

# 基于 DeepSort 的动态车辆多目标跟踪方法研究\*

何维堃 彭育辉 黄炜 姚宇捷 陈泽辉

(福州大学,福州 350116)

**【摘要】**为了提高汽车对外界环境信息的感知能力和对动态车辆目标行为的预测能力,采用 YOLOX 作为前端检测器,结合优化的 DeepSort 跟踪算法开展动态车辆多目标跟踪方法研究。在车辆特征匹配过程中,提取 Haar-like 特征对车辆的明暗变化信息进行匹配,提高物体匹配精度;基于 DeepSort 重识别网络,采用改进的 ResNet13 作为特征提取的骨干网络,加入 SENet 调整不同通道维度的特征权重。使用实际道路驾驶采集的视频数据对改进算法与传统算法进行对比,结果表明:相较于传统 DeepSort 算法,改进算法的多目标跟踪精度(MOTA)提高了 1.4 个百分点,平均数比率(IDF1)提升了 7.7 百分点。

**主题词:**多目标跟踪 DeepSort 自动驾驶 Haar-like SENet

**中图分类号:**U462 **文献标识码:**A **DOI:** 10.19620/j.cnki.1000-3703.20220911

## Research on Dynamic Vehicle Multi-Object Tracking Method Based on DeepSort

He Weikun, Peng Yuhui, Huang Wei, Yao Yujie, Chen Zehui

(Fuzhou University, Fuzhou 350116)

**【Abstract】**In order to improve the environmental perceive ability and better predict the behaviors of dynamic vehicle targets, a YOLOX model, as the front-end detector, combined with the optimized DeepSort method was proposed to develop the research of dynamic vehicle Multi-Object Tracking. In the process of features matching for dynamic vehicles, the change information of vehicles in the degree of light and shade was extracted to improve the object matching accuracy by adding Haarlike features. Furthermore, based on the application of DeepSort RE-identification network, the improved ResNet13 model was adopted as the backbone network for feature extraction, and SENet was introduced to adjust the feature weight of different channel dimensions. The results show that, compared to the conventional DeepSort method, by using the real on-road video data, the Multi-Object Tracking Accuracy (MOTA) and IDF1 indicator of the proposed algorithm increased by 1.4 percentage points and 7.7 percentage points respectively.

**Key words:** Multi-object tracking, DeepSort, Autonomous vehicle, Haar-like, SENet

**【引用格式】**何维堃,彭育辉,黄炜,等.基于 DeepSort 的动态车辆多目标跟踪方法研究[J].汽车技术,2023(11):27-33.

HE W K, PENG Y H, HUANG W, et al. Research on Dynamic Vehicle Multi-Object Tracking Method Based on DeepSort [J]. Automobile Technology, 2023(11): 27-33.

## 1 前言

面对复杂多变的交通环境,自动驾驶汽车需不断提高对外部环境的感知能力。其中,检测与跟踪道路行驶车辆并预测其运动轨迹是环境感知研究领域的重要课题。

近年来,多目标跟踪识别网络在智能驾驶方面的应用引起了国内外学者的普遍关注<sup>[1-2]</sup>,解决行驶过程中车辆的身份切换问题是本文研究的重点。在基于 Deep-

Sort 的研究方面,Perera 等将 DeepSort 中的卡尔曼滤波替换为无迹卡尔曼滤波<sup>[3]</sup>;Mu 等将卡尔曼滤波替换为拓展卡尔曼滤波<sup>[4]</sup>,使网络能更好地处理应用场景的非线性问题,对目标未来帧进行有效预测,但依然无法保证预测本身的准确性;裴云成将交并比(Intersection Over Union, IOU)匹配更换成广义交并比(Generalized Intersection Over Union, GIOU)<sup>[5]</sup>,使网络能更有效地判断检测框和预测框之间的权重分配,但无法很好地解决特征匹配时出现失配的问题。在 DeepSort 网络中,重识

\*基金项目:福建省引导性科技计划项目(2022H0007);福建省自然科学基金项目(2021J01559)。

通讯作者:彭育辉(1975—),博士,教授,研究方向为汽车智能驾驶、汽车替代燃料等, pengyuhui@fzu.edu.cn。

别网络提取的深度特征用于前、后帧物体的特征匹配,在防止身份切换时能起到很好的作用。金立生等将检测器更换为 Gaussian YOLO V3,并用重识别网络的中心损失函数代替交叉熵函数,提高对车辆特征的提取性能与跟踪网络的判断能力<sup>[6]</sup>,但当光照发生变化时,网络无法适应环境的改变,从而导致身份切换;Wu等将重识别网络替换为 ShuffleNet V2以适用于计算能力较低的设备<sup>[7]</sup>,但在复杂场景下,无法很好地提取物体的特征;Liu等将原有的重识别网络替换为 ResNet50<sup>[8]</sup>,加深了网络深度,并结合特征金字塔网络 (Feature Pyramid Networks, FPN) 提取物体的多层特征,但复杂的网络结构增加了计算负担。

