DOI:10.11918/202507041

面向高效车辆轨迹提取的轻量化目标追踪算法

李 岩[1],魏笑笑[1],陈桂珍[1],李昱燃[1],马 静[2]

(1.长安大学 运输工程学院,西安 710064;2.长安大学 电子与控制工程学院,西安 710064)

摘 要:针对路侧设备部署的算法对车辆轨迹提取的效率及精度难以同时保障问题,提出一种高效的轻量化目标跟踪算法DeepSORT-CKFaster。该算法通过Cutmix数据增强方法缓解标签类别不平衡问题,提高YOLOv5目标检测算法的泛化能力;提出一种自适应反馈式的轻量特征提取网络,动态调整特征提取策略,减少冗余计算和频繁的内存访问,实现网络轻量化的同时提升特征提取效率;在训练过程中,提出自适应卡尔曼滤波学习率调整策略,通过实时监控梯度统计量,自动触发余弦下降或线性上升机制,在稳定提取特征的基础上提升模型收敛效果;所提出算法结合YOLOv5目标检测算法,可有效提取车辆轨迹。应用西安市高清视频在两种配置平台验证所提出算法的结果表明:所提出算法的平均目标追踪准确度为92.2%,平均目标追踪精确度为1.106 m,平均目标追踪速度为26.538 帧/s和12.819 帧/s;相比于SORT算法和基准DeepSORT算法,所提出算法的平均目标追踪准确度提高了4.3%~7.1%,平均目标追踪速度提升了4.811~11.789 帧/s。所提出算法特征提取模型的训练时间为6.2 h,计算量为1.376 GFLOPs,为基准DeepSORT算法模型训练时间的72.1%、模型计算量的71.8%。

关键词:交通信息技术;轻量化目标追踪算法;DeepSORT-CKFaster;车辆轨迹提取;YOLOv5

中图分类号:U495 **文献标志码**:A **文章编号**:0367-6234(2026)07-0054-08

Lightweight object tracking algorithm for efficient vehicle trajectory extraction

Li Yan[1], Wei Xiaoxiao[1], Chen Guizhen[1], Li Yuran[1], Ma Jing[2]

(1. School of Transportation Engineering, Chang′an University, Xi′an 710064, China;
2. School of Electronics and Control Engineering, Chang′an University, Xi′an 710064, China)

Abstract: In order to solve the problem that algorithm deployed on roadside equipment are difficult to guarantee the efficiency and accuracy of the vehicle trajectory extraction at the same time, an efficient lightweight object tracking algorithm DeepSORT-CKFaster was proposed. The algorithm alleviates the problem of label category imbalance and improves the generalization ability of the YOLOv5 object detection algorithm through the Cutmix data augmentation method. An adaptive feedback lightweight feature extraction network was proposed, which dynamically adjusted the feature extraction strategy, reduced redundant computing and frequent memory access, and realized the lightweight network and improved the efficiency of feature extraction. In the training process, an adaptive Kalman filter learning rate adjustment strategy was proposed, which automatically triggered the cosine descent or linear rise mechanism through real-time monitoring of the gradient statistics, so as to improve the convergence effect of the model on the basis of stable feature extraction. Combined with the YOLOv5 target detection algorithm, the vehicle trajectory can be effectively extracted. The results of the proposed algorithm verified by high-resolution video in Xi′an on two configuration platforms indicate that the average target tracking accuracy and average target tracking precision of the proposed algorithm are 92.2%, 1.106 m, respectively. The average target tracking speed of the proposed algorithm are 26.538 frames/s and 12.819 frames/s. Compared with the SORT algorithm and the basic DeepSORT algorithm, the average target tracking accuracy and average target tracking speed of the proposed algorithm are increased by 4.3%–7.1% and 4.811 frames/s–11.789 frames/s, respectively. The training time of the feature extraction model of the proposed algorithm is 6.2 h, and the computation amount is 1.376 GFLOPs, which is 72.1% of the training time and 71.8% of the model computation of the basic DeepSORT algorithm.

Keywords: traffic information technology; lightweight target tracking algorithm; DeepSORT-CKFaster; vehicle trajectory extraction; YOLOv5

收稿日期:2025-07-17;**录用日期**:2025-08-28;**网络首发日期**:2025-09-01
网络首发地址:https://link.cnki.net/urlid/23.1235.t.20250901.1058.002
基金项目:陕西省自然科学基础研究计划(2025JC-YBMS-367,2025JC-YBQN-524);福建省交通运输科技计划重点科技项目(ZD202405,YB202431)
作者简介:李 岩(1983—),男,教授,博士生导师
通信作者:马 静,majing123@chd.edu.cn

高精度的车辆行驶轨迹数据是交通流特性分析、交通行为分析及交通安全评价等的基础[1]。传统轨迹数据采集方法存在检测范围小、精度低、成本高等局限性[2]，随着计算机图像技术的快速发展，路侧高清视频成为车辆轨迹数据的一种重要数据源[3]。路侧视频一般需回传至中央集群端处理，需要大量传输、计算及存储资源才可得到高精度车辆轨迹信息。然而，可在路侧设备端部署的轨迹提取算法受限于设备算力等因素，提取效率和精度均较难满足交通管控的需求[4]，故亟需一种高效、轻量、可部署于路侧设备的车辆轨迹提取算法，在路侧设备提取高精度的车辆行驶轨迹后，将轨迹数据上传至中央平台。由于无需传输大量的视频数据，可在通信带宽受限的环境下，实现交通信息的高效感知与智能获取。

