

DOI: 10.16652/j.issn.1004-373x.2023.05.009

引用格式: 郑繁亭, 邢关生. 基于改进 DeepSort 的行人多目标跟踪算法[J]. 现代电子技术, 2023, 46(5): 40-46.

基于改进 DeepSort 的行人多目标跟踪算法

郑繁亭, 邢关生

(青岛科技大学 自动化与电子工程学院, 山东 青岛 266061)

摘要: 针对现有的行人多目标跟踪算法在遮挡、人群密集和光线差等情况下表现不佳的问题, 提出一种改进 YOLOv4 与改进 DeepSort 算法相结合的行人多目标跟踪算法。首先, 为增强检测网络的特征提取能力, 在 YOLOv4 中嵌入 ECANet 注意力模块, 提高检测精度; 其次, 在改进 DeepSort 的跟踪算法中, 由卡尔曼滤波算法预测多个行人目标在图像中的轨迹之后, 使用 GhostNetV1 替换 DeepSort 中的重识别网络来生成行人的外观特征, 提高行人重识别网络的性能; 进而, 采用匈牙利算法对检测框和预测框进行最优匹配, 对未匹配成功的检测框采用 DIOU 代替 IOU (交并比) 进行二次匹配, 提高 DeepSort 网络的跟踪性能; 最后, 开展了新跟踪算法与原 DeepSort 算法的对比实验, 结果表明新算法的误检、漏检现象变少, 鲁棒性增强, 跟踪性能得到提高, MOTA 提升了 18.8%, IDF1 提升了 18.2%, 身份编号转换次数降低了 84 次。

关键词: 多目标跟踪; 改进 DeepSort; 轨迹预测; 外观特征生成; 图像处理; 对比实验

中图分类号: TN911.73-34; TP391

文献标识码: A

文章编号: 1004-373X(2023)05-0040-07

Pedestrian multi-target tracking algorithm based on improved DeepSort

ZHENG Fanting, XING Guansheng

(College of Automation and Electronic Engineering, Qingdao University of Science and Technology, Qingdao 266061, China)

Abstract: In allusion to the poor performance of existing pedestrian multi-target tracking algorithms in the case of occlusion, dense crowd and poor light, a pedestrian multi-target tracking algorithm based on improved YOLOv4 and improved DeepSort algorithms is proposed. In order to enhance the feature extraction ability of the detection network, ECANet attention module is embedded into YOLOv4 to improve the detection accuracy. In the improved tracking algorithm of DeepSort, the Kalman filter algorithm is used to predict the trajectories of multiple pedestrian targets in the image, and then GhostNetV1 is used to replace the re-recognition network in DeepSort to generate the appearance features of pedestrians, so as to improve the performance of the re-recognition network. The Hungarian algorithm is used to optimize the matching of detection boxes and prediction boxes, and DIOU (distance IOU loss) is used to replace IOU (overlap and union ratio) for the second matching of unmatched detection boxes, so as to improve the tracking performance of DeepSort network. The contrast experiment between the new tracking algorithm and the original DeepSort algorithm was carried out. The results show that the new tracking algorithm has less false detection and missing detection, its robustness is enhanced, its tracking performance is improved, its MOTA is increased by 18.8% its IDF1 is increased by 18.2%, and the number of identity number conversion is decreased by 84.

Keywords: multi-target tracking; improved DeepSort; trajectory prediction; appearance feature generation; image processing; contrast experiment

0 引言

行人跟踪是指检测行人以确定其在图像中的位置, 给每个行人分配不同的编号, 进而生成行人轨迹, 实现持续跟踪。行人跟踪在很多领域具有重要的研究价值和广泛的应用前景, 例如无人驾驶、公共设施安全等^[1]。

目前, 国内外的目标跟踪算法研究工作主要分为三类: 基于相关滤波的跟踪算法、基于特征匹配的跟踪算法和基于检测的跟踪算法。基于相关滤波的跟踪算法利用滤波模板与目标候选区域之间的相关运算求得响应值, 通过响应值的大小来确定目标在视频每一帧中的具体位置, 其计算速度较快, 但在遮挡和环境复杂等情况下

