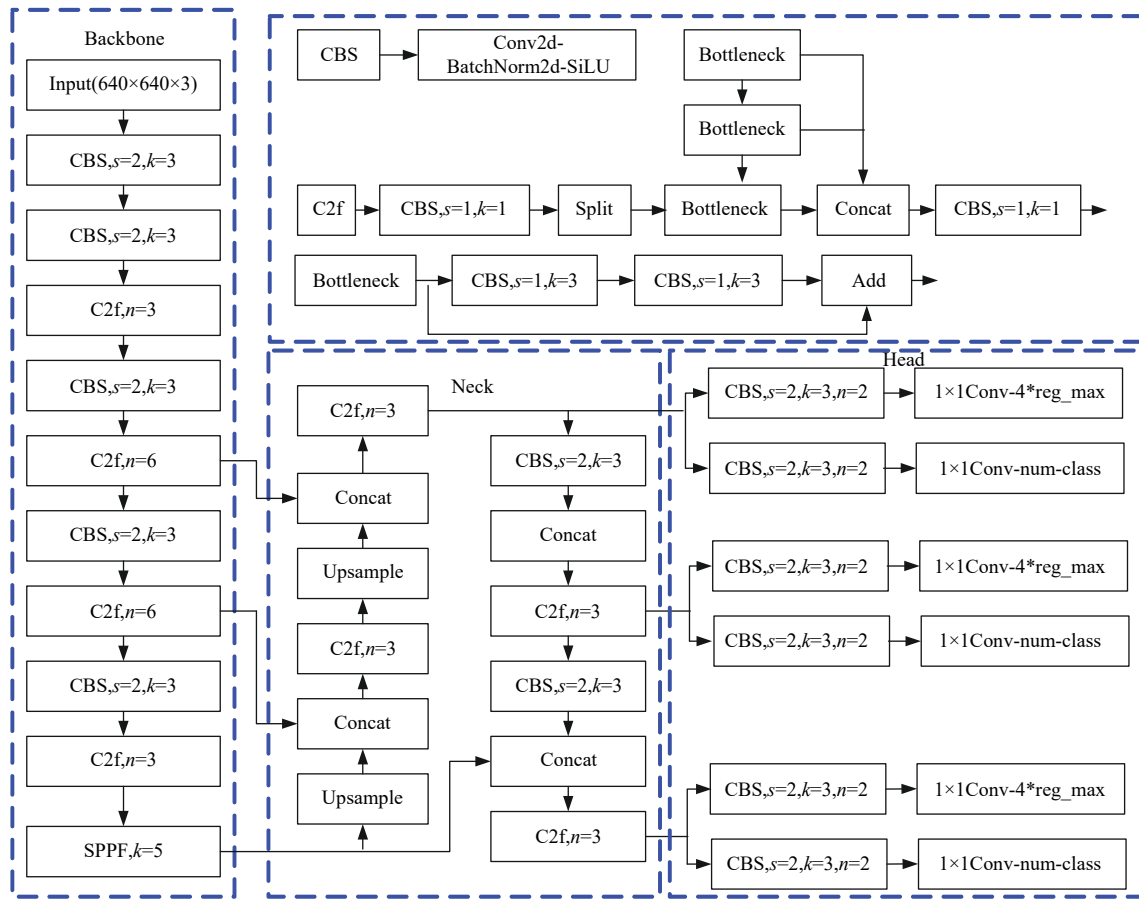


图 1 YOLOv8 网络结构

Fig. 1 YOLOv8 network structure

分辨率的目标检测体系以及基于 YOLACT 的实例分割系统。Backbone 以及 Neck 部分参照了 YOLOv7 ELAN 的开发理念,将 C3 架构转变为梯度流相对全面的 c2f 架构,同时针对各尺度模型优化了对应的通道量。Head 部分相较于 YOLOv5 进行了全面优化,改为当前常见的解耦头模型,将分类与检测分离,并由 Anchor-Based 转变为 Anchor-Free。损失函数计算部分, YOLOv8 的 Loss 计算包括两个部分:分类分支和回归分支。分类分支依然采用 BCE Loss,回归分支则结合了分布式聚焦损失函数(DFL)的积分表示法,同时还使用了 CIoU(Complete intersection over union)Loss。