基于计算机视觉的车辆轨迹提取技术包括目标检测和目标跟踪两个部分[5]。在将常用算法轻量化部署于路侧设备端时，两部分中均存在影响识别准确度和运行效率的因素。

基于深度学习的方法分为两阶段目标检测算法和单阶段目标检测算法。单阶段的目标检测算法，如 YOLO 系列算法[6]，直接利用锚框检测目标的类别和位置，其效率相比两阶段算法更高。与其他单阶段目标检测算法相比，YOLOv5 算法[7-8]具备较强的实时检测能力和较小的模型体积，在保持较高检测精度的同时，显著降低了计算资源需求，适合部署于路侧设备中。然而，在城市道路系统等复杂混合交通流系统中，车辆类型分布不均，存在目标类别不平衡现象。YOLOv5 算法容易受到不平衡数据集影响[7]，在训练时更倾向于拟合数量较多的类别，从而影响模型的检测精度。因此，需针对样本不平衡问题改进 YOLOv5 目标检测算法，以实现高效、精确的目标检测。

基准 DeepSORT 目标追踪算法[9]引入表观特征和级联匹配策略对 SORT 算法改进，提高了目标被遮挡时的跟踪性能，并克服了目标 ID 易于跳变的问题[10]。但其外观特征提取网络结构复杂、计算成本较高，难以满足路侧设备对实时性和高效性的要求。FasterNet 网络[11]作为轻量化神经网络架构，在降低参数量和计算复杂度的同时具备良好的性能和部署优势，适用于边缘设备。然而，其在训练过程中仍存在收敛速度慢、计算策略刚性等问题，影响特征提取的稳定性与准确性。因此，亟需对 DeepSORT 算法的特征提取模块优化，以提升其在复杂场景下的部署性能和资源利用效率。

学习率是深度学习训练中的关键超参数，直接影响模型的收敛速度和性能[12]。基准 DeepSORT 模型采用预先设定好的固定学习率训练，在整个训练过程中学习率保持不变，能够保证训练过程的稳定性，但缺乏对训练状态的动态响应，导致模型训练初期收敛慢、后期易陷入局部最优，影响训练效率和效果。因此，有必要设计动态的学习率调整策略，实现学习率的精准调控，提高模型训练的鲁棒性和适应性。

针对上述问题，拟构建一种可在路侧设备部署的高效、轻量化的目标追踪算法，以在路侧端高效地提取车辆轨迹，在减少视频数据回传所需网络传输带宽的同时，提升系统的实时性与资源利用效率，为交通行为分析及交通安全评价等研究奠定基础。

1　DeepSORT-CKFaster 算法

1.1　研究框架

基于计算机视觉的车辆轨迹提取办法主要包括目标检测和目标跟踪两个部分。如图 1 所示，以视频帧序作为输入，通过 YOLOv5 目标检测器获取每帧视频中车辆的位置、大小等信息，将其作为 DeepSORT 目标追踪算法的输入，为不同帧出现的同一车辆命名为相同 ID，以实现同一车辆状态随时间空间变化的更新，最终输出视频帧序、车辆 ID、车辆位置等信息。在 DeepSORT 目标追踪算法中，将数据集输入搭建好的模型框架中进行模型训练，依据评价指标判断模型训练结果是否符合期望，若不符合期望则需调整后重新训练模型，直至符合期望输出目标追踪模型。

为提升路侧设备对车辆轨迹的提取效率，主要应优化目标跟踪算法，在保证目标跟踪精度的同时，提升目标跟踪的速度。据此确定优化内容如下：1）针对标签类别不平衡问题，在 YOLOv5 目标检测算法中嵌入 CutMix 数据增强方法，将两个原图像样本的部分区域进行组合，生成新的训练样本和标签，从而提高模型的泛化性能。2）针对特征提取速度慢的问题，提出一种自适应反馈式的轻量化特征提取网络，在 FasterNet 特征提取网络中嵌入扩展卡尔曼滤波（extended Kalman filter，EKF），通过对目标状态估计结果反馈，实现高效的特征提取与资源动态分配，提升 DeepSORT 算法的速度与精度。3）考虑模型稳定性和收敛速度两方面的因素，提出自适应卡尔曼滤波学习率调整策略，通过实时监控训练过程中的梯度统计量，自适应地触发余弦下降或线性上升机制，使学习率处于动态变化的过程，保证模型稳定性的同时加快模型的收敛速度。综上，为实现车辆轨迹提取的轻量化追踪，将整体流程优化如图 1 所示。

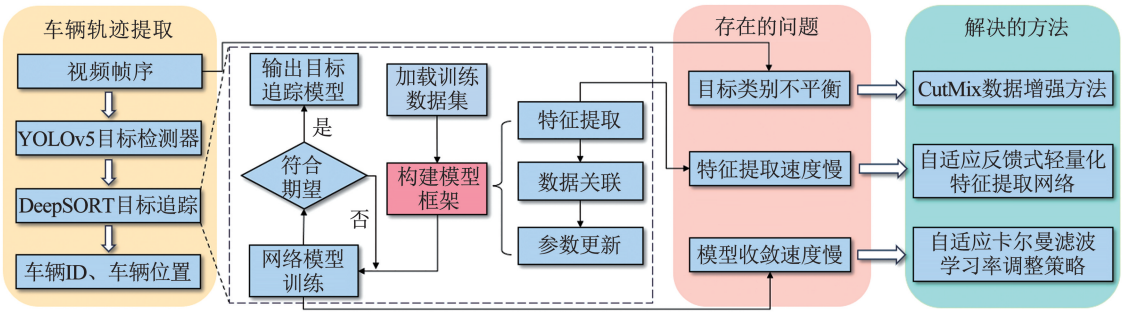


图 1 车辆轨迹提取流程

Fig. 1 Procedures on extracting vehicle trajectory

1.2 CutMix 数据增强方法

CutMix 数据增强方法在输入图像中随机剪切一个矩形区域,将另一幅图像的矩形区域与原图像进行混合,生成新的训练样本,同时改变新图片的标签,可提高较少类别样本的采样概率,较好地解决样本不均的情况,从而提高模型的泛化性能。CutMix 数据增强方法通过线性插值获得新图像样本及对应标签,具体按照剪切区域的面积加权来实现,操作过程如式(1)所示:

$$\begin{cases}\tilde{x}=M\odot x_{\mathrm{A}}+(1-M)\odot x_{\mathrm{B}}\\ \tilde{y}=\lambda y_{\mathrm{A}}+(1-\lambda)y_{\mathrm{B}}\end{cases}\tag{1}$$