收稿日期: 2022-07-31

修回日期: 2022-08-22

基金项目: 国家自然科学基金资助项目: 空地机器人网络的同时视觉目标定位与分布式运动规划(61503118); 国家自然科学基金资助项目: 基于脑认知与机器感知联合决策的脑-机融合性能优化研究(62006135)

的精度较低。基于特征匹配的跟踪算法主要利用人为指定的特征进行特征匹配,其最终的效果主要取决于特征的选取,鲁棒性较差。文献[2]建立具有 RGB 颜色特征的特征模板,使用局部约束编码的形式实现行人追踪。文献[3]以行人目标运动过程中外接矩形的长宽比作为特征模板,使用卡尔曼滤波作为跟踪器。基于检测的跟踪算法是先利用检测网络检测出行人目标,再利用跟踪器实现行人追踪。由于目标检测和跟踪能够形成互补,因此基于检测的跟踪算法成为当前主流的跟踪算法。

目前的检测网络可分为一阶段算法和二阶段算法。一阶段算法主要包括 YOLO 系列和 SSD^[4]等,二阶段算法主要包括 R-CNN、Fast R-CNN 和 Faster R-CNN 等。相比于二阶段算法,一阶段算法的速度更快,而 YOLOv4^[5]作为一阶段算法的典型代表,具有速度快、精度高的特点。

在跟踪算法方面,文献[6]提出 Sort 算法,采用卡尔曼滤波进行预测,由匈牙利算法完成匹配,跟踪速度快,但是没有考虑目标特征,容易发生身份编号转换(ID Switch)现象。文献[7]在 Sort 算法基础上进一步提出 DeepSort 算法,参考了行人重识别的思想,加入 128 维外观特征提取网络,采用级联匹配算法,大大降低了身份切换。文献[8]使用 YOLOv4 网络和 GIOU (Generalized IOU Loss) 对 DeepSort 算法进行改进,但在密集环境下存在漏检问题。文献[9]使用 Gaussian YOLOv3 网络和中心损失函数对 DeepSort 算法进行改进,但在复杂环境下存在鲁棒性不足问题。

针对长时间遮挡、光线差等情况下可能导致的漏检、错检和 ID Switch,本文提出了改进的 YOLOv4 与改进 DeepSort 相结合的行人多目标跟踪算法。为了提高检测网络在密集场景和遮挡场景下的检测精度,引入注意力机制来增强 YOLOv4 的特征提取能力。为了更好地提取行人的外观特征,提高模型的跟踪性能,本文采用 GhostNet^[10]模块来改进重识别网络。为了更好地度量检测框与预测框之间的匹配度以达到更好的跟踪效果,本文对匹配不成功的轨迹和检测结果进行 DIOU (Distance IOU Loss)^[11]关联匹配。

1 YOLOv4 算法的改进

1.1 YOLOv4 算法原理

YOLOv4 是一阶段目标检测算法的典型代表,其兼顾了良好的准确性与实时性。算法主要由主干特征提取网络(Backbone)、颈部特征融合网络(Neck)和头部预测网络(YOLO Head)组成。YOLOv4 网络结构如图 1 所示,Backbone 为 CSPDarkNet53 网络,由一个卷积层和一系列残差结构组成,通过对输入特征图进行下采样来获取图像特征和语义信息;Neck 由 SPP 模块和 PANet 模块组成,SPP^[12]模块利用多组不同池化核尺寸的最大池化操作有效提高了特征图的感受野,PANet 模块为特征金字塔结构,通过上采样、下采样、Concat 和卷积等操作进行特征融合,进一步增强了特征提取效果;头部预测网络为 YOLO Head,用于输出网络的预测结果。

