



JOURNAL OF IMAGE AND GRAPHICS

主办: 中国科学院空天信息创新研究院
中国图象图形学学会
北京应用物理与计算数学研究所

中国图象 图形学报

2021
01
VOL.26

ISSN1006-8961
CN11-3758/TB





第26卷第1期 (总第297期)

2021年1月16日

中国精品科技期刊
中国国际影响力优秀学术期刊
中国科技核心期刊
中文核心期刊

版权声明

凡向《中国图象图形学报》投稿, 均视为同意在本刊网站及CNKI等全文数据库出版, 所刊载论文已获得著作权人的授权。本刊所有图片均为非商业目的使用, 所有内容, 未经许可, 不得转载或以其他方式使用。

Copyright

All rights reserved by Journal of Image and Graphics, Institute of Remote Sensing and Digital Earth, CAS. The content (including but not limited text, photo, etc) published in this journal is for non-commercial use.

主管单位 中国科学院

主办单位 中国科学院空天信息创新研究院

中国图象图形学学会

北京应用物理与计算数学研究所

主 编 吴一戎

编辑出版 《中国图象图形学报》编辑出版委员会

通信地址 北京市海淀区北四环西路19号

邮 编 100190

电子信箱 jig@radi.ac.cn

电 话 010-58887035

网 址 www.cjig.cn

广告发布登记号 京朝工商广登字20170218号

总 发 行 北京报刊发行局

订 购 全国各地邮局

海外发行 中国国际图书贸易集团有限公司

(邮政信箱: 北京399信箱 邮编: 100048)

印刷装订 北京科信印刷有限公司

Journal of Image and Graphics

Title inscription: Song Jian

Monthly, Started in 1996

Superintended by Chinese Academy of Sciences

Sponsored by Aerospace Information Research Institute, CAS

China Society of Image and Graphics

Institute of Applied Physics and Computational Mathematics

Editor-in-Chief Wu Yirong

Editor, Publisher Editorial and Publishing Board of Journal of Image and Graphics

Address No. 19, North 4th Ring Road West, Haidian District, Beijing, P. R. China

Zip code 100190

E-mail jig@radi.ac.cn

Telephone 010-58887035

Website www.cjig.cn

Distributed by Beijing Bureau for Distribution of Newspapers and Journals

Domestic All Local Post Offices in China

Overseas China International Book Trading Corporation

(P.O.Box 399, Beijing 100048, P.R.China)

Printed by Beijing Kexin Printing Co., Ltd.

CN 11-3758/TB

ISSN 1006-8961

CODEN ZTTXFZ

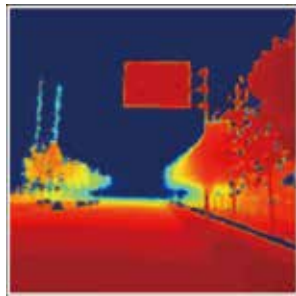
国外发行代号 M1406

国内邮发代号 82-831

国内定价 60.00元



面向智能驾驶的平行视觉感知：基本概念、框架与应用
(第0067页)



Edge-guided GAN: 边界信息引导的深度图像修复
(第0186页)



利用边缘计算的多车协同激光雷达SLAM(第0218页)

序言	王飞跃, 张云泉 I
编者按	II

综述

面向智能驾驶测试的仿真场景构建技术综述	
任秉韬, 邓伟文, 白雪松, 李江坤, 纵瑞雪, 朱冰, 丁娟	0001
自动驾驶软件测试技术研究综述	
冯洋, 夏志龙, 郭安, 陈振宇	0013
强化学习的自动驾驶控制技术研究进展	
潘峰, 鲍泓	0028
自动驾驶地图的数据标准比较研究	
詹骄, 郭迟, 雷婷婷, 屈宜琪, 吴杭彬, 刘经南	0036
视觉感知的端到端自动驾驶运动规划综述	
刘旖菲, 胡学敏, 陈国文, 刘士豪, 陈龙	0049

平行驾驶

面向智能驾驶的平行视觉感知：基本概念、框架与应用	
李轩, 王飞跃	0067
平行视觉的基本框架与关键算法	
张慧, 李轩, 王飞跃	0082

目标检测与跟踪

自动驾驶场景的尺度感知实时行人检测	
徐歆恺, 马岩, 钱旭, 张龔	0093
实时视觉目标跟踪与视频对象分割多任务框架	
李瀚, 刘坤华, 刘嘉杰, 张晓晔	0101
提升预测框定位稳定性的视频目标检测	
郝腾龙, 李熙莹	0113
道路结构特征下的车道线智能检测	
张翔, 唐小林, 黄岩军	0123
适用全速域大曲率路径的自动驾驶跟踪算法	
张龔, 郑颖, 鲍泓	0135
伪3D卷积神经网络与注意力机制结合的疲劳驾驶检测	
庄员, 戚湧	0143
基于眼部自商图—梯度图共生矩阵的疲劳驾驶检测	
潘剑凯, 柳政卿, 王秋成	0154

自动驾驶场景感知与仿真

结合局部平面参数预测的无监督单目图像深度估计	
周大可, 田径, 杨欣	0165
深度纯追随的拟人化无人驾驶转向控制模型	
单云霄, 黄润辉, 何泽, 龚志豪, 景民, 邹雪松	0176
Edge-guided GAN: 边界信息引导的深度图像修复	
刘坤华, 王雪辉, 谢玉婷, 胡坚耀	0186
引入概率分布的深度神经网络贪婪剪枝	
胡骏, 黄启鹏, 刘嘉昕, 刘威, 袁准, 赵宏	0198