综上,复杂场景下跟踪检测目标时,跟踪器的准确性及稳定性较差,现有的重识别网络无法有效地对不同场景下的车辆进行特征提取,前、后帧中目标的匹配效果较差。为此,本文采用 YOLOX 作为检测器,对基于 DeepSort 研究中重识别网络的特征匹配方法进行改进,添加 Haar-like 特征反映车辆的光暗变化,对原有的特征匹配机制进行补充,在协调匹配速度的同时加深重识别网络层数并添加 SENet 注意力机制强化提取车辆外观特征的能力,最后通过实际道路采集的视频数据集对算法的有效性进行验证。

## 2 多特征融合匹配

### 2.1 距离度量

DeepSort 是基于匈牙利算法和卡尔曼滤波算法的单假设跟踪方法,引入卡尔曼滤波对轨迹运动状态进行预测,将检测结果经重识别网络提取的特征与轨迹储存的特征共同传入匈牙利算法,进行关联匹配。在匹配过程中,对检测结果框、轨迹预测框的马氏距离及深度特征的余弦距离信息进行整合判断。

余弦距离主要用于衡量检测结果与轨迹的特征向量间的差距。对于物体被遮挡后再出现的情况,网络将轨迹中储存的特征集与物体在重新出现时的特征向量进行计算,得出其最小余弦距离:

$$d^0(i,j) = \min\{1 - \mathbf{r}_j^T \mathbf{r}_k^i | \mathbf{r}_k^i \in \mathbf{R}_i\} \quad (1)$$

式中,  $d^0(i,j)$  为第  $j$  个检测框内的特征向量与第  $i$  个轨迹的特征向量之间的最小余弦距离;  $\mathbf{r}_j$  为第  $j$  个检测框内的特征向量;  $\mathbf{r}_k^i$  为跟踪器存储的第  $i$  个轨迹的特征向量集,其包含第  $i$  个轨迹过去成功匹配的  $k$  个特征向量;  $\mathbf{R}_i$  为第  $i$  个轨迹的特征向量库。

马氏距离用于衡量上一帧中目标经过卡尔曼滤波

预测的位置信息与检测结果的位置信息的差距:

$$d^2(i,j) = \min\{(\mathbf{d}_j - \mathbf{y}_i)^T \mathbf{S}_i^{-1} (\mathbf{d}_j - \mathbf{y}_i)\} \quad (2)$$

式中,  $d^2(i,j)$  为第  $j$  个检测框和第  $i$  个由跟踪器预测的物体框之间的马氏距离;  $\mathbf{d}_j$  为第  $j$  个检测框的位置信息;  $\mathbf{y}_i$  为第  $i$  个由跟踪器预测的物体框位置信息;  $\mathbf{S}_i$  为第  $i$  条轨迹在检测空间的协方差矩阵。

将基于外观特征的余弦距离和基于运动信息的马氏距离进行线性加权获得最终的度量:

$$c_{ij} = (1 - \lambda)d^0(i,j) + \lambda d^2(i,j) \quad (3)$$

式中,  $c_{ij}$  为最终度量标准;  $\lambda$  为权重系数。

交通场景中的车辆运动轨迹是非线性的,而卡尔曼滤波采用线性状态方程,因此预测能力较弱<sup>[9]</sup>。通过物体的运动状态预测出的运动信息估计量会存在一定偏差,导致马氏距离关联方法易失效,从而造成身份跳转的现象。在最终度量中设置  $\lambda=0$ ,使马氏距离在级联匹配过程中仅为判定值。当 2 个物体的对象特征一致且两者的马氏距离小于一定阈值,则判定匹配成功。

### 2.2 Haar-like 特征匹配

自动驾驶所涉及的场景多变且复杂,在车辆的多目标跟踪中,车辆之间的相互遮蔽、场景变换、光下的阴影变换都将导致追踪目标失败。本文融合传统特征与深度特征信息,结合各自的优点解决跟踪过程中身份跳转的问题。

