

# 基于DeepSORT算法的前方道路车辆跟踪研究

郭薇<sup>1</sup>, 朱泽德<sup>2</sup>, 王道斌<sup>1</sup>, 李宸翔<sup>1</sup>

<sup>1</sup>武汉科技大学汽车与交通工程学院, 湖北 武汉

<sup>2</sup>安徽工业技术创新研究院六安院, 安徽 六安

收稿日期: 2023年10月30日; 录用日期: 2023年12月22日; 发布日期: 2023年12月29日

## 摘要

针对部分遮挡、目标漏检等导致的车辆ID频繁切换以及跟踪精度低的问题, 本文提出了一种优化DeepSORT跟踪器的方法。一方面设计了S2Net36重识别网络: 首先加深重识别网络构建ResNet36网络, 提取更深层次的车辆外观特征; 其次构建SER模块提取目标关键特征以及构建SE-Res2Net模块提取目标区域特征; 最后基于ResNet36网络分别融合SER模块与SE-Res2Net模块得到S2Net36重识别网络。另一方面, 引入三元组损失函数拉近相同目标不同样本之间的特征距离, 通过提取更具有辨别力的车辆外观特征用于数据关联, 进而提升对前方道路车辆的跟踪能力。实验结果表明, 相比于DeepSORT原始算法, 改进的算法的MOTA提高了1.18%, IDF1提升了0.80%, 提高了对前方道路车辆的跟踪精度与稳定性, 有望为自动驾驶车辆提供技术支持。

## 关键词

道路车辆跟踪, DeepSORT, S2Net36, 三元组损失

# Research on Vehicle Tracking on the Road Ahead Based on DeepSORT Algorithm

Wei Guo<sup>1</sup>, Zede Zhu<sup>2</sup>, Daobin Wang<sup>1</sup>, Chenxiang Li<sup>1</sup>

<sup>1</sup>School of Automobile and Traffic Engineering, Wuhan University of Science and Technology, Wuhan Hubei

<sup>2</sup>Lu'an Branch, Anhui Institute of Innovation for Industrial Technology, Lu'an Anhui

Received: Oct. 30<sup>th</sup>, 2023; accepted: Dec. 22<sup>nd</sup>, 2023; published: Dec. 29<sup>th</sup>, 2023

## Abstract

This paper proposed a method to optimize the DeepSORT tracker aiming to solve the problems of

文章引用: 郭薇, 朱泽德, 王道斌, 李宸翔. 基于 DeepSORT 算法的前方道路车辆跟踪研究[J]. 运筹与模糊学, 2023, 13(6): 7806-7816. DOI: 10.12677/orf.2023.136763

frequent target ID switch and low tracking accuracy caused by target missed detection and occlusion. Firstly, a re-identification network called S2Net36 was designed to extract deeper vehicle appearance features. To achieve this, the re-identification network is deepened to build the ResNet36 network. Then, a SER module is constructed to extract the key features of targets and a SE-Res2Net module is constructed to extract the regional features of targets. Finally, the SER module and the SE-Res2Net module were embedded in the ResNet36 network to obtain the S2Net36 re-identification network. Secondly, a triplet loss function was introduced to shorten the feature distance of different samples of the same target. Extracting more discriminating vehicle appearance features for data correlation to improve the tracking ability for the vehicles ahead. The experimental results show that the proposed algorithm could improve the MOTA by 1.18% and IDF1 by 0.80% compared with the original DeepSORT algorithm. This improvement in tracking accuracy and stability of road vehicles ahead is expected to provide technical support for autonomous vehicles.

## Keywords

Road Vehicle Tracking, DeepSORT, S2Net36, Triplet Loss

Copyright © 2023 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

## 1. 引言

多目标跟踪任务旨在输出给定视频中的目标运动轨迹, 并维持各个目标的身份信息(ID) [1], 目标的遮挡、漏检、误检等现象造成的多目标之间的 ID 切换问题是多目标跟踪的研究难点。随着目标检测算法的快速发展, 基于目标检测的跟踪范式(TBD, Tracking-By-Detection)成为多目标跟踪领域的主流框架[2]。该框架下的多目标跟踪方法将视频多目标跟踪任务分解为两个独立的子任务: 目标检测和数据关联, 首先通过目标检测算法对视频中的每一帧图像进行检测, 随后将属于同一目标且来自不同帧中的检测结果关联起来, 进而得到目标的轨迹[3]。

基于 TBD 的跟踪范式, Bewley [4]等提出 SORT 算法, SORT 融合目标检测输出的位置信息与卡尔曼滤波(Kalman Filter)预测的运动信息作为相似度矩阵, 结合匈牙利算法(Hungarian Algorithm)实现目标跨帧的关联匹配, 但该算法未考虑跟踪过程中由于长时间遮挡导致的目标 ID 频繁切换问题。Wojke [5]等在 SORT 基础上提出 DeepSORT 算法, 结合行人重识别技术引入行人重识别(Re-identification)外观模型, 融合外观信息和运动信息作为关联度量, 采用级联匹配策略进一步提升了目标匹配的精度, 在长距离匹配中取得良好的效果。

