【电子与信息科学 / Electronics and Information Science】

基于 DCN-Mobile-YOLO 模型的多车道车辆计数

文 奴^{1,2}, 郭仁忠^{1,2}, 贺 彪^{1,2,3}

1) 深圳大学建筑与城市规划学院, 深圳大学智慧城市研究院, 广东深圳 518061; 2) 粤港澳智慧城市联合实验室, 广东深圳 518061; 3) 国土资源部城市土地资源监测与仿真重点实验室, 广东深圳 518034

摘 要: 单一目标检测方法无法实现目标计数的准确统计, 且模型的检测精度和速度难以同步提升. 以 YOLO v4 目标检测框架为基础, 提出一种移动端的目标追踪和多车道车辆计数模型 DCN-Mobile-YOLO. 使用可变形卷积网络(deformable convolutional networks, DCNs) v2 卷积核和移动端卷积网络 MobileNet v3 框架分别代替 YOLO v4 的常规卷积核和主干网络, 结合 DeepSORT 算法实现对多目标的跟踪和计数, 建立自适应车道检测规则并实现车道内车辆的精确计数. 在 VOC2007 + 2012 数据集和 GoPro 采集数据上验证 DCN-Mobile-YOLO 模型的有效性. 结果表明, DCN-Mobile-YOLO 模型的平均精度均值相比主干网络为 MobileNet v3 和 CSPDarkNet 的 YOLO v4 算法分别提升了 13.19% 和 6.63%, 目标检测平均帧率为 12 帧/s. DCN-Mobile-YOLO 模型不仅提高了目标检测模型的检测精度, 且达到了移动端实时检测的速度.

关键词: 人工智能; 视频目标检测; 多目标跟踪; YOLO v4; 车流量; 深度学习; 卷积神经网络; 目标计数

中图分类号: TP183

文献标志码: A

doi: 10.3724/SP.J.1249.2021.06628

Multi-lane vehicle counting based on DCN-Mobile-YOLO model

WEN Nu^{1,2}, GUO Renzhong^{1,2}, and HE Biao^{1,2,3}

1) School of Architecture & Urban Planning, Research Institute for Smart Cities, Shenzhen University, Shenzhen 518061, Guangdong Province, P. R. China

2) Guangdong-Hong Kong-Macau Joint Laboratory for Smart Cities, Shenzhen 518061, Guangdong Province, P. R. China

3) Key Laboratory of Urban Land Resources Monitoring and Simulation, Ministry of Natural Resources, Shenzhen 518034, Guangdong Province, P. R. China

Abstract: The single object detection method cannot achieve the accurate statistics of object counts and the accuracy and speed of model detection cannot be improved simultaneously. To solve these problems, we propose a mobile-side object tracking and multi-lane vehicle counting method-DCN-Mobile-YOLO based on the YOLO v4 object detection framework. This method uses the DCNs v2 convolution kernel and MobileNet v3 framework to replace the conventional convolution kernel and backbone network of YOLO v4, respectively, and uses the DeepSORT algorithm to achieve multi-object tracking and counting. At the same time, an adaptive lane detection rule is proposed to achieve the accuracy of vehicles in the lane count. Finally, the effectiveness of DCN-Mobile-YOLO method is verified on the

Received: 2021-06-15; **Accepted:** 2021-07-30; **Online (CNKI):** 2021-10-08

Foundation: The Guangdong Science and Technology Strategic Innovation Fund (2020B1212030009); The Open Fund of Key Laboratory of Urban Land Resources Monitoring and Simulation, MNR, China (KF-2018-03-031)

Corresponding author: Professor HE Biao. E-mail: whu_hebiao@hotmail.com

Citation: WEN Nu, GUO Renzhong, HE Biao. Multi-lane vehicle counting based on DCN-Mobile-Yolo model [J]. Journal of Shenzhen University Science and Engineering, 2021, 38(6): 628-635. (in Chinese)

VOC2007 + 2012 data set and GoPro collected data set. The experiment results show that the mAP of DCN-Mobile-YOLO method is improved by 13.19% and 6.63% respectively compared with YOLO v4 algorithm of MobileNet v3 and CSPDarkNet as the backbone network, and the average frame rate of object detection is 12 frame/s. Our DCN-Mobile-YOLO method not only improves the detection accuracy of the object detection model but also achieves the speed of real-time detection on the mobile terminal.

