

余海燕, 霍爱清, 冯若水. 基于 DeepSort 的多目标车辆跟踪优化算法[J]. 智能计算机与应用, 2025, 15(5): 68-74. DOI: 10.20169/j. issn. 2095-2163. 24121701

基于 DeepSort 的多目标车辆跟踪优化算法

余海燕, 霍爱清, 冯若水

(西安石油大学 电子工程学院, 西安 710065)

摘要: 针对密集交通场景中车辆跟踪所面临的漏检率高、误检频发以及跟踪精度低等问题, 提出一种基于 DeepSort 多目标车辆跟踪优化算法, 旨在提升其在复杂环境下的跟踪性能。首先, 优化卡尔曼滤波, 通过增加自适应调制噪声尺度的机制, 动态调整噪声协方差, 更准确地预测目标的运动轨迹, 克服因噪声水平波动导致的预测偏差和不稳定现象。随后, 采用 ResNest50 作为主干网络, 并结合 YOLOv5 检测器, 对外观特征提取网络进行了改进, 增强对车辆外观特征的精细提取能力, 达到准确检测跟踪场景中的多个目标车辆的目的。为了全面评估所提算法的性能, 基于 *MOTA*、*IDSW* 和 *MOTP* 三种指标, 进行了综合评估实验并辅以可视化对比实验。经过实验验证, 所提算法与其他模型相比在 *MOTA*、*MOTP* 上平均提升了 3.460%、3.103%, 同时使 *IDSW* 平均下降了 45.907%, 提高了多目标车辆检测的鲁棒性及跟踪精度。

关键词: DeepSort; NSA 卡尔曼滤波; ResNeSt50; 密集场景; 多目标车辆跟踪; YOLOv5

中图分类号: TP391

文献标志码: A

文章编号: 2095-2163(2025)05-0068-07

Multi-target vehicle tracking optimization algorithm based on DeepSort

YU Haiyan, HUO Aiqing, FENG Ruoshui

(School of Electronic Engineering, Xi'an Shiyou University, Xi'an 710065, China)

Abstract: Aiming at the problems of high miss detection rate, frequent false detection and low tracking accuracy in vehicle tracking in dense traffic scenes, an optimization algorithm for multi-target vehicle tracking based on DeepSort is proposed to improve its tracking performance in complex environments. Firstly, Kalman filter is optimized to dynamically adjust noise covariance by adding adaptive modulation noise scale mechanism, so as to predict target trajectory more accurately and overcome prediction bias and instability caused by noise level fluctuation. Then, ResNest50 is used as the backbone network, combined with YOLOv5 detector, the appearance feature extraction network is improved to enhance the fine extraction ability of vehicle appearance features, so as to accurately detect multiple target vehicles in the tracking scene. In order to evaluate the performance of the proposed algorithm comprehensively, a comprehensive evaluation experiment is carried out based on *MOTA*, *IDSW* and *MOTP*, supplemented by visual contrast experiments. Compared with other models, the proposed algorithm improves *MOTA* and *MOTP* by 3.460% and 3.103% respectively, and reduces *IDSW* by 45.907%, which improves the robustness and tracking accuracy of multi-target vehicle detection.

Key words: DeepSort; NSA Kalman filter; ResNeSt50; dense scene; multi-target vehicle tracking; YOLOv5

0 引言

近年来, 交通环境变得日益复杂, 这也给跟踪算法的准确性提出了更高要求。为应对这一难题, 在轻量化与高效算法探索方面, 欧阳博等学者^[1]提出

了 FA-SORT 轻量化多目标跟踪算法, 展现了在多车辆跟踪任务中的优越性, 但未充分考虑行人对车辆行驶产生的潜在影响。Li 等学者^[2]进行了目标检测和跟踪算法轻量化设计, 降低了算法的复杂度。在算法改进与适应性提升方面, 赵晋芳等学者^[3]对

基金项目: 陕西省科技厅一般项目(2020GY-152)。

作者简介: 余海燕(2000—), 女, 硕士研究生, 主要研究方向: 深度学习, 目标检测; 冯若水(1998—), 女, 硕士研究生, 主要研究方向: 人工智能, 目标检测。