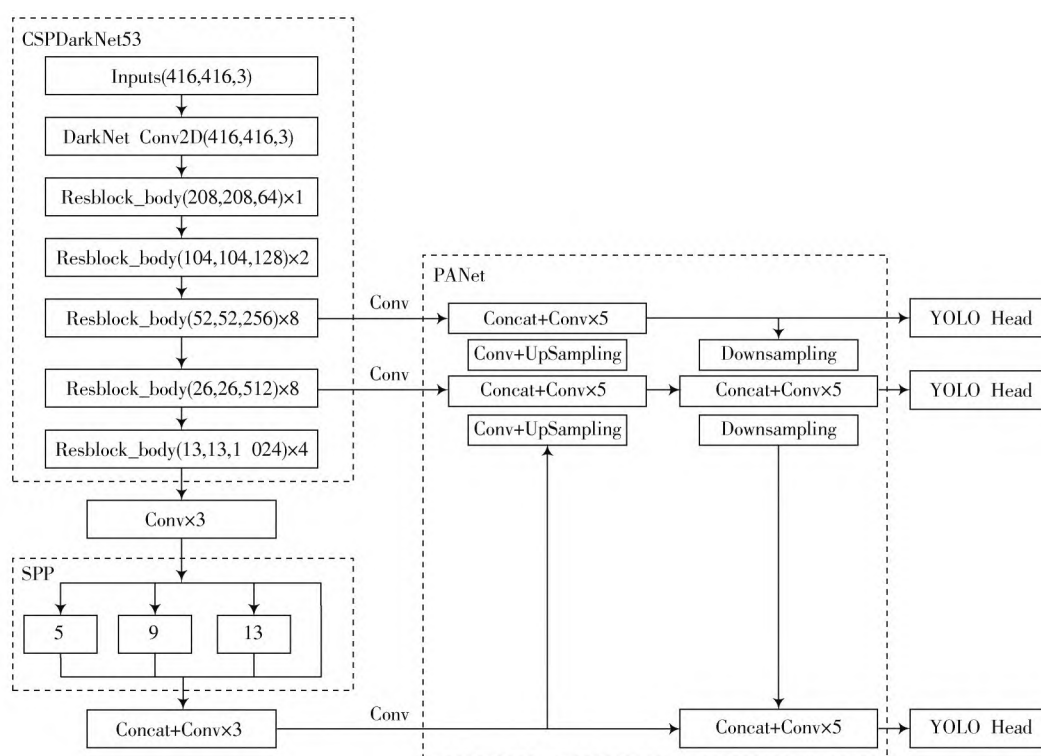


图1 YOLOv4 网络结构

1.2 注意力机制

注意力机制可以视为一种能够提升特征提取网络精度的插件。传统的特征提取网络在提取特征时不会考虑不同特征通道对于待检测目标的重要程度,而注意力模块则会先通过学习得到权重向量,来描述各个特征通道的重要性程度,然后利用权重向量对特征图进行重建。目前最常用的注意力机制模块是SENet^[13]。SENet结构简单,能够在只增加少量计算量的代价下显著提高现有特征提取网络的性能。但是SENet对特征先降维再映射到高维的操作破坏了特征通道与注意力权重之间的直接对应关系。文献[14]提出了一种不降维的局部跨信道交互策略和自适应选择一维卷积核大小的方法ECANet。ECANet同样是一个轻量化模块,只需要很小的计算开销便能带来特征提取网络性能的提升,其结构如图2所示。

ECANet模块的基本步骤为:对于给定的输入 x ,先在空间维度上进行平均池化,将空间信息压缩到通道中,再采用长度为 K 、步长为1的一维卷积进行特征融合,最后经过Sigmoid激活函数生成权重向量。