高精地图构建与SLAM

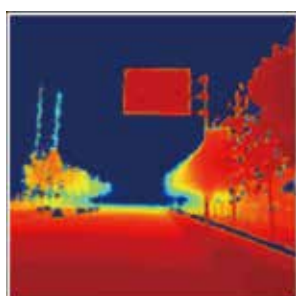
开放道路中匹配高精度地图的在线相机外参标定	
廖文龙, 赵华卿, 严骏驰	0208
利用边缘计算的多车协同激光雷达SLAM	
崔明月, 钟仕鹏, 刘思瑶, 李博洋, 吴成昊, 黄凯	0218

CONTENTS

JOURNAL OF IMAGE AND GRAPHICS



Parallel visual perception for intelligent driving: basic concept, framework and application (P0067)



Edge-guided GAN: a depth image inpainting approach guided by edge information (P0186)



Cooperative LiDAR SLAM for multi-vehicles based on edge computing (P0218)

Review

- Technologies of virtual scenario construction for intelligent driving testing
Ren Bingtao, Deng Weiwen, Bai Xuesong, Li Jiangkun, Zong Ruixue, Zhu Bing, Ding Juan 0001
- Survey of testing techniques of autonomous driving software
Feng Yang, Xia Zhilong, Guo An, Chen Zhenyu 0013
- Research progress of automatic driving control technology based on reinforcement learning
Pan Feng, Bao Hong 0028
- Comparative study on data standards of autonomous driving map
Zhan Jiao, Guo Chi, Lei Tingting, Qu Yiqi, Wu Hangbin, Liu Jingnan 0036
- Review of end-to-end motion planning for autonomous driving with visual perception
Liu Yifei, Hu Xuemin, Chen Guowen, Liu Shihao, Chen Long 0049

Parallel Driving

- Parallel visual perception for intelligent driving: basic concept, framework and application
Li Xuan, Wang Feiyue 0067
- The basic framework and key algorithms of parallel vision
Zhang Hui, Li Xuan, Wang Feiyue 0082

Object Detection and Tracking

- Scale-aware EfficientDet: real-time pedestrian detection algorithm for automated driving
Xu Xinkai, Ma Yan, Qian Xu, Zhang Yan 0093
- Multitask framework for video object tracking and segmentation combined with multi-scale inter-frame information
Li Han, Liu Kunhua, Liu Jiajie, Zhang Xiaoye 0101
- Video object detection method for improving the stability of bounding box
Hao Tenglong, Li Xiying 0113
- Intelligent detection of lane based on road structure characteristics
Zhang Xiang, Tang Xiaolin, Huang Yanjun 0123
- Autonomous vehicle tracking algorithm for high curvature path in full speed range
Zhang Yan, Zheng Ying, Bao Hong 0135
- Driving fatigue detection based on pseudo 3D convolutional neural network and attention mechanisms
Zhuang Yuan, Qi Yong 0143
- Fatigue driving detection based on ocular self-quotient image and gradient image co-occurrence matrix
Pan Jiankai, Liu Zhengqing, Wang Qiucheng 0154

Scene Perception and Simulation of Autonomous Driving

- Unsupervised monocular image depth estimation based on the prediction of local plane parameters
Zhou Dake, Tian Jing, Yang Xin 0165
- Human-like steering model for autonomous driving based on deep pure pursuit method
Shan Yunxiao, Huang Runhui, He Ze, Gong Zhihao, Jing Min, Zou Xuesong 0176
- Edge-guided GAN: a depth image inpainting approach guided by edge information
Liu Kunhua, Wang Xuehui, Xie Yuting, Hu Jianyao 0186
- Greedy pruning of deep neural networks fused with probability distribution
Hu Jun, Huang Qipeng, Liu Jiaxin, Liu Wei, Yuan Huai, Zhao Hong 0198

High-Definition Map Construction and SLAM

- Online extrinsic camera calibration based on high-definition map matching on public roadway
Liao Wenlong, Zhao Huaqing, Yan Junchi 0208
- Cooperative LiDAR SLAM for multi-vehicles based on edge computing
Cui Mingyue, Zhong Shipeng, Liu Siyao, Li Boyang, Wu Chenghao, Huang Kai 0218

中图法分类号: TP391 文献标识码: A 文章编号: 1006-8961(2021)01-0093-08

论文引用格式: Xu X K, Ma Y, Qian X and Zhang Y. 2021. Scale-aware EfficientDet: real-time pedestrian detection algorithm for automated driving. Journal of Image and Graphics, 26(01):0093-0100(徐歆恺, 马岩, 钱旭, 张龔. 2021. 自动驾驶场景的尺度感知实时行人检测. 中国图象图形学报, 26(01):0093-0100)[DOI:10.11834/jig.200445]