通信作者: 霍爱清(1966—), 女, 博士, 教授, 主要研究方向: 导向钻井控制。Email: aqhuo@xsyu.edu.cn。

收稿日期: 2024-12-17

SORT 算法进行了改进,但在处理极端遮挡、快速移动等情况时仍面临挑战。马庆禄等学者^[4]提出了面向车流量的智能检测 YOLOv7 算法改进,但在实际应用中仍存在误检、复检等问题。此外,在特殊场景下的算法研究方面, Ma 等学者^[5]针对弱光场景下的多目标跟踪进行了深入研究,为解决光照不足条件下的跟踪难题提供了新思路。Yuan 等学者^[6]对基于滤波的单目标跟踪算法进行了系统总结,为不同应用场景下的算法选择和优化策略提供了重要参考。

尽管上述研究在多目标跟踪领域取得了显著进展,但仍面临两大核心难题:一是密集场景下目标车辆易遮挡,导致跟踪准确性下降,甚至目标丢失;二是复杂环境中难以完全避免误检与重复检测的问题。

目标检测算法的准确性和效率直接决定了跟踪系统的性能。在目标检测算法方面, SSD^[7] 算法更适用于静态图像的目标检测任务,在多目标跟踪过程中,易造成漏检、误检等问题。SORT 算法是一种基于卡尔曼预测与匈牙利算法匹配的多目标追踪算法^[8],但在跟踪目标遇到遮挡时,易丢失目标,导致 ID 频繁切换。JDE^[9] 算法在处理长时间遮挡或严重遮挡的情况下,位置信息提取不足造成目标 ID 切换。DeepSort 算法采用了重识别技术,改善了目标被遮挡时的身份切换问题,提高了追踪的准确性^[10]。

鉴于上述局限性,本文提出一种基于 DeepSort^[11] (Deep Simple online and realtime tracking) 多目标车辆跟踪优化算法,旨在密集场景下提高目标车辆的跟踪性能,减少漏检和误检现象。主要工作如下:

(1) 通过改进卡尔曼滤波器提升预测准确度,减少遮挡丢失。

(2) 引入 ResNest50 增强车辆细节捕捉,提升密集场景性能。

(3) 融合 YOLOv5 算法,强化跟踪系统整体性能与鲁棒性。

1 基于 DeepSort 的算法改进

DeepSort 算法结合了深度学习模型的目标检测和卡尔曼滤波器的目标状态预测和数据关联,故主要从以下方面展开对算法的优化。

1.1 卡尔曼滤波优化

卡尔曼滤波虽融合先验知识与实时测量,但受线性与高斯噪声限制,难以满足在实时变化的环境

下对数据的处理需求。选择 NSA 卡尔曼滤波^[12],在保留卡尔曼滤波基本框架的基础上,通过自适应机制自动调节卡尔曼滤波算法中的数值参数提高滤波效果。卡尔曼滤波优化过程如图 1 所示。参照图 1,这里将对卡尔曼滤波优化过程展开研究分述如下。

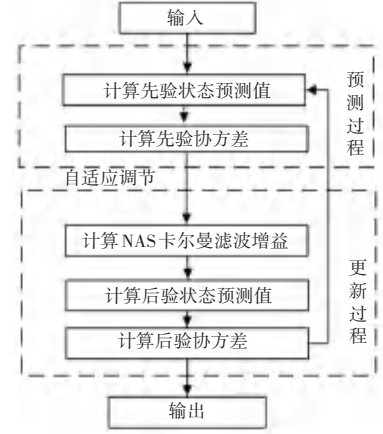


图 1 卡尔曼滤波优化过程

Fig. 1 Kalman filter optimization process

(1) 预测过程。卡尔曼滤波预测过程,通过估计其预测框的初始轨迹,基于 $t-1$ 时刻的状态估计进行 t 时刻的跟踪框状态预测,更新预测的不确定性。预测过程可以表示为:

$$x_t = \phi x_{t-1} \quad (1)$$

$$P_t = \phi P_{t-1} \phi^T + Q \quad (2)$$

其中, x_t 表示 t 时刻的预测框; P_t 表示 t 时刻预测协方差矩阵; Q 表示滤波器的过程噪声矩阵; ϕ 表示状态转移矩阵,可表示为:

$$\phi = \begin{bmatrix} 1 & 0 & 0 & 0 & dt & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & dt & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & dt & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & dt \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} \quad (3)$$