2 YOLOv8 模型改进

ASF-YOLO^[19]模型是一种针对细胞图像的一阶段实例分割模型,该模型在 YOLO 框架中集成了注意力尺度序列融合模块(ASF)。本研究参照 ASF-YOLO 模型,将其 TFE 模块引入 YOLOv8n 架构,并修改其 Neck 部分。

ASF-YOLO 提出的 TFE 模块旨在增强对密

集细胞的小目标检测能力,通过在空间维度上拼接 3 种不同尺度(大、中、小)的特征来提取小目标的详细信息。为了识别密集重叠的小目标,模型通过放大图像来比较不同尺度下的形状或外观变化。由于骨干网络的不同特征层尺度不同,常规的 FPN 融合机制仅对小尺寸特征图进行上采样,然后将其拆分或添加到前一层的特征中,忽略了大尺寸特征层的丰富详细信息。而 TFE 模块则将大、中、小特征进行拆分,引入大尺寸特征图,并通过特征放大以增强详细特征信息。TFE 模块的结构如图 2 所示。

在特征编码之前,首先调整特征通道的数量,使其与主尺度特征一致。对大尺寸特征图(Large)进行处理时,先将通道数调整为 1C,然后使用最大池化与平均池化的混合结构进行下采样,以保留高分辨率特征的细节和多样性。对于小尺寸特征图,同样先引入卷积模块对通道数进行优化,然后使用最近邻插值方法进行上采样,以保持低分辨率图像的局部特征丰富性,并防止小目标特征信息丢失。最后,对 3 个尺度一致的大、

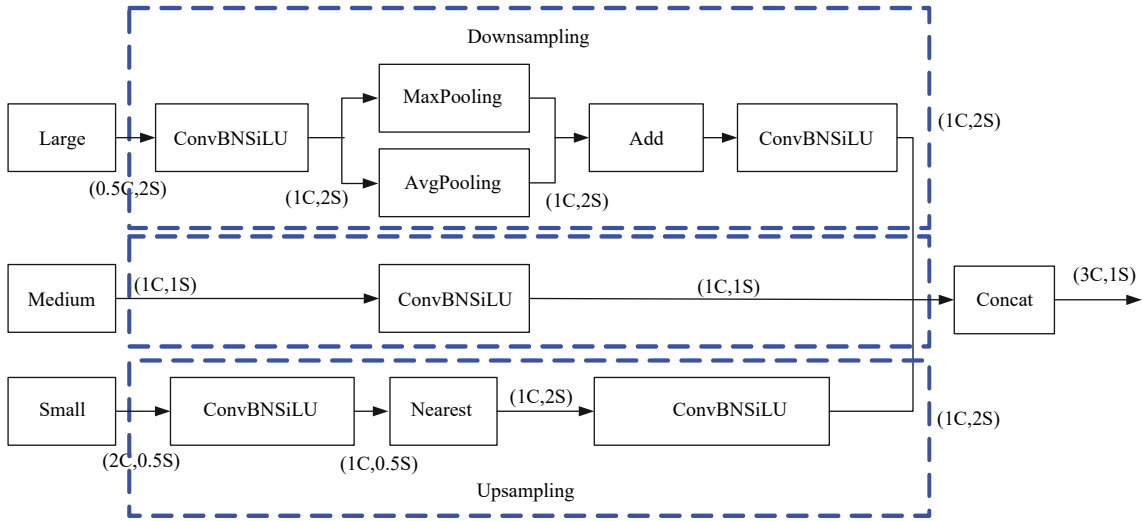


图 2 TFE 模块结构

Fig. 2 TFE module structure

中、小尺度特征图进行卷积处理,之后在通道维度上进行拼接,该操作可表示为:

$$F_{TFE} = \text{Concat}(F_l, F_m, F_s) \quad (1)$$

式中: F_{TFE} 为 TFE 模块输出的特征图; F_l, F_m, F_s 分别为大、中和小尺寸的特征图; F_{TFE} 由 F_l, F_m 和 F_s 拼接得到,具有与 F_m 相同的分辨率,并且通道数是 F_m 的 3 倍。

