

文章编号: 1005-0523(2024)04-0064-09



基于多尺度和双源运动感知的车辆检测方法

李晓晗¹, 刘石坚¹, 刘建华¹, 戴宇晨¹, 邹 峥²

(1. 福建省大数据挖掘与应用技术重点实验室, 福建 福州 350118; 2. 福建师范大学计算机与网络空间安全学院, 福建 福州 350117)

摘要:【目的】车辆检测是城市智能交通研究的重要部分,以交通监控图像作为输入,以运动车辆的检测作为目标,围绕其中的小目标问题、高密度问题以及运动属性问题开展研究。【方法】在无锚框 CenteNet 的基础上提出一种基于多尺度双源运动感知的检测方法。首先,引入坐标注意力,并融合网络抽象层的多尺度和全局上下文特征,多层次多阶段地补充信息,提高模型对车辆和场景的理解力;其次,借助代表车辆实际运动特征的模糊纹理和代表车辆通用运动特征的光流知识,构建模型对运动车辆的感知能力。【结果】实验数据来源于公共数据集 UA-DETRAC,以均值平均精度(mAP)和帧率(FPS)作为精确度和速度评价指标,将文章方法与已有主流方法进行比较,结果表明文章方法的 mAP 和 FPS 分别为 72.46% 和 30 frame/s,在对比方法中具有最佳的速度与准确率均衡性。【结论】文章方法能够胜任运动车辆的检测任务。

关键词: 运动车辆检测; CenterNet; 模糊纹理; 光流; 多任务学习

中图分类号: TP311

文献标志码: A

本文引用格式: 李晓晗, 刘石坚, 刘建华, 等. 基于多尺度和双源运动感知的车辆检测方法[J]. 华东交通大学学报, 2024, 41(4): 64-72.

Study on Multi-Scale Feature and Dual-Source Motion Perception for Vehicle Detection

Li Xiaohan¹, Liu Shijian¹, Liu Jianhua¹, Dai Yuchen¹, Zou Zheng²

(1. Fujian Provincial Key Laboratory of Big Data Mining and Applications, Fuzhou 350118, China;
2. College of Computer and Cyber Security, Fujian Normal University, Fuzhou 350117, China)

Abstract: 【Objective】Vehicle detection is critical for urban intelligent transportation. Focusing on small target problems, high-density problems, and motion attribute problems, this study takes traffic surveillance images as input and aims to detect moving vehicles. 【Method】Based on the anchor-free CenteNet, a detection method of multi-scale features and dual-source motion perception was proposed. Firstly, coordinate attention was introduced to the multi-scale and global context features of the network's abstraction layer, so as to supplement information in multiple stages and at multiple levels and improve the model's understanding of vehicles and scenes. Secondly, through fuzzy textures representing actual motion features of vehicles and optical flow knowledge representing general motion features of vehicles, the model's perception ability of moving vehicles was constructed. 【Result】The experimental data came from the public dataset UA-DETRAC. The mean average precision (mAP) and frames per second (FPS) were used as the evaluation metrics of accuracy and speed. Experiment results

收稿日期: 2024-02-05

基金项目: 国家自然科学基金项目(62172095); 福建省科技厅自然科学基金项目(2022J01932); 福建省教育厅科技项目(JAT210283, JAT220052); 福建省创新资金项目(2022C0022); 湖南省自然科学基金区域联合基金项目(2024JJ7549)

show that the mAP and FPS of the proposed method are 70% and 30 frame/s respectively, which have the best balance between speed and accuracy among other compared methods. **【Conclusion】**It maintains that the proposed method is competent in the task of moving vehicle detection.

Key words: moving vehicle detection; CenterNet; fuzzy texture; optical flow; multi-task learning

Citation format: LI X H, LIU S J, LIU J H, et al. Study on multi-scale feature and dual-source motion perception for vehicle detection[J]. Journal of East China Jiaotong University, 2024, 41(4): 64–72.

【研究意义】车辆检测是交通管理、自动驾驶等智能交通系统^[1]的重要部分。随着人工智能技术的发展,基于深度学习的车辆检测研究成为当下的热点。**【研究进展】**现有车辆检测方法大多基于锚框。该类方法会生成大量冗余的候选框,影响系统效率;当多个目标高度重叠时,候选框的筛选过程易造成漏检。而无锚框方法则可有效避免上述问题,是当前的研究趋势。**【创新特色】**本文以运动车辆为研究对象,目标是从监控图像中精确检测出行驶的车辆,该研究可为车辆追踪^[2]、车辆违规行为判定^[3]、驾驶环境评估^[4]等下游任务提供客观的评判依据。

【关键问题】由于拍摄角度、距离、光照、行车环境的动态变化,精确识别运动车辆仍存在以下挑战:①小目标问题:距离拍摄设备较远的车辆所占图像像素少,容易因信息不足而被漏检;②高密度问题:拥挤交通场景下的监控图像中,常存在大量重叠遮挡的车辆;③运动属性问题:本工作关注行驶车辆,然而停泊或等候交通信号灯等静态车辆也具有车辆外形特征,容易造成混淆,导致误检。以上因素会造成不同程度的识别误差并逐渐累积,最终降低识别精确度。针对以上难点,本文提出一种基于无锚框策略的车辆检测方案,实现运动车辆的精准检测。