现阶段, 目标跟踪的研究可分为前端的目标检测器与后端的 DeepSORT 跟踪器两部分。对于目标检测器的研究, 殷远齐[6]等在 YOLOv4 [7]检测器中嵌入 CBAM 注意力机制[8], 融合特征空间信息与通道信息, 进而提取目标的关键特征并弱化非重要特征信息; Gao [9]等提出了 Res2Net 多尺度模块, 该模块可以便捷地嵌入到目标检测网络中且不增加网络计算开销, 通过提取多种尺度特征, 进而有效提取目标全局以及局部特征; 顾立鹏[10]等在 Res2Net 模块中嵌入  $1 \times 1$  卷积和 SE 注意力机制[11]构建新模块, 并用构建的新模块替换目标检测器 CenterNet [12]的原有残差单元, 以提升网络对目标区域特征的提取能力。

对于后端 DeepSORT 跟踪器的研究, 一方面, DeepSORT 现有的重识别网络较为简单, 难以获取丰富的车辆外观信息; 另一方面, 由于道路车辆外观差异较小且车辆在行驶过程中, 角度不断变化会造成

车辆姿态多样,从而导致不同 ID 车辆前后帧的外观特征差异小于同一 ID 车辆前后帧的外观差异,导致提取的车辆外观信息没有辨别力。针对现有的重识别网络获取外观信息不足的问题,何维堃[13]等进一步加深重识别网络,进而更好地提取目标的形状等信息;尤晓雨[14]等将重识别网络替换为 Darknet53 网络[15],通过 Darknet53 网络结构的五次下采样操作来扩大感受野,进而获取目标深层特征。但提取的车辆外观信息缺乏辨别力主要是因为 DeepSORT 的交叉熵损失函数更适用于学习类间信息,而对特征相近的目标进行分类较为困难[6],金立生[16]等基于此引入中心损失函数,其可以拉近同一目标的不同样本的特征距离,通过结合中心损失函数与交叉熵损失函数进行重识别训练,进而能够提取更具辨别力的车辆外观特征用于数据关联。

现有研究表明加深重识别网络能够丰富车辆外观信息,优化损失函数能够增强特征的辨别能力,但对于后端 DeepSORT 跟踪器的研究仍存在一定不足,主要表现在两个方面:1) 针对现有的重识别网络较为简单难以提取丰富的外观信息,大多数方法主要通过加深网络层数提取更深层特征,方法较为单一;2) 现有针对跟踪的研究通常是从加深重识别网络或优化损失函数单一方向进行研究,从改进重识别网络和优化损失函数两个方向同时对跟踪模型进行改进的研究仍比较缺乏。因此,本文针对上述问题对 DeepSORT 跟踪器进行了优化,主要贡献如下:

1) 通过加深重识别网络层数构建 ResNet36 网络,用于提取更深层的车辆外观特征。

2) 基于 ResNet36 网络进一步构建 S2Net36 重识别网络,在 ResNet36 中分别嵌入构建的 SER 模块与 SE-Res2Net 模块,用于提取目标关键特征及区域特征。

3) 结合三元组损失函数与交叉熵损失函数对重识别网络进行训练,提升网络对特征相近目标的分类预测性能,减少车辆 ID 切换和提高跟踪精度。

## 2. 跟踪模型框架

本文改进的跟踪模型框架图如图 1 所示,通过对改进的 S2Net36 重识别网络进行训练得到车辆外观特征提取模型,YOLOv5 检测到的车辆目标信息会通过车辆外观特征提取模型提取出车辆外观信息,继而结合卡尔曼滤波预测的车辆轨迹的运动信息进行级联匹配,最后通过匈牙利算法为每个目标分配不同的跟踪 ID。级联匹配未成功的目标会继续进行 IOU 匹配,匹配成功后也将由匈牙利算法完成 ID 分配,对于 IOU 匹配仍未成功且连续 100 帧未匹配成功的车辆轨迹将从库中删除。

### 级联匹配

级联匹配是根据目标检测框、轨迹预测框的余弦距离及马氏距离信息进行对目标进行匹配。余弦距离衡量的是目标检测框与轨迹预测框在外观特征上的差距,外观相似度越高,其距离越近。对于遮挡后重新出现的目标,网络会将该目标的特征分别与存储的特征集进行计算,获得最小的余弦距离:

$$d^{(1)}(i, j) = \min \left\{ 1 - r_j^T r_k^{(i)} \mid r_k^{(i)} \in R_i \right\} \quad (1)$$

式中,  $d^{(1)}(i, j)$  为检测框的外观特征向量与轨迹预测框的外观特征向量之间的最小余弦距离;  $r_j$  为第  $j$  个检测框的外观特征向量;  $R^i$  为轨迹特征向量库,其存储了不超过 100 帧轨迹的外观特征向量。马氏距离用于度量检测目标框与轨迹预测框之间的运动特征信息,实现运动信息的匹配:

$$d^{(2)}(i, j) = (d_j - y_i)^T S_i^{-1} (d_j - y_i) \quad (2)$$

式中,  $d^{(2)}(i, j)$  为第  $i$  个跟踪器预测框与第  $j$  个检测目标框之间的马氏距离;  $d_j$  为第  $j$  个检测目标的位置;  $y_i$  为第  $i$  个跟踪器预测的目标位置;  $S_i^{-1}$  为第  $i$  个跟踪器的协方差预测矩阵。