式中:$\tilde{x}$ 为新的训练样本,$\tilde{y}$ 为新样本对应的标签;x_{A} 和 y_{A} 为样本 A 及其所对应的标签,x_{B} 和 y_{B} 为样本 B 及其所对应的标签;$\odot$ 代表逐像素相乘,$M\in\{0,1\}^{(W\times H)}$ 表示剪切和粘贴区域的二进制掩码,代表图像删除和填充的位置,对于样本,$M=1$ 部分的图像被保留,对于样本 B,$M=0$ 部分的图像被保留,W 和 H 为图像的宽和高;$\lambda\in[0,1]$,取值服从 Beta 分布 $\beta(\alpha,\alpha)$,令 $\alpha=1$ 则 λ 服从(0,1)的均匀分布。

选取裁剪区域 $C=(r_x,r_y,r_w,r_h)$ 对掩码 M 采样,其中 r_x、r_y 表示裁剪区域左上角的随机坐标点,r_w 和 r_h 表示裁剪区域的宽和高。裁剪区域的边界框采样公式为

$$\begin{cases}r_x\sim\mathrm{Unif}(0,W),r_w=W\sqrt{1-\lambda}\\ r_y\sim\mathrm{Unif}(0,H),r_h=H\sqrt{1-\lambda}\end{cases}\tag{2}$$

由式(2)可知,裁剪区域 C 的面积与样本 B 原图像的面积比例为

$$\frac{r_w r_h}{WH}=\frac{W\sqrt{1-\lambda}\times H\sqrt{1-\lambda}}{WH}=1-\lambda\tag{3}$$

保持裁剪区域与原图像的面积比例为 $\frac{r_w r_h}{WH}=1-\lambda$ 确定裁剪区域 C,将掩码 M 中区域 C 的掩码设置为 0,其余区域的掩码设置为 1,即移除样本 A 的区域 C,保留样本 B 的区域 C,从而完成掩码 M 的采样。将样本 B 的裁剪区域 C 填充到样本 A,并按照式(1)计算可得到新样本的标签。

1.3 自适应反馈式轻量化特征提取网络

为提高目标追踪算法的速度与精度,提出一种实时反馈式的轻量化特征提取网络,DeepSORT-CKFaster 算法的框架如图 2 所示。

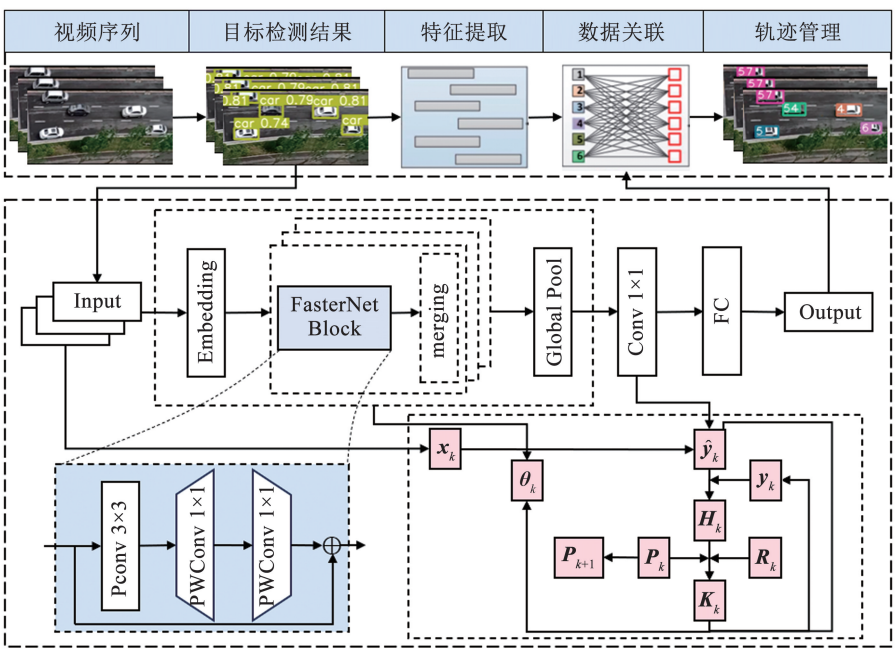


图 2 DeepSORT-CKFaster 算法的框架

Fig. 2 The framework of the DeepSORT-CKFaster algorithm

该特征提取网络以轻量化的 FasterNet 特征提取网络为主干结构，结合扩展卡尔曼滤波器的状态估计结果，构建闭环反馈机制，可实现在保持目标追踪精度的同时，有效提升算法的实时性与资源利用效率。

FasterNet 网络主要应用逐点卷积（pointwise convolution，PWConv）和部分卷积（partial convolution，PConv），以减少冗余计算和内存访问，提升特征提取速度。为进一步提升训练效率与资源利用率，在 FasterNet 模块中嵌入扩展卡尔曼滤波，构建状态感知驱动的动态计算机制。