其中, dt 表示时间帧差。

(2) 更新过程。卡尔曼滤波更新过程是建立在预测过程基础上,首先计算 t 时刻的目标车辆增益 K_t , 计算公式如下:

$$K_t = \frac{\bar{P}_t H^T}{H \bar{P}_t H^T + R} \quad (4)$$

其中, $\mathbf{R}$ 为观测噪声的协方差矩阵, 表示观测过程中的不确定性; $\mathbf{P}_t$ 表示 t 时刻预测协方差矩阵; $\mathbf{H}$ 表示当前系统状态到观测数值的转换矩阵, 具体如下:

$$\mathbf{H} = \begin{bmatrix} \hat{e}_1^1 & 0 & 0 & 0 \\ \hat{e}_1^0 & 1 & 0 & 0 \\ \hat{e}_0^1 & 0 & 1 & 0 \\ \hat{e}_0^0 & 0 & 0 & 1 \\ \hat{e}_0^0 & 0 & 0 & 0 \\ \hat{e}_0^0 & 0 & 0 & 0 \\ \hat{e}_0^0 & 0 & 0 & 0 \end{bmatrix} \quad (5)$$

为减轻噪声波动对目标状态更新的影响, NSA 卡尔曼滤波采用自适应调节, 依据检测结果的置信度动态调整噪声协方差矩阵: 高置信度时减小噪声协方差矩阵的值, 低时则增大。获得噪声协方差矩阵表达式, 如下所示:

$$\tilde{\mathbf{R}}_t = (1 - x_t) \mathbf{R}_t \quad (6)$$

其中, x_t 表示 t 时刻检测置信度。

调整后的增益表达式, 如下所示:

$$K_t = \frac{\tilde{\mathbf{P}}_t \mathbf{H}^T}{\mathbf{H} \tilde{\mathbf{P}}_t \mathbf{H}^T + \tilde{\mathbf{R}}_t} \quad (7)$$

其中, $\mathbf{R}_t$ 表示预测的噪声协方差矩阵。

利用 NSA 卡尔曼滤波调整后的增益进行目标后验状态的更新, 后验状态预测值 $\hat{\mathbf{Y}}_t$ 为:

$$\hat{\mathbf{Y}}_t = \bar{\mathbf{Y}}_t + K_t (\mathbf{Z}_t - \mathbf{H} \bar{\mathbf{Y}}_t) \quad (8)$$

其中, $\mathbf{Z}_t$ 表示目标实际观测值。

计算后验协方差为:

$$\mathbf{P}_t = (\mathbf{I} + K_t \mathbf{H}) \tilde{\mathbf{P}}_t \quad (9)$$

其中, $\mathbf{I}$ 表示单位矩阵。

1.2 外观特征提取网络改进

DeepSort 通常采用 CNN (Convolutional Neural Network) 网络进行外观特征提取, 在处理复杂场景和细粒度特征时具有一定的局限性。为此选择 ResNeSt50^[13] 网络进行多目标跟踪。

ResNeSt50 融合了选择性核网络^[14] (Selective Kernel Networks, SKNet) 和压缩与激励网络^[15] (Squeeze-and-Excitation Networks, SENet) 的思想, 通过特征分离和分组技术, 不仅强化了跨通道特征提取的能力, 还利用统一化操作有效降低了模型复杂度, 并通过引入分割注意力块模块 (Split Attention Block, SAB), 使注意力能够跨特征图组。SAB 结构示意如图 2 所示。

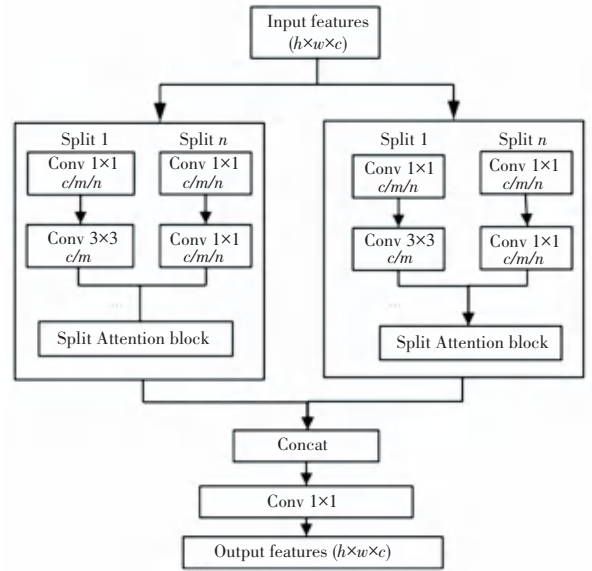


图 2 SAB 结构示意图

Fig. 2 SAB structural representation

SAB 模块通过特征图的双重分组 (基数组与切片组) 策略, 凭借基数组 K 与切片组 R 捕捉通道特征进行特征细化, 提高模型表达能力和计算效率。

同时, 采用类似 SKNet 的结构实现多路结构到统一操作的转换。通过 2 个全连接层和 1 个激活函数的组合, 计算出各分组的权重。对于第 n 个基数组的第 m 个切片, 其权重计算过程可以表示为:

$$S_C^n = \frac{1}{H \times W} \sum_{i=1}^H \sum_{j=1}^W U_C^n(i, j) \quad (10)$$