根据余弦距离和马氏距离来综合衡量检测目标与轨迹预测目标之间的相似性:

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

式中,  $c_{i,j}$  为目标相似性度量, 超参数  $\lambda$  为两种度量方式在级联匹配中所占比重。

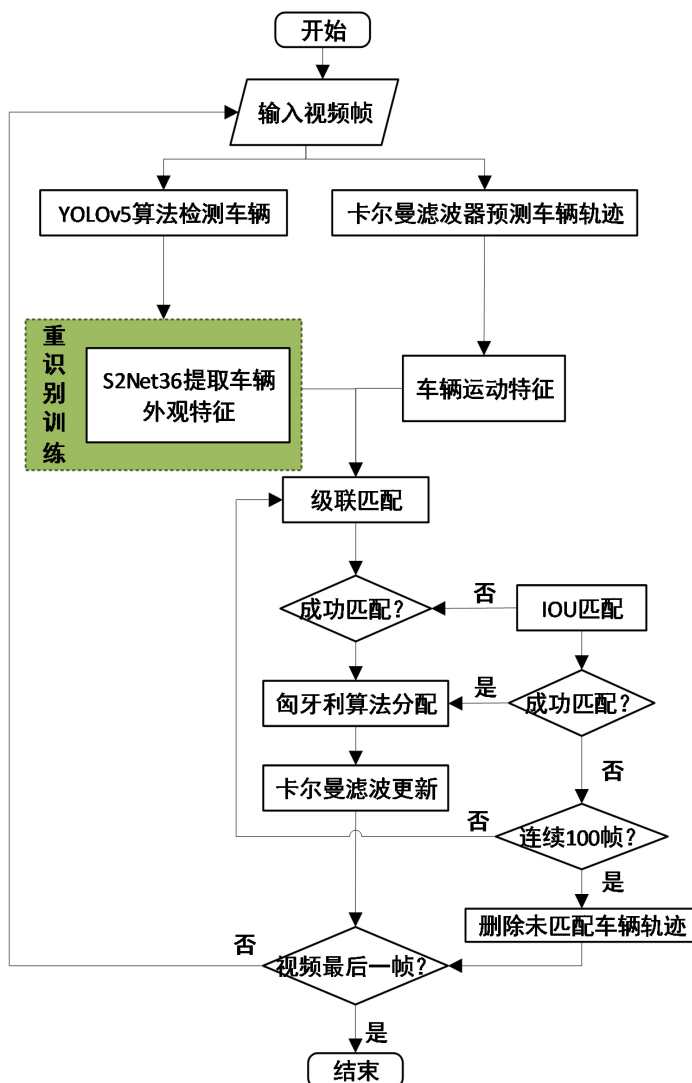


Figure 1. Diagram of the improved tracking frame

图 1. 改进的跟踪框架图

### 3. 重识别网络优化

#### 3.1. ResNet36

DeepSORT 原始重识别网络结构如图 2 所示, 其最终输出 128 维特征, BasicBlock 为其残差层的单元残差结构。图 3 为改进后的 ResNet36 重识别网络, 其在原本重识别网络的基础上将残差层扩增至五层, 且每层残差层通过叠加不同数量的 BasicBlock 结构加深网络, L1~L5 五层残差层叠加的 BasicBlock 结构数量分别 4, 5, 3, 1, 1, 通过批处理和归一化输出 512 维特征。

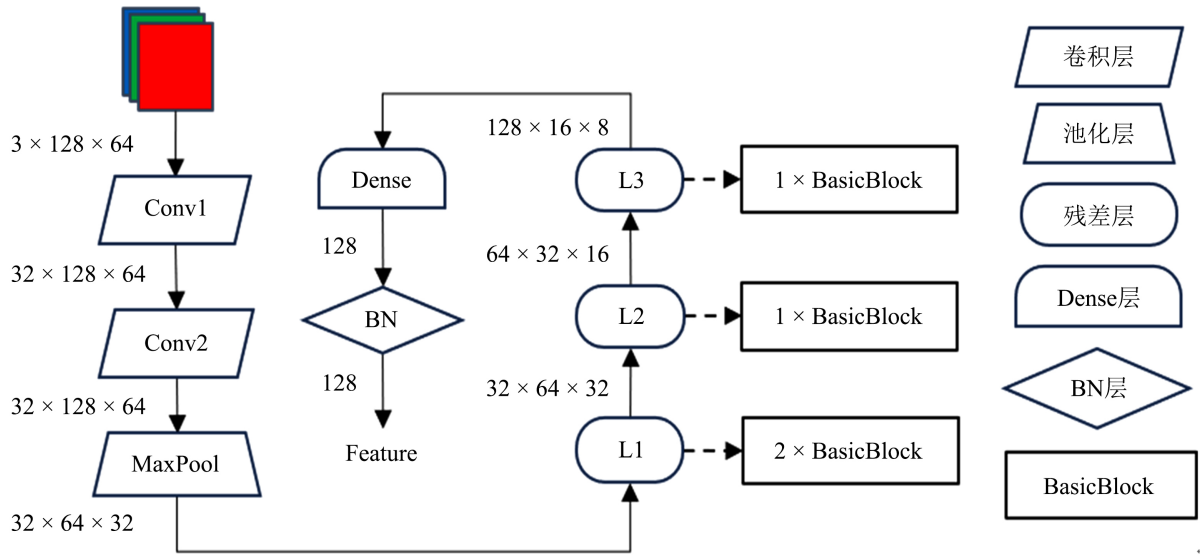


Figure 2. Diagram of the DeepSORT raw re-identification network structure  
图 2. DeepSORT 原始重识别网络结构图