传统特征相较于深度特征,在精度方面的表现较差,且泛化能力不足。深度特征的提取受训练集影响较大,训练过程需要大量的场景图像数据,而传统特征的提取作为人工认知驱动的方法,具有解释性,不需要训练,且对算力的要求较低。因此,本文算法以原有的深度特征作为主要匹配依据,融合传统特征 Haar-like 作为补充匹配。

Haar-like<sup>[10]</sup>是计算机视觉和图像处理中用于目标检测的特征描述算子。特征模板内有白色和黑色 2 种矩形,矩形的位置关系不同,影响着所提取的特征信息,如边缘特征、线条特征、中心特征、对角线特征。通过特征模块大小的变换和在图片子窗口的移动得到所需的特征,如图 1 所示。其中,图片特征提取框为 YOLOX 输出图片,而非整张图,可明显缩短特征提取的时间。

传统特征中,Haar-like 能够表示图像像素值明暗变化信息,能更好地体现车辆行驶状态的特征<sup>[11]</sup>。在原有的 DeepSort 中加入 Haar-like,将深度神经网络提取的车辆抽象特征与 Haar-like 信息相融合<sup>[12]</sup>,可有效增强网络对不同场景的适应性。图 2 所示为融入 Haar-like 特征的 DeepSort 网络结构。首先,对检测器输入结果进行

深度特征提取和 Haar-like 特征提取。在匹配过程中,将确认态轨迹中所保留的深度特征与重识别网络提取的深度特征进行匹配,再将未匹配的轨迹和检测物体进行 Haar-like 特征匹配,最后进行 IOU 匹配。

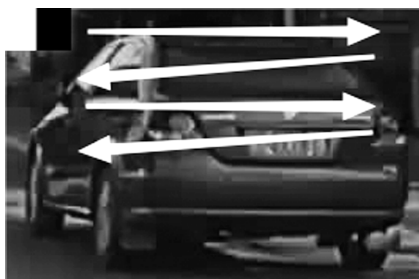


图1 特征模块移动方式

### 3 重识别网络的优化

车辆目标由于遮挡或未被检测,会出现与之前的目标断开联系且在短时间内重新出现在检测器中的现象。DeepSort 利用重识别网络提取车辆外观信息,再将外观信息放入级联匹配。通过计算不同帧物体的外观

信息的余弦距离,进行物体的相似度配对,将同一辆车重新建立联系。

车辆作为刚体,在不同视角下所呈现的姿态、大小多变<sup>[13]</sup>。在自动驾驶过程中,当检测车辆与摄像头的距离发生改变,或检测车辆处于半遮挡及转弯的状态,检测车辆的外观都会发生变化。在原网络中,重识别网络结构简单,无法准确提取行驶过程中实时变化的车辆外观信息,在重识别车辆时具有明显的局限性。

#### 3.1 物体框尺寸的计算

无人驾驶车辆行驶过程中,对象车辆由于距离变化及遮挡的原因,其宽和高在视频中的形变量较大。经过 YOLOX<sup>[14]</sup>检测器识别后,检测器会将车辆的物体框传入追踪网络。重识别网络对车辆特征进行提取的过程中,需要重新调整图像大小。原网络中定义了图片的高为 128,宽为 64,适用于对行人的特征提取,但无法满足对车辆的特征提取<sup>[15]</sup>。为避免过大幅度地调整图片的高和宽,导致识别误差,应寻找合适的车辆物体框的高和宽。

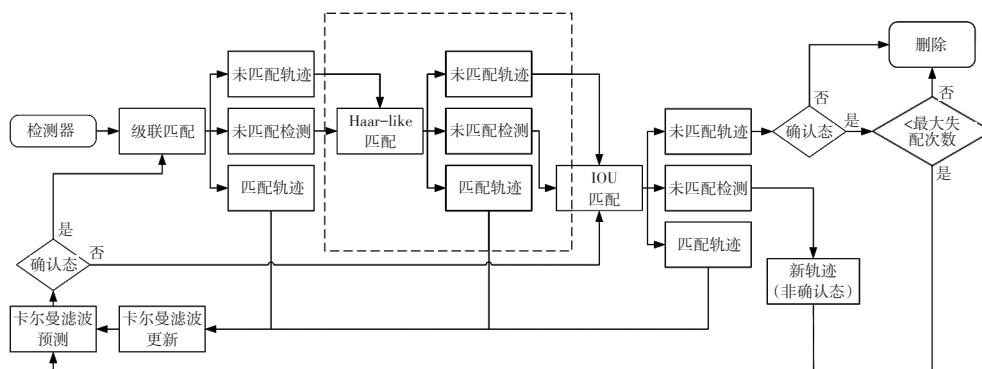


图2 融合 Haar-like 的 DeepSort 目标跟踪算法流程

本文在自制的道路数据集上进行标注,截取视频中不同场景下(无遮挡、被其他车辆、树木和路障等遮挡)不同车型(挂车、皮卡和轿车等)大小及形态不同的物体框,对识别的物体框的宽和高进行均值处理,求得最优的物体框尺寸为高 96、宽 128。