其中, C, H, W 分别表示输出特征图的各维度大小。

第 n 个基数组为 $U^n = \sum_i^{Rk} U_j$, 其中 $U^n \in R^{H \times W \times C/K}$, $n \in 1, 2, \dots, K$ 。

每个特征图通道通过切片加权组合的方式形成, 如第 c 个通道的权重计算过程表达式具体如下:

$$V_c^K = \sum_{i=1}^R \alpha_i^K(c) U_{R(k-1)+i} \quad (11)$$

$$\alpha_i^K(c) = \begin{cases} \frac{\exp(g_i^c(s^k))}{\sum_{j=0}^{R-1} \exp(g_j^c(s^k))}, & R > 1 \\ \frac{1}{1 + \exp(-g_i^c(s^k))}, & R = 1 \end{cases} \quad (12)$$

其中, g_i^c 表示映射函数; $\alpha_i^K(c)$ 表示分配权重。

如此, ResNeSt50 通过引入平均池化以减少特征信息损失, 提高计算效率; 将 7×7 卷积替换为 3 个 3×3 卷积, 加速网络训练。

1.3 YOLOv5 算法融合

选用 YOLOv5 作为优化的 DeepSort 跟踪算法的车辆位置检测器,通过检测与跟踪的相互反馈以提高系统对复杂场景变化的适应性,实现快速、准确的获取目标的边界框和尺寸信息。

最后,通过级联匹配方法将检测到的边界框与卡尔曼滤波器预测的预测框进行匹配。级联匹配成功,使用 NSA 卡尔曼滤波算法更新该轨迹的状态信息;级联匹配失败,通过 IoU^[16] 匹配确定目标与轨迹

的关系。IoU 匹配后,获得 3 种匹配结果:

(1)匹配成功,则更新轨迹信息。

(2)IoU 值高于阈值但无成功匹配,视为新目标,创建新轨迹。

(3)既无级联匹配、又无 IoU 匹配,则视为误检或噪声,不进行轨迹处理。重复上述流程,最终的跟踪结果输出为所有成功追踪到的轨迹信息。DeepSort 优化跟踪算法的整体结构如图 3 所示。

