

面向车辆盲区防撞系统的交通目标智能检测

瓢正泉¹ 唐林波¹ 赵保军¹ 曾涛^{1,2}

(1. 北京理工大学信息与电子学院雷达技术研究所, 北京 100081; 2. 嵌入式实时信息处理技术北京市重点实验室, 北京 100081)

摘要: 交通目标智能检测是车辆盲区智能防撞系统中的基础技术, 该技术的研究和应用对降低交通事故损失具有重要意义。本文面向车辆盲区防撞系统设计的交通目标智能检测, 其在基础模型中融合了两个性能提升策略。将该模型应用于国内和国外道路场景检测数据集, 以验证模型在所有范围和近距离目标的检测性能。实验结果表明该模型可以对近距目标表现出较高的检测精度, 且具有较高的检测速度, 因此该模型可适用于车辆低速启动或者转弯时智能盲区防撞系统对交通目标的检测需求。

关键词: 道路场景; 目标检测; 深度网络; 智能防撞

中图分类号: TP751.1 **文献标识码:** A **DOI:** 10.16798/j.issn.1003-0530.2021.02.009

引用格式: 瓢正泉, 唐林波, 赵保军, 等. 面向车辆盲区防撞系统的交通目标智能检测[J]. 信号处理, 2021, 37(2): 242-247. DOI: 10.16798/j.issn.1003-0530.2021.02.009.

Reference format: PIAO Zhengquan, TANG Linbo, ZHAO Baojun, et al. Intelligent Detection of Traffic Targets for Vehicle Blind Zone Collision Avoidance System[J]. Journal of Signal Processing, 2021, 37(2): 242-247. DOI: 10.16798/j.issn.1003-0530.2021.02.009.

Intelligent Detection of Traffic Targets for Vehicle Blind Zone Collision Avoidance System

PIAO Zhengquan¹ TANG Linbo¹ ZHAO Baojun¹ ZENG Tao^{1,2}

(1. Radar Research Lab, School of Information and Electronics, Beijing Institute of Technology, Beijing 100081, China; 2. Beijing Key Laboratory of Embedded Real-time Information Processing Technology, Beijing 100081, China)

Abstract: Intelligent detection of traffic targets is the basic technology in the intelligent collision avoidance system for vehicle blind spots. The research and application of this technology is of great significance to reduce the loss of traffic accidents. In this paper, a detection model for vehicle blind zone collision avoidance system design, which combines two performance improvement strategies in the basic model. The model is applied to domestic and foreign road scene detection data sets to verify the detection performance of the model in all ranges and short-range targets. Experimental results show that the model can show high detection accuracy for close targets and has a high detection speed. Therefore, the model can be applied to the detection requirements of the intelligent blind zone collision avoidance system for traffic targets when the vehicle starts at low speed or when turning.

Key words: road scene; object detection; deep networks; intelligent collision avoidance

1 引言

交通事故引发巨大损失。研究发现非人为因素(如视觉盲区)和驾驶员自身因素(如不文明驾驶

行为等)是造成众多交通事故的主要因素。而根据 Daimler-Benz 公司的研究^[1]: 如果驾驶员能够提前 0.5 s 得到预警就可以避免 60% 的后方碰撞发生, 而如果能够提前 1 s 得到预警则可以避免 90% 的后

方碰撞。因此,针对由视觉盲区和驾驶员自身因素所导致的碰撞问题,开展面向视觉盲区的车辆智能预防碰撞相关技术研究,对降低由视觉盲区和驾驶员不文明驾驶行为所造成的交通事故发生频次、减少经济损失、保障人民的生命安全具有重要意义。