#### 3.2 重识别网络加深

在重识别网络中,提取的语义信息随着卷积层数的加深而越来越明显。加深网络层数,会带来更加强大的非线性表达能力和复杂的变换,使提取的信息愈加复杂、抽象<sup>[16]</sup>。在特征提取网络中,底层网络会提取物体的边缘信息,较深层的网络会提取物体的纹理特征,更深层的网络则会提取目标的形状,能更好地表达车辆的整体信息。原始的重识别网络由 6 个残差网络组成,最后的特征维度为 128。广泛应用于行人的重识别网络由 8 个残差网络组成,特征维度为 512,增加了特征维度

使得分类更加细致,但其仍无法进行复杂情况下的车辆特征提取。为了提取车辆更深层次的信息,以满足实际行车中的追踪要求,本文在保证其时效性的同时加深了网络的深度<sup>[17]</sup>,称为 ResNet13,网络结构如表 1 所示。

表1 ResNet13 网络结构

网络层	输出	网络层	输出
卷积层 1	32×96×128	残差层 6×2	128×12×16
卷积层 2	32×96×128	残差层 7×2	256×6×8
最大池化层 3	32×48×64	残差层 8×2	512×6×8
残差层 4×2	32×48×64	平均池化层 9	512×1×1
残差层 5×2	64×24×32	批处理和 L2 正则化	512

### 4 重识别网络融合 SENet

SENet<sup>[18]</sup>在通道维度上增加 1 个注意力机制,并利用损失函数的自我学习,调整不同通道维度的特征权重,

增大有效特征通道的权重。根据SENet所提供的各特征通道的重要程度,网络能自主地关注重要的特征通道。其包括挤压、激励和特征重标定3个部分,如图3所示。

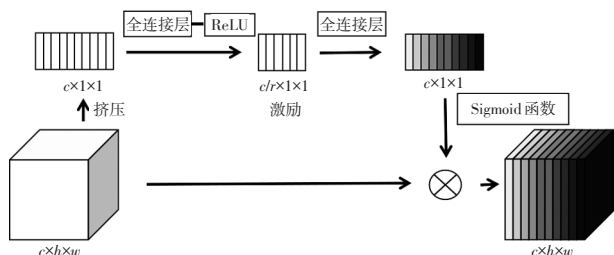


图3 SENet结构

挤压是将不同特征通道数( $c$ )的特征图集的全局信息嵌入,聚合成 $c \times 1 \times 1$ 的特征描述符。在不同特征通道数的特征图集中加入SENet,将网络中的特征图集聚合成 $32 \times 1 \times 1$ 、 $64 \times 1 \times 1$ 、 $128 \times 1 \times 1$ 、 $256 \times 1 \times 1$ 、 $512 \times 1 \times 1$ 的特征描述符。在不同层次上获得每个特征图集的全局描述符:

$$z_c = \frac{1}{w \times h} \sum_{i=1}^w \sum_{j=1}^h u_c(i, j) \quad (4)$$

式中, $z_c$ 为特征图挤压聚合得到的特征描述; $u_c$ 为特征图; $w$ 为特征图的宽; $h$ 为特征图的高。

获取不同深度的特征描述符后,激励通过2个全连接层获取通道间的关系,再经由Sigmoid函数转化为不同网络深度(32、64、128、256、512)下每个通道的权重值:

$$s = \sigma(W_2 \delta(W_1 z)) \quad (5)$$

式中, $s$ 为所有特征图的权重集; $\sigma$ 为Sigmoid函数; $\delta$ 为线性整流(Rectified Linear Unit, ReLU)函数; $W_1$ 为第1个全连接层操作; $W_2$ 为第2个全连接层操作; $z$ 为所有特征图挤压聚合所得的特征描述符集。