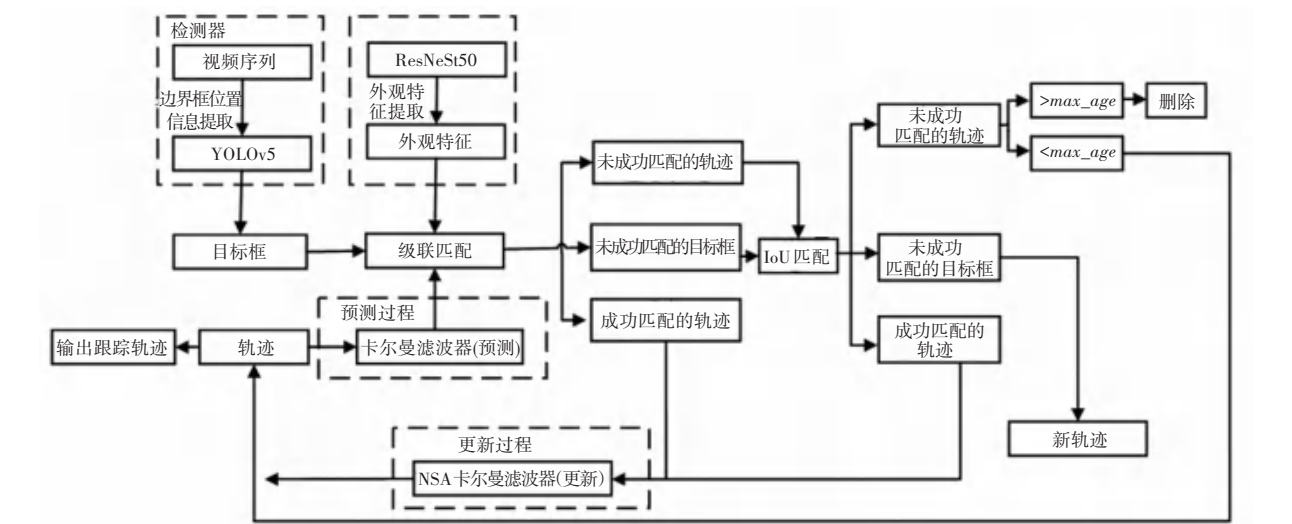


图3 改进 DeepSort 跟踪模型整体结构图

Fig. 3 Overall structure diagram of improved DeepSort tracking model

2 实验结果与分析

2.1 实验环境配置

本文实验平台配置见表 1。

表 1 实验平台配置

Table 1 Experimental platform configuration	
软件/硬件	版本
操作系统	Windows 10
CPU	Intel Core i7-10510H
GPU	NVIDIA CUDA 10.0
深度学习框架	Pycharm

2.2 数据集及评估指标

2.2.1 数据集

本文选取 UA-DETRAC 数据集^[17](包含 4 个不同地点,8 250 车辆、14 万帧画面、121 万个标记的对象边界框)展开实验。

2.2.2 综合评估指标

从跟踪准确性、稳定性及实时性三个关键维度,选用跟踪准确率^[18](Multi-Object Tracking Accuracy, *MOTA*)、轨迹身份 ID 切换次数^[19]

(IDentity SWitch, *IDS**W*) 和跟踪精确率^[20](Multi-Object Tracking Precision, *MOTP*) 三种性能评价指标对算法进行了综合评估实验。

(1) *MOTA*。针对误检、漏检及匹配失误,采用 *MOTA* 确保检测跟踪的一致性与准确性。*MOTA* 的计算公式为:

$$MOTA = 1 - \frac{\sum_t (TN_t + AP_t + IDS W_t)}{\sum_t CT_t} \quad (13)$$

其中, *TN* 表示漏检情况,真实框与预测框未正确匹配;*CT* 表示当前帧中实际目标的数量;*AP* 表示误检情况,无真实框但模型错误地预测出了目标框。

(2) *IDS**W*。用于统计整个跟踪序列中目标身份发生错误切换的总次数。

(3) *MOTP*。用于评估跟踪器锁定目标的位置精度,值越高、跟踪效果越佳。*MOTP* 的计算公式为:

$$MOTP = \frac{\sum_{t,j} W_{t,j}}{\sum_t S_t} \quad (14)$$

其中, $W_{t,j}$ 表示在第 t 帧上第 j 个位置的精度; S_t 表示对于 t 帧的匹配数量的总和。

2.3 实验结果对比与分析

2.3.1 多算法的对比实验

在保持参数设置环境一致的前提下,建立对当前主流的多目标跟踪方法 (SORT、JDE)、YOLOv5 + DeepSort、基于 YOLOv5 改进的 DeepSort 跟踪模型的对比实验,实验结果见表 2。

表 2 不同算法的对比结果

算法	$MOTA \uparrow$	$IDSW \downarrow$	$MOTP \uparrow$
SORT	15.37	12 875	78.10
DeepSort	19.20	16 790	79.80
JDE	20.11	14 640	79.83
YOLOv5+DeepSORT	19.50	16 784	80.10
本文	20.50	7 890	81.80

实验结果表明,DeepSort 算法相较于 SORT,尽管在某些情况下 $IDSW$ 增加了 30.407%。但在 $MOTA$ 和

$MOTP$ 上,相比 SORT 算法均提升了 24.918%、2.176%。同时,与 SORT、JDE、YOLOv5+DeepSort 对比时,在 $MOTA$ 上分别提升了 3.33%、1.93% 和 5.12%;在 $MOTP$ 上也分别提升了 4.73%、2.46% 和 2.12%;在身份切换次数 ($IDSW$) 上相比其他模型分别降低了 38.72%、46.10% 和 52.90%。

2.3.2 可视化实验

根据截取 UA-DETRAC 数据集测试视频的不同序列帧,建立可视化实验,可视化结果如图 4 所示。

图 4 展示了目标车辆的 ID 标签,其中 ID 为 2、6、7、9 的车辆持续被跟踪。在图 4(c) 中, ID 为 7 的车辆因机动车行驶导致遮挡而丢失跟踪,但算法通过保存表观特征在图 4(d) 中恢复了跟踪。然而,图 4(d) 中因车辆密度大,部分车辆如 ID 为 100 的丢失了跟踪。在图 4(e) 中,该车辆被误判为新目标并赋予 ID137,但系统在图 4(f) 中通过外观特征匹配成功修正,恢复了 ID 为 100。这体现了跟踪系统在复杂环境下的漏检处理和 ID 自动修正能力。