近些年国内外在盲区防撞上的相关应用研究主要包括前向防撞系统设计和盲区检测。其中,防撞系统设计主要包括通过利用不同传感器采集信号,然后对其进行处理和分析实现车辆前向防撞的预警。盲区检测则主要用于车辆变道时,通过对盲区中的目标进行检测和识别以达到防撞的目的。(1)国外的防撞系统设计和盲区检测研究包括:Wei Pan 等^[2]采用激光雷达与可见光摄像头融合的方式进行道路交通指示塔的检测识别,以防止车辆进入禁止区域;Angel M 等人^[3]利用单目摄像头并结合聚类和分类算法实现对盲区车辆的检测。针对后方盲区防撞的问题,Ra M^[4]利用鱼镜头检测后方的来车,当检测到车辆出现在盲区时向驾驶员发出警报信息。Zhao 等^[5]将盲区检测问题转化为分类问题,通过设计轻量深度神经分类网络对摄像头采集的盲区图像进行检测。(2)国内的防撞系统设计和盲区检测研究包括:Yutao Ba 等人^[6]对多种传感器采集到的车辆动态、测量的距离、驾驶行为和生理特征数据进行判别分析,通过监督学习的方式预测车辆发生碰撞的风险。Zhuang Jiayuan 等人^[7]采用雷达传感器进行障碍目标的检测,并通过检测结果调整车辆航向和前进速度实现防撞的目的。Liu 等^[8]利用毫米波雷达对车辆后部盲区进行探测,所提出方法对白天和夜晚场景均有效。杨炜等^[9]设计一种车载高清红外视觉传感器的右侧盲区预警系统,该系统利用深度学习模型进行行人和骑行者的识别。杨思思^[10]则采用单目视觉技术对城市道路上车辆后视镜盲区的骑行车辆进行检测和对高速公路弯道进行识别。

视觉目标检测是智能防撞系统中的关键技术,近些年单目视觉中的目标检测领域取得了较大发展。单目视觉目标检测当前的主要研究现状是:按照所在的不同时期,目标检测主要可以分为非深度学习的检测和基于深度学习的检测方法。2012 年以前的目标检测主要基于对候选窗特征提取进行识别范式,不同检测器间的区别一般在于候选窗的生成方式和特征提取、分类方法的不同。典型成果如基于 HoG 特征和 SVM 判别的行人检测^[11];利用

通道特征的 ICF 检测器^[12]。然而由于手工设计的特征不能很好地描述和区分复杂场景中的目标,因此该阶段目标检测的性能相对较低。随着 2012 年 AlexNet 网络^[13]在图像识别任务上的性能遥遥领先当时最先进的传统识别算法,深度学习应用于目标检测领域的时代正式到来。深度学习技术与先前主流的传统机器学习技术的主要差别是:深度学习可以端到端地从数据中学习逐层抽象的特征描述,避免了人手工设计特征的繁琐和呆板。因为这些深度特征相对于传统的浅层特征具有更强的描述性能和可鉴别能力,所以基于深度学习的目标检测可以达到更高的检测精度,但当前该类方法的一大不足是其依赖大量的有标签数据。基于深度学习的目标检测按照检测模型的框架不同主要可分为两阶段检测^[14,17]、单阶段检测^[15]和无锚点检测^[16]。两阶段检测的优点是检测精度高,缺点是检测速度较慢;单阶段检测则较好地平衡了检测精度与速度;而针对基于锚点检测模型对锚点参数设置敏感和由锚点所产生的大量计算冗余等问题,发展出了无锚点检测的思想。

然而现阶段较少有检测模型直接面向车载智能防撞系统需求,因此当前主流的检测模型一般并不直接与该应用场景相匹配。面向车辆盲区防撞系统的交通目标智能检测需要综合考虑车载硬件平台计算资源的限制,其中主要包括两个方面:一方面模型的检测精度与计算复杂度需要能够较好地折中,另一方面检测网络的基本单元利于硬件平台的实现。由于现有以 Faster RCNN 为代表的两阶段检测器由于计算复杂度较高,而不适用于资源受限的车载终端平台,因此单阶段检测模型更适用于实际场景。众多单阶段检测器中,YOLOv3 可以较好地取得检测精度和较高的检测性能。此外,现阶段一些较新的网络所包含的基本单元(如深度可分离卷积、反卷积,空洞卷积等)不利于硬件平台(如华为海思 AI 芯片、FPGA 等)实现,而 YOLOv3 的网络单元主要包括基本的卷积和池化操作,因而对硬件的实现较为友好。综合上述两方面的因素,YOLOv3 作为车载平台交通目标智能检测的基础模型。在此基础上,结合该应用场景下两个方面的约束,本文构建了道路场景目标检测模型,在原始模型的基础上引入性能提升策略,并分别在国内、国外的道路场景目标检测数据集上验证了模型的计算效率和模型在全距离与近距离检测的性能。