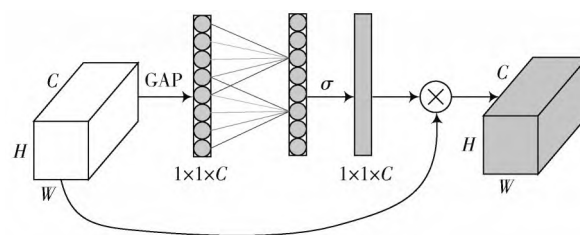


图2 ECANet模块结构

1.3 引入注意力机制的改进YOLOv4

本文通过引入ECANet模块来提高YOLOv4特征提取模块的性能,改进的YOLOv4网络结构如图3所示。首先对输入的图像进行预处理,将其尺寸调整为416×416,然后将其输入到CSPDarkNet53中进行特征提取,再把经过特征提取后得到的三个尺寸分别为52×52, 26×26, 13×13的特征图分别输入到ECANet中进行特征图的重新校准,将重新校准后特征图的通道数通过卷积、SPP和PANet操作降低至18(3×5+C),其中C为待检测目标的类别数,最后通过YOLO Head完成行人检测。

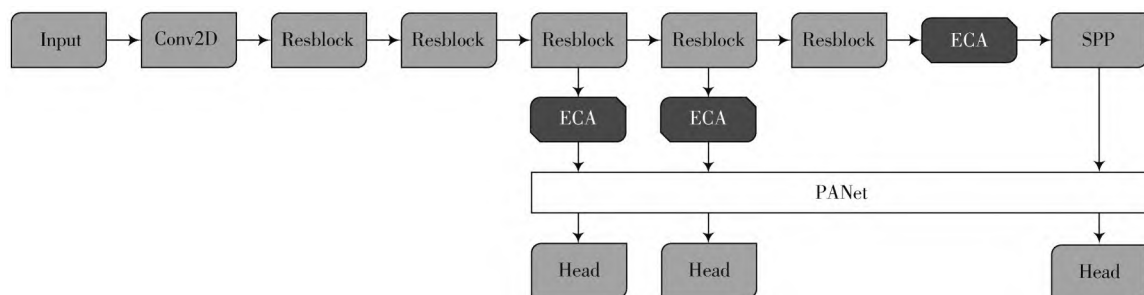


图3 改进的YOLOv4网络结构

2 DeepSort算法的改进

2.1 DeepSort算法原理

DeepSort是Sort算法的改进版本,为了降低跟踪目标身份编号转换(ID Switch)次数,DeepSort引入了外观特征和级联匹配,以目标检测网络检测结果的检测框坐标和置信度作为输入,通过卡尔曼滤波算法进行轨迹预测,匈牙利算法将预测轨迹与当前帧进行级联匹配,最后进行卡尔曼滤波更新。DeepSort使用八维向量 $(u, v, r, h, \dot{u}, \dot{v}, \dot{r}, \dot{h})$ 进行计算,其中 (u, v) 为边界框中心坐标, r 为矩形框的纵横比, h 为矩形框的高度, $(\dot{u}, \dot{v}, \dot{r}, \dot{h})$ 为 (u, v, r, h) 的速度。

DeepSort采用运动和外观两个维度的信息进行目标匹配,通过计算跟踪器预测框与行人检测框的马氏距离作为运动信息度量,其表达式为:

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

式中: $d^{(1)}(i, j)$ 为第 j 个检测框和第 i 条轨迹之间的运动匹配度; s_i 是由卡尔曼滤波器得到的当前时刻观测空间协方差矩阵; y_i 是轨迹在当前时刻的预测框状态; d_j 是第 j 个检测框的状态 $[u, v, r, h]^T$ 。通过马氏距离对检测框进行筛选,采用9.487 7作为阈值 $t^{(1)}$,若关联的马氏距离 $d^{(1)} < t^{(1)}$,则表示运动状态关联成功。

对于外观信息,引入重识别模型来提取目标的外观特征。采用最小余弦距离对检测框进行筛选,该距离小于一定阈值则代表跟踪成功。最小余弦距离公式如式(2)所示:

$$d^{(2)}(i, j) = \min \{ 1 - r_j^T r_k^{(i)} \}, \quad r_k^{(i)} \in R_i \quad (2)$$