Key words: artificial intelligence; video object detection; multi-object tracking; YOLO v4; vehicle flowrate; deep learning; convolutional neural networks; target counting

目标计数方法按照技术路线可分为基于传统特征提取的方法^[1-3]和卷积神经网络(convolutional neural network, CNN)的方法^[4-6]. 传统特征提取方法有基于哈尔(Haar-like, Haar)特征、局部二值模式(local binary pattern, LBP)特征和方向梯度直方图(histogram of oriented gradient, HOG)特征等,这类算法通常以边缘特征提取为主实现对目标个体的检测和计数,因而对边缘检测的精确度要求很高,不适合目标间有重叠或密集目标的计数. 随着深度学习的发展,基于CNN提取目标特征进而实现目标检测和计数逐步成为目标计数的主流算法. 基于CNN的目标检测算法通常分为单级检测器(one-stage detection)和二级检测器(two-stage detection),一般来说,前者在速度上占优,后者在准确率和定位精度上占优. YOLO(you only look once)系列是单级检测器的代表,不仅推理速度快且检测和定位精度高. YOLO v4^[7]是基于YOLO v3的目标检测架构,在数据处理与增强、主干网络、网络训练、激活函数和损失函数等方面都进行了优化,在保证运行效率的同时极大提高了目标检测的精度. 其中,在主干网络部分,YOLO v4借鉴了残差网络(如ResNet^[8])提取深层特征信息的优点,提出空间金字塔池化网络(如spatial pyramid pooling-net, SSP-Net^[9]和 atrous spatial pyramid pooling-net, ASSP-Net^[10])的设计思路,通过拼接任意大小的特征图并转换成固定输出大小的特征向量,实现一次性输出多尺度目标检测的结果. 若直接使用CNNs方法对图像目标进行计数,可能存在检测漏检与重复计数的问题. DeepSORT方法^[11]是一种基于检测的跟踪(tracking-by-detection)方法,是SORT(simple online and realtime tracking)方法^[12]的升级版,算法结合传统的卡尔曼滤波(Kalman filter, KF)和匈牙利算法(Hungarian algorithm)实现检测目标的跟踪,通过加入特征提取网络(如重识别模块 ReID^[13])实现对特征的匹配和更新. 该方法在一定程度上解决了单独使用目标检测方法存在的漏检和重复计数

问题.

本研究采用YOLO v4作为目标检测算法的基本框架,以DeepSORT方法实现检测目标的跟踪和计数. 针对YOLO v4主干网络CSPDarkNet参数量大无法适用于移动端进行实时处理的问题,引入深度可分离卷积^[14],实现网络轻量化. 同时,使用可变形卷积^[15]代替常规卷积增强感兴趣区域(region of interest, ROI)的范围,提高目标定位的精度.

1 相关工作

1.1 传统特征提取方法

传统特征提取包括基于梯度信息^[1]、颜色特征^[2]和帧间差异等^[3]特征信息方法. 其中,基于梯度信息包括Haar特征、LBP特征和HOG特征等,它们利用方向梯度算子对图像边缘特征进行提取以实现目标的检测与计数. 这类方法普适性较差且检测精度受目标个体和背景差异影响明显,对重叠目标的检测效果受限. 基于颜色特征的提取方法,对图像尺寸、旋转和平移等因素影响较小,可以提高特征提取的鲁棒性,但光照变化容易导致颜色空间分布信息混乱,影响目标检测和计数的准确性. 基于帧间差异的特征提取方法有光流法^[16]和背景差分法^[17]. 光流法无法对单帧图像进行建模和目标检测;背景差分法计算过程简单,计算量小,可满足实时性的要求,解决帧差法提取目标存在内部空洞问题,但是对外部环境和场景具有高度敏感性.

1.2 基于卷积神经网络的方法

现有的基于深度学习的目标检测与识别算法大致分为两大类:①基于区域建议的目标检测与识别算法,如Faster R-CNN^[18]和Mask R-CNN^[19];②基于回归的目标检测与识别算法,如SSD^[20]和YOLO^[21]. 前者是由算法生成一系列样本候选框,再通过CNN进行样本分类,属于二级检测器;后者则直接将目标边框定位的问题转化为回归问题处理,属于单级检测器. 两种方法各有优势,前者在

检测准确率和定位精度上占优, 后者在算法速度上占优. 因此, 研究准确率与速度同步提升的目标检测方法, 是研究学者关注的重要方向.