3 结合时空模型的多帧融合

3.1 时空模型

地铁车厢内部可以大致划分为 2 个不同的区域:车门以及车厢。列车门通常为均匀对称的排布形式,因此研究一道门和毗邻门所处区域称作 S_i ,其中涵盖了车门及其相邻的车厢区域(见图 3)。

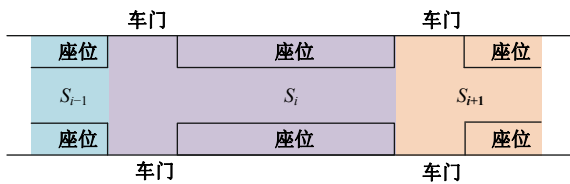


图 3 车厢空间模型图

Fig. 3 Carriage space model diagram

从时间上看,列车运行时间可分为到站时段 T_d 与行驶时段 T_s (见图 4)。

列车抵达车站之后,乘客开始流动,车厢内乘

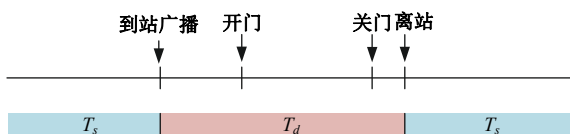


图 4 时间分割模型图

Fig. 4 Time division model diagram

员数量随之开始变动;运行期间,乘客往往车厢内的活动相对较少,因此人群也大多比较固定。

3.2 多帧融合处理

若视频数据共有 f 帧,则在第 i 帧,时间 t_i 为:

$$t_i = \frac{i}{f} \quad (2)$$

车厢区域 S_j 内检测到的人数为 $D(t_i, S_j)$,而实际人数为 $N(t_i, S_j)$,则误差 $E(t_i, S_j)$ 为:

$$E(t_i, S_j) = D(t_i, S_j) - N(t_i, S_j) \quad (3)$$

E 多被用于表示漏检,即没有被检测到的目标。漏检的主要原因为采用的检测模型性能不达标或者目标受到遮挡。将改进后的 YOLOv8 模型作为单帧检测的基础,利用扩大样本等方式提升检测性能;对于目标被遮挡问题,则采用融合多帧检测进行处理。

设置一个阈值 d , 1 s 内逐帧检测到的人数最大值与最小值的差值 ΔD 为:

$$\Delta D = D_{\max} - D_{\min} \quad (4)$$

若 $\Delta D \geq d$, 那么判定此时处于 T_d 阶段,即 $t_i \in T_d$, 设定 1 s 为最小检测区间,结合 1 s 内测得的信息,那么 m 秒检测到的人数 $D'(m, S_j)$ 为:

$$D(m, S_j) = \max \begin{pmatrix} D(t_i, S_j) \\ D(t_i + 1, S_j) \\ \vdots \\ D(t_i + f, S_j) \end{pmatrix} \quad (5)$$

若 $\Delta D < d$, 那么判定此时处于 T_s 阶段,即 $t_i \in T_s$, 结合 1 s 内测得的信息,那么 m 秒检测到的人数 $D'(m, S_j)$ 为:

$$D'(m, S_j) = \text{mode} \begin{pmatrix} D(t_i, S_j) \\ D(t_i + 1, S_j) \\ \vdots \\ D(t_i + f, S_j) \end{pmatrix} \quad (6)$$