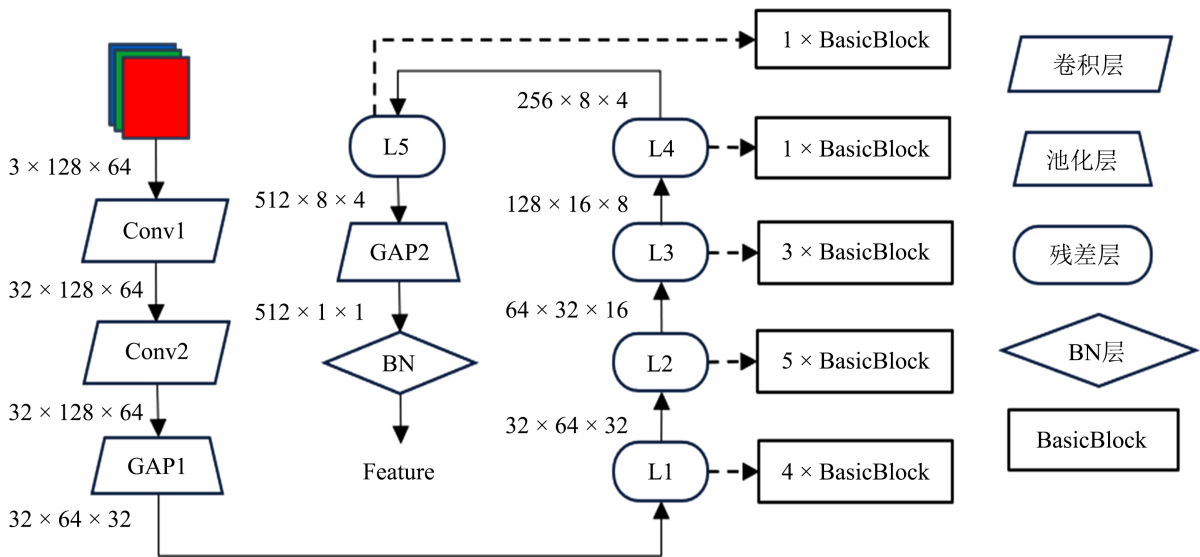


Figure 3. Diagram of the ResNet36 network structure  
图 3. ResNet36 网络结构图

### 3.2. SER 模块

融合 SE 注意力机制与 BasicBlock 构建特征提取 SER 模块, 如图 4 所示。经过 BasicBlock 的两层  $3 \times 3$  卷积生成尺寸为  $H \times W \times C$  的特征图, 进而将其输入 SE 注意力机制中进行 Squeeze、Excitation 及 Scale 操作, 实现特征图不同通道信息的融合。Squeeze 操作是将提取的特征图通过 Global pooling 从空间维度进行压缩; Excitation 操作是通过两层 FC、ReLU 及 Sigmoid 激活函数捕获通道之间的相互关系, 进而获得每个通道的权值; Scale 操作将经卷积操作得到的特征图与每个通道权值进行加权, 输出 SER 模块的残差映射(Residual Mapping)。最后将其与原始特征图  $X$  的恒等映射(Identity Mapping)进行加和得到 SER 模块的输出  $\tilde{X}$ 。

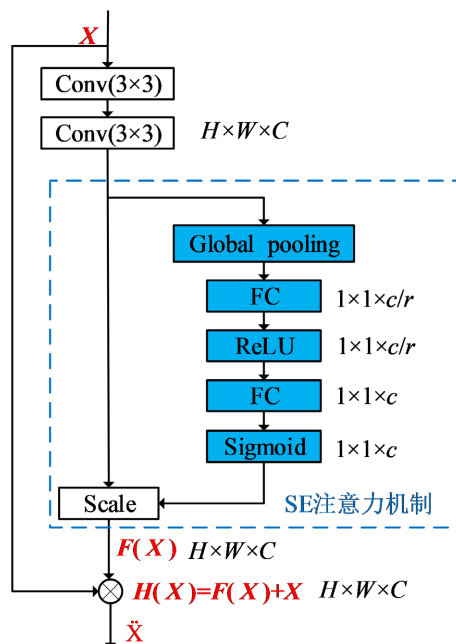


Figure 4. Diagram of the SER module  
图 4. SER 模块图

### 3.3. SE-Res2Net 模块

基于 Res2Net 与 SE 注意力机制构建 SE-Res2Net 模块, 如图 5 所示。Res2Net 输入的特征图为  $x \in R^{W \times H \times C}$ , 其中  $W \times H$  为特征图尺寸,  $C$  为特征通道数。将特征图平均划分为 4 个特征子集  $x_i$ ,  $x_i \in R^{w \times h \times C/4}$  ( $i=1,2,3,4$ ), 然后通过  $3 \times 3$  卷积滤波器对  $x_i$  进行卷积操作, 获得  $y_i$ , 如式(4)所示。