2 方法

2.1 出发点

考虑到道路场景的复杂性和计算平台的资源限制,检测模型既需要保持较高的计算效率和检测性能,又需要有利于现有硬件平台的实现。根据先前引言部分介绍选择 YOLOv3^[24] 作为原始检测模型,在此基础上分别于模型的训练阶段引入了其他策略,使得在不提高计算复杂度的情况下,进一步提升检测模型的性能。通过在模型训练阶段采用新型的数据增强方式,以增大训练数据集的多样性;通过使用性能更好的目标定位损失计算方式,以帮助模型更好地学习。由于所引入的策略分别作用于模型的训练阶段的数据增强和损失计算,因而不会在预测阶段引入额外的计算消耗,且可以保持原有模型的硬件移植特性。

2.2 模型概述

本文选择 YOLOv3^[24] 作为原始检测模型,模型的结构示意图如下图 1。

模型的损失函数为:

$$\begin{aligned} \text{loss} = & \lambda_{\text{coord}} \sum_i (1 - \text{GIoU}(\text{gbox}_i, \text{pbox}_i)) + \\ & \lambda_{\text{cls}} \sum_i p_{ic} \log(\hat{p}_{ic}) + \lambda_{\text{obj}} \sum_i 1_i^{\text{obj}} (c_i - \hat{c}_i)^2 + \\ & \lambda_{\text{noobj}} \sum_i 1_i^{\text{noobj}} (c_i - \hat{c}_i)^2 \end{aligned} \quad (1)$$

其中, λ 为权重系数; p_{ic} 表示类别编码; c_i 为预测置信度; $1^{\text{obj}}(\cdot)$ 则为示性函数。

2.3 提升策略

(1) Mosaic 数据增强

数据增强(广)是一种常见的提升模型性能的手段,其主要思想是通过额外扩增训练数据的多样性以提高模型的泛化性能。数据增强一般只作用于模型的训练阶段,因而该策略不会对测试模型引入附加的计算复杂度。现有的数据增广方式一般

针对单张图像进行几何和颜色空间等变换来扩增训练样本,最近,一些同时融合多张图像进行数据增强的方法被相继提出,比如 Mixup^[18] 和 CutMix^[19]。而不同于 Mixup 和 CutMix 同时融合两张图像, Mosaic 的优点则是可以同时融合四张图像,因此 Mosaic 具有更大的组合空间,因而增强数据的多样性更强,检测目标的背景变化也更加丰富, Mosaic 已被应用于 YOLOv4^[20] 中。Mosaic 以随机的方式将四张图像拼接成一幅固定大小的图像,即拼图上各子区域划分是随机确定的,被拼接的每张图像对应于其中某一个子区域,拼接时不改变原图像的尺寸和宽高比,而将原图像上超出拼接子区域的部分直接截断。

(2) 利用 GIoU 计算定位损失

传统基于 L2 范数表示的定位误差不能真实地衡量预测框与真实框之间的位置偏差,针对该问题已有工作利用 IoU 计算预测框与真实框间的偏差,以更好地进行检测框的定位。但由于存在两个没有相交区域的目标框所对应的 IoU 值为 0 的情况,而该情况在训练阶段无法直接利用 IoU 产生定位损失,进而无法通过梯度信息进行框坐标预测的调整。针对该问题, GIoU 度量准则^[21] 被提出,该准则在 IoU 的基础上考虑了两框的最小外接矩形区域大小,以克服相离检测框之间 IoU 为 0 的问题。公式 (2) 给出了 GIoU 的计算方法。

$$\begin{aligned} \text{IoU} &= \frac{|\text{box}_a \cap \text{box}_b|}{|\text{box}_a \cup \text{box}_b|} \\ \text{GIoU} &= \text{IoU} - \frac{|\text{box}_{a*b} \setminus (\text{box}_a \cup \text{box}_b)|}{|\text{box}_{a*b}|} \end{aligned} \quad (2)$$

其中, $|\cdot|$ 表示检测框 box 的区域面积; $A \setminus B$ 表示在 A 中去除 B 区域; box_{a*b} 则表示包围 box_a 与 box_b 的最小面积矩形。