自动驾驶场景的尺度感知实时行人检测

徐歆恺^{1,2}, 马岩², 钱旭¹, 张龔¹

1. 中国矿业大学(北京), 北京 100083; 2. 北京联合大学北京市智能机械创新设计服务工程技术研究中心, 北京 100101

摘要: **目的** 行人检测是目标检测中的一个基准问题,在自动驾驶等场景有着较大的实用价值,在路径规划和智能避障方面发挥着重要作用。受限于现实的算法功耗和运行效率,在自动驾驶场景下行人检测存在检测速度不佳、遮挡行人检测精度不足和小尺度行人漏检率高等问题,在保证实时性的前提下设计一种适合行人检测的算法,是一项挑战性的工作。**方法** 本文旨在解决自动驾驶场景中耗时长、行人遮挡和小尺度行人检测结果精度低的问题,提出了一种尺度注意力并行检测算法(scale-aware and efficient object detection, Scale-aware EfficientDet):在特征提取与检测中使用了 EfficientDet 的主干网络,保证算法效率和功耗的平衡;在行人遮挡方面,为了提高模型对遮挡现象的检测精度,引入了可以增强行人与其他物体之间特征差异的损失函数;在提高小目标行人检测精度方面,采用 scale-aware 双路网络算法来增加对小目标行人的检测精度。**结果** 本文选择 Caltech 行人数据集作为对比数据集,选取 YOLO(you only look once)、YOLOv3、SA-FastRCNN(scale-aware fast region-based convolutional neural network)等算法进行对比,在运行效率方面,本文算法在连续输入单帧图像的情况下达到了 35 帧/s,多图像输入时达到了 70 帧/s 的工作效率;在模型精度测试中,本文算法也略胜一筹。本文算法应用于 2020 年中国智能汽车大赛中,在安全避障环节皆获得满分。**结论** 本文设计的尺度感知的行人检测算法,在 EfficientDet 高性能检测器的基础上,通过结合损失函数、scale-aware 双路子网络的改进,进一步提升了本文检测器的鲁棒性。

关键词: 自动驾驶;行人检测;目标检测;EfficientDet;卷积神经网络

Scale-aware EfficientDet: real-time pedestrian detection algorithm for automated driving

Xu Xinkai^{1,2}, Ma Yan², Qian Xu¹, Zhang Yan¹

1. China University of Mining and Technology (Beijing), Beijing 100083, China; 2. Beijing Engineering Research Center of Smart Mechanical Innovation Design Service, Beijing Union University, Beijing 100101, China

Abstract: Objective Pedestrian detection is a crucial safety factor in autonomous driving scenarios. Consistent pedestrian detection results play a particular role in path planning and pedestrian collision avoidance. In recent years, pedestrian detection algorithms have become a research hotspot in the field of autonomous driving. For the pedestrian detection task, several problems need to be solved. 1) Pedestrian occlusion in traffic scenes. Pedestrian occlusion is a challenging driving safety problem in autonomous driving scenarios. Pedestrians who are obscured by other objects (such as buildings, vehi-

收稿日期:2020-08-03;修回日期:2020-10-23;预印本日期:2020-10-30

基金项目:国家自然科学基金重点项目(61932012);北京联合大学科研项目(ZK80202003)

Supported by: State Key Program of National Natural Science Foundation of China(61932012); Academic Research Projects of Beijing Union University(ZK80202003)

cles, and other pedestrians) are difficult to detect. 2) Small pedestrian detection accuracy needs to be improved. In an autonomous driving environment, the accuracy of pedestrian detection plays a crucial role in vehicle control systems based on vision algorithms. When the vehicle speed is fast, the pedestrians at a long distance need to be detected accurately. With the need for low algorithm power consumption and good operating efficiency, designing an algorithm suitable for pedestrian detection to maintain excellent detection performance under the premise of achieving real-time performance is a difficult problem. **Method** This paper proposed a real-time pedestrian detection algorithm called scale-aware and efficient object detection (Scale-aware EfficientDet) based on EfficientDet, which achieves state-of-the-art performance in object detection. Our approach aimed to solve the problems of high time consumption, pedestrian occlusion, and low accuracy of small pedestrian detection results in autonomous driving scenarios. Most of the computing power and running time of the existing object detection algorithms are consumed in the visual feature extraction stage, so the use of a lightweight feature extraction network is a crucial factor in improving the efficiency of the algorithm. Our method uses EfficientDet in feature extraction to ensure the algorithm's computational efficiency and power consumption balance. Our approach aimed to observe occluded pedestrians precisely. The loss function was introduced to improve the model's detection accuracy of occlusion phenomena. The function can enhance the feature difference between pedestrians and other objects, and reduce the feature difference between occlude pedestrians and normal pedestrians. In terms of improving the accuracy of small target pedestrian detection, we use the scale-aware mechanism to enhance the algorithm's detection accuracy for small target pedestrians. **Result** The Caltech pedestrian dataset was used for model comparison. You only look once (YOLO), YOLOv3 scale-aware fast region-based convolutional neural network (fast R-CNN), and other algorithms are selected for comparison. In terms of operating efficiency, our algorithm achieves 35 frame/s with continuous input of a single frame image and a working efficiency of 70 frame/s with multi-image input. In the test of model accuracy, our algorithm is more accurate than YOLOv3, SA-FastRCNN (scale-aware fast region-based convolutional neural network), EfficientDet, and other algorithms. In the preliminaries and finals of the China Intelligent Vehicle Championship (CIVC) 2020, the safety and obstacle avoidance links all received full marks. **Conclusion** To address the problems of detection speed in pedestrian detection in autonomous driving, this paper designs the scale-aware EfficientDet real-time pedestrian detector, which is based on the efficient and high-precision EfficientDet. Our method solved the insufficient detection accuracy for occluded pedestrians and the high missed detection rate of small-scale pedestrians. In accordance with the occlusion characteristics of pedestrians, the loss function with repulsive force is used to solve the problem of pedestrian occlusion. Considering the significant differences in visual appearance and extracted feature maps between small-scale and large-scale pedestrians, scale-aware networks are used separately to minimize the missed detection rate of small-scale pedestrians. The improvements in these two aspects further improve the robustness of the designed detector. In future work, our methods can be adjusted to improve detection performance, find optimization methods, and improve neural networks. The detection performance and detection accuracy can be further improved to promote its better application in the field of autonomous driving.