式中: $d^{(2)}(i, j)$ 表示第 j 个检测框与第 i 个轨迹的最小余弦距离; r_j 表示第 j 个检测框 d_j 的外观描述符,设置 $\|r_j\| = 1$; $r_k^{(i)}$ 表示第 i 个轨迹的外观描述符; R_i 表示第 i 个轨迹的外观信息库,保存最新的100条目标成功匹配的

外观信息。

当面对复杂情况时运动和外观两个度量指标可以互补,来增强网络的跟踪能力。运动度量可以提供更多的位置信息,当位置信息偏弱时,外观度量可以提供更多的外观信息,最后使用线性加权的方式将两个度量信息关联起来,关联公式如式(3)所示:

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

式中 λ 为权重值。

2.2 引入 DIOU 改进 DeepSort

DeepSort 在进行目标匹配时会交并比(IOU)作为一个关联指标,对于经过级联匹配之后未成功匹配的预测框,DeepSort 会将其与原检测框进行 IOU 匹配来确定最终的匹配状态。但是 IOU 仅仅考虑了两框的重合程度,并不能反映两框的距离,因此本文引入 DIOU 代替 IOU,如图4所示。

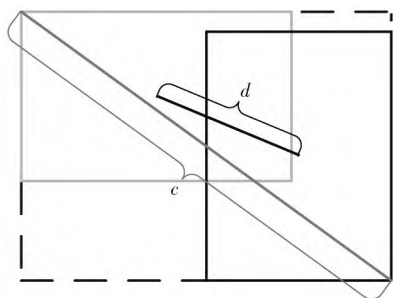


图4 DIOU

对于两个矩形水平边界框, DIOU 的计算式为:

$$\text{DIOU} = \text{IOU} - \frac{\rho^2(b, b^{\text{gt}})}{c^2} = \text{IOU} - \frac{d^2}{c^2} \quad (4)$$

$$\text{DIOU}_{\text{loss}} = 1 - \text{DIOU} \quad (5)$$

式中: $\rho^2(b, b^{\text{gt}})$ 表示两框中心点的欧氏距离 d , b 表示预测框(图4中的黑色矩形), b^{gt} 表示真实框(图4中的灰色矩形); c 为两框最小外接矩形的对角线长度。

从计算原理上可以看出 DIOU 能够更好地反映两个框的相交情况,将其用于计算网络的损失函数可以直接最小化两框之间的距离,收敛速度更快。

2.3 行人重识别网络

DeepSort 之所以能有效解决遮挡问题,是因为其搭建了一个重识别神经网络 WRN。WRN 主要由残差块构成,参数多、计算量大,为了增强重识别网络的速度和精度,引入 Ghost 模块代替 WRN 中的残差模块。Ghost 模块主要由 Ghost module 和 DWConv 组成,其结构如图5a)所示,其中 Ghost module 的结构如图5b)所示。首先对输入的特征图进行 1×1 的卷积,将通道数压缩为原来的一半,然后对输出的特征图进行逐层卷积,最后将两部分的输出进行拼接。

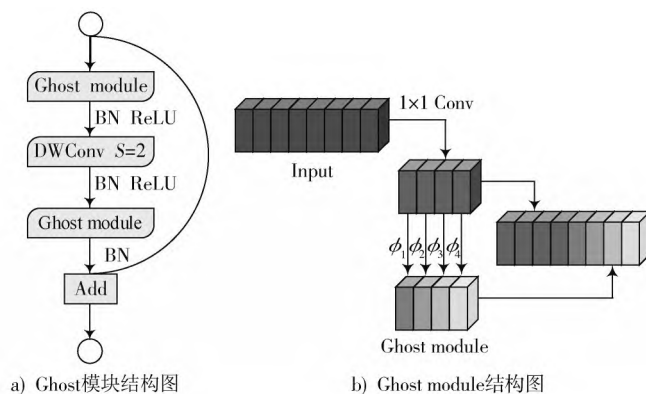


图5 模块结构图

GhostNetV1 行人重识别模型如表1所示,网络的输入尺寸为 $32 \times 128 \times 64$,通过堆叠 Ghost 模块来提取特征,最后经过平均池化和 FC 层生成 128 维外观特征向量,用于进行外观度量的计算。