最后,特征重标定将激励获得的通道权重逐通道加权到不同深度的原始特征图通道上,减小不重要的特征图的权重,提高包含重要分类信息的特征图权重:

$$\tilde{X}_c = s_c u_c \quad (6)$$

式中, $\tilde{X}_c$ 为特征图与对应权重相乘得到的特征图; $s_c$ 为特征图的权重。

ResNet13采用残差块结构<sup>[19]</sup>,将输入的 $X$ 与由多个卷积级联输出的 $F(X)$ 相加。残差结构作为小型网络, $X$ 为输入的观测值, $H(X)$ 为输出的残差网络的预测值, $F(X)$ 为观测值与预测值之间的差,即残差。该网络结构有利于保护信息的完整性,且在加深网络层数的同时,能有效解决梯度消失和梯度爆炸的问题。在ResNet13中,随着网络层数的增加,特征图的深度不断扩张。将SENet加入残差网络中,可以对网络中不同深度的特征图进行重构,如图4所示。

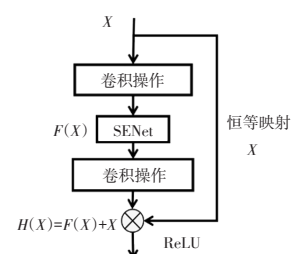


图4 残差网络结合SENet结构

## 5 试验

试验采用开源的PyTorch 1.6.0深度学习框架,CPU配置为Intel® Aeon® Silver 4108,主频1.80 GHz,显卡为GeForce GTX 1080Ti,内存容量为64 GB,系统为Ubuntu16.04 LTS,包含Python 3.7、CUDA 10.1、CuDNN 7.6.5、OpenCV 4.4.0环境。数据集的采集设备为长虹CXC-X6行车记录仪。

### 5.1 重识别网络训练

将ResNet13和SENet应用于复杂场景下对车辆特征的提取,以VeRi-776<sup>[20-21]</sup>数据集训练所得的准确率(Accuracy)和损失(Loss)为基准判断算法的质量。对重识别网络进行数据集的训练时,为保证数据的准确性,3组网络训练采用相同的参数:学习率为0.001,经过20代(epoch),学习率乘以0.1。

准确率是用于对重识别网络训练效果进行度量的主要指标,计算重识别模型正确分类的样本数与总样本数之比,本文利用误差(Error)对准确率进行表达。图5展示了原网络6层残差网络、8层残差网络,以及ResNet13并加入SENet网络预测期间的误差曲线。

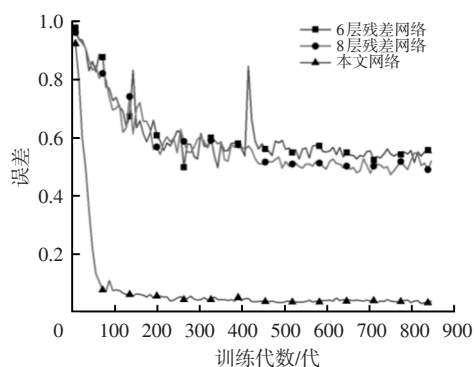


图5 误差对比

误差公式为:

$$\alpha = 1 - \frac{T_p + T_N}{T_p + T_N + F_p + F_N} \quad (7)$$

式中, $\alpha$ 为误差率; $T_p$ 为真正例数量; $T_N$ 为真负例数量; $F_p$ 为假正例数量; $F_N$ 为假负例数量。

损失函数用于表达实际输出和期望输出之间的差

值,使网络更新模型参数,减小误差。图6展示了原网络6层残差网络、8层残差网络,以及ResNet13并加入SENet网络预测期间的损失。

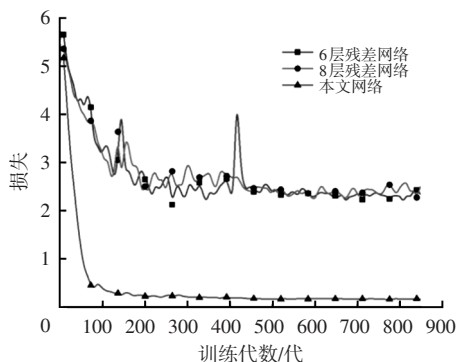


图6 损失对比

损失公式为:

$$H(p, q) = -\sum_{i=1}^n (p(x_i)) \log(q(x_i)) \quad (8)$$

式中,  $q(x_i)$  为实际输出的各类别的概率分布;  $p(x_i)$  为期望输出各类别的分布;  $n$  为类别总数;  $H(p, q)$  为交叉熵损失。

图5和图6中,原始网络的6层残差网络与本文的重识别网络都进行850代的训练对比。在2个网络训练到第80代时,误差和损失都迅速减少,随后下降速度逐渐减缓,直至第200代时趋于平稳。但原始网络在第200代后,训练的模型精度较差,波动较为明显。在经历850代训练后,误差和损失依然不理想,误差为0.55,损失为2.35。本文网络在训练过程中比原始网络更加平稳,误差和损失也更加精确,误差为0.03,损失为0.16。因此,本文网络在车辆外观提取方面,比原始网络具有更好的性能。