YOLO v4 是一个能够应用于实际工作环境中的快速目标检测系统, 且能够被并行优化, 它在识别精度和效率之间取得平衡. YOLO 系列算法是最接近工程应用的目标检测算法, 被广泛用于交通标志识别与检测^[22]、行人与车辆检测^[23]、工业零件识别和缺陷检测^[24]等研究领域.

以上方法在不同方面的应用都取得较好的结果, 但由于神经网络算法的网络结构复杂且参数量大, 需强大的图形处理器 (graphics processing unit, GPU) 运算能力方能实现实时检测和计数, 这对于移动设备或嵌入式设备的移植应用与实时计数带来较大挑战. 对于多车道的目标检测, 目前的目标检测方法容易出现漏检和重复检测的问题, 无法确定目标计数的唯一性. 因此, 本研究提出一种轻量化 DCN-Mobile-YOLO 网络模型, 并结合 DeepSORT 进行视频多目标跟踪, 进而实现城市道路交叉口潮汐车道的车辆计数.

2 DCN-Mobile-YOLO 多目标跟踪算法

本研究基于 YOLO v4 目标检测方法, 通过修改主干网络框架, 并使用可变形卷积网络 (deformable convolutional networks, DCN) v2 卷积以扩展特征层的 ROI, 从而增强目标检测的精度. 同时, 为解决单一目标检测框架在视频检测中存在的漏检与误检问题, 结合 DeepSORT 追踪技术与目标检测, 进一步提升视频检测的召回率. 采用 MobileNet 主干网络实现算法在移动端的移植, 提升多目标跟踪算法的检测效率.

2.1 网络基本框架

DCN-Mobile-YOLO 模型的网络框架可分为主干网络 DCN-MobileNet、特征增强的瓶颈 neck 和检测输出 head 三个部分, 如图 1. 其中, neck 主要包括空间金字塔池化深度卷积网络 (spatial pyramid pooling in deep convolutional networks, SPP-Net) 和金字塔注意力网络 (pyramid attention network, PANet).

图 1 DCN-Mobile-YOLO 模型的网络框架

Fig. 1 Network Architecture of DCN-Mobile-YOLO model

DCN-MobileNet 沿用 YOLO v4 中的 DarkNet 网络, 为减少网络结构的参数量, 采用 MobileNet 主

体结构作为主干网络的核心卷积层. 为避免可分离卷积导致的网络精度下降问题, 用可变形卷积代替

网络框架中的所有常规卷积. 通过引入跨阶段局部网络(cross stage partial network, CSPNet), 将梯度的变化集成到特征图中, 既保证了准确率又减少了计算量. 激活函数则用更加平滑的 mish() 函数代替 relu() 函数. 同时, 为实现对不同尺度的目标都有较好的检测效果, 根据数据集真实(ground-truth, GT)边界框的数据对锚定框进行聚类, 设计出最接近真实框的 3 类锚定框. 网络的特征输出采用大、中和小 3 个尺度的特征图, 根据输入图像尺寸 $416 \times 416 \times 3$ 像素, 输出尺寸为 52×52 、 26×26 和 13×13 像素的 3 种网格.

瓶颈块 neck 旨在增强网络特征提取的能力. SPP-Net 采用了 3 个不同尺寸的卷积, 以增加特征图对上层输出的感受野, 增强网络的目标识别能力. 通过引入差异化池化策略, 使这类多尺度训练方法, 不仅避免了网络过拟合的风险还可以实现固定尺寸的图像特征输出. PANet 在特征金字塔网络(feature pyramid networks, FPN)的基础上, 采用降采样和上采样方法同时对不同尺度特征图进行融合, 使映射和融合后的输出层特征更丰富, 提高网络对浅层特征信息和深层语义信息的表达能力.

输出特征图 head 对应主干网络最后 3 个特征层, 具有多特征层目标检测能力, 3 个特征层尺寸分别为 $(52, 52, 75)$ 、 $(38, 38, 75)$ 和 $(19, 19, 75)$. 其中, 前 2 个维度表示特征图网格的大小, 可实现对不同大小的待检目标的提取; 第 3 个维度为所使用的网络训练 VOC2007 + 2012 数据集的类别, 本研究设为 20; 位置信息 $([x, y, h, w])$ 加上类别信息(class)的维度合计为 5. 由于不同尺寸的锚定框(anchor)都有 3 种, 故共有 $3 \times (20 + 5) = 75$ 个锚定框.

