



西南交通大学学报

Journal of Southwest Jiaotong University

ISSN 0258-2724, CN 51-1277/U

《西南交通大学学报》网络首发论文

题目：顾及小目标特征的视频人流量智能统计方法
作者：朱军，张天奕，谢亚坤，张杰，李闯农，赵犁，李维炼
收稿日期：2020-07-06
网络首发日期：2021-05-14
引用格式：朱军，张天奕，谢亚坤，张杰，李闯农，赵犁，李维炼. 顾及小目标特征的视频人流量智能统计方法. 西南交通大学学报.
<https://kns.cnki.net/kcms/detail/51.1277.U.20210514.0931.006.html>



网络首发：在编辑部工作流程中，稿件从录用到出版要经历录用定稿、排版定稿、整期汇编定稿等阶段。录用定稿指内容已经确定，且通过同行评议、主编终审同意刊用的稿件。排版定稿指录用定稿按照期刊特定版式（包括网络呈现版式）排版后的稿件，可暂不确定出版年、卷、期和页码。整期汇编定稿指出版年、卷、期、页码均已确定的印刷或数字出版的整期汇编稿件。录用定稿网络首发稿件内容必须符合《出版管理条例》和《期刊出版管理规定》的有关规定；学术研究成果具有创新性、科学性和先进性，符合编辑部对刊文的录用要求，不存在学术不端行为及其他侵权行为；稿件内容应基本符合国家有关书刊编辑、出版的技术标准，正确使用和统一规范语言文字、符号、数字、外文字母、法定计量单位及地图标注等。为确保录用定稿网络首发的严肃性，录用定稿一经发布，不得修改论文题目、作者、机构名称和学术内容，只可基于编辑规范进行少量文字的修改。

出版确认：纸质期刊编辑部通过与《中国学术期刊（光盘版）》电子杂志社有限公司签约，在《中国学术期刊（网络版）》出版传播平台上创办与纸质期刊内容一致的网络版，以单篇或整期出版形式，在印刷出版之前刊发论文的录用定稿、排版定稿、整期汇编定稿。因为《中国学术期刊（网络版）》是国家新闻出版广电总局批准的网络连续型出版物（ISSN 2096-4188，CN 11-6037/Z），所以签约期刊的网络版上网络首发论文视为正式出版。

DOI: 10.3969/j.issn.0258-2724.20200425

顾及小目标特征的视频人流量智能统计方法

朱军¹ 张天奕¹ 谢亚坤¹ 张杰¹ 李闯农¹ 赵犁² 李维炼¹

(1.西南交通大学地球科学与环境工程学院, 四川 成都 611756; 2.四川豪格远景市政建设有限公司, 四川 成都 610036)

摘要: 人流量统计对于智能安防等领域具有重要研究价值。针对视频监控系统中人流量统计准确率较低的问题, 提出一种顾及小目标特征的视频人流量智能统计方法, 重点研究用于小目标检测的 faster R-CNN 改进算法、运动目标关联匹配以及双向人流量智能统计等关键技术, 根据人头目标的小尺度特点对 faster R-CNN 网络结构进行适应性改进, 利用图像浅层特征提高网络对于小目标的特征提取能力, 通过基于轨迹预测的跟踪算法实现运动目标的实时追踪, 同时设计双向人流量智能统计算法以实现人流量高准确率统计。为证明方法的有效性, 在密集程度不同的场景中进行测试, 结果表明: 改进的目标检测算法相较于原始算法, 在 Brainwash 测试集和 Pets2009 基准数据集上的平均准确率分别提高了 7.31%、10.71%, 视频人流量智能统计方法在多种场景下的综合评价指标 F 值均能达到 90%以上, 相较于近年来优秀方法: SSD-Sort 算法、Yolov3-DeepSort 算法, 其 F 值提高了 1.14%~3.04%。

关键词: 监控视频; 人流量统计; Faster R-CNN 改进算法; 目标检测; 目标跟踪

中图分类号: TP391.4

文献标识码: A

Intelligent Statistic Method for Video Pedestrian Flow Considering Small Object Features

ZHU Jun¹ ZHANG Tianyi¹ XIE Yakun¹ ZHANG Jie¹ LI Chuangnong¹ ZHAO Li² LI Weilian¹

(1. Faculty of Geosciences and Environmental Engineering, Southwest Jiaotong University, Chengdu 611756, China; 2. Sichuan Haoge Yuanjing Municipal Construction Co., Ltd., Chengdu 610036, China)

Abstract: Pedestrian flow statistics have an important value of research in the fields like intelligent security. In view of the low accuracy in pedestrian flow statistics of video surveillance systems, an intelligent statistic method of video pedestrian flow is proposed with small object features considered. Key technologies are focused in this work such as the improved faster R-CNN algorithm for small object detection, moving object association and matching, and intelligent statistics of bidirectional pedestrian flow. More efforts include adapting the faster R-CNN network structure according to the small-scale characteristics of the head object, improving the feature extraction ability of the network for small objects by using shallow features of images, and realizing real-time tracking of moving objects through the tracking algorithm based on trajectory prediction. Meanwhile, an intelligent statistic algorithm for bidirectional pedestrian flow is developed to achieve accurate statistics of pedestrian flow. To prove the effectiveness of the proposed method, the experiments were conducted in scenes with different levels of density. The results show that compared with the original algorithm, the improved object detection algorithm improves the mean average precision by 7.31% and 10.71% on the Brainwash test set and Pets2009 benchmark data set, respectively. For the intelligent statistic method of video pedestrian flow, the comprehensive evaluation index F value in various scenes can reach above 90%, which is 1.14%-3.04% higher than the excellent methods in recent years.