基于上述研究,改进 YOLO 算法的检测流程如图 5 所示。

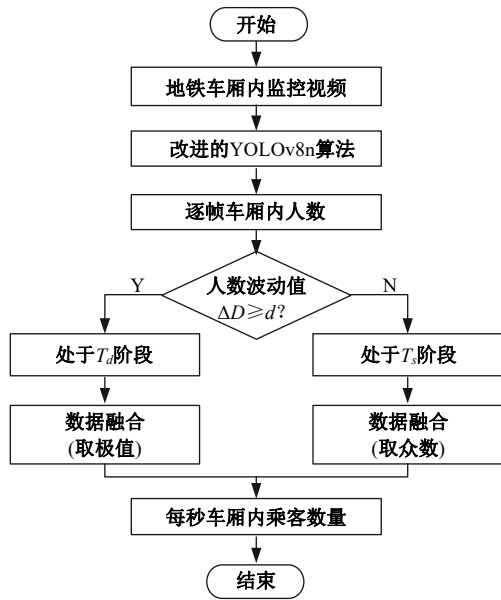


图 5 优化 YOLO 算法的多帧融合客流分析流程图

Fig. 5 Flow chart of multi frame fusion passenger flow analysis for optimizing YOLO algorithm

4 实验结果及分析

实验数据为上海地铁某节车厢内的真实监控视频数据(见图 6),视频帧率 25 帧/s,逐帧检测结果如图 7 所示。



图 6 车厢视频单帧图片示例

Fig. 6 Example of a single frame image of a carriage video



图 7 检测结果示例

Fig. 7 Example of detection results

4.1 非高峰期客流分析

非出行高峰期车厢内人流分析结果如图 8 所示。

这一片段涵盖了常规运行以及列车到站两类情形。在到站前列车正常行驶阶段,改进的 YOLOv8n 模型表现较好,检测偏差主要由人流靠近监控设备导致的形变和人流彼此的遮挡引起。列车到站期间,人流活动会造成短期遮挡,导致结果