## 5.2 跟踪算法试验

分别采用DeepSort的原始版本和本文改进的版本对车辆进行追踪,利用车辆在不同场景(城市、乡村)下的视频集进行试验,验证算法的可行性。本文的试验场景为自制的双车道的来往车辆视频,包括城市和乡镇,具有一定的普遍性。

图7所示为原始网络在数据集中前、后帧的变化情况。由图7可知,身份标签为10的车辆在视频中发生了身份跳转。

图8所示为添加 Haar-like 特征匹配网络后,试验的前、后帧变化情况。当车辆处于视频的远端角落时,会发生形变及特征不清晰的情况。原始重识别网络提取的深度特征不足以应对复杂的车辆行驶情况,从而发生身份跳转。添加 Haar-like 特征匹配网络以提取目标车辆的光暗变化,从而实现原始网络在深度特征匹配之

后的补齐。将深度特征无法匹配的车辆进行 Haar-like 特征再匹配,抑制了身份跳转的发生。

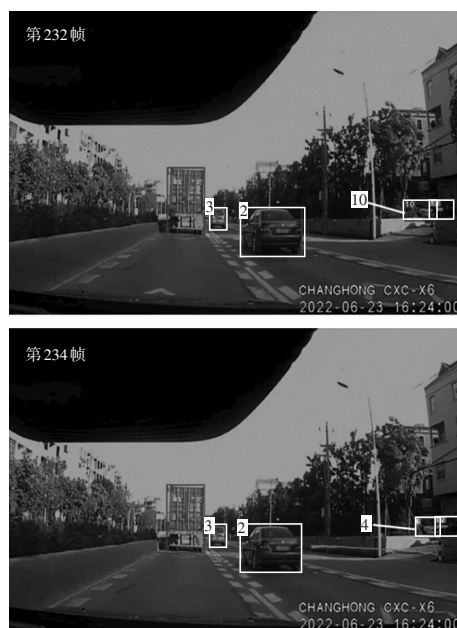


图7 原始网络的前、后帧变化



图8 添加 Haar-like 特征匹配网络的前、后帧变化

图9所示为原始网络在数据集中前、后帧的变化情况。车辆失去识别后再次出现时,身份标签为3的车辆发生了身份跳转。

图10所示为改进后的 ResNet13 以及添加了注意力机制网络后,视频前、后帧的变化情况。在相同视频帧下,车辆的身份并没有发生切换;在车辆处于半遮挡状态时,ResNet13 可提高网络对车辆的特征提取效果。结果表明:SENet 可自动学习到不同通道特征的重要程度,注意力机制提高了目标特征的权重,使网络专注于