收稿日期: 2020-07-06

基金项目: 国家自然科学基金(U2034202, 41871289), 四川省科技计划资助(2020JDTD0003)

作者简介: 朱军(1976-), 男, 教授, 博士, 研究方向为三维地理信息系统与计算机视觉, E-mail: zhu.jun@swjtu.edu.cn

Key words: surveillance video; pedestrian flow statistics; improved faster R-CNN algorithm; object detection; object tracking

智能视频监控设备作为安防的重要手段而广泛存在于商场、医院、学校、景点等公共场所。人流量统计是智能视频监控的重要研究内容之一,对于智能安防、智慧旅游、灾后救援、交通规划等领域的研究具有重要价值^[1-3]。

在机器学习技术出现之前,监控视频中的人流量统计这一任务多由人工完成,但这不仅需要耗费大量的人力物力且视频中的大量信息不能得到很好理解。随着计算机视觉和机器学习技术的不断突破,基于浅层学习的人流量统计方法不断涌现^[4-6]。由于此类方法采用的是人工设计和提取特征,特征冗余度较高,且无法智能地抽取有用信息,导致统计精度较低。而构建深层次神经网络则可以通过自动学习的方式有效抽取图像深层次特征,进一步解决以上问题。

随着人工智能时代的到来,基于深度学习的人流量统计方法不断被提出^[7-8]。其与传统方法的最大不同是它能从数据中自动学习得到特征,不需要人工设计和提取特征且无需进行前景分割。Zhang 等人^[9]通过卷积神经网络模型生成密度图的方式实现了跨场景计数,相比于手工设计特征的方式具有高效性和鲁棒性。虽然此类方法不需要检测每个行人,可以直接通过回归或对人群密度图积分的方式得到人数,但也正是因为无

法获得每个行人的具体信息,使得无法进行后续行人轨迹分析等方面的研究。近年来,由于深度学习在目标检测任务中不断取得突破,基于深度学习检测的人流量统计方法也相继被提出^[10-11]。曹诚等人^[12]提出了一种基于卷积神经网络的视频行人分析方法,实现了监控区域中的人流量统计,但当行人之间相互遮挡时会导致漏检。张天琦^[13]基于 SSD(single shot multibox detector)算法和 KCF(kernel correlation filter)算法对人头目标进行检测和跟踪,进而分析目标轨迹实现了双向计数,其通过检测人头目标有效解决了行人之间的遮挡问题,但由于该方法采取了从行人正上方检测头部的策略,使得算法的应用场景在很大程度上受到了限制。尽管上述方法取得了一定效果,但由于视频中行人姿态、尺度各异且存在不同程度的遮挡情况,使得现有方法在不同场景下的人流量统计准确度较低。

为解决上述问题,本文拟开展顾及小目标特征的视频人流量智能统计方法研究。通过研究用于小目标检测的 Faster R-CNN 改进算法、探讨基于轨迹预测的目标跟踪技术、设计双向人流量智能统计算法,并构建实验环境开展实验分析,期望实现视频人流量精确智能统计。