1 相关工作

1.1 挑战应对策略

针对上述三大挑战,已有方法分别采取不同策略予以应对。首先,对于小目标问题,一种朴素的想法是使用更高清晰度的输入图像,但这也意味着更重的计算负担。为了取得精度与速度的平衡,Hamzenejadi等^[5]提出一种基于YOLOv5的改进模型。然而,更换高质量输入数据的方案已超出本文的讨论范围。另一种策略是增强目标特征。为此,

Luo等^[6]将多层特征信息与上下文信息相结合,以提升模型应对小目标的鲁棒性,同时指出,该策略同样有益于解决遮挡问题。受此启发,本文方法同样关注多尺度特征及上下文信息,但在特征提取和融合方式上具有自己的特色。

在基于锚框的目标检测方法中,多尺度、高密度目标意味着算法将产生大量形状不一且高度重叠的候选框,因此锚框配置和候选框筛选策略决定了模型的检测性能。隋昕^[7]和Mao等^[8]分别针对上述两点进行优化。具体来说,隋昕^[7]使用K-means方法针对名为UA-DETRAC^[9]的公共数据集进行分析,以获取锚框的优化配置。然而该方法的有效性严重依赖数据集,泛化能力有待考察。鉴于传统的非极大值抑制(non-maximum suppression, NMS)候选框筛选策略容易造成漏检,Mao等^[8]提出一种名为软NMS的方法。而本文方法采用无锚框策略,因此不存在锚框的配置和筛选问题。

其次,对于高密度问题,Kang等^[10]指出:当遮挡出现时,目标边缘区域存在极大的不确定性;相较之下,目标中心区域更可靠。基于此论断,作者使用模糊信息熵对上述不确定性进行建模,以提升检测性能。而本文方法采用的无锚框策略更关注目标中心点的检测,具有类似的底层逻辑,因此同样有利于解决高密度问题。

最后,针对运动目标的检测,传统方法大多依赖背景差分策略,将其运用于深度学习模型中的报道并不多见,例如Fu等^[11]使用一个循环神经网络(recurrent neural network, RNN)来表征背景差分思想,用于过滤输入数据。本文方法则借助光流法和运动模糊思想对车辆的运动特征进行描述。

1.2 模型架构

根据是否使用锚框作为划分依据,可将基于深度学习的车辆检测方法分为有锚框和无锚框两

类。例如, Luo等^[6]和徐浩等^[12]所用方法的基础模型分别为Faster R-CNN^[13]和SSD^[14], 而文献方法^[5, 10, 15-17]则是建立在YOLO^[18]的各种变体之上, 它们均属于有锚框的方法。然而, 正如上文所述, 有锚框方法的准确率受锚框大小、比例以及NMS超参数的影响较大, 于是当无锚框思想被提出时, 基于无锚框的目标检测受到学术界的广泛关注。

相比有锚框方法, 无锚框方法具有时间复杂度低、召回率高等特点, 为此, 黄跃珍等^[19]和赵帅豪^[20]分别基于CenterNet^[21]和YOLOX^[22]对车辆检测进行研究, 但均未考虑车辆的运行属性等问题。表1展示了上述相关工作的更多信息, 包括来源、发表年份、基础模型等。

本文前期针对安全帽检测的工作^[23]展示了无锚框方法的优越性。本文方法选择CenterNet作为基础网络, 结合上一节所提出的思想对其进行改进, 提出一种基于多尺度双源运动感知的车辆检测方法。

2 方法

2.1 网络架构

图1展示为本文方法的网络架构。相对基础模型, 本文方法的核心改进主要体现在图1中虚线框所示的两个环节, 即主干网中名为“多尺度全局上下文提升”(multi-scale global context promotion, MSGCP)的模块, 以及网络头部名为“双源运动感知”(dual source motion perception, DSMP)的模块。