信息量大的通道特征,减弱不重要信息的通道特征,将 ResNet13 与 SENet 相结合可达到更好的匹配结果。

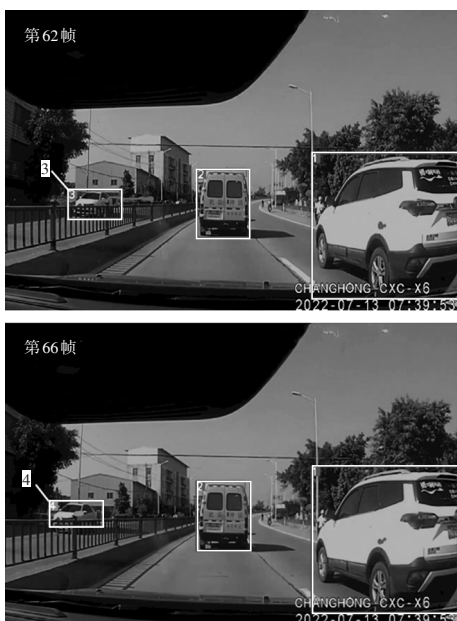


图9 原始网络的前、后帧变化

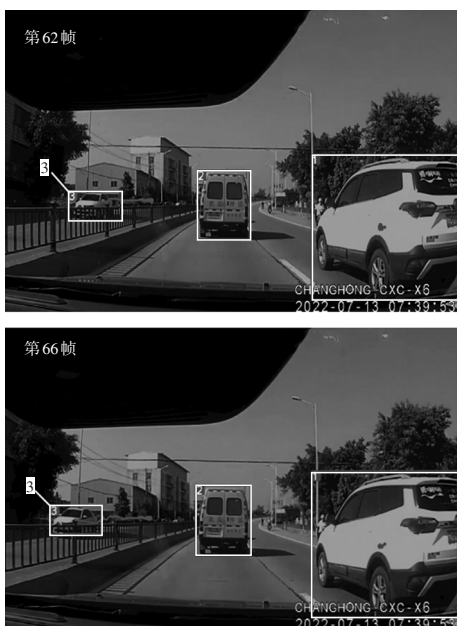


图10 加深网络和加入SENet的前、后帧变化

### 5.3 追踪结果测试与分析

采用多目标跟踪精度 (Multiple Object Tracking Accuracy, MOTA)、平均数比率 IDF1 (Identification F1) 和帧速率 3 个评价指标对跟踪算法进行评估。MOTA 衡量了跟踪器在检测物体和保持轨迹时的性能,相比 IDF1 更专注于检测器的性能。IDF1 用于判断跟踪器在视频中长时间地对同一事物进行准确跟踪的能力,着重于计算物体的初始身份号码占其在视频中出现的身份号码的比例。帧速率反映了处理速度。对比算法

的评价结果,如表 2 所示。

表 2 评价结果

算法	MOTA/%	IDF1/%	帧速率/帧·s <sup>-1</sup>
DeepSort	73.5	72.9	42.3
DeepSort+Haar-like	73.6	78.3	38.0
DeepSort+(均值处理框+更新 REID)	74.2	75.6	36.2
改进 DeepSort	74.9	80.6	34.4

由表 2 可知,在跟踪网络中加入 Haar-like 特征且对重识别网络进行改进,可有效提升跟踪网络的效果。

## 6 结束语

本文针对自动驾驶汽车,面向外界动态车辆多目标跟踪的场景,提出以 YOLOX 为前端检测器的 DeepSort 改进算法,利用 Haar-like 算法提取车辆的光暗变化特征,实现对网络原有匹配机制的补充。此外,通过 ResNet13 网络对重识别网络进行更换、添加注意力机制,从而提高重识别网络对车辆深度特征的提取能力。采用实际道路驾驶采集的视频数据,对本文算法和传统算法进行对比验证,结果表明:相较于传统 DeepSort 算法,改进后算法的 MOTA 提高了 1.4 个百分点, IDF1 提升了 7.7 百分点。

### 参考文献

- [1] 陈龙, 朱程铮, 蔡英凤, 等. 基于自查询的车载多目标跟踪算法研究[J]. 汽车工程, 2021, 43(11): 1587-1593+1610.  
CHEN L, ZHU C Z, CAI Y F, et al. Research on Vehicle Multi-Target Environment Aware Tracking Algorithm Based on Self-Query[J]. Automotive Engineering, 2021, 43(11): 1587-1593+1610.
- [2] CHEN X, JIA Y, TONG X, et al. Research on Pedestrian Detection and DeepSort Tracking in front of Intelligent Vehicle Based on Deep Learning[J]. Sustainability, 2022, 14 (15).
- [3] PERERA I, SENAVIRATHNA S, JAYARATHNE A, et al. Vehicle Tracking Based on An Improved DeepSORT Algorithm and the YOLOv4 Framework[C]// 2021 10th International Conference on Information and Automation for Sustainability (ICIAIS). Quebec City, Canada: IEEE, 2021: 305-309.
- [4] MU L, ZHAO H, LI Y, et al. Traffic Flow Statistics Method Based on Deep Learning and Multi-Feature Fusion[J]. Computer Modeling in Engineering & Sciences, 2021, 129 (2): 465-483.
- [5] 裴云成. 基于DeepSORT的行人多目标跟踪算法研究与应用[D]. 济南: 齐鲁工业大学, 2022.