Key words: automated driving; pedestrian detection; object detection; EfficientDet; convolutional neural network (CNN)

0 引言

目标检测是计算机视觉中的一个关键问题,旨在检测视频或图像中目标的类别和边界框位置。作为目标检测的基准问题,行人检测在诸如自动驾驶、智能交通、客流统计和监控等场景有着很大的实用价值。虽然行人检测已经取得了很大进展,但在实际场景中,由于姿态多样性、背景复杂、遮挡等原因,行人检测仍然存在较大的性能提升空间,是目前极

具挑战性的问题之一。

本文改进了 Tan 等人(2020)提出的 EfficientDet 算法以检测行人,特别是有遮挡行人,借鉴了 Li 等人(2018)提出的 SA-FastRCNN (scale-aware fast region-based convolutional neural network) 的大小尺度行人检测子网络,并引入 Wang 等人(2018)提出的损失函数 RepLoss (repulsion loss),从而能够灵活应对普通行人以及“鬼探头现象”。以中国智能汽车大赛 (China Intelligent Vehicle Championship, CIVC) 官方提供的驾驶仿真平台为验证平台,本文

算法在2020年的CIVC智能驾驶仿真赛的行人避障测试场景的安全避障环节取得了满分成绩。

1 相关工作

行人检测是计算机视觉中目标检测任务的重要分支,涉及识别特定类别的行人,通常是在城市道路上的行人。目前已经出现了多种不同的行人检测方法,基本都包含候选区域提取、特征提取和区域分类3个阶段,如图1所示。传统的方法如Dalal和Triggs(2005)提出的方向梯度直方图(histogram of oriented gradient, HOG)和Nam等人(2014)提出的去局部相关性的通道特征(local decorrelated channel features, LDCF),Broggi等人(2003)、Dalal等人(2006)、Gavrila和Munder(2007)的方法中通过人工提取行人特征,计算效率较高但精度较低。随着机器学习(特别是深度学习)的探索和发展,卷积神经网络(convolutional neural network, CNN)也广泛应用于行人检测,如Sermanet等人(2013)、Tian等人(2015)和Yang等人(2016)的研究,精度有了显著提升,逐渐成为主流。



图1 行人检测的通用步骤

Fig. 1 Common pipeline for pedestrian detection

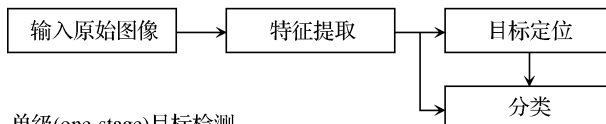
与通用的目标检测算法类似,基于深度卷积网络的行人检测算法,根据其设计模式可大致分为两级(two-stage)检测算法与单级(one-stage)检测算法。

两级检测算法由候选区域提取模块和卷积神经网络(CNN)检测模块组成,首先通过滑窗选取输入图像中所有的候选区域,然后对每个候选区域进行分类和回归,以获得更丰富的特征和更高的准确率。虽然Liu等人(2017)、Li等人(2018)和Zhou等人(2019)等诸多研究者在两级架构上设计出的行人检测器在检测精度上取得了相当大的进步,但由于候选区域冗余检测框过多,导致其检测速度较慢。

为了提高运行效率,出现了以Redmon等人(2016)设计的YOLO(you only look once)系列和Liu等人(2016)设计的SSD(single shot multibox detec-

tor)为代表的单级(one-stage)检测器,不需要预生成候选框,将目标的识别行为和定位行为合二为一,如图2所示,简化了检测流程,从而大大提高了检测速度。单级检测器结构简单、推理速度快,在算法的应用部署阶段可以达到近乎实时的推理速度。随着单级检测算法不断改进,速度更快,精度亦逐渐有赶超两级检测算法的趋势,越来越多的研究者如Milton(2019)、Ahmed等人(2019)、Zhuang等人(2020)开始关注使用单级算法来检测行人。

两级(two-stage)目标检测



单级(one-stage)目标检测

图2 目标检测基本结构

Fig. 2 Architecture of object detection

2 本文行人检测算法

在自动驾驶场景中车速较快(≥ 100 km/h),行人检测速度偏慢的两级检测算法,可能导致检测不及时而相撞的情况发生,本文不做讨论。新一代的高性能目标检测网络除了YOLO系列外,谷歌大脑实验室提出的新型目标检测器EfficientDet也十分优秀,结构简单、模型扩展效率高、性能和发展潜力大。该模型采用Tan和Le(2019)提出的分类模型EfficientNet作为骨架网络,并沿用其复合尺度(compound scaling)方法,在目标检测任务中达到精度和速度的良好平衡。在模型设计上,与YOLO等主流单级检测模型相似,EfficientDet仍采用“特征提取+多尺度特征融合+box/class预测”的架构。

EfficientDet系列模型包含D0~D7共8个模型,数值越大,则所有主干网络、特征网络、边界框/类别预测网络的分辨率/深度/宽度的特征扩大比例越高。但其精确度越高,计算量也会相应增大,因此计算速度也随之由快到慢变化。考虑到真实自动驾驶环境中软硬件的性价比,本文没有选择精度最优的模型,而是选择了EfficientDet-D1进行目标检测(最终检测结果如图3),其处理速度为35帧/s。

在行人检测中,遮挡始终是影响行人检测器性



图3 行人检测效果示例

Fig. 3 Sample of pedestrian detection effect

能的挑战性问题。为改善城市道路场景中的遮挡鲁棒性,减少遮挡影响,本文提出了 Scale-aware EfficientDet(scale-aware and efficient object detection)检测器(以下简称 SAE),以 EfficientDet 为基线检测器,同时根据行人的遮挡特性,使用具有排斥力的损失函数 RepLoss 以解决行人遮挡问题。根据行人识别的任务特性,RepLoss 的组成包括 L_{Attr} 、 L_{RepGT} 和 L_{RepBox} 共 3 个部分,可表示为

$$L = L_{Attr} + \alpha L_{RepGT} + \beta L_{RepBox} \quad (1)$$

式中, α 和 β 是系数,充当权重以平衡辅助损失。 L_{Attr} 是行人检测特征的主要梯度函数,它使得预测框和匹配上的目标框尽可能接近,可采用欧氏距离、 $L1/L2$ 距离和 $Smooth_{L1}$ 距离等计算检测结果中的回归损失。本文采用 $Smooth_{L1}$ 距离,即

$$L_{Attr} = \frac{\sum_{P \in P_+} Smooth_{L1}(P, G_{Attr}^P)}{|P_+|} \quad (2)$$