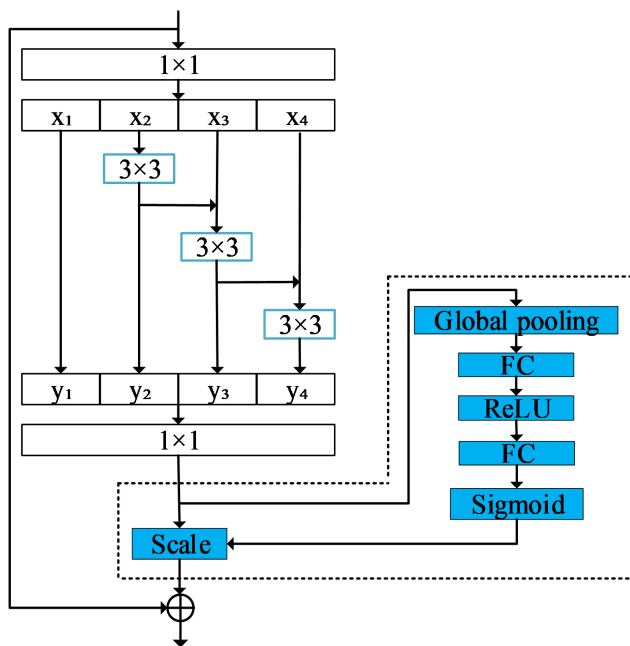


Figure 5. Diagram of the SE-Res2Net module  
图 5. SE-Res2Net 模块图



$$y_i = \begin{cases} x_i, & i = 1 \\ K(x_i), & i = 2 \\ K(x_i + y_{i-1}), & i = 3, 4 \end{cases} \quad (4)$$

$y_i (i=1,2,3,4)$  通过  $1 \times 1$  滤波器可以得到 Res2Net 的输出  $y$ , SE-Res2Net 将  $y$  输入到嵌入的 SE 注意力机制中, 通过 Global pooling 实现特征图的压缩, 通过两层 FC、ReLU 和 Sigmoid 激活函数获得每个通道的权值, 通过 Scale 操作完成特征图与相应通道权值的加权, 进而完成 SE 注意力机制对  $y$  的处理, 获得 SE-Res2Net 的残差映射。最后与输入的原始特征图  $x \in R^{w \times h \times C}$  的恒等映射进行求和获得 SE-Res2Net 的输出。

### 3.4. S2Net36

在 ResNet36 网络的基础上使用 SER 模块分别替换 ResNet36 网络 L3 层的后两个 BasicBlock 结构以及 L4 层的 BasicBlock 结构, 使用 SE-Res2Net 模块替换 L5 层的 BasicBlock 结构, 进而获得新的重识别网络 S2Net36, 如图 6 所示。

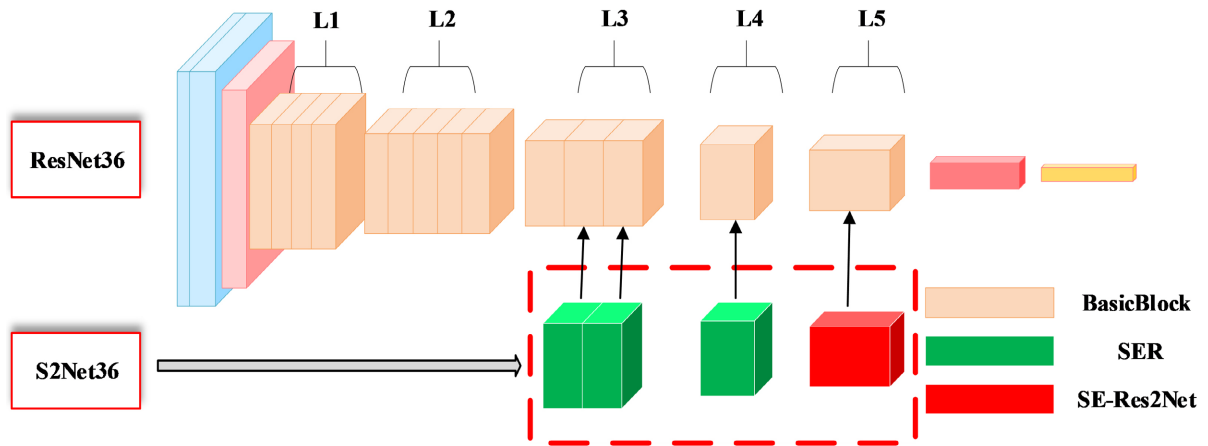


Figure 6. Diagram of the S2Net36 network structure

图 6. S2Net36 重识别网络结构图

## 4. 损失优化

采用交叉熵损失函数(Cross Entropy Loss)对重识别网络分类层获得的分类预测值  $p_i$  进行训练。交叉熵损失函数常用于衡量真实值与网络预测值之间的距离, 判定二者的接近程度, 计算公式如(5)所示。

$$L_{ce} = -\sum_{i=1}^N p_i \log(\hat{p}_i) \quad (5)$$

式中  $\hat{p}$  为输入样本的真实值,  $y$  为真实分类标签,  $p_i$  为分类预测值, 当样本  $i$  的真实类别等于  $y$  时  $p_i$  取 1, 否则取 0。