表1 Ghost 模块网络结构

名称	输出尺寸
Conv1	$32 \times 128 \times 64$
Conv2	$32 \times 128 \times 64$
Ghost3	$32 \times 64 \times 32$
Ghost4	$64 \times 32 \times 18$
Ghost5	$128 \times 16 \times 8$
Average Pool6	$128 \times 1 \times 1$
FC7	128

3 实验

3.1 数据集

本文使用 MOT16 数据集进行网络训练与评估,该数据集是 MOT challenge 系列中一个衡量多目标检测跟踪方法的标准数据集。如图6所示,为了增强行人检测模型的鲁棒性,本文还通过图像增强技术来随机产生相似但不相同的训练样本,以达到扩增数据集的目的。



图6 数据增强

3.2 实验环境

本文使用Pytorch框架搭建神经网络,将Python作为编程语言。硬件平台配置如下:CPU:intel Core i9-9900K;GPU:NVIDIA GeForce GTX2080Ti 11 GB;RAM:32 GB;操作系统:Ubuntu 18.04;软件环境: CUDA 11.0,Python 3.7,Pytorch 1.7.1。

3.3 模型训练及评价指标

3.3.1 模型训练

在训练阶段,网络初始参数配置为:32个样本为一个批次,初始学习率为0.001,训练轮数为100轮。为增强模型的训练效果,采用动态调整的方式更新学习率,在训练迭代次数达到80和90次时,学习率可分别降低至初始学习率的10%和1%。

3.3.2 评价指标

选用平均准确率AP作为目标检测网络的评价指标。选用MOTChallenge定义五个指标评价多目标跟踪效果:

- 1) ID Switch(↓):表示一条轨迹中跟踪目标改变自身编号的次数。
- 2) MOTA(↑):多目标跟踪准确率。
- 3) IDF1(↑):正确识别出的检测数与平均正确检测数和计算出的检测数之比。
- 4) MT(↑):主要跟踪目标百分比。
- 5) ML(↓):主要丢失目标百分比。

其中:↑代表值越高跟踪效果越好,↓代表值越低跟踪效果越好。

3.4 实验分析

3.4.1 检测性能实验

本文以YOLOv4作为基线,采用消融实验的方式来验证本文改进点的有效性。为保证训练过程中变量的一致性,网络均采用相同的训练参数,实验结果如表2所示。

表2 不同检测算法性能对比

算法	AP/%	输入尺寸	Paras	FLOPs	FPS
YOLOv4	82.88	416	64.363	60.334	16.5
YOLOv4-SE	84.78	416	65.052	60.979	15.52
YOLOv4-ECA(our)	89.59	416	64.363	60.336	16

由表2可以看出:引入两个注意力模块后都能提升模型的精度,但是ECANet对比SENet拥有更少的参数和计算量,且对模型精度的提升更大。在对输入的特征图进行处理时ECANet使用自适应一维卷积的方式进行信息融合,而不是和SENet一样采用先降维再映射到高维度的操作,从而保证了信息与通道之间的联系。

为了定性分析改进方案的优缺点,在两种不同场景情况下,将本文算法与原始YOLOv4算法进行对比。两种场景情况依次为人群密集情况和行人遮挡情况,实验结果如图7所示。



图7 不同算法检测结果对比图

图7a)展示了在人群密集情况下原始YOLOv4算法与本文算法的检测效果对比。可以发现在人群密集、路灯遮挡、图像模糊的情况下,原始YOLOv4算法出现了漏检,而本文算法没有出现漏检行人的情况,保持了较高的准确度。图7b)展示了在遮挡情况下原始YOLOv4算法与本文算法的检测效果对比。可以发现在左边树木遮挡的情况下原始YOLOv4算法出现了漏检,而本文算法成功地检测到了被遮挡的行人。