式中, P 代表算法推理一次的所有推理结果,而 P_+ 表示其中与人工标记相关的正样本集合。

损失函数中的 L_{RepGT} 是为了使预测结果 P (prediction) 与预测结果周边的 G (ground-truth) 尽可能远离。其中,与 P 相关的 G 数学表示为

$$G_{Rep}^P = \underset{G \in G_{Attr}^P}{\operatorname{argmax}} IoU(G, P) \quad (3)$$

式中, G_{Attr}^P 表示检测结果周边所有的 ground truth,即人工标注的结果框。为了使 P 与 G 之间的差异更有利于检测, L_{RepGT} 基于 IoU(intersection over union)使用 IoG(intersection over ground-truth)测量 P 与 G 之间的距离,即

$$IoG(P, G) = \frac{\operatorname{area}(P \cap G)}{\operatorname{area}(G)} \quad (4)$$

L_{RepGT} 根据 P 和与 P 相关的 G 之间的预测关系计算,即

$$L_{RepGT} = \frac{\sum_{P \in P_+} Smooth_{ln}(\log(P, G_{Attr}^P))}{|P_+|} \quad (5)$$

式中, $Smooth_{ln}$ 使用超参数 σ 进行调节,这是一个调节 L_{RepGT} 对 G 敏感程度的参数

$$Smooth_{ln} = \begin{cases} -\ln(1-x) & x \leq \sigma \\ \frac{x-\sigma}{1-\sigma} - \ln(1-\sigma) & x > \sigma \end{cases} \quad (6)$$

在行人检测中,各个检测结果在最终结果非极大抑制(non-maximum suppression, NMS)中会对部分遮挡目标带来一定的精度损失。损失函数中的 L_{RepBox} 是拉大某个预测框 P_i 与其周围预测框 P_j 在图像中相对距离的关键梯度。其中 P_i 与 P_j 的配对使用 IoU 进行

$$L_{RepBox} = \frac{\sum_{i \neq j} Smooth_{ln}(IoU(P_i, P_j))}{\sum_{i \neq j} 1[IoU(P_i, P_j) > 0] + e} \quad (7)$$

根据遮挡与未遮挡行人在图像中表现的不同,其检测损失的类间差与类内差各不相同。本文将遮挡与未遮挡行人的类内差缩小,而遮挡行人与背景、车辆的类间差应扩大。

此外,在自动驾驶的典型场景中,经常会出现较小的行人目标,当行驶车速较快时,小目标行人的漏检可能会对车辆的驾驶安全产生较大隐患。据 Dollar 等人(2012)统计,在常用行人检测数据集中,有 15% 的行人目标是小于 30 像素的。考虑 Zhang 等人(2016)提到小尺度与大尺度行人在视觉外观和提取的特征图上存在显著差异,受 SA-FastRCNN 启发,本文针对这两种尺度的行人使用尺度感知(scale-aware)双路子网络分别进行检测,使小尺度行人的漏检率降到最低。

尺度感知机制在推理末端将进行感知融合,其中 w_1 是大尺度子网络的尺度加权因子, w_s 是小尺度子网络的尺度加权因子,它们进行分支网络的推理结果加权。 w_1 计算为

$$w_1 = \frac{1}{1 + e_1 \exp\left(-\frac{h - \bar{h}}{e_2}\right)} \quad (8)$$

式中, h 是输入图像中目标行人的高度, $\bar{h}$ 是数据集行人目标的平均高度, e_1 与 e_2 是用来调节网络权重的参数,且这两个参数可通过反向传播自动学习。

$$\frac{\partial L}{\partial e_1} = - \frac{\exp\left(-\frac{h-\bar{h}}{e_2}\right)}{\left(1 + e_1 \exp\left(-\frac{h-\bar{h}}{e_2}\right)\right)^2} \frac{\partial L}{\partial w_l}$$

$$\frac{\partial L}{\partial e_2} = - \frac{e_1(h-\bar{h}) \exp\left(-\frac{h-\bar{h}}{e_2}\right)}{e_2^2 \left(1 + e_1 \exp\left(-\frac{h-\bar{h}}{e_2}\right)\right)^2} \frac{\partial L}{\partial w_l} \quad (9)$$

对于小尺度子网络的尺度因子 w_s , 计算为

$$w_s = 1 - w_l \quad (10)$$

3 实验及结果分析

3.1 实验设置

本实验硬件环境的处理器为 GPU@1531 MHz Intel(R) Core(TM) i7-5930K CPU @ 3.50 GHz, 内存为 64 GB, 显卡为 GTX NVIDIA TITAN V 12 GB × 2, Ubuntu16.04 LTS 操作系统。

行人检测的数据集众多, 比较重要的有 INRIA 行人数据集 (Dalal 和 Triggs, 2005)、美国加州理工学院行人数据集 Caltech (Dollar 等, 2009)、多目标跟踪数据集 MOT17Det (Milan 等, 2016) 和 CityPersons 行人数据集 (Zhang 等, 2017), 其中 Caltech 是专门为检测自动驾驶背景下的行人而收集的。本文实验基于 Caltech 行人数据集进行。

