



基于轻量化 DeepSort 的人数统计

吴 迪, 宋家豪, 冯晓婉

(沈阳师范大学 物理科学与技术学院, 沈阳 110034)

摘 要: 为了检测学校教室或图书馆的人流量,设计了一种优化的 YOLOv5 和轻量化 DeepSort 的人流量统计系统。为提升目标检测效果,采用 CIOU loss 回归损失和 DIOU-NMS 非极大抑制,在加快目标边界框参数学习的同时可提高定位精度,提升遮挡行人的检测性能;在 DeepSort 外观特征提取网络的基础上,结合轻量级网络 ShuffleNet V2,对特征提取模型重新训练,减小模型的参数网络复杂度并保持良好的精确度,提高了系统的可移植能力;在视频中设置虚拟线,利用行人通过虚拟线的时间差进一步计算行人速度。实验结果表明,采用端到端的方式对行人目标进行高效追踪,缩小后的模型体积仅为原模型的 5%,大大改善了对遮挡行人的检测性能,可以较准确地统计出人流量与行速,并提高了鲁棒性。

关 键 词: 多目标追踪; 目标检测; YOLOv5; DeepSort; 深度学习

中图分类号: TP391.41

文献标志码: A

doi: 10.3969/j.issn.1673-5862.2022.04.007

People counting based on lightweight DeepSort

WU Di, SONG Jiahao, FENG Xiaowan

(College of Physical Science and Technology, Shenyang Normal University, Shenyang 110034, China)

Abstract: Aimed at the flow of people in the classroom or library, an optimized YOLOv5 and a lightweight DeepSort flow statistics system are proposed. In order to enhance the target detection effects, the CIOU loss regression loss and DIOU-NMS non-maximum suppression are used to speed up the learning target bounding box parameters, improve the positioning accuracy and detection performance of occluded pedestrians. Based on DeepSort appearance feature extraction network and ShuffleNetV2 lightweight network, the feature extraction model is retrained to reduce the parameter network complexity and maintain good accuracy, and boost the system portability. Virtual lines are set up in the video to detect the flow of people, and the time difference between pedestrians passing the virtual lines is used to further calculate pedestrian speed. In this paper, the end-to-end method is used to effectively monitor the pedestrian target. Experimental results show that the proposed design represents only 5% of the original network volume, which improves the performance of occluded pedestrian detection, accurately counts the pedestrian flow and speed, and enhances the robustness.

Key words: Multi-target tracking; target detection; YOLOv5; DeepSort; deep learning

0 引 言

随着高校的大量扩招,教育资源的缺乏已经严重影响到学生的学习,尤其是在学校教学楼和图书馆等场所,学生排起长队进入教学楼或图书馆的情况已经被多次报道^[1]。由于公共场所人员众多会发生诸如推搡和踩踏等危险事故^[2],因而在校园中建立一种目标检测与跟踪技术就显得尤为重要。

收稿日期: 2021-11-22

基金项目: 国家自然科学基金资助项目(11804235)。

作者简介: 吴 迪(1979—),男,辽宁沈阳人,沈阳师范大学副教授,博士。

传统的跟踪算法是利用手工方式提取目标表观特征,并在连续帧中进行相似特征的搜索,以实现目标追踪,代表算法有粒子滤波、光流法、Mean Shift 算法等^[3-5]。陆云等^[6]提出了一种目标追踪算法,利用 Mean Shift 对图像间相关信息进行处理,通过计算图像间的相关性实现复杂背景下的目标追踪。然而,传统的跟踪算法依然容易受到光照、遮挡等因素的影响,且对图像的特征信息未能全面利用,随着深度学习技术的高速发展,结合卷积神经网络的检测追踪技术^[7-9]逐渐替代了传统方法。基于深度神经网络识别目标的目标检测算法,如 Fast R-CNN, R-CNN, YOLO 等^[10-12],通过跟踪模型对图像特征的处理,可以实现更好的跟踪效果。虽然复杂的网络算法具有较高的精准度,但庞大的参数量对硬件的要求较高,为了将目标检测与跟踪算法应用于小型设备,曹昭睿等^[13]设计了一种将轻量化的目标检测网络与 KCF 算法相结合的目标识别跟踪方法,该算法在识别精确度与计算速度上达到了较好的平衡。陈畅等^[14]将改进的 SSD 网络与 DeepSort^[15]算法相结合,证明了轻量化跟踪网络在行人检测跟踪领域的有效性。