3.4.2 跟踪性能实验

为了验证本文提出算法的可行性,采用消融实验的方式验证各个改动对跟踪算法的影响,实验结果见表3。

表3 改进模块对跟踪性能的影响

DeepSort	YOLOv4	ECANet	DIOU	GhostNetV1	MOTA/%	IDF1/%	MT/%	ML/%	ID Switch	FPS
√					48.7	51.4	11.6	50.1	485	12.6
√	√				61.7	66.1	16.4	36.0	430	13.2
√	√	√			63.7	67.3	17.0	34.6	402	12.5
√	√	√	√		66.5	68.5	18.0	33.2	381	13
√	√	√	√	√	67.5	69.6	18.0	32.2	401	16

选取 DeepSort 算法作为 Baseline, 引入 YOLOv4 后网络的跟踪性能大幅度提高, MOTA 提高了 13%, IDF1 提高了 14.7%, MT 提高了 4.8%, ML 降低了 14.1%, ID Switch 降低 55, FPS 提升 0.6, 可见更换更好的检测网络可以大幅度提升跟踪网络的性能; 引入 ECANet 模块之后, 网络的特征提取能力增强, MOTA 提升了 2%, 各指标小幅度提升, FPS 降低 0.7; 将 IOU 替换为 DIOU 后, 网络的匹配性能提高, MOTA 提升 2.8%, IDF1 提高了 1.2%, MT 提高了 1%, ML 降低了 1.4%, ID Switch 降低 21, FPS 提升 0.5; 将 WRN 替换为 GhostNetV1 后, 在保持精度的同时, 降低了重识别网络的参数量, MOTA 提升 1%, ID Switch 增加 20, 但是网络的实时性大幅度提升, FPS 提高了 3。

为了进一步展示主要改进点的效果, 将 DeepSort、YOLOv4+DeepSort 和本文算法的可视化结果进行输出对比, 如图 8、图 9 所示。

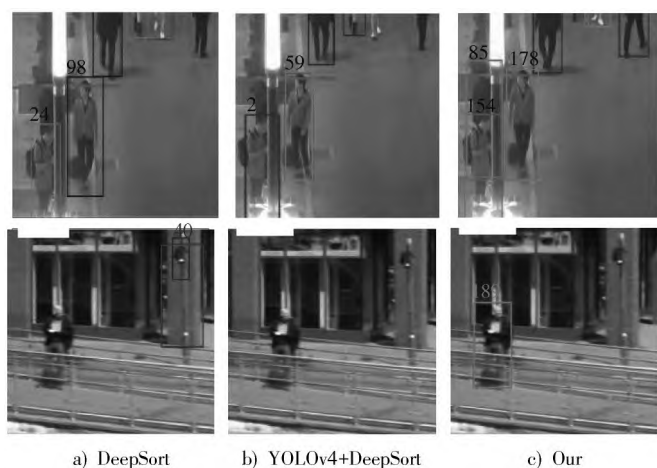


图 8 三种跟踪算法检测和跟踪结果

如图 8 所示, 第一排为遮挡情况下三种跟踪算法的对比结果, 可以看出图中左下角路灯部分, DeepSort 算法和 YOLOv4+DeepSort 算法均没有跟踪到目标, 本文算法能够准确地跟踪到目标。第二排为在相机移动下三种算法的检测结果。可以看出, DeepSort 算法和 YOLOv4+DeepSort 算法均没有跟踪到目标, DeepSort 算法还出现了将红绿灯检测为行人的误检情况, 而本文算法能够准确跟踪到目标。

图 9 为同一场景下, 三种算法的 ID Switch 情况, 从左到右为 80 帧、96 帧、104 帧的跟踪结果。从图 9 可以看出, DeepSort 算法的检测结果显示了跟踪目标 ID 互换的情况 (40 与 53 发生了互换), YOLOv4+DeepSort 算法虽未出现 ID 互换, 但是第 96 帧的跟踪结果出现了误检的情况, 而本文算法的检测结果显示跟踪成功且未出现误检。