该机制将 FasterNet 特征提取网络的输入作为 EKF 的输入向量 $\boldsymbol{x}_k$，将提取的特征作为 EKF 的观测值 $\boldsymbol{y}_k$，其神经网络行为可以表述为非线性离散时间系统：

$$\begin{cases}\boldsymbol{y}_k = g(\theta_k) + \boldsymbol{v}_k \\ \theta_{k+1} = \theta_k + \boldsymbol{\omega}_k\end{cases} \tag{4}$$

式中：θ_k 为平稳过程中神经网络的权重参数；$\boldsymbol{\omega}_k$ 为过程噪声，记 $\mathrm{cov}\{\boldsymbol{\omega}_k\} = \boldsymbol{Q}_k$；$\boldsymbol{y}_k$ 是测量向量；$\boldsymbol{v}_k$ 为测量噪声，记 $\mathrm{cov}\{\boldsymbol{v}_k\} = \boldsymbol{R}_k$；$g(\cdot)$ 是状态的非线性函数。根据泰勒公式，式（4）可围绕状态估计值 $\hat{\theta}$ 展开为

$$y_k = g(\hat{\theta}) + \frac{\partial g(\hat{\theta})}{\partial \hat{\theta}}(\theta - \hat{\theta}) + o(\theta) \tag{5}$$

忽略高阶项，训练过程中的 EKF 的修正及更新过程可以表述为

$$\begin{cases}\hat{\theta}_k = \hat{\theta}_{k-1} + \boldsymbol{K}_k[\boldsymbol{y}_k - g(\hat{\theta}_{k-1})] \\ \boldsymbol{K}_k = \boldsymbol{P}_k\boldsymbol{H}_k(\boldsymbol{R}_k + \boldsymbol{H}_k^{\mathrm{T}}\boldsymbol{P}_k\boldsymbol{H}_k)^{-1} \\ \boldsymbol{P}_{k+1} = \boldsymbol{P}_k - \boldsymbol{K}_k\boldsymbol{H}_k^{\mathrm{T}}\boldsymbol{P}_k\end{cases} \tag{6}$$

式中：$\boldsymbol{K}_k$ 为卡尔曼增益，$\boldsymbol{H}_k$ 为状态向量到测量向量的转换矩阵，$\boldsymbol{P}_k$ 为近似误差协方差矩阵。扩展卡尔曼滤波模型中的 $\boldsymbol{P}$、$\boldsymbol{Q}$ 和 $\boldsymbol{R}$ 均为对角矩阵，依据相关文献[13]设定 $\boldsymbol{P}_0 = 0.1\boldsymbol{I}$，$\boldsymbol{Q} = 1\times10^{-4}\boldsymbol{I}$，$\boldsymbol{R} = 1\times10^{-2}\boldsymbol{I}$，$\boldsymbol{K}_k$ 按照式（6）进行实时更新。

该机制利用协方差矩阵 $\boldsymbol{P}_k$ 衡量预测可信度，从而动态调整特征提取策略：当归一化协方差迹小于设定阈值时（设置不同参数取值对比模型训练状况后，确定阈值为 0.05），状态估计可靠，则应用 EKF 修正的权重参数 θ_k 实现特征提取网络优化，通过感兴趣区域（region of interest，ROI）裁剪的方式启用轻量化特征提取网络，以降低计算资源消耗；当归一化协方差迹超过阈值，则启用完整特征提取模块，以保证特征提取的完整性和精度。该机制实现了特征提取强度与计算资源之间的自适应调节，在保障目标跟踪精度的同时，显著提升了系统的实时性与部署效率。

1.4　自适应学习率调整策略

考虑到模型稳定性和收敛速度两方面的因素，将卡尔曼滤波引入至学习率控制模块，构建基于状态估计的自适应学习率调整策略，实现对传统固定学习率方法的有效改进。

该自适应学习率调整策略将学习率调节建模为一个状态估计问题，当前时刻的观测值 y_t（基于梯度统计量计算的“理想学习率”）及历史状态估计值（上一时刻卡尔曼滤波器输出的学习率估计值 η_{t-1}）作为卡尔曼滤波的输入，利用卡尔曼滤波动态评估学习率的合理性，输出当前时刻的学习率估计值 η_t 和协方差，并根据估计结果触发相应的调整机制。通过上述机制，训练过程中的学习率始终处于动态调控状态，实现了更稳定、高效的训练，提升了模型的收敛速度与泛化能力。自适应学习率调整策略的整体流程如图 3 所示。

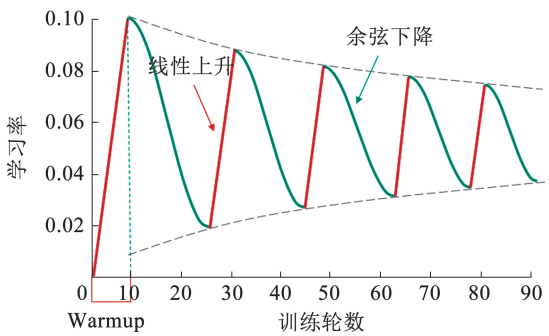


图 3　自适应学习率调整策略

Fig. 3　Adaptive learning rate adjustment strategy

训练初期，系统采用 Warmup 策略逐步提升学习率，帮助模型平稳启动；当滤波器判断学习率偏大、训练波动增大时，启用如式（7）所示的余弦退火策略平滑下降学习率，防止过拟合。

$$\eta(t) = \eta_{\min} + \frac{1}{2}(\eta_0 - \eta_{\min})\left(1 + \cos\frac{T_{\mathrm{cur}}}{T_{\max}}\pi\right) \tag{7}$$

式中：$\eta_{\max}$ 为当前的学习率，$\eta_{\min}$ 为设定的学习率最小值，T_{cur} 为当前 epoch，$T_{\max}$ 为余弦退火周期的最大长度。