使用轻量级目标检测模型 YOLOv5s,结合 DeepSort 目标跟踪算法可以实现人流量计数统计与行速计算。利用轻量化网络 ShuffleNetV2^[16]优化特征提取网络,在行人重识别数据集上重新训练,可实现端到端的检测追踪与统计;使用 CIoU loss 替换原始的目标框损失函数,可提高定位精度并加快训练时模型的收敛速度;使用 DIoU-NMS 替换原有的加权 NMS,可减少行人遮挡所导致的漏检现象。优化后的跟踪方法采用性能优良的 YOLOv5 作为检测器,将检测结果作为 DeepSort 算法的输入,实现了端到端的目标跟踪,并进行了人流量的统计与行速计算。压缩后的模型依然可以达到较好的追踪效果,满足校园内人流量统计的精度需求,且轻量化的追踪系统更易部署边缘设备。

1 优化 YOLOv5 目标检测算法

YOLOv5 是 YOLO 系列最新一代的目标检测网络,具有 4 个适合不同需求的网络模型,在常用数据集上均可获得更快的检测速度,适合实际应用。在 YOLOv5 训练时,损失函数会影响目标框的选取,优良的损失函数可以使模型更快地达到收敛状态,减少模型预测值的误差。为减少行人接近或遮挡而出现的漏检情况,首先选择 CIoU loss 作为训练时的损失函数:

$$\text{Loss}_{\text{CIoU}} = 1 - \text{IoU} + \frac{c^2(b, b^{\text{gt}})}{c^2} + \alpha v \quad (1)$$

相比于 GIoU loss, CIoU loss 更加关注检测框回归的 3 个重要因素:中心点距离、长宽比、重叠面积,以提高训练收敛的速度和精度。将 YOLOv5 中抑制冗余框的非极大抑制法替换成 DIoU-NMS,可以提高遮挡目标检测框的回归,降低漏检率。

2 DeepSort 与轻量化网络 ShuffleNetV2

DeepSort 在多目标跟踪领域是常用的一种算法,在 Sort 基础上,增加一个可以提取行人表观的特征提取模型,使用标准卡尔曼滤波器进行目标运动状态的预测,并通过匈牙利算法将目标运动状态的预测结果与目标检测结果进行信息匹配。在目标追踪过程中,为改善有遮挡情况下的目标追踪效果,通过将运动信息与表观特征相结合的方式能实现最近邻匹配,可以较大程度上降低 ID-Switch 的频率,如图 1 所示。

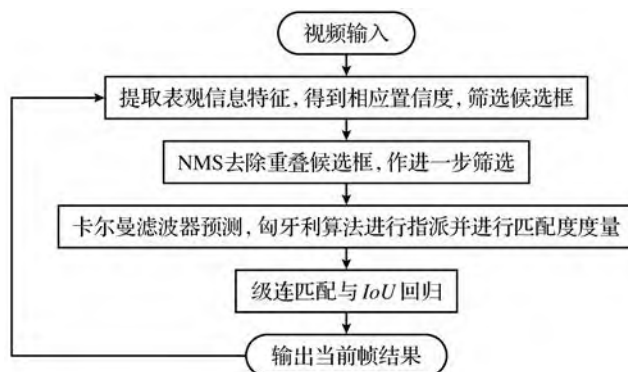


图 1 DeepSort 跟踪流程
Fig. 1 DeepSort tracing process

ShuffleNetV2 是轻量级卷积神经网络,其基于轻量化网络设计的 4 条准则设计,具有较低的网络复杂度和高精度。为降低模型体积,本文结合轻量级卷积神经网络将模型重训练,减少了网络参数,压缩了模型体积,如图 2 所示。