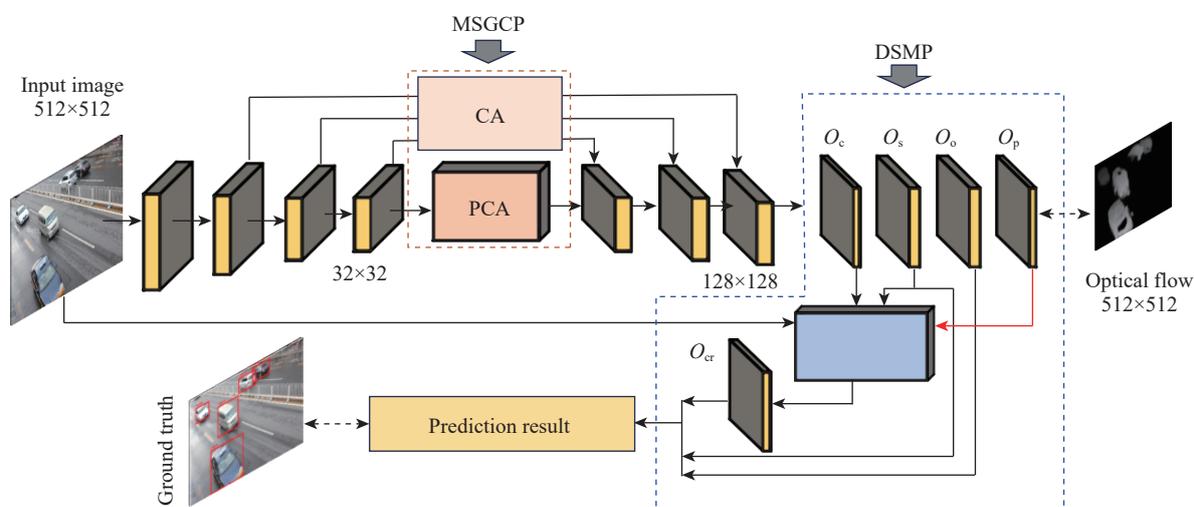


图1 本文方法的网络架构

Fig. 1 The network architecture of the proposed method

表1 车辆检测相关工作列表

Tab.1 Related work list for vehicle detection

| Reference | Source | Year of publication | Base model |
|---------------|----------------------|---------------------|--------------|
| Reference[5] | Expert Syst Appl | 2023 | YOLOv5 |
| Reference[6] | Def Technol | 2021 | Faster R-CNN |
| Reference[7] | UESTC | 2023 | YOLOv5 |
| Reference[8] | Appl Intell | 2020 | YOLOv3 |
| Reference[10] | Expert Syst Appl | 2024 | YOLOv5 |
| Reference[11] | IEEE T Image Process | 2019 | RNN+CNN |
| Reference[12] | CEAJ | 2022 | SSD |
| Reference[15] | CEAJ | 2021 | YOLOv4 |
| Reference[16] | Eng Appl Artif Intel | 2022 | YOLOv5 |
| Reference[17] | Expert Syst Appl | 2023 | YOLOv5 |
| Reference[19] | Journal of SCUT | 2021 | CenterNet |
| Reference[20] | CSA | 2022 | YOLOX |

2.2 多尺度全局上下文提升

为了增强模型的位置定位能力, 本文方法首先向编码器与解码器之间的跨层连接中加入代表位置信息的坐标注意力(coordinate attention, CA)子模块(图1)。该优化操作既实现了浅层特征的去噪, 又补充了语义较弱的目标特征。

此外, 提出一种金字塔上下文关联(pyramid context association, PCA)子模块, 并将其应用于编码器-解码器的最深层。这是因为该层特征的全局性和抽象性最高, 有利于上下文信息的关联融合。如图2所示, 该子模块由“空间多尺度池化关联”

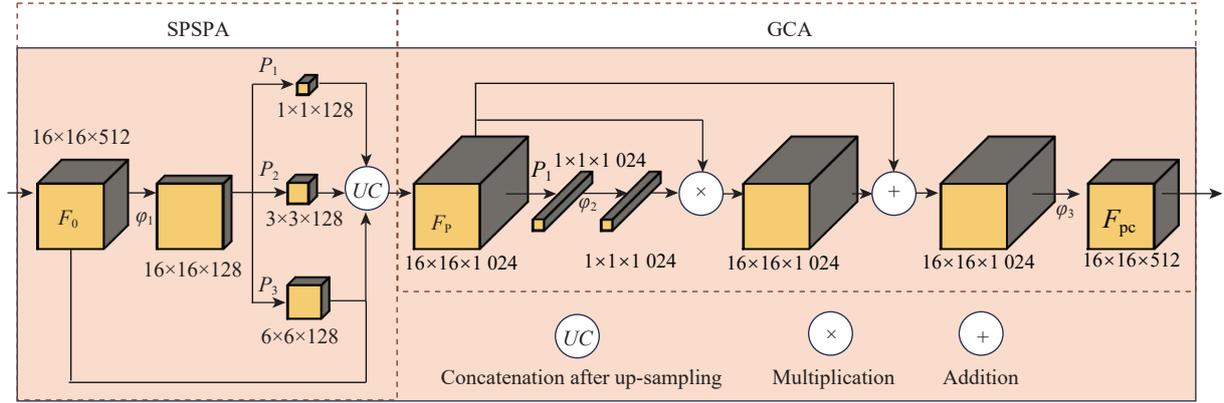


图2 金字塔上下文关联子模块示意图

Fig. 2 Demonstration of the proposed sub-module for pyramid context association

(spatial multi-scale pooling association, SPSPA) 和“全局上下文关联”(global context association, GCA) 两部分组成。

令 F_0 表示 PCA 子模块的第一层特征, 则 SPS-PA 可以表示为

$$F_p = UC\{F_0, P_1(\varphi_1(F_0)), P_2(\varphi_1(F_0)), P_3(\varphi_1(F_0))\} \quad (1)$$