2.2 DeepSORT 目标跟踪

DeepSORT 算法的核心内容包括匈牙利算法和 KF 算法. 使用 DCN-MobileNet 检测可获得当前帧目标检测的结果; 使用匈牙利算法判断当前帧的某个目标与前一帧某个目标是否相同; KF 则可以修正目标检测算法边界框的位置. KF 基于目标前一刻的位置信息预测当前时刻的位置, 因此可获得比检测器更准确的估计目标位置. 目标跟踪的处理流程如下:

1) 采用 DCN-Mobile-YOLO 读取当前帧目标检测框位置, 根据各检测框图像块的深度特征定义唯一的追踪身份证标识号(identity document, ID);

2) 采用 DeepSORT 对检测框的置信度进行排

序, 将不符合阈值的检测框及特征删除;

3) 采用非极大值抑制(non-maximum suppression, NMS)方法对同一目标的多个检测框进行交并比(intersection-over-union, IoU), 使目标具有唯一边界框;

4) 获取当前检测框的状态参量(边界框面积、长宽比、中心点横坐标和纵坐标), 并使用 KF 预测目标在当前帧的位置;

5) 使用匈牙利算法将预测后的追踪器和当前帧中的检测器进行匹配(级联匹配和 IoU 匹配);

6) 更新 Kalman 追踪器参数及特征集, 并进行目标消失和新目标出现的判断.

2.3 自适应车道检测规则

常见的车道车辆总数统计方法有虚拟检测线法和虚拟线框法, 前者计算量小且实时性高, 但在车速较慢时的统计误差大, 后者采用固定检测区域与检测框重叠面积的比值来判定车道车辆数, 准确性更高. 但是虚拟线框法仅适用于拍摄传感器正对车道线的情况, 否则将会出现车道检测的误判.

车道线划分的目的是检测出车道线, 并判定当前车道目标检测的车辆流量. 车辆目标检测边界框与车道划分区域的重叠面积(即 IoU)是判定目标属于某个车道的前提条件, 因此选择合适的车道检测规则对多车道目标检测意义重大. 本研究的车道线检测采用人工标定的方法判别车道划分是否准确, 在车道上设置计数检测区域, 根据车辆行驶规律、车辆与车道长宽比和道路视频拍摄角度与位置, 提出自适应的车道检测规则.

以图 2 所示的车道检测示意图为例, 共有 4 种检测规则: ① 若以车辆检测边界框(Box_A, 红色实线框)与车道检测框(Box_B, 黑色虚线框)的 IoU 计算, 则当车道宽度过大(>3.5 m)或车辆高度过高(>2.5 m)会出现 $\text{IoU} < 0.5$, 导致漏检; ② 若以判定车辆检测边界框的中心点(Center_A)是否落在车道检测框(Box_B, 黑色)中为检测规则, 则当车辆高度过高或车辆行驶偏两侧时, 中心点会落在邻近车道中, 导致误检; ③ 若以车辆投影中心点(Center_C)是否落在车道检测框(Box_B)中为检测规则, 当车辆行驶靠近两侧(由于拍摄位置原因, 图 2 中是偏右侧)时, 中心点会落在右侧邻近车道导致误检; ④ 以车辆投影中心点(Center_C, 考虑车辆高度、实际检测框高度与车道关系)为基础, 判定其是否落在车道实际检测区域(Box_D, 黄色实线框)作为检测计数规则. 根据实验设计要求

和道路视频拍摄情况,本研究选定规则 4 作为车道统计计数的规则。

图 2 车道检测区域与车辆目标检测框

Fig. 2 Vehicle detection area and vehicle object detection bounding box

3 实验结果和分析

实验硬件配置为: 计算机 CPU 为 Intel (R) Core (TM) i9-9900K, 频率为 3.60 GHz, 内存为 64 Gbyte, 显卡型号为 NVIDIA GeForce GTX 2080Ti, 显存 11 Gbyte $\times$ 2. 软件环境为 Ubuntu 16.04, 64 bit. 深度学习框架为 TensorFlow1.13.1 & Keras 2.3.1, 并行计算框架为 CUDA 10.