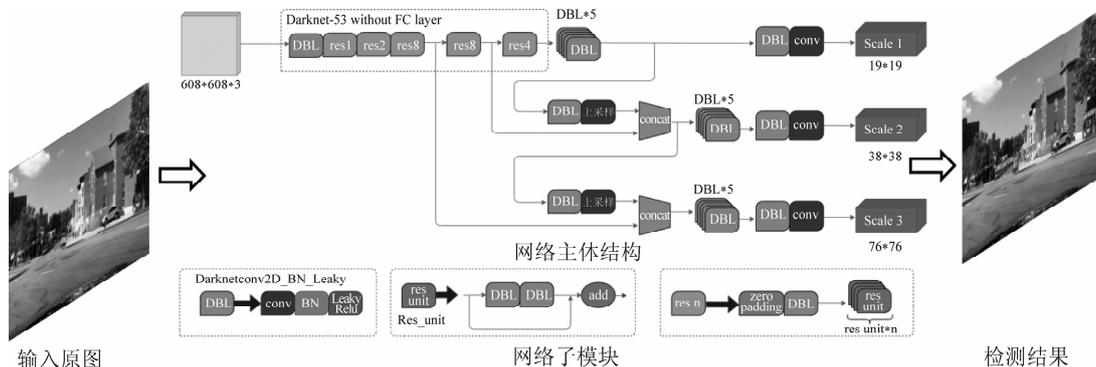


图1 检测模型结构示意图

Fig. 1 Structure of the detection model

3 实验与结果

3.1 实验基础

(1) 数据集

分别选择国内道路场景目标检测数据 ApolloScape^[22] 和国外道路场景目标检测数据 BDD100K^[23] 进行实验验证。其中 ApolloScape 是由百度公司通过在采集车车顶上安装的高分辨率相机以 1 m/帧的速率同步记录采集车周围的场景数据得到, 数据采集于多条不同道路场景和不同天气场景, 图像数据的分辨率为 3384×2710, 处理后包含训练数据 6 万张、测试数据 1237 张。BDD100K 则由伯克利大学 AI 实验室发布, 该数据集包含不同场地、不同天气和不同光照下的多种场景的采样图像, 图像分辨率为 1280×720, 经处理后包含训练数据 7 万张、测试数据 1 万张。

针对车辆低速启动或者转弯时的防撞需求, 需衡量检测模型近距离 (20 m 左右) 目标的检测性能。针对该需求, 需要对原始数据集作出额外处理。该处理的思路为采用手动点选确定需保留的近距离目标, 而那些没有被点选的目标则交由机器剔除。为了保证数据质量, 被选目标与被剔除目标间的相交区域需要机器做出判断, 然后再交由人工涂鸦的方式抹除。

(2) 实验设置

模型训练和测试的平台为单块 Tesla V100 GPU。模型的具体参数设置则主要包括, 初始学习率为 0.00579、训练轮数为 300、模型通过 SGD 方法进行学习、动量置为 0.937、权重衰减系数为 0.000484、批训练大小设置为 64; 此外, 模型训练阶段分别在其第 240 轮和第 270 将学习率衰减为当前的 0.1 倍。

3.2 实验结果

(1) ApolloScape

分别进行图像所有 (距离) 目标的检测测试和近距离目标测试, 结果如图 2 所示。



图 2 ApolloScape 上的检测结果

Fig. 2 The detection results on ApolloScape

表 1 ApolloScape 所有范围目标的检测结果 AP 值统计

	行人	汽车	骑行者	自行车/摩托	mAP
SSD	11.9	26.0	18.9	16.7	18.4
原始 YOLOv3	18.9	35.6	37.2	34.3	31.5
改进后	23.4	47.1	39.1	36.6	36.6

表 2 ApolloScape 近距离 (20 m 内) 检测结果 AP 值统计

	行人	汽车	骑行者	自行车/摩托	mAP
SSD	69.7	86.8	83.8	68.1	77.1
原始 YOLOv3	46.8	88.3	81.0	65.6	70.4
改进后	82.4	87.8	87.8	66.1	80.5

上表 1 和表 2 结果显示, 改进后的网络比原始网络达到更高的 AP 值, 改进后的模型对 ApolloScape 所有范围目标 mAP 提高 5%, 而在近距离的检测中 mAP 提高 10%。为了进一步验证网络的性能, 将其与近期在嵌入式上应用较为广泛的单阶段检测网络 SSD 进行了比较, 其中由上表 2 可以看出结果改进后的 YOLOv3 在近距离目标的检测精度要高出 3% 以上。此外可以看出, 虽然检测模型在该数据集所有范围上的目标检测率较低, 但对近距离目标的检测性能较高。由于远处目标一般尺寸较小且易受光照、干扰物等其他因素的影响而造成检测困难, 所以远处目标检测性能较低。而对于车辆低速启动或者转弯时的防撞场景, 关注的是车辆周围近距离目标的检测情况, 而那些较远距离目标的检测与否并不直接影响近距离防撞的性能。通过上述实验结果可以看出该模型检测可以在一定程度上满足该场景需求。