式中: φ_1 为 1×1 卷积, 将 F_0 的通道数压缩到原来的 $1/4$; P_1, P_2, P_3 均为平均池化操作, 分别生成 $1 \times 1, 3 \times 3$ 和 6×6 共 3 种空间尺寸特征; UC 表示通道拼接, 将多尺度特征融合得到输出特征 F_p 。

GCA 部分以 F_p 作为输入, 经过平均池化 $P_1, 1 \times 1$ 卷积 φ_2 , 逐通道相乘 \otimes 以及逐元素相加 \oplus , 最后通过 1×1 卷积 φ_3 得到与 F_0 相同尺寸的输出特征 F_{pc} , 该过程使用公式表达如下

$$F_{pc} = \varphi_3(F_p \otimes \varphi_2(P_1(F_p)) \oplus F_p) \quad (2)$$

式中: P_1 为对全局特征的表征; φ_2 和 \otimes 一起实现通道注意力; \oplus 起到特征融合的作用; φ_3 确保 F_{pc} 与 F_0 的尺寸一致。

综上, 本模块实现了各层间、空间内、通道间的多重关联, 丰富了特征融合的形式, 补充了位置信息和相关空间信息, 提升了网络的上下文关联能力。

2.3 双源运动感知

运动特征是检测动态目标的关键之一, 为此本文提出一种基于模糊纹理和光流知识的双源运动感知方法。其中, 模糊纹理是使用输入图像结合 CenterNet 热图计算得到的 (图 3 中的“FTME”部分), 代表当前输入数据中车辆的实际运动特征; 而光流知识是模型训练阶段采用光流法进行监督学

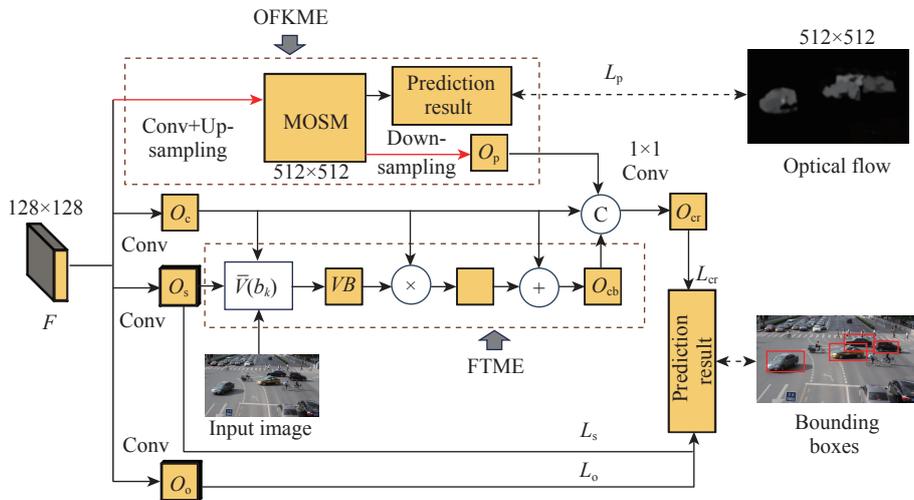


图3 基于模糊纹理和光流的运动注意力模块示意图

Fig. 3 Motion attention module based on fuzzy texture and optical flow

习得到的(图3中的“OFKME”部分),代表车辆的通用运动特征。本文方法将两者相结合,既参考个性特点,又考虑共性规律,因此能够有效构建模型的运动感知能力。

2.3.1 模糊纹理

众所周知,移动物体的成像往往伴随运动模糊现象^[24]。对于图像中的像素点来说,模糊度越大,梯度值越小^[25]。受此启发,本文方法借助梯度所表达的模糊纹理特征,解决运动感知问题。

具体来说,如图3中名为“模糊纹理映射提取”(fuzzy texture mapping extraction, FTME)的部分所示,首先依据中心点热图 O_c 和包围盒尺寸热图 O_s 确定潜在目标包围盒集合 B , 然后对原图上所有 $b_k \in B (1 \leq k \leq N)$ (N 为包围盒的个数)对应的区域计算得到 b_k 的梯度 $G(b_k)$ 为

$$G(b_k) = \frac{\sum_{(x,y) \in b_k} G(x,y)}{|b_k|} \quad (3)$$

式中: $G(x,y)$ 为原图中像素点 (x,y) 对应的梯度值; $|b_k|$ 为 b_k 区域内像素点的个数。接下来分别根据式(4)和式(5)将 $G(b_k)$ 转化为 b_k 的模糊度 $V(b_k)$, 并将其归一化至 $[0, 1]$ 区间, 得到 $\bar{V}(b_k)$ 为

$$V(b_k) = \max G(b_i) - G(b_k), \quad 1 \leq i \leq N \quad (4)$$

$$\bar{V}(b_k) = \frac{1}{1 + e^{-V(b_k)}} \quad (5)$$

然后,根据 $\bar{V}(b_k)$ 生成与原图同尺寸的模糊度特征图 VB 。对于 VB 中的任意像素点 (x,y) 来说,其模糊度 $VB(x,y)$ 可以计算得到

$$VB(x,y) = \begin{cases} \max \bar{V}(b_j), & \text{if } (x,y) \in \{b_j\} (1 \leq j \leq M) \\ 0, & \text{otherwise} \end{cases} \quad (6)$$