图 9 三种跟踪算法 ID Switch 对比结果

4 结 语

遮挡、光线模糊和行人密集等因素容易造成行人多目标跟踪任务出现 ID Switch 现象。本文以 YOLOv4 模型为基础对行人多目标跟踪流程中的目标检测算法进行改进, 通过引入注意力机制, 提升了模型的特征提取能力, 进而提升了行人检测精度。采用 GhostNetV1 重识别网络对行人进行特征提取, 提升了模型的外观特征提取能力, 并在损失函数中引入 DIOU 提高了网络的匹配性能。最后的对比实验及分析表明, 本文提出的跟踪算法在行人密集、遮挡等复杂情况下表现良好, MOTA 达到 67.5%、ID Switch 仅为 401, FPS 可达 16, 相比其他常见的算法具有较好的鲁棒性和精度, 能够更好地应用在行人跟踪领域。

注: 本文通讯作者为邢关生。

参 考 文 献

- [1] 宋欢欢. 基于深度特征共享的行人检测与跟踪研究[D]. 哈尔滨: 哈尔滨工业大学, 2019.
- [2] BIAN C, YANG Z, ZHANG T, et al. Pedestrian tracking from an unmanned aerial vehicle [C]// 2016 IEEE 13th International Conference on Signal Processing. Chengdu: IEEE, 2016: 1067-1071.
- [3] POLANA R, NELSON R. Low level recognition of human motion (or how to get your man without finding his body parts)

- [C]// Proceedings of 1994 IEEE Workshop on Motion of Non-rigid and Articulated Objects. Austin: IEEE, 1994: 77-82.
- [4] LIU W, ANGUELOV D, ERHAN D, et al. Ssd: single shot multibox detector [C]// European Conference on Computer Vision. Cham: Springer, 2016: 21-37.
- [5] BOCHKOVSKIY A, WANG C Y, LIAO H Y M. YOLOv4: Optimal speed and accuracy of object detection [J]. Computer science, 2004, 10934: 2020.
- [6] BEWLEY A, GE Z, OTT L, et al. Simple online and realtime tracking [C]// 2016 IEEE International Conference on Image Processing. San Antonio: IEEE, 2016: 3464-3468.
- [7] WOJKE N, BEWLEY A, PAULUS D. Simple online and realtime tracking with a deep association metric [C]// 2017 IEEE International Conference on Image Processing. Beijing: IEEE, 2017: 3645-3649.
- [8] 陈紫强, 张雅琼. 一种基于YOLOv4的改进DeepSort目标跟踪算法[J]. 桂林电子科技大学学报, 2021, 41(2): 140-145.
- [9] 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.
- [10] HAN K, WANG Y, TIAN Q, et al. Ghostnet: More features from cheap operations [C]// Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. Seattle: IEEE, 2020: 1580-1589.
- [11] ZHENG Z H, WANG P, LIU W, et al. Distance-IoU loss: faster and better learning for bounding box regression [C]// Proceedings of the AAAI Conference on Artificial Intelligence. New York: AAAI, 2020, 34: 12993-13000.
- [12] HE K, ZHANG X, REN S, et al. Spatial pyramid pooling in deep convolutional networks for visual recognition [J]. IEEE transactions on pattern analysis and machine intelligence, 2015, 37(9): 1904-1916.
- [13] HU J, SHEN L, SUN G. Squeeze-and-excitation networks [C]// Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Salt Lake city: IEEE, 2018: 7132-7141.
- [14] WANG Q, WU B, ZHU P, et al. Supplementary material for ECA-Net: Efficient channel attention for deep convolutional neural networks [C]// Proceedings of the 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Seattle, WA, USA: IEEE, 2020: 13-19.

作者简介: 郑繁亭(1998—), 山东滨州人, 硕士研究生, 研究方向为深度学习、计算机视觉。

邢关生(1980—), 河北秦皇岛人, 副教授, 研究方向为智能机器人系统、多机器人协调、机器人视觉。