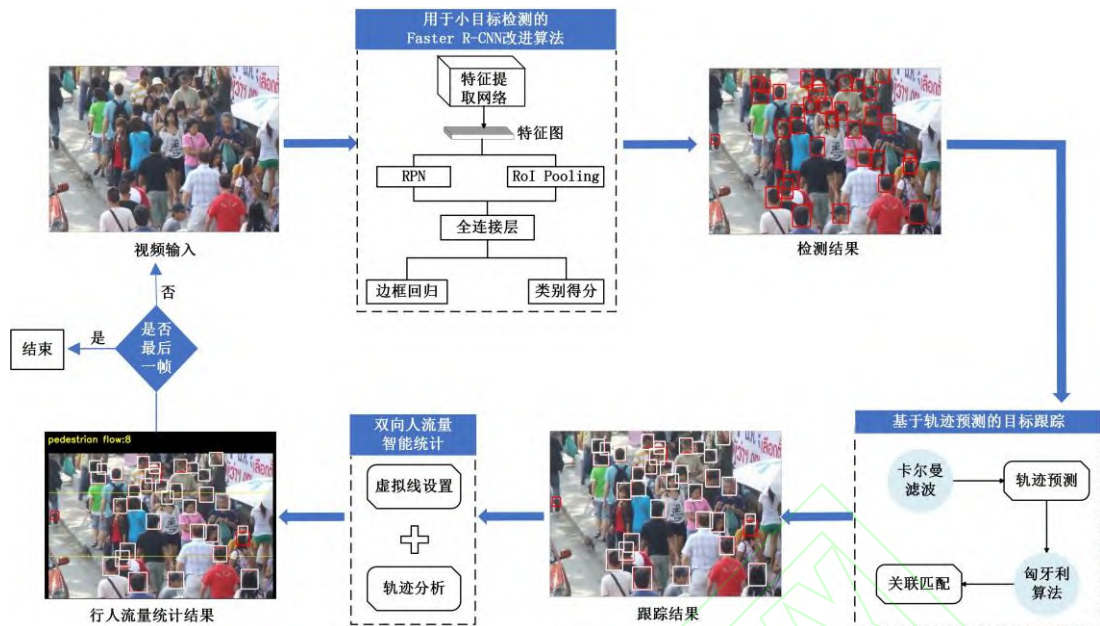


图1 总体框架

Fig.1 Overall framework

1 视频人流量智能统计方法

1.1 总体研究思路

视频人流量智能统计的研究思路如图1所示, 主要包括用于小目标检测的 Faster R-CNN 改进算法、基于轨迹预测的目标跟踪、双向人流量智能统计三个部分。

1.2 用于小目标检测的 Faster R-CNN 改进算法

Faster R-CNN^[14]在通用目标检测领域表现优异, 对于常规目标来说, 其检测精度高且速度较快。但由于本文所选择的检测对象为人头目标, 其在图像上的尺寸较小, 而 Faster R-CNN 对于小目标检测任务精度受限^[15]。因此, 针对人头目标, 本文提出一种用于小目标检测的 Faster R-CNN 改进算法, 算法框架如图2所示。

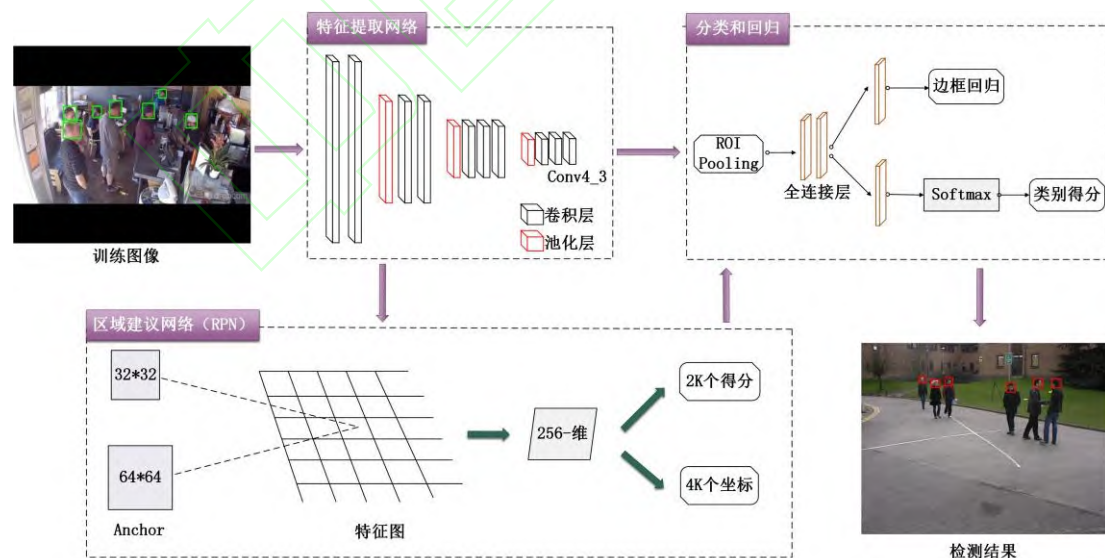


图2 Faster R-CNN 改进算法框架

Fig.2 Framework of improved faster R-CNN algorithm

卷积神经网络不同深度对应不同层次的特征, 深层网络虽然能够较好地表示图像

的语义特征, 但同时也会忽略图像的纹理、边缘等细节特征, 而浅层网络能够很好地表

示这些细节特征^[16]。对于常规目标来说, 使用深层特征图能够提取到较为丰富的语义特征, 进而提高算法的鲁棒性, 但人头目标与常规目标相比, 其信息量少且在图像上占比较小, 对应区域像素反映的信息量有限, 在语义特征较难提取的情况下再使用深层特征图更会使得细节信息缺失。因此, 本文将使用 VGG16^[17]中的 Conv4_3 层代替 Conv5_3 层输入区域建议网络 (Region Proposal Network, RPN) 中生成候选区域 (Regions of Interest, RoI), 既能减少网络复杂度, 又可弥补小目标在深层特征上细节信息的缺失。

Faster R-CNN 中的 RPN 针对需要进行检测的通用目标设置了三种尺度(128, 256, 512)、三种宽高比(1:1, 1:2, 2:1)共 9 种不同大小的锚(Anchor), 但由于人头目标在图像上占比较小, 因此这种设置并不适用于小尺度的人头目标。本文在 RPN 的基础上根据训练数据集中人头目标的宽高比特性, 将 Anchor 的尺度调整为(32, 64), 且由于人头目标多为正方形, 因此只设定一种宽高比例 1:1, 这两种尺度的 Anchor 既能覆盖数据集中大部分的人头目标, 且由于产生的 Anchor 数量与原来相比大幅减少, 导致后续需要进行坐标回归和分类的 Anchor 数量降低。图 3 为改进的 Anchor 与原 RPN 中的 Anchor 在 Brainwash 数据集^[18]中的对比图, 其中红色方框代表改进后的 Anchor, 尺寸分别为 32×32 和 64×64, 蓝色方框代表原 RPN 中 1:1 比例的 Anchor, 尺寸分别为 128×128、256×256 和 512×512。由图 3 可以看出, 原 Anchor(蓝色框)由于尺度过大并不适用于人头目标, 而改进后的 Anchor(红色框)能够对人头目标实现更为准确的覆盖。