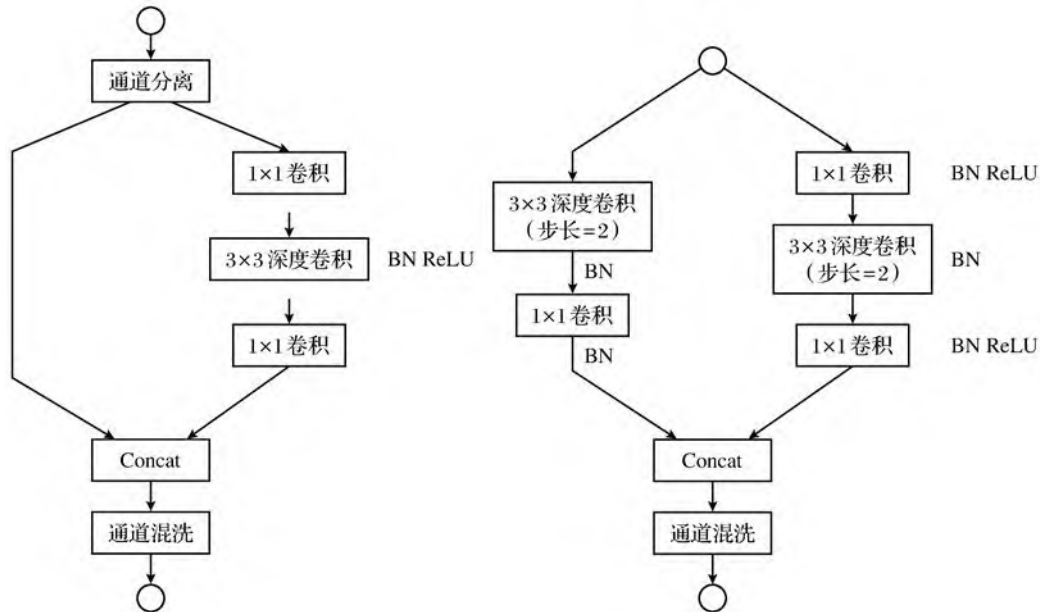


图 2 ShuffleNet V2 结构
Fig. 2 The ShuffleNet V2 structure

2.1 优化 DeepSort

卷积神经网络(convolutional neural networks,CNN),如 ResNet 和 DenseNet,可以有效地提升目标检测的精准度。虽然过于复杂的网络模型可以使追踪的精准度得到提高,但其计算速度较慢,更多地应用在高性能设备上。为了将目标检测与跟踪算法应用在低性能设备上,算法需具有低延迟等特点,即将轻量化模型部署到边缘设备时也可以获得较好的精度与速度。为满足以上需求,本文改进了 DeepSort 的特征训练网络,利用轻量化网络 ShuffleNetV2 将特征提取模型进行重训练。实验结果表明,优化后的轻量化模型体积缩小 18 倍,并依然保持较好的准确度。

基于 ShuffleNetV2 的轻量化卷积网络在行人重识别数据集上将 DeepSort 特征提取模型进行重训练,将已训练的 Shuff-DeepSort 模型用于校园内图书馆的进出人流量统计和人速测量。具体人流量统计系统流程如图 3 所示。