式中: M 为 (x,y) 所属包围盒的个数。

最后,将 VB 与 O_c 进行逐通道相乘后继续与 O_s 进行逐像素相加,实现模糊运动注意力的施加,获得中心点模糊纹理 O_{cb} 。

2.3.2 光流知识

除模糊纹理之外,本文方法还通过网络头部如图3中名为“光流知识映射提取”(optical flow knowledge mapping extraction, OFKME)的分支获取运动特征。由于该特征是在训练时使用光流法结果进行监督学习得到的,故称之为光流知识。

具体来说,以主干网络输出特征 F 作为输入,经过卷积和上采样得到一个与原图同尺寸的运动

目标分割映射(moving object segmentation mapping, MOSM)特征,其具有2个输出分支,分别是:预测运动目标分割结果和下采样得到光流知识 O_p 。随着分割任务监督学习的进行, O_p 也将得到逐步优化,为运动目标检测任务提供有力支撑。

上述整合目标检测任务和分割任务的灵感来自经典实例分割模型 Mask R-CNN^[26]。区别在于 Mask R-CNN 的分割任务是以目标检测任务为基础的,即在包围盒内进行分割;而本文方法中,两者并无从属关系,分割结果的准确性虽不是本文方法的目标,但将其用于辅助目标检测任务,确实有利于提升系统性能。

2.4 双源运动特征融合及损失函数

本文方法遵循 CenterNet 模型头部通过3张热力图确定目标检测结果的方式,不同之处在于:采用双源运动增强输出 O_{cr} 替换基础模型中的中心点输出 O_c , 即采用 O_{cr} , 包围盒尺寸输出 O_s 和偏移量输出 O_o 进行预测。

上述 O_{cr} 是将中心点模糊纹理 O_{cb} , 光流知识 O_p 以及 O_c 三者通道拼接后应用 1×1 卷积得到的单通道特征,是运动特征与中心点特征的融合。

为了进行监督学习,采用如下损失函数 L

$$L = L_p + L_{cr} + 0.1L_s + L_o \quad (7)$$

式中: L_p 为光流知识损失。令 N 表示原图像素点的个数, x_i 和 y_i 分别表示第 $i (1 \leq i \leq N)$ 个点的分割预测结果和真实值,则 L_p 定义为

$$L_p = -\frac{1}{N} \sum_{i=1}^N y_i \log(\text{sigmoid}(x_i)) + (1 - y_i) \log(1 - \text{sigmoid}(x_i)) \quad (8)$$

式中: L_{cr} 为中心点双源运动增强损失,其定义与基础模型中的中心点损失 L_c 定义类似; L_s 和 L_o 分别为包围盒尺寸损失和偏移量损失。它们的定义可参考文献[21]。

3 实验与结果

为了验证本文方法的有效性和性能,将其与权威方法进行对比。实验结果既包括以均值平均精度(mean average precision, mAP)和帧率(frames per second, FPS)作为精确度和速度评价指标的量化结果,也包括典型样本上的可视化结果。最后,对关键模块进行消融实验来展示各模块对系统性能贡献情况。所有实验均运行于一台显卡型号

为Nvidia GeForce RTX 3060(6 GB)的笔记本电脑。

3.1 实验数据和超参数设置

本文实验数据来自UA-DETRAC公共数据集^[9]。由于其原始数据时长共约10 h,视频帧率为25帧/s,本实验按照AVSS2017挑战赛^[27]的标准,以10帧为间隔从中抽取帧图片构建得到共10 314张图片的实验数据,其中训练集(Training set)、验证集(Validation set)、测试集(Testing set)分别为7 856、601张以及1 857张图片(表2)。数据集依据场景复杂性和挑战性(如遮挡、光照和交通繁忙程度等)被分为简单(Easy)、中等(Medium)、困难(Hard)3种难度,不同难度在测试集中的分布如表3所示。模型训练所采用的学习率(Learning rate)、迭代次数(Epoch)等主要超参数设置详见表4。

3.2 评价指标

本文采用mAP作为评价模型性能的关键指标。其通常定义为每个类别下平均精确度(average accuracy, AP)的平均值。由于本文方法仅有一

表2 数据集划分

Tab.2 Distribution of the dataset

| Dataset | Scale |
|----------------|--------|
| Training set | 7 856 |
| Validation set | 601 |
| Testing set | 1 857 |
| Total | 10 314 |

表3 测试集不同困难等级分布

Tab.3 Distribution of different difficulty levels in the test set

| Difficulty | Scale |
|------------|-------|
| Easy | 477 |
| Medium | 920 |
| Hard | 460 |
| Total | 1 857 |