为验证本研究提出的移动端潮汐车道车辆计数方法的运行速度和检测效果,选取早高峰潮汐车道拍摄视频作为验证集验证算法的时效性,同时在 VOC2007 + 2012 数据集中分析算法的检测性能。

3.1 检测性能分析

1) 精度 (P) 与召回率 (R): 精度是衡量目标检

测模型正样本检出量占总检出样本的比例,召回率代表检测出的正样本占有所有正样本的比例,计算公式如式(1)和式(2)。

$$P = TP / (TP + FP) \quad (1)$$

$$R = TP / (TP + FN) \quad (2)$$

其中, TP 为正样本被正确检索的概率; FP 为负样本被误检成正样本的概率; FN 为正样本被误检成负样本的概率。

2) 平均精度 (average precision, AP): AP 是某个类在 P - R 曲线上的平均精度值,常用式(3)计得。好的目标检测模型意味着召回率增加的同时精度也增加。

$$AP = \int_0^1 p(r) dr \quad (3)$$

其中, r 为取值在 0 ~ 1 的召回率; $p(r)$ 为精度。

3) 平均精度均值 (mean average precision, mAP): mAP 是计算所有类别 P - R 曲线下面积的平均值,即

$$mAP = \frac{1}{n} \sum_{k=1}^n AP_k \quad (4)$$

其中, n 为类别数。

为验证 DCN-Mobile-YOLO 模型的检测效果,分别采用 MobileNet、CSPDarkNet 和 DCN-MobileNet 三个不同主干网络的 YOLO v4 模型,计算对 VOC2007 + 2012 数据集中所有类别检测后的 AP 值,结果如图 3。由图 3 可见,DCN-Mobile-YOLO 相对于主干网络为 MobileNet 和 CSPDarkNet 的 YOLO v4 网络, mAP 值分别提升了 13.19% 和 6.63%。

图 3 不同主干网络的 YOLO 模型 AP 值

Fig. 3 AP values of YOLO models with different backbone networks

4) 参数量. DCN-MobileNet 借鉴了 MobileNet 的可分离卷积思想, 将标准卷积分解为 depthwise 卷积和 pointwise 卷积各 1 个, 分别起到滤波和线性组合的作用, 大量减少了参数量和计算量. 同时, 在常规卷积操作中, 以轻微牺牲参数数量的代价, 采用 DCN v2 卷积扩张 ROI 进而提高目标检测模型的精度. 表 1 列举了不同主干网络模型的参数量. 由表 1 可见, DCN-MobileNet 的参数量比 MobileNet 仅多 0.48%, 是 CSPDarknet 参数量的 17%. 设训练批次 (batchsize) 为 32, 总训练周期 (epoch) 为 200, 对比算法的训练时间发现, 当 MobileNet 作为主干网络时, 单个训练周期可以节省 1/2 的训练时间.

5) 检测效率与精度. 图 4 为不同模型目标检测和车道车辆计数的结果. 需要特别说明的是, 检

测帧率与当前检测出的数量和模型相关, 相同模型下检测数量越多检测帧率越低. 由图 4 可见, DCN-Mobile-YOLO 模型的当前帧检测数量和总检测数量计数比 MobileNet-YOLO 和 CSPDarknet-YOLO 模型都要多, 且实时帧率速度更快.

表 1 不同主干网络的 YOLO v4 模型参数量
Table 1 Model parameters of YOLO v4 models different with backbone networks

主干网络	非训练 参数	非训练 参数	总参数	训练 时间/min
MobileNet	62 672	11 405 969	11 468 641	7
CSPDarknet	66 304	64 363 101	64 429 405	16
DCN-MobileNet	62 672	11 461 322	11 523 994	7

图 4 不同主干网络的 YOLO v4 模型检测结果

Fig. 4 Model test results of YOLO v4 models with different backbone networks

3.2 潮汐车道流量分析

为研究潮汐车道的通行能力和各车道流量数据, 利用 GoPro 运动摄像机采集 2020-01-09 中国深圳市深南-北环立交早高峰时间段 (07:05-09:00) 的视频数据 (07:00-07:05 的数据因安装设备, 未

采集). 以 5 min 为一个时间周期, 将早高峰时间段划分为 23 个时段, 得到不同车道目标检测和车流量结果分别见表 2 和图 5. 其中, 1~23 为时段序号; 车道 1 为潮汐车道; 车道 2 为左转车道; 车道 3~7 为直行车道; 车道 8 为右转车道.