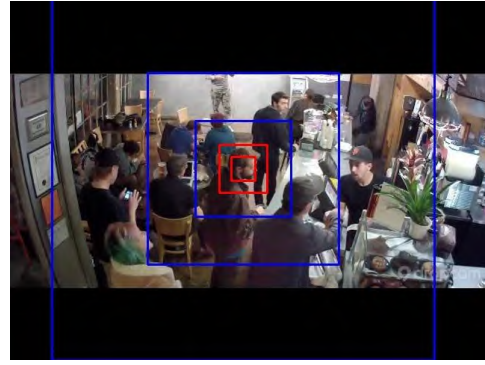


图 3 多尺度下 Anchor 尺寸对比

Fig.3 Contrast of multi-scale anchor size

1.3 基于轨迹预测的目标跟踪

基于轨迹预测的目标跟踪算法^[19]主要包含轨迹预测和数据关联两部分, 如图 4 所示, 首先利用卡尔曼滤波预测出每个目标在下一帧的位置, 进而通过匈牙利算法将实际检测位置与预测位置进行对比关联。详细阐述如下。

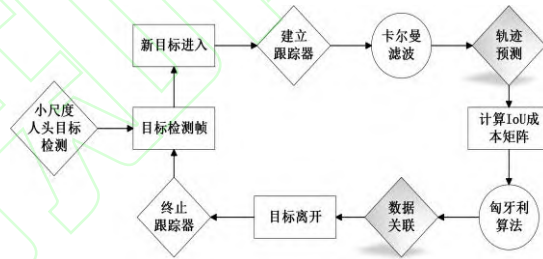


图 4 目标跟踪算法流程

Fig.4 Flow chart of object tracking algorithm

卡尔曼滤波用于预测目标轨迹状态, 其将目标的帧间运动近似为线性运动, 并认为目标间、目标与相机间运动独立, 那么运动目标的状态可以用式(1)表示。

$$\mathbf{x} = [u, v, s, r, m, n, o]^T \quad (1)$$

其中, (u, v) 表示目标在水平和竖直方向上的中心坐标, s 表示目标的面积, r 表示目标的长宽比, (m, n) 表示目标在水平和竖直方向上的运动速度, o 表示目标的变化比率。若检测框关联到目标后, 用检测框更新目标状态, 同时速度分量利用卡尔曼滤波进行优化求解; 若目标没有与检测框相关联, 则用线性速度模型对目标位置进行预测。

匈牙利算法用于前后帧数据关联。首先通过运动模型得到前一帧中每个待跟踪目标的预测框, 其次计算当前帧中每个目标的检测框和所有预测框之间的交并比

(Intersection over Union, IoU)作为成本矩阵, 最后使用匈牙利算法对矩阵进行优化求解。此外应拒绝检测框与目标预测框重叠小于 \min 的分配, 本文实验中 \min 取值为 0.3。

当某个目标的检测框和所有现有目标预测框之间的 IoU 都小于阈值时则认为进入新的待跟踪目标, 此时使用检测框信息初始化新目标的位置信息。若连续 T 帧没有检测框和目标预测框的 IoU 匹配, 则认为目标消失, 本文实验中 T 取值为 1。这种基于轨迹预测的在线跟踪方式能够解决目标前后帧运动过快以至于匹配失败的问题, 将下一

帧预测的轨迹状态与检测目标进行关联, 使得目标跟踪成功的机率大大提高。

1.4 双向人流量智能统计

传统人流量统计方法主要基于单线法实现, 但单线法存在无法灵活设定目标区域、无法准确判断行人运动方向等问题, 为解决上述问题, 本文提出一种双向人流量智能统计方法。相比于传统单线法, 通过双虚拟线对需要监视的目标区域进行准确设定, 将复杂背景干扰因素排除在目标区域之外, 为目标检测和跟踪创造有利条件, 提高人流量统计精度, 算法流程如图 5 所示。

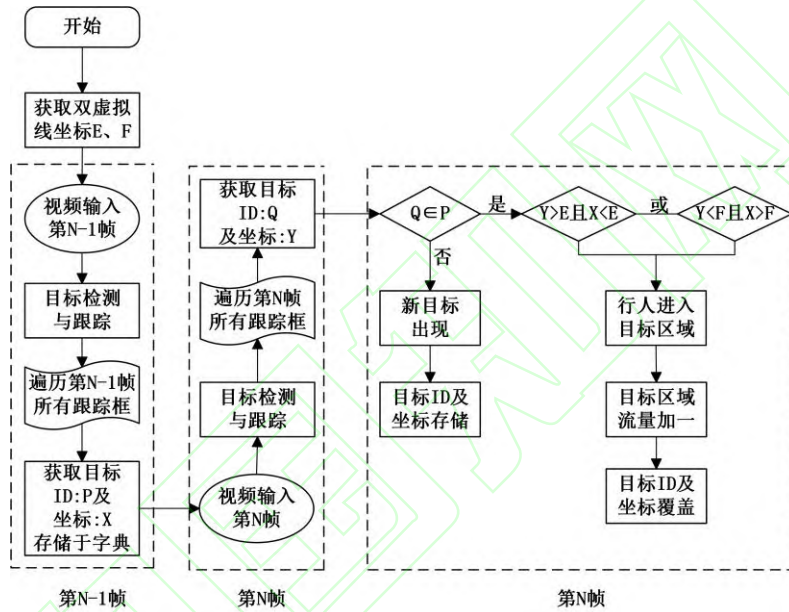


图 5 1.4 双向人流量智能统计算法流程

Fig.5 Flow chart of intelligent statistic algorithm of bidirectional pedestrian flow