表5 本方法与主流通用目标检测方法的对比结果

Tab.5 Comparison results between mainstream general object detection methods and the proposed method

| Method | mAP/% | | | | FPS/(frame/s) |
|--------------|--------------|--------------|--------------|--------------|---------------|
| | Easy | Medium | Hard | Overall | |
| Ours | 85.89 | 75.54 | 63.17 | 72.46 | 30 |
| Faster R-CNN | 82.75 | 63.05 | 44.25 | 58.45 | 7 |
| YOLOv5 | 71.45 | 58.59 | 38.58 | 57.69 | 64 |
| CenterNet | 72.48 | 69.17 | 57.56 | 64.22 | 33 |
| YOLOX | 76.91 | 66.71 | 56.18 | 62.32 | 35 |

表4 超参数设置

Tab.4 The settings of super parameters

| Name | Value |
|---------------------------|-----------|
| Learning rate | $5e^{-4}$ |
| Input size | 512×512 |
| Epoch | 140 |
| Batch size | 8 |
| Learning rate decay step | 90,120 |
| Learning rate decay ratio | 0.1 |
| IoU | 0.7 |

个车辆类别,因此mAP在数值上等同于AP,AP的计算式为

$$AP = \int_0^1 p(r)dr \quad (9)$$

式中: $p(r)$ 为召回率 r 对应的精确率。

3.3 对比实验

所选对比方法既包括主流通用目标检测方法,也有表1所列举的代表性车辆检测方法(详见表5和表6)。

如表5所示,本文方法(Ours)在处理简单、中等和困难模式时的 mAP 指标分别为85.89%、75.54%和63.17%,且与次优方法差距均超过3%以上;若不区分难度,本文方法的 mAP 为72.46%,也优于所有对比方法。速度方面YOLOv5最佳,Faster R-CNN表现最差,而本文方法在具有最高准确率的同时,速度依然可达30 frame/s。

相对车辆检测的主流方法而言,如表6所示,基于RNN和CNN的改进方法FG-BR_Net^[11]具有最高的准确率,本文方法次之,基于Faster R-CNN的改进方法^[6]排名第3。然而,方法YOLO-FA^[10]和NAS-Faster R-CNN^[6]均是以牺牲速度为代价的。在速度方面,YOLOv5的轻量化改进方法^[17]表现最佳,同样基于YOLOv5的改进方法^[16]和YOLO-FA^[10]排名二、三。若以30 frame/s作为实时性方法的评判标准,

表6 本方法与主流车辆检测方法的对比结果

Tab.6 Comparison results among mainstream vehicle detection methods and the proposed method

| Method | $mAP/\%$ | $FPS/(frame/s)$ |
|----------------|--------------|-----------------|
| Ours | 72.46 | 30 |
| Reference[6] | 68.25 | 7 |
| Reference [10] | 63.80 | 66 |
| Reference [11] | 79.96 | 10 |
| Reference [16] | 55.60 | 68 |
| Reference [17] | 42.90 | 72 |

则本文方法是满足实时性条件中准确率最高的方法。

综上,本文方法在保证准确率的前提下,最大程度地提高了检测速度,在准确性和速度方面具有最优的均衡性。

3.4 实验结果可视化

为直观展示本文方法与对比方法的性能差异,从测试集中选取2张典型图片进行测试,并可视化为图4的两行结果。第1行场景涵盖了多尺度且存在相互遮挡的车辆目标;第2行场景包含大量小目