偏差。此外,乘客的姿态等也对检测精度产生影响。总的来看,原始 YOLOv8n 模型平均精度为

67.4%,改进后的模型提升至 80.0%,多帧融合处理后则达到了 88.4%,效果提升显著。

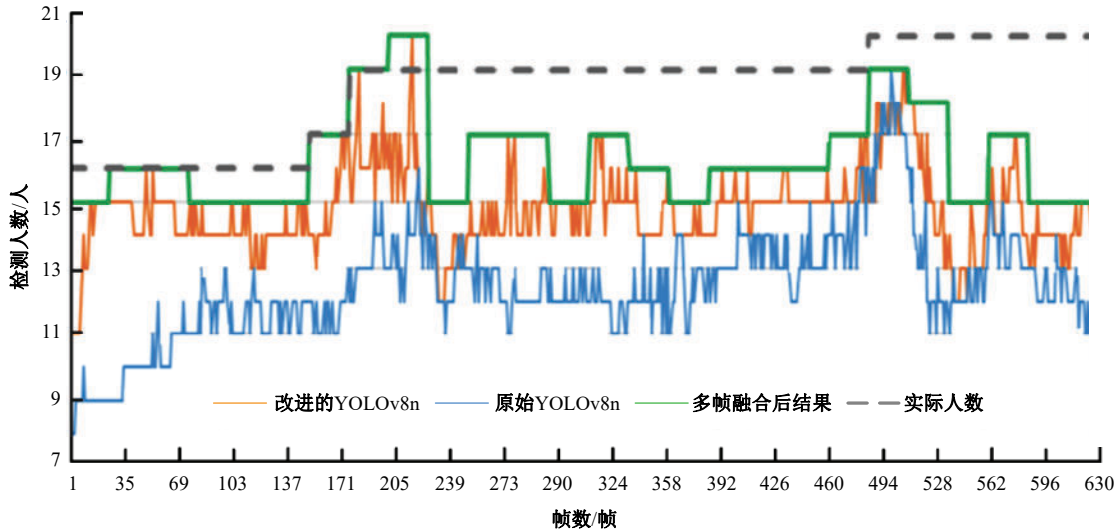


图 8 非出行高峰期客流检测结果

Fig. 8 Passenger flow detection results during non peak travel hours

4.2 高峰期到站客流检测分析

出行高峰期时,列车到站时期地铁车厢内客流的检测结果如图 9 所示。

该片段包含列车到站前后的整个过程。列车到站前,车厢内乘客数量较少,此时检测精度较高;当列车到站时,瞬间涌入多名乘客,乘客

走动导致严重的遮挡和形变,使得检测精度下降;在列车重新出发后,乘客基本停止移动,检测精度有所回升。总的来看,原始 YOLOv8n 模型平均精度为 61.5%,改进后的模型提升至 71.6%,多帧融合处理后则达到 79.6%,总体提升效果显著。

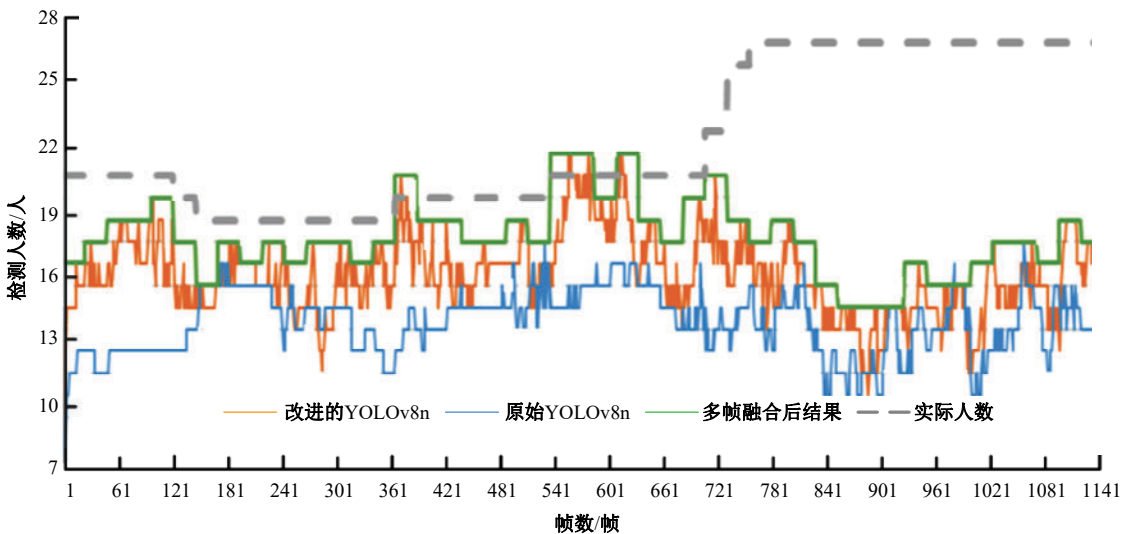


图 9 出行高峰期到站时期客流检测结果

Fig. 9 Passenger flow detection results during peak travel hours and arrival at stations

4.3 高峰期行驶时期客流检测

列车到站后出发时车厢内客流检测结果如图 10 所示。车厢内乘客相对较多,遮挡带来的影响更为显著,除了中途有两名乘客去往其他车厢外,

其他时间车厢内人数稳定。但从结果来看,原始 YOLOv8n 模型平均精度为 42.1%,改进后的模型提升至 57.0%,多帧融合处理后为 62.9%,效果有所提升。