首先, 预先设定虚拟线并获取其坐标; 其次, 经过目标检测和跟踪, 第 $N-1$ 帧和第 N 帧的每个目标都被设定了唯一 ID, 将第 $N-1$ 帧的目标以 ID 为键、坐标为值存储于字典中; 再次, 遍历第 N 帧中的所有目标并依次判断其是否在字典内, 如不存在, 则在字典中增加一条记录, 如存在, 则判断目标坐标和字典中相同 ID 的目标坐标与虚拟线坐标之间的大小关系, 具体判断过程视实际场景而定; 最后, 遍历所有视频帧并重复上一步骤, 完成目标区域内的人流量统计。

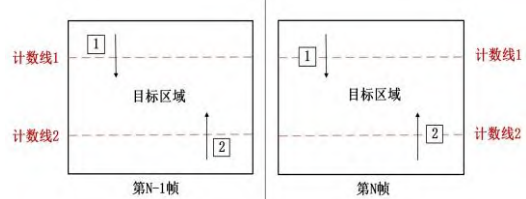


图 6 人流量智能统计过程

Fig.6 Intelligent statistic process of pedestrian flow

在实际视频场景中利用双虚拟线进行人流量统计的过程如图 6 所示, 在视频中预先设置两条虚拟计数线, 其与视频左右边界围成的区域即为要监测的目标区域。如 1 号目标在第 $N-1$ 帧中的坐标小于计数线 1 的坐标, 而在第 N 帧中其坐标大于计数线 1 的坐标, 表明行人此时进入目标区域, 目标区域人流量加一。在实际应用中, 根据场景的具

体情况灵活设置虚拟线的位置即可。本文设计的双向人流量智能统计方法不仅算法复杂度低, 可实现快速计数, 且通过准确判断行人运动方向实现了人流量高准确率统计。

2 实验结果与分析

2.1 实验环境与参数设置

实验所用的硬件配置为: 操作系统为 Windows 10, CPU 为 Inter(R) i5-8300H @2.30GHz, GPU 为 NVIDIA GTX1060, 内存为 8G。模型训练的平台基于深度学习开发工具 TensorFlow, 编程语言为 Python。训练过程中参数设置为: 学习率为 0.001, Batch_size 为 128, 最大迭代次数为 20000。

2.2 数据集

实验训练集采用倾斜向下的摄像机拍摄的 Brainwash 人头数据集, 可有效地处理遮挡问题并且不会受到应用场景的限制, 此数据集涵盖了不同光照、不同遮挡程度的场景, 以及不同尺度、不同颜色头发的行人头部样本, 数据集中用于训练的样本有 10917 幅, 验证样本和测试样本数量各 500 幅。

为检验本文方法在不同场景中的鲁棒性, 本文选取多种场景下的图像及视频进行测试, 并分别设置目标检测实验和人流量统计实验, 以测试本文方法的可靠性和先进性。

目标检测实验选用两种数据集进行测试, 其一为 Brainwash 基准数据集中的 500 幅测试图像, 其二为 Pets2009 基准数据集^[20]中 S2L1 的四种场景下随机选取的 509 幅图像, 如图 7 所示。



(a) 场景一

(b) 场景二



(c) 场景三



(d) 场景四

图 7 目标检测实验测试集所包含场景

Fig.7 Scenes for test set in object detection experiments

人流量统计实验采用了互联网下载的三段密集程度不同的监控视频进行测试, 测试视频具体信息如表 1 所示。

表 1 人流量统计实验视频信息

Tab.1 Experimental video information of pedestrian

flow statistics				
视频序列	时长/s	帧率/fps	人数	密集程度
视频一	222	25	65	稀疏
视频二	35	25	9	稀疏
视频三	46	30	63	密集

2.3 评价标准

本文使用平均准确率(mean Average Precision, mAP)作为目标检测算法的评价标准, 使用召回率(Recall)(式 2)、精确率(Precision)(式 3)、和 F 值(式 4)作为人流量统计算法的评价标准, 其中 Recall 衡量算法对于正样本的覆盖能力, Precision 衡量算法判别正样本的准确程度, F 值则为权衡 Recall 和 Precision 的综合指标。

$$R = \frac{TP}{TP+FN} \quad (2)$$

$$P = \frac{TP}{TP+FP} \quad (3)$$

$$F = \frac{2 \cdot P \cdot R}{P+R} \quad (4)$$

其中, TP 指实际为行人且被正确统计的数量, 即正检人数; FP 指实际为非行人但被错误统计为行人的数量, 即误检人数; FN 指实际为行人但没有被统计的数量, 即漏检人数。