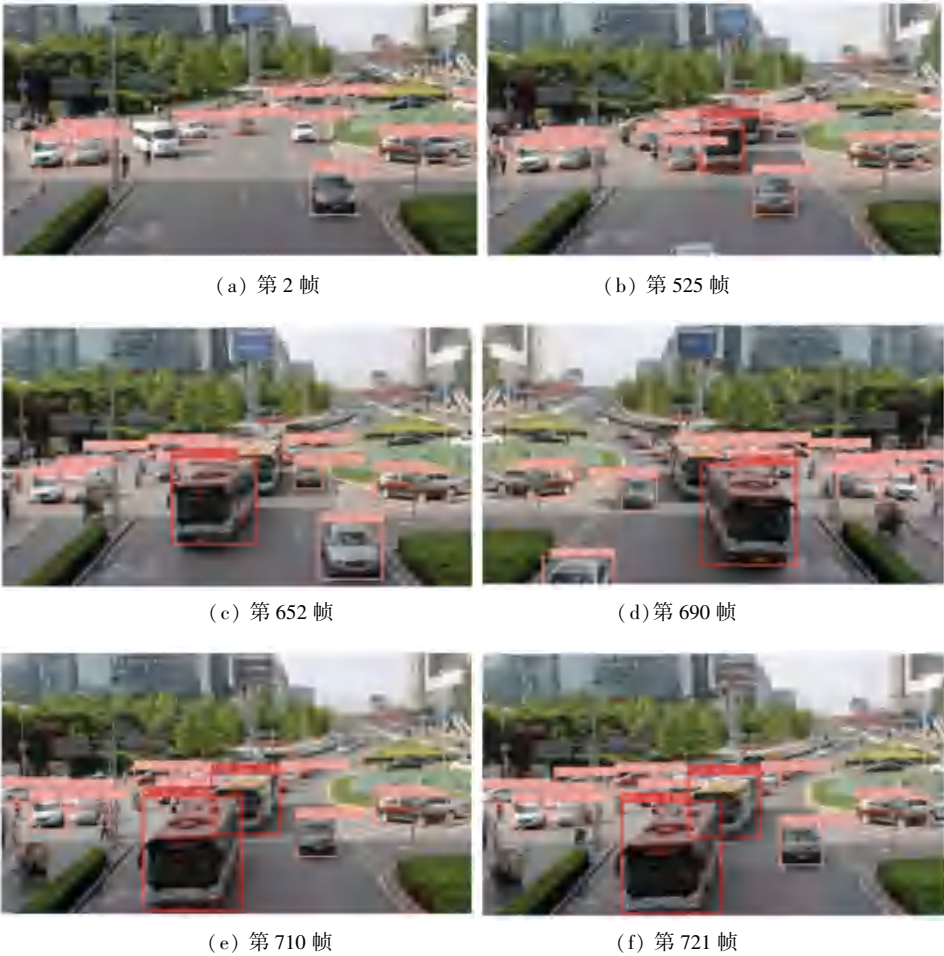


图 4 密集场景下车辆跟踪的可视化结果

Fig. 4 Visualization of vehicle tracking in dense scenes

2.3.3 跟踪效果比对实验

将 YOLOv5+DeepSort 与本文的优化算法展开可视化对比实验以检验算法的跟踪准确性。图 5 展示了优化前后的跟踪算法性能对比结果,左侧图像为原算法的应用效果,右侧为本文所提出跟踪算法结果。

从第 250 帧至 1 043 帧,改进前的算法对 ID 为

5 的车辆跟踪错误,且错误地将行人与机动车纳入跟踪轨迹。改进后的算法则能稳定准确地跟踪目标车辆,避免误跟踪。在最后一帧中,原算法因频繁误检和 ID 跳变导致跟踪目标数量远超实际,而优化算法仅标注到第 49 辆车,只比实际车辆总数多出 22 辆,显著减少了这一数量,有力证明了其在缓解 ID 标签频繁切换问题上的优势。