若学习率偏小，则激活如式（8）所示的线性上升机制增强搜索能力，避免陷入性能瓶颈。

$$\eta(t) = \eta_0 + kt \tag{8}$$

式中：k 为线性增长斜率，η_0 为当前的学习率，增长持续至估计值附近。

2　实验设计

2.1　数据采集及预处理

采用西安市路侧高空安装的高清摄像设备在

2023 年 9 月 23 日—2023 年 9 月 30 日采集的数据构建模型测试数据集。该数据集包含机动车通行的快速路、设置主辅路或机非隔离的主干路以及机非共行的低等级城市道路,基本覆盖了主要城市交通场景,均为连续拍摄视频,视频分辨率为 1 920 × 1 080像素,帧率为 30 帧/s。提取各摄像头 6:00—22:00 的视频并对其切分处理,每个视频时长为 10 min。从视频中每隔 3 帧提取出总计 2 700 000 帧关键帧。根据道路类型将视频分为 3 类,每种道路类型均涵盖不同交通饱和度的道路,3 种道路类型拍摄的视频个数、道路等级以及平均交通饱和度的统计信息如表 1 所示。

表 1 数据集信息

Tab. 1 Dataset information

类别	道路类型	视频总时长/min	平均饱和度
类别 1	快速路	1 500	0.83
类别 2	主干路	1 500	0.76
类别 3	次干路	1 500	0.68

对从视频中提取的关键帧图像进行透视变换,使目标检测后车辆检测框像素坐标更能体现车辆真实的物理坐标位置。原始数据的图片进行透视变换以后,部分区域的车辆会产生非刚性形变,导致训练出的模型难以在透视变换后的图像上保持稳定性能[14]。故利用目标检测标注工具 LabelImg 对图像样本进行矩形框标注。数据集的检测类别包括:小型客车(car)、公交车(bus)和载货汽车(truck),其平均占比分别为 79%,15% 和 6%,随机选取数据集的 80% 作为训练集训练模型,20% 作为验证集验证模型性能。

2.2 实验环境及模型训练

本实验设置了两种硬件环境,以分析不同计算资源对目标检测与追踪算法性能的影响。两种实验环境的软件配置保持一致,均使用 Windows10 系统的 Visual Studio Code 为集成开发工具,PyTorch 为深度学习框架所搭建的深度学习平台,其主要硬件配置如表 2 所示。实验环境一为高性能台式计算机,主要用于模型训练与参数优化;实验环境二为配置较低的便携式计算机,更贴近实际路侧设备的计算能力。实验流程:先在环境一中完成模型训练,再将模型迁移至环境二进行部署测试,以评估算法在不同硬件条件下的适应性和运行效率,验证所提轻量化目标追踪方法的实际可行性与稳定性。

表 2 硬件环境配置对比

Tab. 2 Hardware configuration comparison of experimental environments

实验环境	设备类型	处理器	GPU
实验环境一	台式机	Intel(R) Core(TM) i7-9700	NVIDIA GeForce RTX 3060
实验环境二	笔记本	Intel(R) Core(TM) i7-10510U	NVIDIA GeForce MX250

根据网络深度和特征图宽度,YOLOv5 目标检测模型有 YOLOv5s、YOLOv5m、YOLOv5l 和 YOLOv5x。为提高模型训练的收敛速度,从而减小模型训练的计算量,以 YOLOv5 模型在 COCO(common object in context)数据集上的预训练权重为基础,在自制的数据集上训练模型。比较 4 个目标检测模型的性能,确定所选用的目标检测模型作为目标跟踪算法的检测器,并利用目标跟踪效果评价指标客观分析 SORT 算法、DeepSORT 算法及 DeepSORT-CKFaster 算法的跟踪性能,验证所提出车辆行驶轨迹提取方法的有效性。

2.3 评价指标

依据轨迹提取模型常用的评价指标,选取相应的性能指标对目标检测与目标跟踪模型的效果进行综合评估[15]。目标检测模型采用平均精度均值(mean average precision,mAP)和每秒处理的图像帧数(frames per second,FPS)两个指标来评价,mAP 和 FPS 分别为

$$\mathrm{mAP} = \frac{\sum_{i=1}^{K} \mathrm{AP}_i}{K} \tag{9}$$

$$\mathrm{FPS} = \frac{\mathrm{frameNum}}{\mathrm{elapsedTime}} \tag{10}$$

式中:K 为数据集的检测类别,AP 为各个类别的检测精度。

目标追踪模型的性能采用多目标跟踪准确度(multi-object tracking accuracy,MOTA)、多目标跟踪精确度(multi-object tracking precision,MOTP)、参数量、计算量、运行速率 FPS 和训练时间等 6 个性能评价指标衡量。MOTA 和 MOTP 可分别为

$$\mathrm{MOTA} = 1 - \frac{\mathrm{FN} + \mathrm{FP} + \mathrm{IDSW}}{\mathrm{GT}} \tag{11}$$

$$\mathrm{MOTP} = \frac{\sum_{t,i} d_{t,i}}{\sum_{t} c_t} \tag{12}$$

式中:FN 为漏检个数,FP 为误检个数,IDSW 为 ID 号发生转变个数,GT 为 Ground Truth 物体数量。c_t

为 t 帧时匹配的目标数量，$d_{t,i}$ 表示 t 帧时目标 i 的边界框与真实边界框间的距离。

3　实验结果与分析

3.1　目标检测结果分析

分别统计 4 种目标检测模型根据所采集数据集训练后的 mAP 和 FPS 指标的结果如表 3 所示。

表 3　目标检测算法实验结果

Tab. 3　Experimental results of object detection algorithm

模型	平均检测精度 mAP/%	检测速度/(帧·s^{-1})
YOLOv5s	94.2	21.753
YOLOv5m	93.4	21.281
YOLOv5l	93.7	23.297
YOLOv5x	93.5	23.159