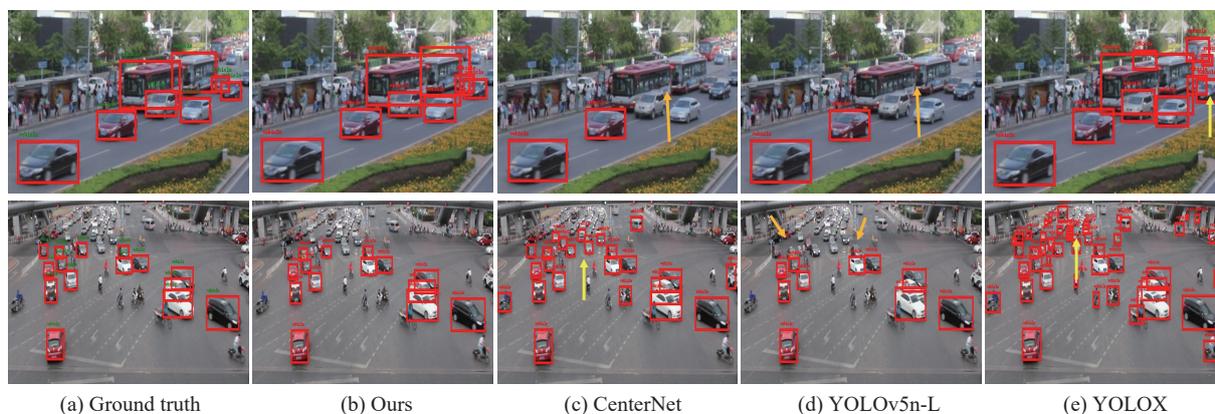


图4 可视化对比结果图

Fig. 4 Visualization of the comparison results

标和动、静态车辆混合的情况。

从图4中可知,在处理多尺度高密度场景时,CenterNet和YOLOv5n-L在黄色箭头处的车辆重叠区存在漏检,YOLOX则存在较多区域误检,本文方法则可以较好应对遮挡问题;此外,在处理小目标和动态对象时,虽然CenterNet和YOLOX比YOLOv5n-L能捕获更多小目标,但由于缺乏动态感知能力,两者都会误检远处待行的静态车辆。由于上下文关联获取能力不足,YOLOX甚至会误检行人目标。得益于第2节所提出的多尺度全局上下文提升模块和双源运动感知模块,本文方法能更精确地检测动态车辆目标。

3.5 消融实验

为评估多尺度全局上下文提升模块(Cont)和双源运动感知模块(Mov)对系统性能的贡献度,开展如表7所示的消融实验,其中√表示启用,×表示禁用。可见,当两者均禁用时,方法准确率仅为64.22%;相较双源运动感知模块,上下文模块对性能的改善稍大;当两者均启用时,准确率最高。此

表7 消融实验结果

Tab.7 Results of the ablation experiment

| Experiment | Cont | Mov | $mAP/\%$ | $FPS/(frame/s)$ |
|------------|------|-----|----------|-----------------|
| 1 | √ | √ | 72.46 | 30 |
| 2 | √ | × | 71.29 | 31 |
| 3 | × | √ | 69.75 | 32 |
| 4 | × | × | 64.22 | 33 |

外,各模块的加入对检测速度的影响较小。

4 结论

1) 优化主干网络的特征提取和融合能力。包括引入坐标注意力子模块,以及增加金字塔上下文关联子模块,使模型具有多尺度上下文感知能力,提升系统检测小目标以及高密度目标的准确率。

2) 通过运动特征建模,提出基于模糊纹理和光流知识的双源运动感知机制,实现运动车辆的检测功能。

3) 与其他主流车辆检测方法的对比表明,本文方法具有准确率高以及检测速度较快等特点,可以

胜任复杂交通环境下动态对象的实时检测。

4) 未来将进一步探索更优的运动特征提取策略及轻量化方法,并通过低照度增强等技术来缓解光照不足等问题,实现本文方法在准确度、速度及泛化能力方面的提升。

参考文献:

- [1] 杜玲丽,胡骥,赵怀明,等.考虑出行者偏好和经验的路线选择行为研究[J].公路交通科技,2019,36(5):138-144.
DU L L, HU J, ZHAO H M, et al. Study on route choice behavior considering traveler's preference and experience [J]. Journal of Highway and Transportation Research and Denelopment, 2019, 36(5): 138-144.
- [2] 郭秋蕊,李建良,田垚,等.基于改进KCF算法和多特征融合的车辆跟踪研究[J].电子测量与仪器学报,2022,36(4):231-240.
GUO Q R, LI J L, TIAN Y, et al. Research on vehicle tracking based on improved KCF algorithm and multi-feature fusion[J]. Journal of Electronic Measurement and Instrumentation, 2022, 36(4): 231-240.
- [3] 姜明新,王洪玉,沐方顺,等.基于视频监控的违章停车检测算法[J].计算机工程,2012,38(19):151-153.
JIANG M X, WANG H Y, MU F S, et al. Illegal parking detection algorithm based on video surveillance[J]. Computer Engineering, 2012, 38(19): 151-153.
- [4] 李卓,陈雨人.基于决策树方法的偏远地区山区公路驾驶安全性研究[J].华东交通大学学报,2017,34(2):29-36.
LI Z, CHEN Y R. Analysis of driving safety on mountain highway in remote areas based on decision tree method [J]. Journal of East China Jiaotong University, 2017, 34 (2): 29-36.
- [5] HAMZENEJADI M H, MOHSENI H, et al. Fine-tuned YOLOv5 for real-time vehicle detection in UAV imagery: Architectural improvements and performance boost [J]. Expert Systems with Applications, 2023: 120845.
- [6] LUO J, FANG H, SHAO F, et al. Multi-scale traffic vehicle detection based on Faster R-CNN with NAS optimization and feature enrichment[J]. Defence Technology, 2021, 17(4): 1542-1554.
- [7] 隋昕.基于深度学习的视频监控中车辆检测方法研究[D].成都:电子科技大学,2023.
SUI X. Research on vehicle detection method in video surveillance based on deep learning[D]. Chengdu: University of Electronic Science and Technology of China, 2023.
- [8] MAO Q C, SUN H, ZUO L, et al. Finding every car: A traffic surveillance multi-scale vehicle object detection method[J]. Applied Intelligence, 2020, 50(10): 3125-3136.
- [9] WEN L, DU D, CAI Z, et al. UA-DETRAC: A new benchmark and protocol for multi-object detection and tracking[J]. Computer Vision and Image Understanding, 2020, 193: 102907.
- [10] KANG L, LU Z, MENG L, et al. YOLO-FA: Type-1 fuzzy attention based YOLO detector for vehicle detection[J]. Expert Systems with Applications, 2024, 237: 121209.
- [11] FU Z, CHEN Y, YONG H, et al. Foreground gating and background refining network for surveillance object detection[J]. IEEE Transactions on Image Processing, 2019, 28(12): 6077-6090.
- [12] 徐浩,杨德刚,蒋倩倩,等.基于SSD的轻量级车辆检测网络改进[J].计算机工程与应用,2022,58(12):209-217.
XU H, YANG D G, JIANG Q Q, et al. Improvement of lightweight vehicle detection network based on SSD[J]. Computer Engineering and Applications, 2022, 58(12): 209-217.
- [13] 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 & Machine Intelligence, 2017, 39(6): 1137-1149.
- [14] LIU W, ANGUELOV D, ERHAN D, et al. SSD: Single shot multibox detector[C]//Amsterdam: Computer Vision-ECCV, 2016.
- [15] 王滢暄,宋焕生,梁浩翔,等.基于改进的YOLOv4高速公路车辆目标检测研究[J].计算机工程与应用,2021,57(13):218-226.
WANG Y X, SONG H S, LIANG H X, et al. Highway vehicle object detection based on improved YOLOv4 method[J]. Computer Engineering and Applications, 2021, 57(13): 218-226.
- [16] DONG X, YAN S, DUAN C. A lightweight vehicles detection network model based on YOLOv5[J]. Engineering Applications of Artificial Intelligence, 2022, 113: 104914.
- [17] BIE M, LIU Y, LI G, et al. Real-time vehicle detection algorithm based on a lightweight You-Only-Look-Once