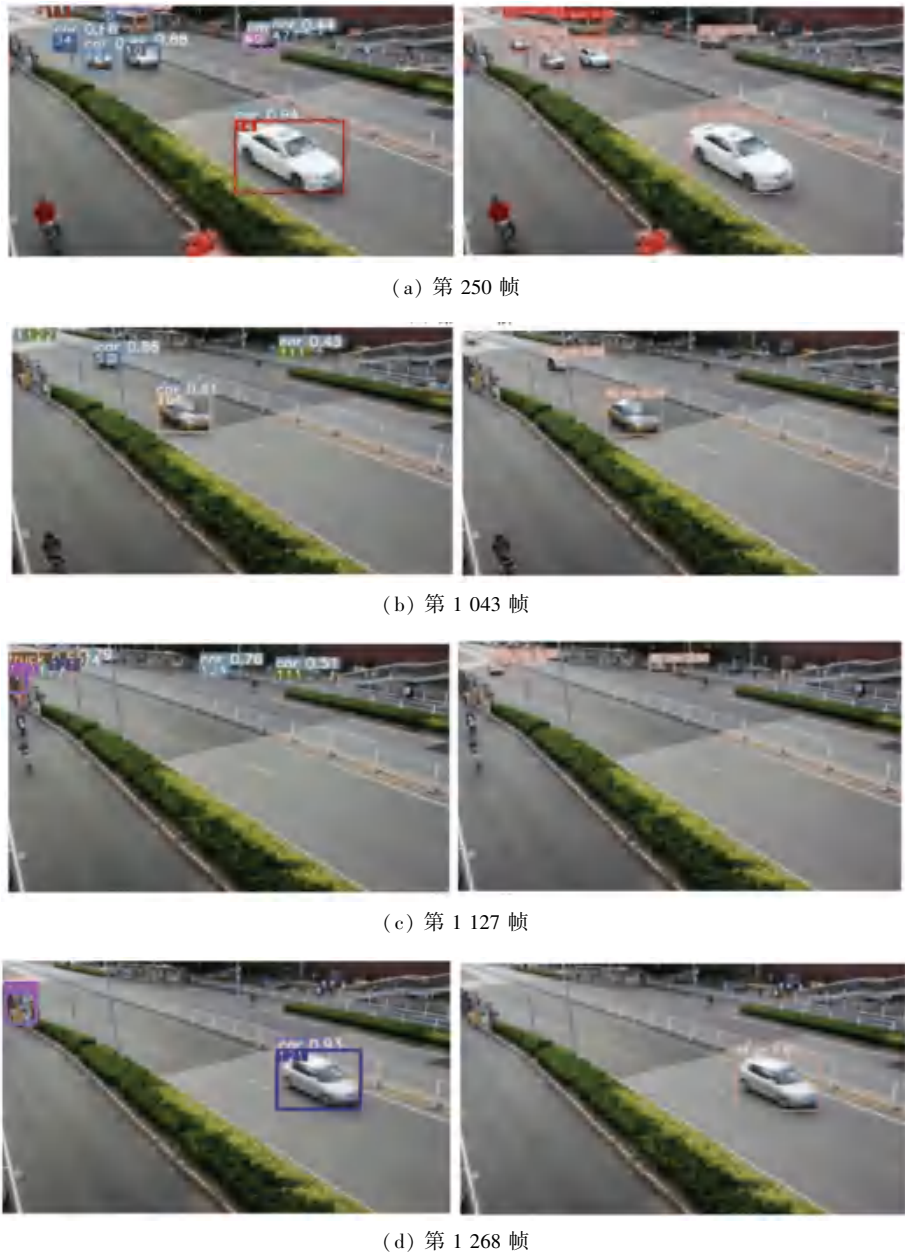


图 5 跟踪算法性能可视化对比结果

Fig. 5 Visual comparison results of tracking algorithm performance

3 结束语

本文针对密集场景下车辆跟踪的精度低、漏检、误检等问题,提出了基于 DeepSort 的多目标跟踪优

化算法。该算法通过采用 NSA 卡尔曼滤波自适应调整协方差,与 ResNeSt50 网络和 YOLOv5 检测器相结合,显著提升了多目标跟踪精度和稳定性。与其他算法相比,在 $MOTA$ 、 $MOTP$ 和 $IDSW$ 评估中,所

提算法分别平均提升了 3.460%、3.103%，并使 *IDSW* 平均下降了 45.907%，有效提高了多目标跟踪的准确性和鲁棒性。

参考文献

[1] 欧阳博,朱勇建,杨礼康,等. FA-SORT:轻量化的多车辆跟踪算法[J]. 计算机工程与应用,2024,60(9):122-134.