在训练过程中, 实验使用了数据增强方法的镜像翻转, 其中的翻转几率为 0.5, 尺度感知结构如图 4 所示。

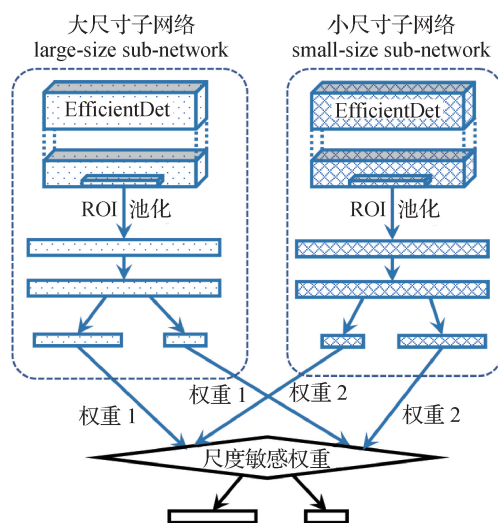


图 4 尺度感知双路子网络

Fig.4 Dual-path sub-network of scale-aware

本实验的网络结构搭建与代码实现基于 Pytorch 深度学习算法平台。算法的数据集在算法的训练过程中, 使用了随机梯度下降优化算法 (stochastic gradient descent, SGD), 动量 0.9, 权重衰减 0.0005。初始的学习率为 0.001, 当训练经过 4 个 epoch 后, 学习率衰减 0.1, 训练经过 9 个 epoch 结束。

每个批次的输入变量为 80 幅随机抽取的图像, 其中 20 幅图像具有特定规格。这 20 幅图像中, 每幅图像的正样本 (行人目标) IoU 值大于 0.5, 剩余的 60 幅图像中, 每幅图像均有可检测的目标。网络中的 e_1 与 e_2 可通过反向传播自动学习。

3.2 实验结果和分析

在算法推理精度上, 本文所采用的 EfficientDet 算法在 MS COCO 数据集上精度表现较好, 如表 1 所示。但其在小、中型目标的检测精度并不理想。由于采用模型符合的层级特征融合, EfficientDet 的 D0—D7 网络随着精度的提升, 其模型的参数相应增加, 运行速度也相应降低。为达到实时性检测的目的, 本文以 EfficientDet D1 网络作为基础特征提取网络, 其运行速度达到了 35 帧/s (如表 2 所示)。

表 1 EfficientDet 模型精度比较
(基于 MS COCO 2017 数据集)

Table 1 Comparison of EfficientDet model accuracy
(MS COCO 2017 datasets)

模型	AP ^{test}	AP ₅₀	AP ₇₅	AP _s	AP _m	AP _L	AP ^{val}	参数/ M	FLOPs/ B
D0	33.8	52.2	35.8	12.0	38.3	51.2	33.5	3.9	2.54
D1	39.6	58.6	42.3	17.9	44.3	56.0	39.1	6.6	6.10
D2	43.0	62.3	46.2	22.5	47.0	58.4	42.5	8.1	11.0
D3	45.8	65.0	49.3	26.6	49.4	59.8	45.9	12.0	24.9
D4	49.4	69.0	53.4	30.3	53.2	63.2	49.0	20.7	55.2
D5	50.7	70.2	54.7	33.2	53.9	63.2	50.5	33.7	135.4
D6	51.7	71.2	56.0	34.1	55.2	64.1	51.3	51.9	225.6
D7	53.7	72.4	58.4	35.8	57.0	66.3	53.4	51.9	324.8

平均正确率 (average precision, AP) 是目标检测算法精度检测的重要指标。表 1 中 AP^{test} 表示算法在 MS COCO 2017 数据集中测试集的检测分数; AP^{val} 表示算法在该数据集验证集中的检测分数; AP₅₀ 表示当交并比取 50% 的算法检测分数; AP₇₅ 表

表 2 EfficientDet 模型速度比较
Table 2 Comparison of EfficientDet model speeds

模型	AP	batch1 latency/ms	batch1 throughput/(帧/s)	batch8 throughput/(帧/s)
D0	33.8	20.4	48	104
D1	39.6	27.2	35	70
D2	43.9	35.9	27	47
D3	45.8	59.5	12	28
D4	49.4	86.4	10	12
D5	50.7	149.8	6	8
D6	51.7	190.5	5	—
D7	53.7	251.7	3.9	—

注:“—”表示“<1 帧/s”,达不到实时性要求。

示当 IoU 取 75% 的算法检测分数; AP_s 表示算法在该数据集中小目标的检测分数; AP_m 表示算法在该数据集中中等目标的检测分数; AP_l 表示算法在该数据集中大目标的检测分数; FLOPs (floating point operations), 指浮点运算数, 常用来衡量算法/模型的复杂度。

表 2 中 batch1 latency 表示当模型的输入图像数量为 1 时, 模型推理一次所使用的时间; batch1 throughput 表示当模型的输入图像数量为 1 时, 模型推理一次时的帧率; batch8 throughput 表示当模型的输入图像数量为 8 时, 模型推理一次的帧率。

本文所提出的行人检测算法采用由 Dollar 等人 (2012) 提出的 Caltech pedestrian detection 行人检测评估指标, 即根据未命中率 (miss rate, MR) 和每幅图像的误报数 (false positives per image, FPPI) 之间的折衷来评估算法的性能。在算法的检测性能上, 本文算法与基准算法 (EfficientDet D1) 相比, 行人检测精度提升明显, 如图 5 所示。

3.3 CIVC 仿真驾驶平台与比赛

作为中国的顶级科技赛事, 中国智能汽车大赛 (CIVC) 拥有国内最权威的智能汽车测评平台, 其指定的仿真驾驶平台 PanoSim (郭姣, 2017) 支持汽车智能驾驶的实时仿真, 集交通模型、车载环境传感模型和汽车行驶环境模型于一体。