- (YOLOv5n-L) approach[J]. Expert Systems with Applications, 2023, 213: 119108.
- [18] REDMON J, DIVVALA S, GIRSHICK R, et al. You only look once: Unified, real-time object detection[C]// Las Vegas: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2016.
- [19] 黄跃珍, 王乃洲, 梁添才, 等. 基于改进 CenterNet 的车辆识别方法[J]. 华南理工大学学报(自然科学版), 2021, 49(7): 94-102.
- HUANG Y Z, WANG N Z, LIANG T C, et al. Vehicle recognition method based on improved CenterNet[J]. Journal of South China University of Technology(Natural Science Edition), 2021, 49(7): 94-102.
- [20] 赵帅豪. 基于 YOLOX 的车辆检测[J]. 计算机系统应用, 2022, 31(12): 195-202.
- ZHAO S H. Vehicle detection based on YOLOX[J]. Computer Systems & Applications, 2022, 31(12): 195-202.
- [21] ZHOU X, WANG D, KRÄHENBÜHL P. Objects as points[EB/OL]. (2019-04-16) [2024-02-05]. <https://arxiv.org/abs/1904.07850>.
- [22] GE Z, LIU S, WANG F, et al. YOLOX: Exceeding YOLO series in 2021[EB/OL]. (2021-07-18) [2024-02-05]. <https://arxiv.org/abs/2107.08430>.
- [23] 黄品超, 刘石坚, 徐戈, 等. 关联增强改进的 CenterNet 安全帽检测方法[J]. 计算机工程与应用, 2023, 59(17): 250-256.
- HUANG P C, LIU S J, XU G, et al. Helmet wearing detection method based on improved CenterNet with enhanced associations[J]. Computer Engineering and Applications, 2023, 59(17): 250-256.
- [24] 黄彦宁, 李伟红, 崔金凯, 等. 强边缘提取网络用于非均匀运动模糊图像盲复原[J]. 自动化学报, 2021, 47(11): 2637-2653.
- HUANG Y N, LI W H, CUI J K, et al. Strong edge extraction network for non-uniform blind motion image deblurring[J]. Acta Automatica Sinica, 2021, 47(11): 2637-2653.
- [25] ZHOU L, MIN W, LIN D, et al. Detecting motion blurred vehicle logo in IOV using Filter-DeblurGAN and VL-YOLO[J]. IEEE Transactions on Vehicular Technology, 2020, 69(4): 3604-3614.
- [26] HE K, GKIOXARI G, DOLLÁR P, et al. Mask R-CNN [C]//Honolulu: Proceedings of the IEEE International Conference on Computer Vision, 2017.
- [27] LYU S, CHANG M C, DU D, et al. UA-DETRAC 2017: Report of AVSS2017 & IWT4S challenge on advanced traffic monitoring[C]// Lecce: 14th IEEE International Conference on Advanced Video and Signal Based Surveillance (AVSS), 2017.



第一作者: 李晓晗(1999—), 女, 硕士研究生, 研究方向为深度学习。E-mail: 307667945@qq.com。



通信作者: 刘石坚(1983—), 男, 副教授, 博士, 硕士生导师, 研究方向为计算机图形学、深度学习。E-mail: liusj2003@fjut.edu.cn。

(责任编辑: 吴海燕)