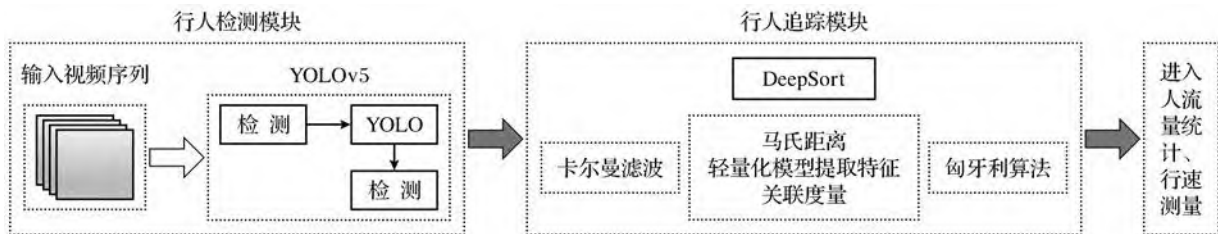


图 3 人流量统计算法流程图
Fig. 3 Flow chart of traffic statistics algorithm

2.2 行速检测

将测量的实际距离映射到图像像素中,并在测试图像中设置虚拟线,通过行人跨越虚拟线的时间差计算出行人的行速,其计算公式为

$$v = \frac{\Delta d}{\Delta t} = \frac{\Delta d}{(t_1 - t_0) * T} = \frac{\Delta d}{(t_1 - t_0) * \left(\frac{1}{30}\right)} \quad (2)$$

其中: Δd 是映射到图像像素中的实际距离,可直接测量获得; t_0 和 t_1 为经过虚拟线时的时间帧,视频为 30 帧/s。

3 实验结果及分析

3.1 实验准备

行人检测数据集:本文采用 CrowdHuman^[17]数据集,在密集人群的情况下,人体存在各种各样的遮挡,每个人体实例都需对人体全身体边界框、头部边界框和人体可见区域边界框进行标记。由于本系统只需对行人目标进行检测,所以需对数据集进行预处理,将标注信息文件转化为 YOLO 训练需要的 TXT 格式,并剔除头部边界框,只保留行人的标注信息,将其按照 9:1:1 比例划分,用于训练、测试和验证。

校园图书馆视频测量帧率为 30 帧/s,帧像素为 $1\,280 \times 720$ 。实验测试采用普通笔记本,处理器为 Intel Core i5,图像处理器为 NVIDIA GTX 960 m。

3.2 行人跟踪及人流量统计

通过判断行人跨线的方式可实现人流量统计,用轻量化特征模型提取出每个行人的表观特征,并比对目标之间的相似度以防止行人反复跨线而重复计数。行人行速采用传统的速度计算公式,利用像素距离和实际距离之间的映射关系,根据行人通过上(下)行线与中间线的时间帧差计算行速。分别采用原 DeepSort 模型与本文 Shuff-DeepSort 模型对行人视频进行处理,截取 2 段视频中的第 1 228 帧和第 1 012 帧,得到的对比效果如图 4 所示。改进后的算法可以在很大程度上检测出遮挡行人,提升了遮挡目标的检测效果,并较为准确地统计出人流量、进出人数和行人行速。



图 4 不同模型的行人追踪效果图

Fig. 4 Same frame pedestrian tracking effect

根据校园图书馆视频人工计算实际的行人数量,不同模型的人流量计数的检测性能对比见表 1。轻量化网络训练的特征提取模型体积很小,并具有良好的准确度,但是在行人出现较大遮挡或外观变化情况下可能会出现一定的漏检现象。

表 1 人流量检测模型性能对比

Table 1 Performance comparison of different human flow detection models

	实际行人数量	检测行人数量	准确度/%	模型体积/MB
DeepSort	42	41	97	43.9
Shuff-DeepSort	42	40	95	2.53

4 实验结果及分析

本文基于 YOLOv5 和 DeepSort 对图书馆进行人流量统计,用 CIoU loss 损失函数与 DIoU-NMS 替换原始 GIoU loss 和加权 NMS,并在遮挡行人数据集上重新训练模型,提高遮挡行人的检测能力;采用 ShuffleNet V2 网络优化 DeepSort 算法,获得一个轻量化的特征提取模型,使用卡尔曼滤波器对目标