由表 3 可知，目标检测模型的平均检测精度 YOLOv5s 最高，YOLOv5l、YOLOv5x、YOLOv5m 分别比 YOLOv5s 少 0.5%、0.7%、0.8%。目标检测模型的检测速度由快到慢依次为 YOLOv5l、YOLOv5x、YOLOv5s、YOLOv5m，YOLOv5x、YOLOv5s、YOLOv5m 分别比 YOLOv5l 慢 0.138 帧/s、1.544 帧/s、2.016 帧/s。由此可见，YOLOv5s 的目标检测精度最高，但检测速度相对较慢；YOLOv5l 的目标检测速度最快，且其目标检测精度仅比 YOLOv5s 低 0.5%。故在综合考虑检测速度与精确度因素时，宜选择 YOLOv5l 模型作为目标追踪器的输入。

3.2　Cutmix 数据增强实验

Cutmix 数据增强方法的超参数主要有 α 和 p，其中 α 控制标签与特征间的差值强度，常设置为0.5 和 1，p 表示增强的概率，常设置为 0.8、0.9 和 1[16]。基于 YOLOv5l 作为预训练权重训练的目标检测模型，利用平均检测精度 mAP 和损失值作为指标，比较引入 Cutmix 数据增强方法前后目标检测模型的性能，以及不同 α 和 p 取值组合下模型的性能，当 $\alpha=1$，$p=0.9$ 时，检测效果最好，实验结果如表 4 所示。

表 4　Cutmix 数据增强实验结果

Tab. 4　Results of Cutmix data augmentation experiment

模型	Cutmix 超参数	mAP/%	Loss
YOLOv5l	无	93.7	0.366
YOLOv5l	$\alpha=0.5, p=0.8$	94.6	0.323
YOLOv5l	$\alpha=0.5, p=0.9$	95.2	0.309
YOLOv5l	$\alpha=0.5, p=1$	94.2	0.341
YOLOv5l	$\alpha=1, p=0.8$	95.8	0.287
YOLOv5l	$\alpha=1, p=0.9$	96.5	0.235
YOLOv5l	$\alpha=1, p=1$	96.1	0.248

根据表 4 的实验结果得出，原 YOLOv5l 目标检测模型的平均目标检测精度为 93.7%，损失值为 0.366；引入 Cutmix 数据增强方法后目标检测模型的精度可提升至 96.5%，损失值为 0.235，说明该 Cutmix 数据增强方法可有效提高目标检测精度，降低损失。

3.3　DeepSORT-CKFaster 模型优化实验

使用训练后的 YOLOv5l 算法作为检测器，选取工程中广泛应用的 SORT 算法、DeepSORT 算法[17]及提出的 DeepSORT-CKFaster 算法进行目标跟踪实验。取各项评价指标的平均值作为衡量该算法跟踪性能的指标，其测试结果如表 5 所示。

表 5　目标追踪实验结果

Tab. 5　Results of object tracking experiment

算法	实验环境	数据	MOTA/%□	MOTP/m□	检测速度/(帧·s^{-1})□
SORT	实验环境 1 实验环境 2	类别 1/2/3/平均值	85.4/81.8/88.2/85.1	1.144/1.988/0.881/1.338	22.071/18.142/19.530/19.914 8.837/7.283/7.906/8.008
DeepSORT	实验环境 1 实验环境 2	类别 1/2/3/平均值	88.6/84.8/90.3/87.9	1.097/1.973/0.802/1.291	15.519/13.824/14.905/14.749 6.735/5.591/5.894/6.073
DeepSORT + FasterNet	实验环境 1 实验环境 2	类别 1/2/3/平均值	91.5/89.0/94.1/91.5	0.997/1.905/0.588/1.163	25.315/22.673/23.809/23.932 11.108/9.716/10.190/10.338
DeepSORT-CKFaster	实验环境 1 实验环境 2	类别 1/2/3/平均值	92.3/89.8/94.6/92.2	0.941/1.869/0.507/1.106	28.032/24.968/26.615/26.538 13.850/11.923/12.685/12.819

从表 5 可知，两种实验环境中各算法的 MOTA 和 MOTP 保持一致，说明在使用相同模型权重和测试数据的前提下，推理精度不受运行硬件环境影响。各对比模型中，SORT 目标追踪模型的 MOTA 最低，为 85.1%。基准 DeepSORT 算法的 MOTA 相对于 SORT 算法提升了 2.8%，但平均目标追踪速度降低了 1.935 帧/s。而 DeepSORT-CKFaster 算法的 MOTA 可达 92.2%，相对于基准 DeepSORT 算法提升 4.3%，MOTP 为 1.106 m，降低了 14.3%；与特征提取网络为 FasterNet 的 DeepSORT 算法相比，该算法

的 MOTA 值增加了 0.7%，MOTP 降低了 4.9%。实验环境一中，DeepSORT-CKFaster 算法的目标追踪速度可达 26.538 帧/s，比基准 DeepSORT 算法提升 11.789 帧/s，即 DeepSORT-CKFaster 算法可在保证目标追踪精度的同时，明显加快目标追踪速度。

在实验环境一中，类别 3 中数据的 MOTA 最高，类别 2 的 MOTA 最低；类别 1 的 FPS 最大，类别 2 的 FPS 最小。总体来看，高峰时段城市主干路与快速路车流密集、易发生拥堵，相较于快速路和次干路，目标追踪算法在主干路场景中的轨迹提取速度更慢，精度相对较低。对比两种实验环境可发现，实验环境一中各算法的平均帧率均高于实验环境二，说明在更高性能的硬件支持下，目标追踪算法的运行速度显著加快。3 种目标跟踪算法特征提取模型的训练时间及模型参数量、计算量指标的对比结果如表 6 所示。