2.4 实验结果分析

2.4.1 目标检测实验分析

由于目标检测的精度直接影响人流量统计结果, 实验首先对目标检测算法进行定量精度评定。将本文算法与原始 Faster R-CNN 算法以及近年来流行的人头目标检测算法在 Brainwash 基准数据集上进行对比测试, 结果如表 2 所示。文献[18]通过提出一种 ReInspect 算法解决了密集场景中目标检测的遮挡问题, 文献[21]提出了一种卷积

层信息融合的方法,通过联合不同尺度的卷积特征信息以提高小目标检测准确率,文献[22]提出了一种基于锚点的完全卷积的FCHD 头部检测算法,文献[23]提出了一种基于改进 Yolov3-tiny 的轻量人头检测算法MKYolov3-tiny,实现了人头目标的快速检测。由对比结果可以看出,本文算法在Brainwash 测试集上的平均准确率达到79.58%,平均准确率不仅高出原始算法7.31%,相比于其余三种模型,整体精度提升了1.58%-37.75%。

表 2 不同目标检测算法对比结果

Tab.2 Results of comparing different object detection

algorithms	
算法	平均准确率/%
文献[18]	78.00
文献[21]	74.20
文献[22]	70.00
文献[23]	41.83
文献[14]	72.27
本文算法	79.58

为测试本文算法在多种不同场景下的有效性,选用涵盖多种场景的Pets2009 数据集进行测试,除原始算法和本文算法外,增加对比算法进行测试,其将Conv3_3 层代替Conv5_3 层输入RPN 中,Anchor 尺寸设置与本文算法保持一致,结果如表3所示。由结果可以看出,本文目标检测算法所训练模型的平均准确率达到74.24%,平均检测每张图像的时间为0.168s,与原始算法相比,整体精度提升了10.71%,而单张图像检测速度相差仅为0.002s。另一方面,本文算法检测精度明显高于对比算法,这说明Conv4_3 层相较于Conv3_3 层针对小目标能够提取到更加有效的特征。总体看来,经过本文方法改进之后的Faster R-CNN 通过极少量的时间代价显著提高了小目标检测精度,有效地平衡了检测精度与速度。

表 3 模型对比结果

Tab.3 Results of comparing models

算法	平均准确率/%	时间/s
原始算法	63.53	0.166
对比算法	67.19	0.184

本文算法	74.24	0.168
------	-------	-------

针对小目标检测问题,一些文献利用多尺度特征融合方式对特征提取网络进行改进,以提高检测精度^[24-25]。为进一步证明本文改进算法的先进性,采用通道叠加方式对多层特征图进行融合实验,结果如表4所示。由结果可以看出,采用多尺度特征融合方式所训练模型的平均准确率最高可达到69.46%,虽然与原始算法相比有所提升,但比本文算法低出4.78%。这是由于行人头部相比于行人全身所占像素小得多,且数据集中每个行人头部在图像上占比较为平均,因此行人头部的多尺度特性表现不明显。另外,越低层特征图虽然边缘细节信息更为丰富,但同时其包含的无关杂乱信息也就越多,因此Conv4_3 相比于多尺度融合而成的特征图能够提取到细节信息更为准确、语义信息更为丰富的特征,从而本文方法相较于多尺度特征融合方法精度更高,充分证明了本文改进方案是行之有效的。

表 4 多尺度特征融合方法精度

Tab.4 Mean average precision of multi-scale

feature fusion methods				
Conv 2_2	Conv 3_3	Conv 4_3	Conv 5_3	平均准确 率/%
	√	√	√	64.92
√		√	√	68.43
√	√	√		69.46

2.4.2 人流量统计实验分析

人流量统计测试结果如图8所示,图中红色框为检测结果,白色框为跟踪结果,左上角黄色文字为人流量实时统计结果,两条黄色虚拟计数线所围成的区域即为目标区域。图8(a)与(b)为普通场景下的监控视频,从图中可以看出本文方法能够实现行人的全部检测和跟踪,并能对人流量进行准确统计,图8(c)为密集场景下的监控视频,从图中可以看出本文方法亦具有较好的效果,此外对于人头密集及存在部分重叠的情况下亦可以实现高精度的检测与跟踪(图8(c)中蓝色框选部分)。由此可以证明本文方法具有较好的精度且具有一定的鲁棒性。

表 5 视频人流量统计结果

Tab.5 Results of video pedestrian flow statistics

视频序列	召回率/%	精确率/%	F 值/%
视频一	90.77	93.65	92.19
视频二	88.89	100.00	94.12
视频三	92.06	93.55	92.80

为定量评价本文视频人流量智能统计方法的性能,通过人工统计出了测试视频中的正检人数、误检人数、漏检人数,并由以上统计量计算出召回率、精确率和 F 值,结果如表 5 所示。由结果可以看出,虽然在密集场景中由于误检人数较多导致在精确率标准上低于普通场景,但整体而言,本文方法在三个视频中的召回率、精确率和 F 值标准上表现较好,皆达到了 90% 以上。