[2] LI Xinxin, LI Zuojun, HU Chaofang, et al. Light-weight multi-target detection and tracking algorithm based on M3-YOLOv5 [C]//2023 42nd Chinese Control Conference (CCC). Tianjin, China:CAA, 2023:8159-8164.

[3] 赵晋芳,李权,赵晋利. 基于改进 SSD-MobileNetV3 网络和 SORT 的车辆识别与跟踪[J]. 自动化与仪器仪表,2023(11):16-19.

[4] 马庆禄,吴跃川,张梓轩,等. 面向车流量智能检测的 YOLOv7 算法改进与应用[J]. 公路,2024,69(1):242-249.

[5] MA L, MARIANO V Y. Construction and research of multi-target tracking method in low light scene based on PP-YOLOE+m and PP-tracking[C]//2023 4th International Conference on Intelligent Computing and Human-Computer Interaction (ICHCI). Piscataway,NJ:IEEE, 2023:208-214.

[6] YUAN Zhengdao, SHAN Dongri. Algorithms based on correlation filter in target tracking: A survey [C]//2021 IEEE International Conference on Electrical Engineering and Mechatronics Technology (ICEEMT). Piscataway,NJ: IEEE, 2021:786-790.

[7] 彭林聪,王克瑞,周浩,等. 基于改进 SSD 的目标检测算法[J]. 激光杂志,2024,45(11):71-76.

[8] 叶浩,徐令强,黄杰. 基于改进 YOLOv5 和 SORT 算法的车辆跟踪系统[J]. 自动化应用,2024,65(13):12-15.

[9] 王俊,王鹏,李晓艳,等. 融合多阶语义增强的 JDE 多目标跟踪算法[J]. 西北工业大学报,2022,40(4):944-952.

[10] 柴园园,朱民耀,曹向明,等. 基于 YOLOv5s+DeepSORT 的公共场所行人体温监测方法[J]. 工业控制计算机,2024,37(11):40-41.

[11] RAKOTONIAINAZ A T, CHELBI N E, GINGRAS D, et al. LIV-DeepSORT: Optimized DeepSORT for multiple object tracking in autonomous vehicles using camera and LiDAR data

fusion[C]//2023 IEEE Intelligent Vehicles Symposium (IV). Piscataway, NJ: IEEE, 2023:1-7.

[12] DU Yunhao, WAN Junfeng, ZHAO Yanyun, et al. Giaotracker: A comprehensive framework for mcmot with global information and optimizing strategies in visdrone, 2021 [C]//Proceedings of the IEEE/CVF International Conference on Computer Vision. Piscataway, NJ: IEEE, 2021: 2809-2819.

[13] ZHANG Hang, WU Chongrao, ZHANG Zhongyue, et al. ResNest: Split-attention networks[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. Piscataway,NJ:IEEE,2022:2735-2745.

[14] LI X, WANG W, HU X, et al. Selective kernel networks[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. Piscataway, NJ: IEEE, 2019: 510-519.

[15] LIU Runyu. Pedestrian detection based on SENet with attention mechanism [C]// 2023 IEEE 7th Information Technology and Mechatronics Engineering Conference (ITOEC). Piscataway, NJ: IEEE, 2023:616-619.

[16] 孔维彬,游思思,姚涵涛. 自适应混合相似度的多目标匹配算法[J/OL]. 天津理工大学学报. (2024-06-21). <https://doi.org/10.20169/j.issn.2095-2163.24110802>.

[17] HUANG Gao, LIU Zhuang, LAURENS V D M, et al. Densely connected convolutional networks [C]// 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Piscataway,NJ:IEEE,2017: 2261-2269.

[18] QIN W, DU H, ZHANG X, et al. End to end multi-object tracking algorithm applied to vehicle tracking [C]//2022 Asia Conference on Algorithms, Computing and Machine Learning (CACML). Piscataway,NJ:IEEE, 2022:367-372.

[19] SAETCHNIKОВI V, SKAKUN V, TCHERNIAVSKAIA E A. Deep neural network-based dynamical object recognition and robust multi object tracking technique for onboard unmanned aerial vehicle's computer vision-based systems[J]. IEEE Journal on Miniaturization for Air and Space Systems, 2023,4(3):250-256.

[20] XU Yihong, ŠEP A, BAN Yutong, et al. How to train your deep multi-object Tracker2020 [C]//IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Piscataway, NJ:IEEE, 2020:6786-6795.