表 6　目标跟踪算法对比

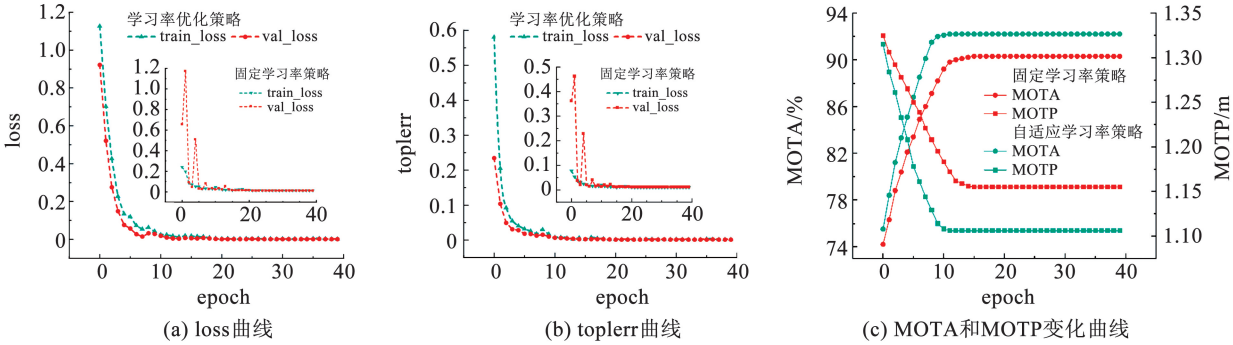
Tab. 6　Comparison of target tracking algorithms

算法	参数量/MB □	计算量/GFLOPs □	训练时间/h □	MOTA/% □	MOTP/m □	检测速度/(帧·s^{-1}) □ (实验环境 1/2)
SORT	38.5	1.054	6.1	85.1	1.338	19.914/8.008
DeepSORT	43.1	1.916	8.6	87.9	1.291	14.749/6.073
DeepSORT + FasterNet	40.9	1.128	5.3	91.5	1.163	23.932/10.338
DeepSORT-CKFaster	41.5	1.376	6.2	92.2	1.106	26.538/12.819

根据表 6 的实验结果，SORT 算法的模型参数量和 GFLOPs 均最小，MOTA 为 85.1%，即该模型最轻量化，计算量最小，但其目标追踪精度最低。与基准 DeepSORT 算法相比，DeepSORT-CKFaster 算法特征提取模型的参数量减少了 1.6 MB，训练时间减少了 2.4 h，计算量减少了 0.540 GFLOPs，即 DeepSORT-CKFaster 算法的模型更轻量化，精度较高，可满足在路侧部署以从道路视频监控数据高效提取车辆行驶轨迹的需求。

3.4　自适应卡尔曼滤波学习率调整策略实验

为提升目标跟踪模型对车辆的跟踪性能，需对 DeepSORT 模型重训练，而训练的学习率会影响模型的目标跟踪性能。合理的学习率调度和优化策略可加速收敛过程，提高模型的性能和稳定性。通过对不同学习率调度策略下的模型进行重训练，比较其对收敛速度与性能的影响，以确定最优的学习率调度方案。各策略下 DeepSORT-CKFaster 算法训练过程的 loss 曲线与 toplerr 曲线如图 4(a) 和 4(b) 所示，子图为固定学习率实验结果，主图为自适应卡尔曼滤波学习率调整策略实验结果；MOTA 和 MOTP 变化曲线如图 4(c) 所示。



(a) loss曲线　(b) toplerr曲线　(c) MOTA和MOTP变化曲线

图 4　不同学习率策略对比

Fig. 4　Comparison of various learning rate strategies

从图 4(a) 和 4(b) 可看出，固定学习率策略下模型训练过程不稳定，在训练进行到第 1 轮和第 4 轮时，验证集的 loss 和 toplerr 突然增大，存在一定的偶然性。与固定学习率相比，自适应卡尔曼滤波学习率调整策略可减少训练过程中的剧烈震荡，加速收敛过程，提高模型稳定性。从图 4(c) 可看出，与固定学习率策略的 DeepSORT-CKFaster 算法相比，采用自适应卡尔曼滤波学习率的 DeepSORT-CKFaster 算法收敛速度更快，其趋于稳定时的 MOTA 在 90.3% 的基础上增加了 1.9%，而 MOTP 降低了 4.1%，说明模型的追踪准确性和定位精度均得到优化。由此可见，自适应卡尔曼滤波学习率

策略能有效地引导模型训练过程，提高模型的性能和稳定性。

4　结　论

1）为提高路侧设备的轨迹提取效率，提出了一种高效轻量化的车辆目标追踪方法。应用西安市实拍视频的测试结果表明，所提出方法在较低性能计算机上亦可快速有效地提取高精度车辆行驶轨迹。随着硬件性能的提升，车辆轨迹提取速度可大幅提高。

2）应用西安市视频监测数据的测试结果表明：DeepSORT-CKFaster 算法的 MOTA 可达到 92.2%，两种实验环境下该算法的检测速度可分别达到 26.538 帧/s和 12.819 帧/s，特征提取模型的训练时间为 6.2 h，计算量为 1.376 GFLOPs。与基准 DeepSORT 算法相比，平均目标追踪准确度增加了 4.3%，平均目标追踪速度提升了 6.746 ~ 11.789 帧/s，特征提取模型的训练时间减少 2.4 h，计算量减少了 0.540 GFLOPs。