采用三元组损失函数(Triplet loss) [17]对重识别网络输出的 512 维特征进行训练。给定三元组训练数据  $\{x_a, x_p, x_n\}$ ,  $x_a$  和  $x_p$  为 ID 相同车辆的不同样本,  $x_a$  和  $x_n$  互为不同 ID 车辆的样本, 三元组函数公式如(6):

$$L_{trip} = \max(d(a, p) - d(a, n) + \text{margin}, 0) \quad (6)$$

其中,  $d(a,p)$  为正样本对之间的距离,  $d(a,n)$  为负样本对之间的距离, margin 为阈值, 用以控制正负样本的距离。

## 5. 实验

实验基于 Windows10 操作系统, 处理器配置为 12 vCPU Intel(R) Xeon(R) Platinum 8255C CPU@ 2.50 GHZ, GPU 为 NVIDIA GeForce GTX2080Ti (11 GB), 深度学习框架为 PyTorch1.9.0。

### 5.1. 重识别训练

采用 VeRi-776 [18] 车辆重识别数据集分别对 ResNet36 网络与 S2Net36 进行训练。VeRi-776 为大规模城市交通监控车辆再识别图像数据集, 该数据集包含 20 个监控摄像机采集的 776 辆车的图像, 共 50,000 多张图片。

训练过程设置单次传递给模型用以训练的数据(batchsize)设置为 64, 采用随机梯度下降(SGD)作为优化器, SGD 的 momentum 为设置为 0.9, 模型训练超参数中学习率设置为 0.01, 三元组损失函数的 margin 为 0.3, 分别使用准确率(Accuracy)与损失(Loss)用以衡量模型的准确度与性能。

准确率是指模型预测结果与真实结果的匹配度, 准确率越高模型预测结果更准, 计算公式如式(7)所示:

$$Acc = 1 - \frac{T_P + T_N}{T_P + T_N + F_P + F_N} \quad (7)$$

损失是指模型预测结果与真实结果之间的差距, 损失越小模型性能越好, 本文结合交叉熵损失函数与三元组损失函数对重识别网络进行训练, 训练过程需使得交叉熵损失值与三元组损失值尽量小, 损失计算公式如式(8)所示:

$$Loss = L_{ce} + L_{trip} \quad (8)$$

S2Net36 网络的训练准确率与损失变化如图 7 所示, 可以发现当训练迭代次数为 40 时, S2Net36 网络的准确率和损失变化开始趋于平稳, 其中准确率达到到了 94.7%, 损失则降低到 0.97, 表明网络此时已具有较好的分类预测能力。

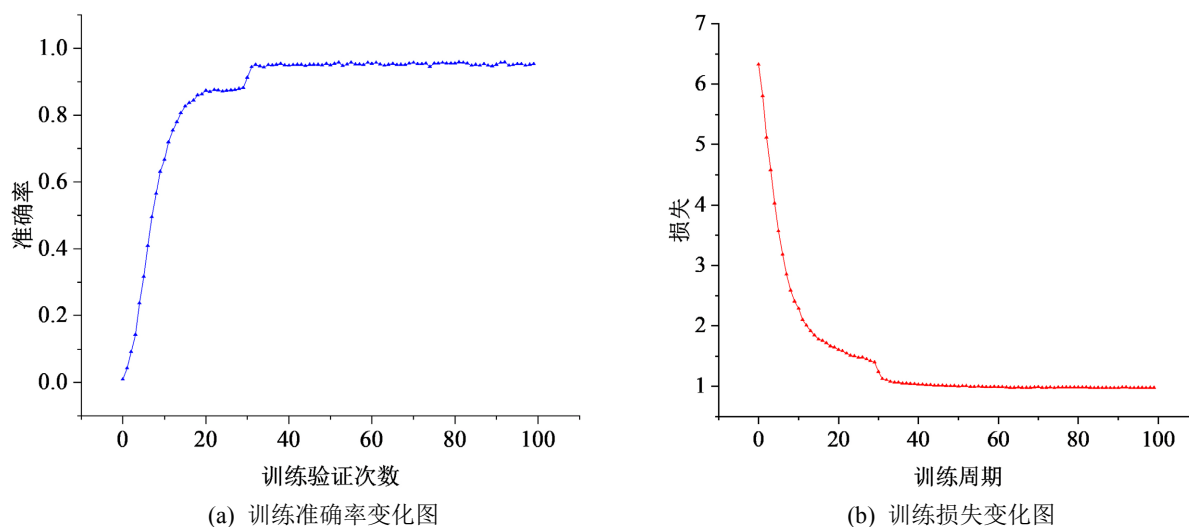


Figure 7. Diagram of the S2Net36 network re-identification training  
图 7. S2Net36 重识别训练图



## 5.2. 跟踪评价

本文采用 MOTA、MOTP、IDF1 作为多目标跟踪的评价指标，其中 MOTA 衡量跟踪算法在检测物体与保持轨迹时的性能，MOTP 用于衡量目标位置确定的精确程度，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。DeepSORT 原始算法与本文改进算法在多目标跟踪评价指标上的对比结果如表 1 所示。