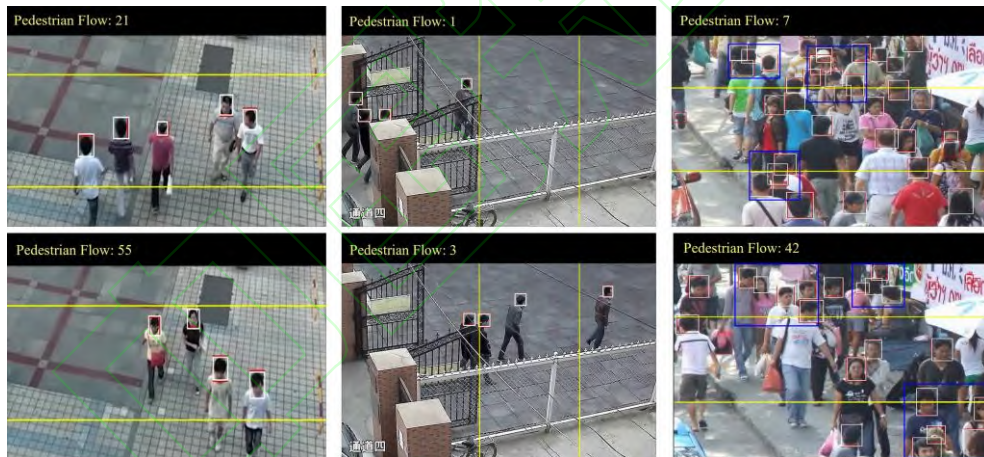
我们统计了近年来表现优异的视频人流量统计方法,其算法框架主要有基于 SSD-Sort 检测跟踪框架^[26]以及基于 Yolov3-DeepSort 检测跟踪框架^[27]等。将以

上两种人流量统计算法框架与本文方法在相同的数据集下进行对比,结果如表 6 所示。由对比结果可以看出,相较于上述两个算法,本文方法在召回率标准上提高了 1.46%-3.65%,这证明本文方法在普通场景和密集场景下的漏检情况有所降低。本文方法在精确率标准上相较于 SSD-Sort 算法提高了 2.38%,与 Yolov3-DeepSort 算法基本持平。总体而言,综合指标 F 值提高了 1.14%-3.04%,充分证明了本文方法在视频人流量统计应用上的鲁棒性和先进性。

表 6 不同算法对比结果

Tab.6 Results of comparing different algorithms for pedestrian flow statistics

算法	召回率 /%	精确率 /%	F 值/%
SSD-Sort	87.59	91.60	89.55
Yolov3-DeepSort	89.78	93.18	91.45
本文方法	91.24	93.98	92.59



(a) 视频一人流量统计结果 (b) 视频二人流量统计结果 (c) 视频三人流量统计结果

图 8 不同场景下视频人流量统计结果

Fig.8 Results of video pedestrian flow statistics in different scenes

3 结语

(1) 提出了用于小目标检测的 Faster R-CNN 改进算法,检测准确率在 Brainwash 数据集和 Pets2009 数据集上分别比原始算法高出 7.31%、10.71%,相较于传统多尺度特征融合方法效果更优。

(2) 设计了双向人流量智能统计算法,人流量统计指标 F 值在多种场景下均达到了 90% 以上,相较于以往优秀算法得到了一

定提升,实现了多场景下视频人流量的精确智能统计。

参考文献:

- [1] 鲁工圆, 马驹, 王坤, 等. 城市轨道交通线路客流控制整数规划模型[J]. 西南交通大学学报, 2017, 52(02): 319-325.
- LU Gongyuan, MA Si, WANG Kun, et al. Integer programming model of passenger flow assignment for congested urban rail lines[J]. Journal of Southwest Jiaotong University, 2017,