3）后续可考虑在不同气候条件、交通管制区域等特殊场景下进行跨场景测试与对比分析，进一步提升复杂交通场景下车辆轨迹提取的精度与速度。

参考文献

[1] Chen X Q, Li Z B, Yang Y S, et al. High-resolution vehicle trajectory extraction and denoising from aerial videos [J]. IEEE Transactions on Intelligent Transportation Systems, 2021, 22(5): 3190

[2] Mu L, Zhao H, Li Y, et al. Traffic flow statistics method based on deep learning and multi-feature fusion [J]. CMES-Computer Modeling in Engineering & Sciences, 2021, 129(2): 465

[3] Li Y, Xu J H, Li Y R, et al. Estimation and prediction of freeway traffic congestion propagation using tagged vehicle positioning data [J]. Transportmetrica B: Transport Dynamics, 2024, 12(1): 2297143

[4] Zhang W B, Yu Y H, Qi Y, et al. Short-term traffic flow prediction based on spatio-temporal analysis and CNN deep learning [J]. Transportmetrica A: Transport Science, 2019, 15(2): 1688

[5] 游峰，梁健中，曹水金，等. 面向多目标跟踪的密集行人群轨迹提取和运动语义感知[J]. 交通运输系统工程与信息，2021，21(6)：42
You Feng, Liang Jianzhong, Cao Shuijin, et al. Dense pedestrian crowd trajectory extraction and motion semantic information perception based on multi-object tracking [J]. Journal of Transportation Systems Engineering and Information Technology, 2021, 21(6): 42

[6] Xing Y, Lü C, Cao D P. Personalized vehicle trajectory prediction based on joint time-series modeling for connected vehicles[J]. IEEE Transactions on Vehicular Technology, 2020, 69(2): 1341

[7] 黄凯雯，房宵杰，梅林，等. MPE-YOLOv5：面向边缘计算的轻量化 YOLOv5 手势识别算法[J]. 哈尔滨工业大学学报，2023，55(5)：1
Huang Kaiwen, Fang Xiaojie, Mei Lin, et al. MPE-YOLOv5: lightweight YOLOv5 gesture recognition algorithm for edge computing [J]. Journal of Harbin Institute of Technology, 2023, 55(5): 1

[8] 苏亚鹏，陈高曙，赵彤. 不平衡样本下的 SA-YOLO 自适应损失目标检测算法[J]. 中国科学院大学学报(中英文)，2024，41(3)：411
Su Yapang, Chen Gaoshu, Zhao Tong. SA-YOLO: self-adaptive loss object detection method under imbalance samples[J]. Journal of University of Chinese Academy of Sciences, 2024, 41(3): 411

[9] Wojke N, Bewley A, Paulus D. Simple online andrealtime tracking with a deep association metric [C]//2017 IEEE International Conference on Image Prcessing. Beijing: IEEE, 2017: 3645

[10] 马永杰，马芸婷，程时升，等. 基于改进 YOLOv3 模型与 DeepSORT 算法的道路车辆检测方法[J]. 交通运输工程学报，2021，21(2)：222
Ma Yongjie, Ma Yunting, Cheng Shisheng, et al. Road vehicle detection method based on improved YOLOv3 model and deepSORT algorithm [J]. Journal of Traffic and Transportation Engineering, 2021, 21(2): 222

[11] Chen J, Kao S, He H, et al. Run, don't walk: chasing higher FLOPS networks[C]// Proceedings for of Faster the Neural IEEE/CVF Conference on Computer Vision and Pattern Recognition. Vancouver: IEEE, 2023: 12021

[12] Xu J H, Lu W B, Li Y R, et al. A multi-directional recurrent graph convolutional network model for reconstructing traffic spatiotemporal diagram[J]. Transportation Letters, 2024, 16(5): 405

[13] 唐宏志，姜金辉. 扩展卡尔曼滤波对时变参数追踪性能的影响研究[J]. 南京航空航天大学学报，2022，54(2)：304
Tang Hongzhi, Jiang Jinhui. Research on tracking performance of extended Kalman filter for time varying parameters[J]. Journal of Nanjing University of Aeronautics & Astronautics, 2022, 54(2): 304. DOI:10.16356/j.1005 - 2615.2022.02.017

[14] 刘宏利，王雨林，邵磊，等. 基于检测框下边沿的单目视觉车辆测距研究[J]. 计算机科学，2024，51(增刊 1)：427
Liu Hongli, Wang Yulin, Shao Lei, et al. Study on monocular vision vehicle ranging based on lower edge of detection frame[J]. Computer Science, 2024, 51(S1): 427

[15] Zhang Q. Multi-object trajectory extraction based on YOLOv3-DeepSort for pedestrian-vehicle interaction behavior analysis at non-signalized intersections [J]. Multimedia Tools and Applications, 2023, 82(10): 15223

[16] 耿鹏志，唐云祁，樊红兴，等. 基于 CutMix 算法和改进 Xception 网络的深度伪造检测研究[J]. 激光与光电子学进展，2022，59(16)：358
Geng Pengzhi, Tang Yunqi, Fan Hongxing, et al. Deep forgery detection using CutMix algorithm and improved Xception network [J]. Laser & Optoelectronics Progress, 2022, 59(16): 358

[17] 贾志，李茂军，李婉婷. 基于改进 YOLOv5 + DeepSort 算法模型的交叉路口车辆实时检测[J]. 计算机工程与科学，2023，45(4)：674
Jia Zhi, Li Maojun, Li Wanting. Real-time vehicle detection at intersections based on improved YOLOv5 + DeepSort algorithm model[J]. Computer Engineering & Science, 2023, 45(4): 674