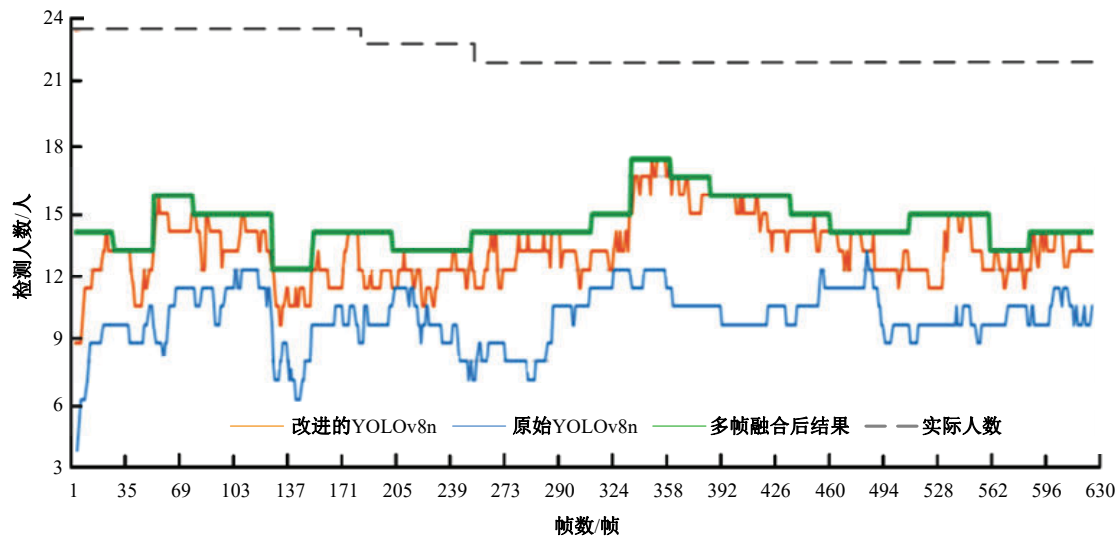


图 10 出行高峰期行驶时期客流检测结果

Fig. 10 Passenger flow detection results during peak travel periods

5 结束语

将 ASF-YOLO 中用于增强小目标检测的 TFE 模块引入 YOLOv8n 中,有效提高了车厢内乘客检测精度,再结合时空模型融合多帧检测结果,一定程度上减小了遮挡等带来的影响。在本文的实验中,原始 YOLOv8n 模型精度均值为 57.0%,优化以后达到了 69.1%,多帧融合处理后为 76.6%。基于检测结果获得的乘客流量数据可以为后续站点运营决策管理、智能运维、乘客出行决策提供支持。在本文研究中,长期遮挡引起的误差较为突出,后续将通过姿态骨架识别等方法加以改善。

参考文献:

- [1] Aguilera V, Allio S, Benezech V, et al. Using cell phone data to measure quality of service and passenger flows of Paris transit system[J]. Transportation Research Part C: Emerging Technologies, 2014, 43(1):198-211.
- [2] Zhong G, Wan X, Zhang J, et al. Characterizing passenger flow for a transportation hub based on mobile phone data[J]. IEEE Transactions on Intelligent Transportation Systems, 2016, 18(6): 1507-1518.
- [3] Niu Q, Wu H F, Gao C Y, et al. Laser-based bidirectional pedestrian counting via height map guided regression and voting[J]. Signal, Image and Video Processing, 2017, 11(5): 897-904.
- [4] 余燕. 基于红外 CCD 与激光测距仪融合的行人检测技术研究[D]. 长春: 吉林大学电子科学与工程学院, 2008.

院, 2008.

Yu Yan. Study on the pedestrian detection technology based on the information fusion of infrared CCD and laser measurement system[D]. Changchun: College of Electronic Science and Engineering, Jilin University, 2008.

- [5] 王连震, 王晖, 罗孟德. 基于压力传感器的出租车客流检测系统[J]. 交通科技与经济, 2018, 20(5): 29-31.

Wang Lian-zhen, Wang Hui, Luo Meng-de. A detection system of taxi passenger flow based on pressure sensor[J]. Technology & Economy in Areas of Communications, 2018, 20(5): 29-31.

- [6] Sidenbladh H, Black M J. Learning image statistics for Bayesian tracking[C]//Proceedings Eighth IEEE International Conference on Computer Vision, Vancouver: Canada, 2001: 709-716.

- [7] 方晨晨, 周继彪, 董升, 等. 基于 BP 神经网络的地铁车厢拥挤度预测方法[J]. 交通信息与安全, 2018, 36(6): 47-53.