(2) BDD100K

所有 (距离) 目标的检测测试和近距离目标测试结果如下图 3 所示。



图 3 BDD100K 上的检测结果

Fig. 3 The detection results on BDD100K

由表3和表4可以看出,改进后模型对汽车、人和自行车的检测精度要高于改进前模型的检测性能,且对摩托车的检测性能与基础模型检测性能接近,但平均精度 mAP 要明显高于原始模型。与经典检测网络 SSD 相比,改进后的 YOLOv3 在 BDD 近距离目标的检测精度要高出 15% 左右。另一方面,近距离检测精度明显高于所有范围目标的检测性能。但由于 BDD100K 数据包含了不同天气、不同天吋和不同场地的数据,即其数据复杂度要明显高于 ApolloScape 数据集,所以 BDD100K 上的检测精度低于 ApolloScape 上的检测精度。此外还可以看出,在这两个数据集上该模型对近距汽车类目标的检测率均在 90% 以上。

表3 BDD100K 数据集所有范围目标的检测结果 AP 值统计
Tab.3 AP value statistics of detection results of all range targets on BDD100K

	人	汽车	摩托车	自行车	mAP
SSD	27.7	56.2	28.7	25.6	34.5
原始 YOLOv3	39.0	56.5	29.5	27.5	38.1
改进后	59.3	75.6	42.7	36.0	53.4

表4 BDD100K 近距离(20 m 内)检测结果 AP 值统计

Tab.4 Close distance (within 20 m) test results AP value statistics in BDD100K

	人	汽车	摩托车	自行车	mAP
SSD	56.9	90.9	57.1	58.1	65.7
原始 YOLOv3	68.5	90.3	62.2	54.2	68.8
改进后	85.7	99.0	62.0	57.9	76.2

(3) 检测速度

在单块 V100 GPU 平台上进行检测速度统计,得到单张图像的检测耗时情况如表5所示,从中可以看出其检测帧率可达 87 FPS,可满足实时性检测需求。

表5 检测速度统计

Tab.5 Detection speed statistics

图像尺寸/ pixels	网络预测/ ms	NMS/ ms	总共/ ms	帧率/ FPS
608×608	9.8	1.6	11.4	87

4 结论

针对车载智能防撞系统中道路场景目标检测的需求,构建综合考虑计算复杂度和精度的检测模型。为了提升检测性能,模型结合了 Mosaic 数据增

强策略和 GIoU 的目标定位损失计算方式。此外,分别通过国内、国外道路场景检测数据集进行了实验验证,实验结果表明该检测模型可以具有较高的检测实时性,同时对近距离目标检测的性能也可达到较高的精度,因而适用于车辆低速启动或者转弯时智能避撞时的检测需求。

参考文献

- [1] BOARD U. Vehicle-and Infrastructure-based Technology for the Prevention of Rear-end Collisions, ser[J]. Special Investigation Report. National Transportation Safety Board, 2001.
 - [2] PAN WEI, LUCAS C, TASMIA R, et al. LiDAR and Camera Detection Fusion in a Real Time Industrial Multi-Sensor Collision Avoidance System [J]. Electronics, 2018, 7(6): 84.
 - [3] SOTELO M Á, BARRIGA J. Blind spot detection using vision for automotive applications[J]. Journal of Zhejiang University-Science A, 2008, 9(10): 1369-1372.
 - [4] RA M, JUNG H G, SUHR J K, et al. Part-based vehicle detection in side-rectilinear images for blind-spot detection[J]. Expert Systems with Applications, 2018, 101: 116-128.
 - [5] ZHAO Yiming, LIN Bai, et al. Camera-Based Blind Spot Detection with a General Purpose Lightweight Neural Network[J]. Electronics, 2019, 8(2): 233.
 - [6] BA Y, ZHANG Wei, WANG Qinhu, et al. Crash prediction with behavioral and physiological features for advanced vehicle collision avoidance system[J]. Transportation Research Part C: Emerging Technologies, 2017, 74: 22-33.
 - [7] ZHUANG J, ZHANG L, ZHAO S, et al. Radar-based collision avoidance for unmanned surface vehicles [J]. China Ocean Engineering, 2016, 30(6): 867-883.
 - [8] LIU G, ZHOU M, WANG L, et al. A blind spot detection and warning system based on millimeter wave radar for driver assistance [J]. Optik-International Journal for Light and Electron Optics, 2017, 135: 353-365.
 - [9] 王敏,周树道,彭文星,等. 基于超声波传感器的汽车盲区检测系统研究[J]. 自动化技术与应用, 2017, 36(3):110-112.
- WANG Min, ZHOU Shudao, PENG Wenxing, et al. Research on Vehicle Blind Spot Detection System Based on Ultrasonic Sensor[J]. Automation Technology and Appli-