PEI Y C. Research and Application of Pedestrian Multi-汽车技术

- Target Tracking Algorithm Based on DeepSORT[D]. Jinan: Qilu University of Technology, 2022.
- [6] 金立生, 华强, 郭柏苍, 等. 基于优化DeepSort的前方车辆多目标跟踪[J]. 浙江大学学报(工学版), 2021, 55(6): 1056–1064.  
JIN L S, HUA Q, GUO B C, et al. Multi-Target Tracking of Vehicles Based on Optimized DeepSort[J]. Journal of Zhejiang University (Engineering Science), 2021, 55(6): 1056–1064.
- [7] WU Y, WANG Y, ZHANG D, et al. Research on Vehicle Tracking Method Based on UAV Video[C]// 2nd International Conference on Internet of Things and Smart City (IoTSC 2022). Xiamen, China: SPIE, 2022, 12249: 801–806.
- [8] LIU H Y, PEI Y C, BEI Q C, et al. Improved DeepSORT Algorithm Based on Multi Feature Fusion[J]. Applied System Innovation, 2022, 5(3).
- [9] 弋英民, 王柯颖, 苑易伟, 等. 基于扩展卡尔曼滤波的固定翼无人机姿态解算方法[J/OL]. 小型微型计算机系统: 1–9 (2022–08–12)[2022–11–12]. <http://kns.cnki.net/kcms/detail/21.1106.TP.20220729.1118.014.html>.  
GE Y M, WANG K Y, WAN Y W, et al. Attitude Algorithm Method of Fixed-Wing UAV Based on Extended Kalman Filter[J/OL]. Journal of Chinese Computer Systems: 1–9 (2022–08–12)[2022–11–12]. <http://kns.cnki.net/kcms/detail/21.1106.TP.20220729.1118.014.html>.
- [10] LIENHART R, MAYDT J. An Extended Set of Haar-Like Features for Rapid Object Detection[C]// International Conference on Image Processing. Rochester, NY, USA: IEEE, 2002.
- [11] 黎向锋, 赵伟康, 豆心愿, 等. 基于改进AdaBoost+Haar的车辆检测算法[J]. 测控技术, 2019, 38(2): 42–45.  
LI X F, ZHAO K W, DOU X Y, et al. Vehicle Detection Algorithm Based on Improved AdaBoost and Haar[J]. Measurement & Control Technology, 2019, 38(2): 42–45.
- [12] 储慧芳. 基于深度特征与传统特征融合的行人再识别研究[D]. 合肥: 合肥工业大学, 2018.  
CHU H F. Research on Person Re-identification Based on Fusion of Deep Features and Traditional Features[D]. Hefei: Hefei University of Technology, 2018.
- [13] 敖为能. 交通场景下基于深度学习的车辆检测方法研究[D]. 长沙: 湘潭大学, 2021.  
AO W N. Research on Vehicle Detection Method Based on Deep Learning in Traffic Scenarios[D]. Changsha: Xiangtan University, 2021.
- [14] GE Z, LIU S T, WANG F, et al. YOLOX: Exceeding YOLO Series in 2021[EB/OL]. (2021–08–06)[2022–11–12]. <https://arxiv.org/abs/2107.08430>.
- [15] 李永上, 马荣贵, 张美月. 改进YOLOv5s+DeepSORT的监控视频车流量统计[J]. 计算机工程与应用, 2022, 58(5): 271–279.  
LI Y S, MA R G, ZHANG M Y. Traffic Monitoring Video Vehicle Volume Statistics Method Based on Improved YOLOv5s + Deep SORT[J]. Computer Engineering and Applications, 2022, 58(5): 271–279.
- [16] MONTUFAR G F, PASCANU R, CHO K, et al. On the Number of Linear Regions of Deep Neural Networks[J]. Advances in Neural Information Processing Systems, 2014, 27.
- [17] BENGIO Y, LECUN Y. Scaling Learning Algorithms Towards AI[J]. Large-Scale Kernel Machines, 2007, 34(5): 1–41.
- [18] 赵璐璐, 王学营, 张翼, 等. 基于YOLOv5s融合SENet的车辆目标检测技术研究[J]. 图学学报, 2022, 43(5): 776–782.  
ZHAO L L, WANG X Y, ZHANG Y, et al. Vehicle Target Detection Based on YOLOv5s Fusion SENet[J]. Journal of Graphics, 2022, 43(5): 776–782.
- [19] 李守彪, 武志斐. 基于多尺寸分解卷积的车道线检测[J]. 汽车技术, 2022(8): 32–37.  
LI S B, WU Z F. Lane Detection Based on Multi-Size Factorized Convolution[J]. Automobile Technology, 2022 (8): 32–37.
- [20] LIU X, LIU W, MEI T, et al. A Deep Learning-Based Approach to Progressive Vehicle Re-Identification for Urban Surveillance[C]// European Conference on Computer Vision. Amsterdam, Netherlands: Springer, Cham, 2016: 869–884.
- [21] LIU X C, LIU W, MEI T, et al. PROVID: Progressive and Multimodal Vehicle Reidentification for Large-Scale Urban Surveillance[J]. IEEE Transactions on Multimedia, 2017, 20(3): 645–658.

(责任编辑 斛 畔)

修改稿收到日期为2022年9月11日。