- 52(02): 319-325.
- [2] 张君军, 石志广, 李吉成. 人数统计与人群密度估计技术研究现状与趋势[J]. 计算机工程与科学, 2018, 40(02): 282-291.
- ZHANG Junjun, SHI Zhiguang, LI Jicheng. Current researches and future perspectives of crowd counting and crowd density estimation technology[J]. Computer Engineering and Science, 2018, 40(02): 282-291.
- [3] XIE Y, ZHU J, CAO Y, et al. Efficient video fire detection exploiting motion-flicker-based dynamic features and deep static features[J]. IEEE Access, 2020, 8(1): 81904-81917.
- [4] 蔡泽彬. 基于视频分析的行人检测及统计方法研究[D]. 广州: 华南理工大学, 2015.
- [5] 李航, 张涛, 李菲. 一种基于智能视频分析的人流量统计算法[J]. 信息工程大学学报, 2018, 19(03): 373-378.
- LI Hang, ZHANG Tao, LI Fei. Pedestrian volume counting algorithm based on intelligent video analysis[J]. Journal of Information Engineering University, 2018, 19(03): 373-378.
- [6] 徐超, 高梦珠, 查宇峰, 等. 基于 HOG 和 SVM 的公交乘客人流量统计算法[J]. 仪器仪表学报, 2015, 36(2): 446-452.
- XU Chao, GAO Mengzhu, ZHA Yufeng, et al. Bus passenger flow calculation algorithm based on HOG and SVM[J]. Chinese Journal of Scientific Instrument, 2015, 36(2): 446-452.
- [7] 彭山珍, 方志军, 高永彬, 等. 基于多尺度全卷积网络特征融合的人群计数[J]. 武汉大学学报(理学版), 2018, 64(03): 249-254.
- PENG Shanzhen, FANG Zhijun, GAO Yongbin, et al. Crowd counting based on feature fusion of multi-scale fully convolutional networks[J]. Journal of Wuhan University (Natural Science Edition), 2018, 64(03): 249-254.
- [8] CHAN A B, LIANG Z J, VASCONCELOS N. Privacy preserving crowd monitoring: Counting people without people models or tracking[C]//IEEE Conference on Computer Vision and Pattern Recognition. Anchorage: IEEE, 2008: 1-7.
- [9] ZHANG C, LI H, WANG X, et al. Cross-scene crowd counting via deep convolutional neural networks[C]//IEEE Conference on Computer Vision and Pattern Recognition. Boston: IEEE, 2015: 833-841.
- [10] 秦方. 基于计算机视觉的行人检测与人数统计算法研究[D]. 成都: 电子科技大学, 2018.
- [11] HE M, LUO H, HUI B, et al. Pedestrian flow tracking and statistics of monocular camera based on convolutional neural network and kalman filter[J]. Applied Science-Basels, 2019, 9(8): 1-13.
- [12] 曹诚, 卿粼波, 韩龙玫, 等. 城市量化研究中视频人流统计分析[J]. 计算机系统应用, 2018, 27(04): 88-93.
- CAO Cheng, QING Linbo, HAN Longmei, et al. Human traffic analysis based on video for urban quantitative research[J]. Computer Systems and Applications, 2018, 27(04): 88-93.
- [13] 张天琦. 基于深度学习的行人流量统计算法研究[D]. 哈尔滨: 哈尔滨工业大学, 2017.
- [14] REN S, HE K, GIRSHICK R, et al. Faster R-CNN: Towards real-time object detection with region proposal networks[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2017, 39(6): 1137-1149.
- [15] ZHANG L, LIN L, LIANG X, et al. Is Faster R-CNN doing well for pedestrian detection?[C]//European Conference on Computer Vision. Amsterdam: Springer, Cham, 2016: 443-457.
- [16] GATYS L A, ECKER A S, BETHGE M. Image style transfer using convolutional neural networks[C]//IEEE Conference on Computer Vision and Pattern Recognition. Las Vegas: IEEE, 2016: 2414-2423.
- [17] SIMONYAN K, ZISSERMAN A. Very deep convolutional networks for large-scale image recognition[J]. Computer Science, 2014, arXiv: 1409.1556.
- [18] STEWART R, ANDRILUKA M, NG A Y. End-to-end people detection in crowded scenes[C]//IEEE Conference on Computer Vision and Pattern Recognition. Las Vegas: IEEE, 2016: 2325-2333.

- [19] BEWLEY A, ALEX Z, OZZ L, et al. Simple online and realtime tracking[C]//IEEE International Conference on Image Processing. Phoenix: IEEE, 2016: 3464-3468.
- [20] FERRYMAN J, SHAHROKNI A. PETS2009: Dataset and challenge[C]//Twelfth IEEE International Workshop on Performance Evaluation of Tracking and Surveillance. Snowbird: IEEE, 2009: 1-6.
- [21] 吕俊奇, 邱卫根, 张立臣, 等. 多层卷积特征融合的行人检测[J]. 计算机工程与设计, 2018, 39(11): 3481-3485.
- LIU Junqi, QIU Weigen, ZHANG Lichen, et al. Multi-scale convolutional feature fusion for pedestrian detection[J]. Computer Engineering and Design, 2018, 39(11): 3481-3485.
- [22] VORA A, CHILAKA V. FCHD: A fast and accurate head detector[J]. arXiv preprint arXiv: 1809.08766, 2018.
- [23] 高玮军, 师阳, 杨杰, 等. 一种改进的轻量人头检测方法[J/OL]. 计算机工程与应用: 1-9[2020-11-14].<http://kns.cnki.net/kcms/detail/11.2127.tp.20200221.0926.008.html>.
- GAO Weijun, SHI Yang, YANG Jie, et al. An improved lightweight head detection method[J/OL]. Computer Engineering and Applications: 1-9[2020-11-14].<http://kns.cnki.net/kcms/detail/11.2127.tp.20200221.0926.008.html>.
- [24] BELL S, ZITNICK C L, BALA K, et al. Inside-Outside Net: Detecting objects in context with skip pooling and recurrent neural networks[C]//IEEE Conference on Computer Vision and Pattern Recognition. Las Vegas: IEEE, 2016: 2874-2883.
- [25] KONG T, YAO A, CHEN Y, et al. HyperNet: Towards accurate region proposal generation and joint object detection[C]//IEEE Conference on Computer Vision and Pattern Recognition. Las Vegas: IEEE, 2016: 845-853.
- [26] 李继秀, 李啸天, 刘子仪. 基于 SSD 卷积神经网络的公交车下车人数统计[J]. 计算机系统应用, 2019, 28(3): 51-58.
- LI Jixiu, LI Xiaotian, LIU Ziyi. Statistics on number of people getting off bus based on SSD convolutional neural network[J]. Computer Systems & Applications, 2019, 28(3): 51-58.
- [27] 杨亦乐, 高玮玮, 马晓峰, 等. 基于深度学习的行人数量统计方法[J]. 软件, 2019, 40(11): 119-122+151.
- YANG Yile, GAO Weiwei, MA Xiaofeng, et al. Pedestrian statistics based on deep learning[J]. Computer Engineering & Software, 2019, 40(11): 119-122+151.