- cation, 2017, 36(3): 110-112. (in Chinese)
- [10] 杨思思. 基于单目视觉的车辆盲区预警系统的研究及实现[D]. 杭州: 浙江大学, 2015.
YANG Sisi. Vehicle Blind Spot Warning Algorithm Research and System Design based on Monocular Vision [D]. Hangzhou: Zhejiang University, 2015. (in Chinese)
- [11] DALAL N, TRIGGS B. Histograms of oriented gradients for human detection [C] // 2005 IEEE Conference on Computer Vision and Pattern Recognition (CVPR'05). Ieee, 2005, 1: 886-893.
- [12] DOLLÁR P, TU Z, PERONA P, et al. Integral channel features [C] // In: Proceedings of the British Machine Vision Conference. BMVC Press, London, 2009: 91.1-91.11.
- [13] KRIZHEVSKY A, SUTSKEVER I, HINTON G E. Imagenet classification with deep convolutional neural networks [C] // Advances in Neural Information Processing Systems, 2012: 1097-1105.
- [14] GIRSHICK R. Fast r-cnn [C] // Proceedings of the IEEE International Conference on Computer Vision, 2015: 1440-1448.
- [15] REDMON J, DIVVALA S, GIRSHICK R, et al. You only look once: Unified, real-time object detection [C] // Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2016: 779-788.
- [16] LAW H, DENG J. Cornernet: Detecting objects as paired keypoints [C] // Proceedings of the European Conference on Computer Vision (ECCV), 2018: 734-750.
- [17] 张天坤, 李汶原, 平凡, 等. 面向多源遥感图像的自适应目标检测[J]. 信号处理, 2020, 36(9): 1407-1414.
ZHANG Tiankun, LI Wenyuan, PING Fan, et al. Adaptive target detection for multi-source remote sensing images [J]. Journal of Signal Processing, 2020, 36(9): 1407-1414. (in Chinese)
- [18] ZHANG H, CISSE M, DAUPHIN Y N, et al. mixup: Beyond empirical risk minimization [J]. arXiv preprint arXiv: 1710.09412, 2017.
- [19] YUN S, HAN D, OH S J, et al. Cutmix: Regularization strategy to train strong classifiers with localizable features [C] // Proceedings of the IEEE International Conference on Computer Vision, 2019: 6023-6032.
- [20] BOCHKOVSKIY A, WANG C Y, LIAO H Y M. YOLOv4: Optimal Speed and Accuracy of Object Detection [J]. arXiv preprint arXiv: 2004.10934, 2020.
- [21] REZATOFIGHI H, TSOI N, GWAK J Y, et al. Generalized intersection over union: A metric and a loss for bounding box regression [C] // Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2019: 658-666.
- [22] HUANG X, CHENG X, GENG Q, et al. The apollo-scope dataset for autonomous driving [C] // Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops, 2018: 954-960.
- [23] YU F, CHEN H, WANG X, et al. BDD100K: A diverse driving dataset for heterogeneous multitask learning [C] // Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2020: 2636-2645.
- [24] REDMON J, FARHADI A. Yolov3: An incremental improvement [J]. arXiv preprint arXiv: 1804.02767, 2018.

作者简介



瓢正泉 男, 1994 年生, 河南固始人。北京理工大学信息与电子学院, 博士研究生, 主要研究方向为目标检测与识别。

E-mail: piaozhengquan@bit.edu.cn



唐林波 男, 1978 年生, 湖北京山人。北京理工大学信息与电子学院, 讲师, 博士, 主要研究方向为智能图像处理。

E-mail: tanglinbo@bit.edu.cn



赵保军 男, 1960 年生, 陕西周至县人。北京理工大学信息与电子学院, 教授, 博士后, 主要研究方向为信号处理。

E-mail: zbj@bit.edu.cn



曾涛 男, 1971 年生, 天津人。北京理工大学信息与电子学院, 研究员, 主要研究方向为雷达信息、信号处理与系统设计。

E-mail: zengtao@bit.edu.cn