轨迹进行预测,并通过匈牙利算法与马氏度量对预测轨迹与目标轨迹进行匹配,连接优化后的 YOLOv5 目标检测端,设计出具有高准确度的轻量化行人检测与追踪系统;在视频序列中设置虚拟线,根据行人的中心点位置判断行人是否通过虚拟线进而计算人流量,进一步利用行人跨越虚拟线的距离与时间差计算行人速度,可提高行人的安全防护。对行人在校园图书馆与道路 2 种场景进行测试,实验结果表明,本文设计的轻量化目标检测与追踪系统可实现行人目标的跟踪计数与行速计算,并降低和减少了追踪中行人遮挡的干扰和特征模型体积,提高了鲁棒性。将该方法应用于图书馆内,能有效地对人流进行统计,也可应用于大型超市等场景。未来将利用模型剪枝进一步减少参数量并提升检测追踪能力,以提高其在边缘设备上的使用性能。

参考文献:

- [1] 吴坤喜. 图书馆监控视频人数统计的研究与开发[D]. 广州: 广东工业大学, 2015.
- [2] 刘德海, 屈丰安, 常成德. 非传统安全视角下我国高校突发事件应急管理研究[J]. 电子科技大学学报(社科版), 2012, 14(1): 53-61.
- [3] 刘鑫, 金恒宏. 四帧间差分与光流法结合的目标检测及追踪[J]. 光电工程, 2018, 45(8): 35-42.
- [4] 朱胜利, 朱善安, 李旭超. 快速运动目标的 Mean Shift 跟踪算法[J]. 光电工程, 2006(5): 66-70.
- [5] 程淑红, 高许, 程树春, 等. 基于计算机视觉的运动车辆检测[J]. 计量学报, 2017, 38(3): 288-291.
- [6] 陆云, 高强, 黄秀清. 复杂背景下运动目标检测与实时跟踪[J]. 软件导刊, 2014, 13(5): 68-70.
- [7] 谷燕飞. 基于改进 YOLOv3+DeepSort 多目标跟踪系统的研究与实现[D]. 沈阳: 辽宁大学, 2020.
- [8] 陈佳倩, 金恒宏, 王文远, 等. 基于 YOLOv3 和 DeepSort 的车流量检测[J]. 计量学报, 2021, 42(6): 718-723.
- [9] 欧阳谷, 钟必能, 白冰, 等. 深度神经网络在目标跟踪算法中的应用与最新研究进展[J]. 小型微型计算机系统, 2018, 39(2): 315-323.
- [10] GIRSHICK R, DONAHUE O J, DARRELL T, et al. Rich feature hierarchies for accurate object detection and semantic segmentation[C]//2014 IEEE Conference on Computer Vision and Pattern Recognition. Columbus: IEEE, 2014: 580-587.
- [11] GIRSHICK R. Fast R-CNN[C]//2015 IEEE International Conference on Computer Vision (ICCV). Santiago: IEEE, 2015: 1440-1448.
- [12] BOCHKOVSKIY A, WANG C Y, LIAO H. YOLOv4: Optimal speed and accuracy of object detection[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Nashville: IEEE, 2020: 10934.
- [13] 曹昭睿, 白帆, 刘凤丽, 等. 基于轻量化神经网络的目标识别跟踪算法研究[J]. 弹箭与制导学报, 2020, 40(1): 19-23.
- [14] 陈畅. 目标检测与跟踪的轻量化网络设计与应用研究[D]. 济南: 山东大学, 2020.
- [15] WOJKE N, BEWLEY A, PAULUS D. Simple online and realtime tracking with a deep association metric[C]//2017 IEEE International Conference on Image Processing (ICIP). Beijing: IEEE, 2017: 3645-3649.
- [16] MA N, ZHANG X, ZHENG H T, et al. Shufflenet V2: Practical guidelines for efficient cnn architecture design[C]//Proceedings of the European Conference on Computer Vision (ECCV). Munich: Springer, 2018: 116-131.
- [17] SHAO S, ZHAO Z, LI B, et al. Crowd human: A benchmark for detecting human in a crowd[EB/OL]. (2018-04-30)[2021-05-07]. <https://arxiv.org/pdf/1805.00123.pdf>.