**Table 1.** Follow-up review results comparison

**表 1.** 跟踪评价结果对比

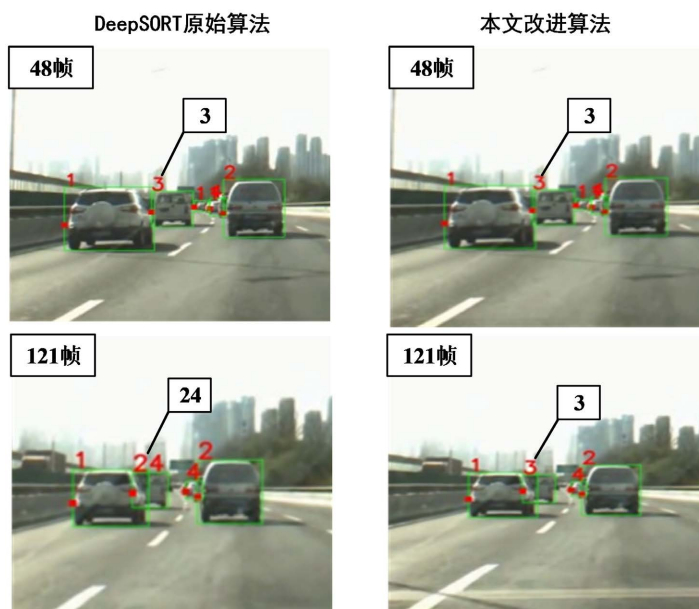
算法	MOTA/%	MOTP/%	IDF1/%
DeepSORT 原始算法	54.74	82.98	60.03
本文改进算法	<b>55.92</b>	82.70	<b>60.83</b>

由表 1 可知，本文改进算法相比于 DeepSORT 原始算法 MOTA 指标提升了 1.18%，MOTP 指标降低了 0.28%，IDF1 指标提升了 0.80%。MOTA 与 IDF1 的提升表明对重识别网络进行改进并优化损失函数，能够提升模型的跟踪效果。

## 5.3. 跟踪结果可视化

在不同的城市道路采集实际数据对比 DeepSORT 原始算法与本文改进算法的跟踪效果。图 8 与图 9 分别为部分遮挡场景和目标漏检场景下两种算法的跟踪效果。其中左侧为 DeepSORT 原始算法的跟踪结果，右侧为本文改进算法的跟踪结果。图中左上角为视频帧数，图中编号为车辆的分配 ID。

根据图 8 可以发现 DeepSORT 原始算法在第 48 帧时检测到目标 ID3，其后由于该车辆局部被遮挡，在第 121 帧时 ID 切换为 24；而采用本文改进算法该车辆的 ID 始终保持为 3，可见本文算法能够在部分遮挡情况下保持对车辆的跟踪能力。



**Figure 8.** Partially occluded scenes

**图 8.** 部分遮挡场景

从图9中可以发现 DeepSORT 原始算法在第 122 帧检测到目标 ID28, 在第 142 帧时该目标发生漏检, 之后在第 148 帧时该目标重新被检测到, 但其 ID 切换为 37; 而采用本文改进算法, 在该车辆同样发生漏检情况下, ID 始终保持为 26, 表明本文改进算法能够在发生漏检情况后依然对车辆保持良好的跟踪能力。