Fang Chen-chen, Zhou Ji-biao, Dong Sheng, et al. A method of forecast congestion of subway carriages based on BP neural network[J]. Journal of Transport Information and Safety, 2018, 36(6): 47-53.

- [8] 杨谱. 基于视频图像处理的地铁站拥挤风险预测与人群疏散[D]. 北京: 北京交通大学机械与电子控制工程学院, 2018.

Yang Pu. Risk prediction and crowd evacuation of subway station congestion based on video image processing[D]. Beijing: School of Mechanical, Electronic and Control Engineering, Beijing Jiaotong University, 2018.

- ty, 2018.
- [9] 谈世磊, 别雄波, 卢功林, 等. 基于YOLOv5网络模型的人员口罩佩戴实时检测[J]. 激光杂志, 2021, 42(2): 147-150.
- Tan Shi-lei, Bie Xiong-bo, Lu Gong-lin, et al. Real time detection of personnel mask wearing based on YOLOv5 network model[J]. Laser Journal, 2021, 42(2): 147-150.
- [10] Gao C Q, Li P, Zhang Y J, et al. People counting based on head detection combining Adaboost and CNN in crowded surveillance environment[J]. Neuro-computing, 2016(208): 108-116.
- [11] Song D, Qiao Y, Corbetta A. Depth driven people counting using deep region proposal network[C]//IEEE International Conference on Information and Automation(ICIA), Macau, China, 2017: 416-421.
- [12] Alqahtani F, Banks J, Chandran V, et al. Investigating the efficiency of a three-dimensional facial tracker using stereo camera arrangements[C]//International Conference on Artificial Intelligence and Big Data, Chengdu, China, 2021: 463-468.
- [13] Tran C K. Face recognition based on similarity feature-based selection and classification algorithms and wrapper model[J]. EJournal Publishing, 2019(3): 357-362.
- [14] Shen G, Jamshidi F, Dong D, et al. Metro pedestrian detection based on mask R-CNN and spatial-temporal feature[C]//IEEE 3rd International Conference on Information Communication and Signal Processing (ICICSP), Shanghai, China, 2020: 173-178.
- [15] 邱芳, 李玉峰, 孔才华. 基于YOLOv5m和注意力机制融合的地铁车厢乘客实时检测[J]. 微处理机, 2022, 43(6): 53-58.
- Qiu Fang, Li Yu-feng, Kong Cai-hua. Real-time detection of passengers in subway carriages based on the fusion of YOLOv5m and attention mechanism[J]. Microprocessors, 2022, 43(6): 53-58.
- [16] Lin S D, Zhu K X, Chen F, et al. Align-Yolact: a one-stage semantic segmentation network for real-time object detection[J]. Journal of Ambient Intelligence and Humanized Computing, 2023, 14(2): 863-870.
- [17] 张馨, 董承梁, 汪晓臣, 等. 基于改进YOLOv5s的列车车厢客流密度检测方法研究[J]. 铁路计算机应用, 2022, 31(10): 10-16.
- Zhang Xin, Dong Cheng-liang, Wang Xiao-chen, et al. Detection method of passenger flow density in train carriage based on improved YOLOv5s[J]. Railway Computer Application, 2022, 31(10): 10-16.
- [18] 周飞, 郭杜杜, 王洋, 等. 基于改进YOLOv8的交通监控车辆检测算法[J]. 计算机工程与应用, 2024, 60(6): 110-120.
- Zhou Fei, Guo Du-du, Wang Yang, et al. Vehicle detection algorithm based on improved YOLOv8 in traffic surveillance[J]. Computer Engineering and Applications, 2024, 60(6): 110-120.
- [19] 葛雯, 姜添元. 改进YOLO与DeepSORT检测跟踪算法的研究[J]. 计算机仿真, 2022, 39(5): 186-190.
- Ge Wen, Jiang Tian-yuan. Research on improved YOLO and deepsort detection and tracking algorithms [J]. Computer Simulation, 2022, 39(5): 186-190.