表 2 2020-01-09 深南-北环立交早高峰各时段各车道车辆计数统计

Table 2 Vehicle count statistics in each lane at Shennan-Beihuan interchange on January 9, 2020 辆

车道	车流量																						
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23
1	40	30	30	28	21	18	31	28	22	36	25	34	21	31	21	40	13	26	33	18	16	14	15
2	10	8	12	13	9	36	12	13	28	10	11	8	12	12	11	23	28	38	27	11	21	18	25
3	85	51	54	63	85	75	62	84	58	60	50	72	55	60	48	64	61	49	57	50	59	44	60
4	76	55	52	64	89	70	62	80	56	58	39	66	54	58	55	62	60	60	48	46	68	46	53
5	69	51	55	47	66	63	57	70	58	60	40	59	43	56	50	66	38	38	35	37	29	30	31
6	72	51	70	59	92	70	60	86	59	50	33	76	47	79	70	65	41	26	27	16	45	29	36
7	61	23	49	54	48	47	37	39	50	63	36	49	29	25	33	35	29	71	67	60	74	55	73
8	29	15	21	16	9	4	3	14	20	20	19	9	18	15	5	14	23	35	16	42	36	53	77
帧率/(帧·s ⁻¹)	12	11	13	11	12	12	13	12	12	10	12	12	11	11	11	11	11	13	11	13	13	12	13

图 5 2020-01-09 深南-北环立交早高峰各时段不同车道流量数据

Fig. 5 Traffic flow data of different lanes at Shennan-Beihuan interchange on 2020-01-09

由表 2 可知,潮汐车道和左转车道 23 个时段内单个车道的平均车流量分别为 26 辆和 17 辆,两者差异不大,直行车道的平均车流量为 55 辆。

将车道按照潮汐 + 左转 (车道 1 和 2)、直行 (车道 3 ~ 7) 和右转 (车道 8) 划分为 3 类,根据表 2 的数据统计对应类型车道在每个时段的平均流量 (时段内总车流量除以车道数),结果如图 5。再统计 3 种类型的车道在整各早高峰时间段 (07:05-09:00) 内每 5 min 内的平均流量,结果表明,潮汐 + 左转、直行和右转车道每 5 min 的车流量分别是 43、274 和 22 辆,占比分别为 2:10:1。可见,若将左转车道设置为直行,潮汐车道设置为左转,可进一步优化车道通行效率,缓解直行车道的通行压力。

结 语

提出一种针对移动端的多车道目标追踪和计数模型 DCN-Mobile-YOLO。为解决 MobileNet-YOLO 模型检测精度有限的问题,在 YOLO v4 主干网络中引入 DCN v2 卷积并替换原来的 CSPDarknet 网络。以 VOC2007 + 2012 数据集作为训练集和验证集,根据 AP 指标分析 YOLO 框架下 MobileNet、CSPDarknet 和 DCN-MobileNet3 种主干网络在验证集中的表现。结果表明,在所有类别的 AP 值实验比较中,DCN-Mobile-YOLO 模型比 MobileNet-YOLO 和 CSPDarknet-YOLO 模型分别提升了 13.19% 和 6.63%。采集早高

峰交通视频数据,使用 DCN-Mobile-YOLO 模型统计左转 (含潮汐车道)、直行和右转的流量数据,通过分析潮汐车道对交通流量的影响,发现早高峰车流量瓶颈在直行方向,提出将左转车道改为可变车道,将潮汐车道设计为左转车道的建议。

基金项目: 广东省科技创新战略专项资金资助项目 (2020B1212030009); 自然资源部城市国土资源监测与仿真重点实验室开放基金资助项目 (KF-2018-03-031)

作者简介: 文 奴 (1989—), 深圳大学博士后研究人员。研究方向: 视频目标检测、图像处理和计算机视觉。
E-mail: wenu1989@126.com

引 文: 文 奴, 郭仁忠, 贺 彪. 基于 DCN-Mobile-YOLO 模型的多车道车辆计数[J]. 深圳大学学报理工版, 2021, 38(6): 628-635.