CIVC 2020 行人避撞场景涉及行人在空旷场景横穿 (如图 6 所示) 和从视觉盲区横穿 (如图 7 所示) 两项测试内容, 能否及时发现行人和及时刹车, 直接影响该项比赛环节的成绩。

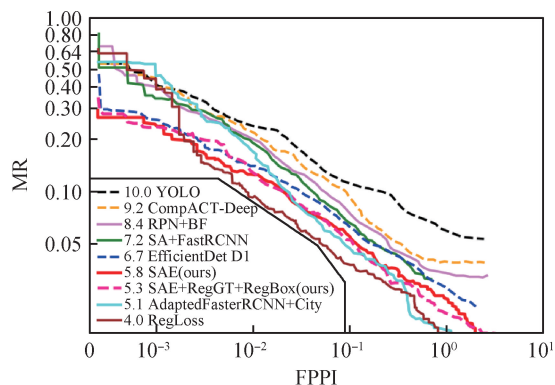


图 5 新 Caltech 测试集上不同方法的比较

Fig. 5 Comparisons with state-of-the-art methods on the new Caltech test subset

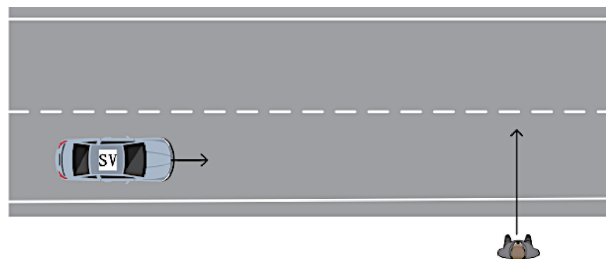


图 6 行人在空旷场景横穿示意

Fig. 6 A pedestrian crossing in an open scene

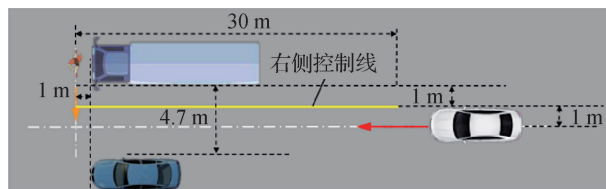


图 7 行人从视觉盲区横穿示意

Fig. 7 A pedestrian crossing from visual blind spot

在 CIVC 2020 的初赛和决赛中, 应用了本文算法后, 行人避撞测试环节的基础分和加分项皆为满分, 团队获得大赛赛道一等奖。

4 结 论

针对自动驾驶领域下行人检测中检测速度不佳、遮挡行人检测精度不足、小尺度行人漏检率高等问题, 本文设计了一种 Scale-aware EfficientDet 实时行人检测器, 以高效、高精度的 EfficientDet 检测器为基准检测器, 根据行人的遮挡特性, 使用具有排斥力的损失函数 RepLoss 解决行人遮挡问题; 并考虑到小尺度与大尺度行人在视觉外观和提取的特征图

上的显著差异,使用尺度感知双路子网络分别进行检测,使小尺度行人的漏检率尽可能降低。这两个方面的改进进一步提高了本文检测器的鲁棒性。

在 Caltech 行人数据集上进行实验,验证了本文方法对自动驾驶领域下行人检测具有一定前景。与其他先进算法相比,本文算法在保证检测速度的情况下,有效提高了遮挡行人和小尺度行人的检测精度,为相关行人检测提供了一种新思路。在参与 CIVC-2020 行人避让场景比赛中,与全国上百支参赛队的算法较量,取得了满分的好成绩。

本文算法使用了两个轻量级主干网络进行特征提取,在空间复杂度仍存在一定的算力浪费,尺度并行检测在轻量级检测算法中仍存在一定的优化空间。在未来工作中,可以将尺度并行检测进行调整以提高检测效率,同时继续寻找优化方法,改进神经网络的特征提取算力浪费的现象,使得检测性能和检测精度得以进一步提高,促进其在自动驾驶领域中获得更好的应用。

参考文献 (References)