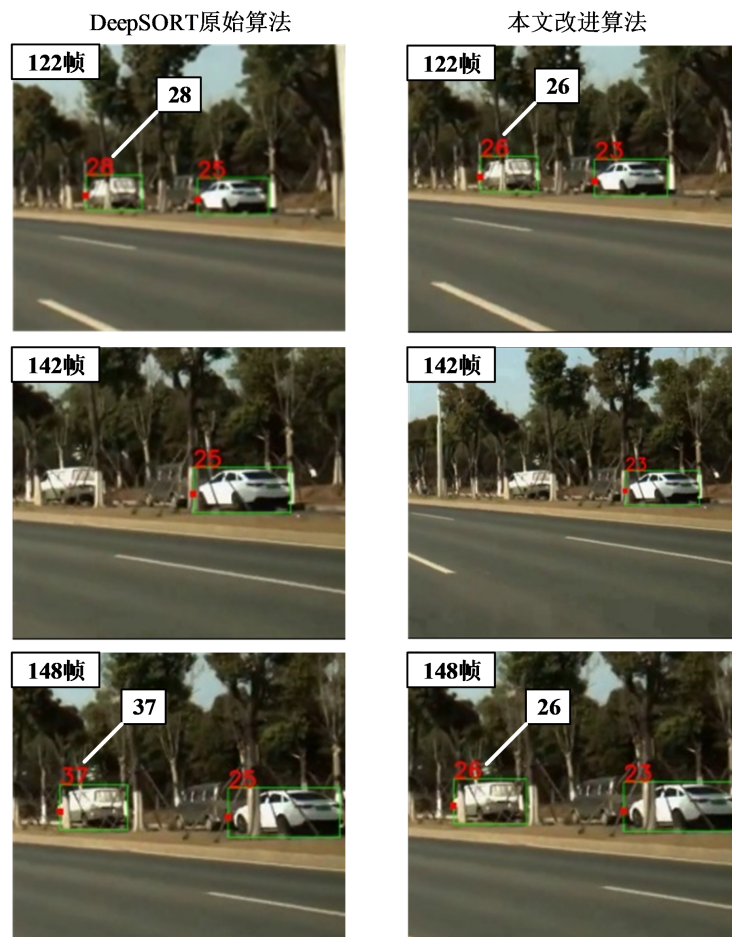


Figure 9. Missed detection scenarios  
图 9. 漏检场景

## 6. 结论

针对部分遮挡、目标漏检等导致车辆 ID 频繁切换的问题, 本文提出了一种基于改进 DeepSORT 算法的道路前方车辆跟踪方法。该方法通过加深网络层数以提取深度特征, 嵌入 SE 注意力机制以关注目标重要通道, 融合 Res2Net 模块以提取目标区域特征, 从而增强重识别网络对车辆外观特征的提取能力。此外, 引入三元组损失函数对重识别网络进行训练, 增强模型获取的车辆外观特征的辨别力。通过对 DeepSORT 原始算法和本文改进算法进行对比实验, 发现本文改进算法能够改善因部分遮挡和目标漏检导致的 ID 切换问题, 提高了对前方道路车辆的跟踪能力, 但在 MOTP 指标的表现上有所不足, 后续将进一步提升目标检测性能及卡尔曼滤波预测的目标位置准确度, 进一步优化目标位置确定的精确程度。

## 参考文献

- [1] 伍瀚, 等. 基于深度学习的视觉多目标跟踪研究综述[J]. 计算机科学, 2023, 50(4): 77-87.

- [2] 贺愉婷, 车进, 吴金蔓. 基于 YOLOv5 和重识别的行人多目标跟踪方法[J]. 液晶与显示, 2022, 37(7): 880-890.
- [3] 储琪. 基于深度学习的视频多目标跟踪算法研究[D]: [博士学位论文]. 合肥: 中国科学技术大学, 2019.
- [4] Bewley, A., Ge, Z., Ott, L., *et al.* (2016) Simple Online and Realtime Tracking. 2016 *IEEE International Conference on Image Processing (ICIP)*, Phoenix, 25-28 September 2016, 3464-3468. <https://doi.org/10.1109/ICIP.2016.7533003>
- [5] Wojke, N., Bewley, A. and Paulus, D. (2017) Simple Online and Realtime Tracking with a Deep Association Metric. 2017 *IEEE International Conference on Image Processing (ICIP)*, Beijing, 17-20 September 2017, 3645-3649. <https://doi.org/10.1109/ICIP.2017.8296962>
- [6] 殷远齐. 基于机器学习的前方车辆行为识别方法研究[D]: [硕士学位论文]. 西安: 长安大学, 2022.
- [7] Bochkovskiy, A., Wang, C. and Liao, H.M. (2020) YOLOv4: Optimal Speed and Accuracy of Object Detection. <https://arxiv.org/pdf/2004.10934.pdf>
- [8] Woo, S., Park, J., Lee, J.Y., *et al.* (2018) Cbam: Convolutional Block Attention Module. In: Ferrari, V., Hebert, M., Sminchisescu, C., Weiss, Y., Eds., *Proceedings of the European Conference on Computer Vision (ECCV)*. Springer, Cham, 3-19. [https://doi.org/10.1007/978-3-030-01234-2\\_1](https://doi.org/10.1007/978-3-030-01234-2_1)
- [9] Gao, S.H., Cheng, M.M., Zhao, K., *et al.* (2019) Res2Net: A New Multi-Scale Backbone Architecture. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **43**, 652-662. <https://doi.org/10.1109/TPAMI.2019.2938758>
- [10] 顾立鹏, 等. 无人车驾驶场景下的多目标车辆与行人跟踪算法[J]. 小型微型计算机系统, 2021, 42(3): 542-549.
- [11] Hu, J., Li, S., Gang, S., *et al.* (2018) Squeeze-and-Excitation Networks. *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, Salt Lake City, 18-23 June 2018, 7132-7141. <https://doi.org/10.1109/CVPR.2018.00745>
- [12] Zhou, X., Wang, D. and Krähenbühl, P. (2019) Objects as Points. arXiv preprint, arXiv:1904.07850.
- [13] 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/OL]. 汽车技术: 1-7. <https://doi.org/10.19620/j.cnki.1000-3703.20220911>, 2023-09-20.
- [14] 尤晓雨. 基于改进的 YOLOv5 和 DeepSort 车辆检测跟踪算法研究[D]: [硕士学位论文]. 西安: 长安大学, 2022.
- [15] Wang, H., Zhang, F. and Wang, L. (2020) Fruit Classification Model Based on Improved Darknet53 Convolutional Neural Network. 2020 *International Conference on Intelligent Transportation, Big Data & Smart City (ICITBS)*, Vientiane, 11-12 January 2020, 881-884. <https://doi.org/10.1109/ICITBS49701.2020.00194>
- [16] 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.
- [17] Schroff, F., Kalenichenko, D. and Philbin, J. (2015) Facenet: A Unified Embedding for Face Recognition and Clustering. *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, Boston, 7-12 June 2015, 815-823. <https://doi.org/10.1109/CVPR.2015.7298682>
- [18] Liu, X., Liu, W., Mei, T., *et al.* (2016) A Deep Learning-Based Approach to Progressive Vehicle Re-Identification for Urban Surveillance. In: Leibe, B., Matas, J., Sebe, N., Welling, M., Eds., *European Conference on Computer Vision*. Springer, Cham, 869-884. [https://doi.org/10.1007/978-3-319-46475-6\\_53](https://doi.org/10.1007/978-3-319-46475-6_53)