参考文献 / References:

- [1] DENIMAL E, MARIN A, GUYOT S, et al. Automatic biological cell counting using a modified gradient Hough transform [J]. Microscopy and Microanalysis, 2017, 23(1): 11.
- [2] LI Daoliang, MIAO Zheng, PENG Fang, et al. Automatic counting methods in aquaculture: a review [J]. Journal of the World Aquaculture Society, 2021, 52(2): 269-283.
- [3] ZHOU Yue, JI Janwei, SONG Kai. A moving target detection method based on improved frame difference background modeling [J]. The Open Cybernetics & Systemics Journal, 2014, 8(1): 970-975.
- [4] HE Kaiming, GKIOXARI G, DOLLÁR P, et al. Mask R-CNN [C]// Proceedings of the IEEE International Con-

- ference on Computer Vision. Shenzhen, China: IEEE, 2017: 2961-2969.
- [5] REDMONJ, FARHADI A. YOLO v3: an incremental improvement [C]// IEEE Conference on Computer Vision and Pattern Recognition. Salt Lake City, USA: IEEE, 2018: 89-95.
- [6] YOO D, PARK S, LEE J Y, et al. AttentionNet: aggregating weak directions for accurate object detection [C]// Proceedings of the IEEE International Conference on Computer Vision. Santiago, Chile: IEEE, 2015: 2659-2667.
- [7] BOCHKOVSKIY A, WANG C Y, LIAO H Y M. YOLO v4: optimal speed and accuracy of object detection [EB/OL]. (2020-04-23) [2021-06-22]. <https://arxiv.org/abs/2004.10934v1>.
- [8] HE Kaiming, ZHANG Xiangyu, REN Shaoqing, et al. Deep residual learning for image recognition [C]// Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Las Vegas, USA: IEEE, 2016: 770-778.
- [9] HE Kaiming, ZHANG Xiangyu, REN Shaoqing, 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.
- [10] LIU Songtao, HUANG Di, WANG Yunhong. Learning spatial fusion for single-shot object detection [EB/OL]. (2019-11-21) [2019-11-25]. <https://arxiv.org/abs/1911.09516>.
- [11] WOJKE N, BEWLEY A, PAULUS D. Simple online and realtime tracking with a deep association metric [C]// IEEE International Conference on Image Processing. Beijing: IEEE, 2017: 3645-3649.
- [12] BEWLEY A, GE Z, OTT L, et al. Simple online and realtime tracking [C]// IEEE International Conference on Image Processing. Phoenix, USA: IEEE, 2016: 3464-3468.
- [13] ZHENG Liang, SHEN Liye, TIAN Lu, et al. Scalable person re-identification: a benchmark [C]// Proceedings of the IEEE International Conference on Computer Vision. Santiago, Chile: IEEE, 2015: 1116-1124.
- [14] HOWARD A, SANDLER M, CHU G, et al. Searching for mobilenet v3 [C]// Proceedings of the IEEE/CVF International Conference on Computer Vision. Seoul: IEEE, 2019: 1314-1324.
- [15] ZHU Xizhou, HU Han, LIN S, et al. Deformable convnets v2: more deformable, better results [C]// Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. Long Beach, USA: IEEE, 2019: 9308-9316.
- [16] LIU Dongfang, CUI Yiming, CHEN Yingjie, et al. Video object detection for autonomous driving: motion-aid feature calibration [J]. Neurocomputing, 2020, 409: 1-11.
- [17] GARCIA-GARCIA B, BOUWMANS T, SILVA A J R. Background subtraction in real applications: challenges, current models and future directions [J]. Computer Science Review, 2020, 35: 100204.
- [18] REN Shaoqing, HE Kaiming, 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.
- [19] HE K, GKIOXARI G, DOLLÁR P, et al. Mask R-CNN [C]// Proceedings of the IEEE International Conference on Computer Vision. Venice, Italy: IEEE, 2017: 2961-2969.
- [20] LIU Wei, ANGUELOV D, ERHAN D, et al. SSD: single shot multibox detector [C] // European Conference on Computer Vision. Amsterdam: Springer, Cham, 2016: 21-37.
- [21] JOSEPH R, SANTOSH D, ROSS G, et al. You only look once: unified, real-time object detection [C]// Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Las Vegas, USA: IEEE, 2016: 779-788.
- [22] JIANG Jinhong, BAO Shengli, SHI Wenxu, et al. Improved traffic sign recognition algorithm based on YOLO v3 algorithm [J]. Journal of Computer Applications, 2020, 40(8): 2472-2478.
- [23] LU Junyan, MA Chi, LI Li, et al. A vehicle detection method for aerial image based on YOLO [J]. Journal of Computer and Communications, 2018, 6(11): 98-107.
- [24] HATAB M, MALEKMOHAMADI H, AMIRA A. Surface defect detection using YOLO network [C]// Proceedings of SAI Intelligent Systems Conference. Berlin: Springer, Cham, 2020: 505-515.

【中文责编: 英 子; 英文责编: 木 柯】