- Ahmed Z, Iniyan R and Madhan M P. 2019. Enhanced vulnerable pedestrian detection using deep learning//Proceedings of International Conference on Communication and Signal Processing (ICCSP). Chennai, India; IEEE: 971-974 [DOI: 10.1109/ICCSP. 2019.8697978]
- Broggi A, Fascioli A, Fedriga I, Tibaldi A and Rose M D. 2003. Stereo-based preprocessing for human shape localization in unstructured environments//IEEE IV2003 Intelligent Vehicles Symposium. Columbus, USA; IEEE: 410-415 [DOI: 10.1109/IVS. 2003.1212946]
- Dalal N and Triggs B. 2005. Histograms of oriented gradients for human detection//Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05). San Diego, USA; IEEE: 886-893 [DOI: 10.1109/CVPR. 2005.177]
- Dalal N, Triggs B and Schmid C. 2006. Human detection using oriented histograms of flow and appearance//Proceedings of the 9th European Conference on Computer Vision. Graz, Austria; Springer: 428-441 [DOI: 10.1007/11744047_33]
- Dollar P, Wojek C, Schiele B and Perona P. 2009. Pedestrian detection: a benchmark//Proceedings of 2009 IEEE Conference on Computer Vision and Pattern Recognition. Miami, USA; IEEE: 304-311 [DOI: 10.1109/CVPR. 2009.5206631]
- Dollar P, Wojek C, Schiele B and Perona P. 2012. Pedestrian detection: an evaluation of the state of the art. IEEE Transactions on Pattern Analysis and Machine Intelligence, 34(4): 743-761 [DOI: 10.1109/TPAMI. 2011.155]
- Gavrila D M and Munder S. 2007. Multi-cue pedestrian detection and tracking from a moving vehicle. International Journal of Computer Vision, 73(1): 41-59 [DOI: 10.1007/s11263-006-9038-7]
- Guo J. 2017. Research on Radar Modeling For Vehicle Intelligence. Changchun: Jilin University (郭姣. 2017. 面向汽车智能化仿真的雷达模拟研究. 长春: 吉林大学)
- Li J N, Liang X D, Shen S M, Xu T F, Feng J S and Yan S C. 2018. Scale-aware fast R-CNN for pedestrian detection. IEEE Transactions on Multimedia, 20(4): 985-996 [DOI: 10.1109/TMM. 2017.2759508]
- Liu J, Gao X K, Bao N Y, Tang J and Wu G S. 2017. Deep convolutional neural networks for pedestrian detection with skip pooling//Proceedings of International Joint Conference on Neural Networks (IJCNN). Anchorage, USA; IEEE: 2056-2063 [DOI: 10.1109/IJCNN. 2017.7966103]
- Liu W, Anguelov D, Erhan D, Szegedy C, Reed S, Fu C Y and Berg A C. 2016. SSD: single shot MultiBox detector//Proceedings of the 14th European Conference Computer Vision. Amsterdam, the Netherlands; Springer: 21-37 [DOI: 10.1007/978-3-319-46448-0_2]
- Milan A, Leal-Taixe L, Reid I, Roth S and Schindler K. 2016. MOT16: a benchmark for multi-object tracking [EB/OL]. [2020-07-23]. <https://arxiv.org/pdf/1603.00831.pdf>
- Milton A A. 2019. Towards pedestrian detection using RetinaNet in ECCV 2018 wider pedestrian detection challenge [EB/OL]. [2020-07-23]. <https://arxiv.org/pdf/1902.01031.pdf>
- Nam W, Dollár P and Han J H. 2014. Local decorrelation for improved pedestrian detection//Proceedings of the 27th International Conference on Neural Information Processing Systems. Montreal, Canada; ACM: 424-432 [DOI: 10.5555/2968826.2968874]
- Redmon J, Divvala S, Girshick R and Farhadi A. 2016. You only look once: unified, real-time object detection//Proceedings of IEEE Conference on Computer Vision and Pattern Recognition. Las Vegas, USA; IEEE: 779-788 [DOI: 10.1109/CVPR. 2016.91]
- Sermanet P, Kavukcuoglu K, Chintala S and Lecun Y. 2013. Pedestrian detection with unsupervised multi-stage feature learning//Proceedings of 2013 IEEE Conference on Computer Vision and Pattern Recognition. Portland, USA; IEEE: 3626-3633 [DOI: 10.1109/CVPR. 2013.465]
- Tan M X and Le Q V. 2019. EfficientNet: rethinking model scaling for convolutional neural networks [EB/OL]. [2020-07-23]. <https://arxiv.org/pdf/1905.11946.pdf>
- Tan M X, Pang R M and Le Q V. 2020. EfficientDet: scalable and efficient object detection//Proceedings of 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Seattle, USA; IEEE: 10781-10790 [DOI: 10.1109/cvpr42600. 2020.01079]
- Tian Y L, Luo P, Wang X G and Tang X O. 2015. Deep learning strong parts for pedestrian detection//Proceedings of 2015 IEEE Interna-

- tional Conference on Computer Vision. Santiago, Chile; IEEE: 1904-1912 [DOI: 10.1109/iccv.2015.221]
- Wang X L, Xiao T T, Jiang Y N, Shao S, Sun J and Shen C H. 2018. Repulsion loss: detecting pedestrians in a crowd//Proceedings of 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Salt Lake City, USA; IEEE: 7774-7783 [DOI: 10.1109/CVPR.2018.00811]
- Yang F, Choi W and Lin Y Q. 2016. Exploit all the layers: fast and accurate CNN object detector with scale dependent pooling and cascaded rejection classifiers//Proceedings of 2016 IEEE Conference on Computer Vision and Pattern Recognition. Las Vegas, USA; IEEE: 2129-2137 [DOI: 10.1109/CVPR.2016.234]
- Zhang L L, Lin L, Liang X D and He K M. 2016. Is faster R-CNN doing well for pedestrian detection?//Proceedings of the 14th European Conference on Computer Vision-ECCV 2016. Amsterdam, The Netherlands; Springer: 443-457 [DOI: 10.1007/978-3-319-46475-6_28]
- Zhang S S, Benenson R and Schiele B. 2017. CityPersons: a diverse dataset for pedestrian detection//Proceedings of 2017 IEEE Conference on Computer Vision and Pattern Recognition. Honolulu, USA; IEEE: 4457-4465 [DOI: 10.1109/CVPR.2017.474]
- Zhou C L, Yang M and Yuan J S. 2019. Discriminative feature transformation for occluded pedestrian detection//Proceedings of 2019 IEEE/CVF International Conference on Computer Vision. Seoul, South Korea; IEEE: 9557-9566 [DOI: 10.1109/ICCV.2019.00965]

- Zhuang C B, Lei Z and Li S Z. 2020. SADet: learning an efficient and accurate pedestrian detector [EB/OL]. [2020-07-26]. <https://arxiv.org/pdf/2007.13119.pdf>

作者简介



徐歆恺, 1980 年生, 男, 博士, 讲师, 主要研究方向为人工智能和人机交互。

E-mail: ldtxinkai@163.com



马岩, 通信作者, 男, 博士, 讲师, 主要研究方向为人工智能和智能矿业。

E-mail: jqrmayan@bnu.edu.cn

钱旭, 男, 教授, 主要研究方向为矿山图像处理、人工智能。

E-mail: xuqian@cumtb.edu.cn

张龔, 男, 博士研究生, 主要研究方向为人工智能。

E-mail: iv_allen@163.com