

基于旋转框的轻量高效安全带检测算法研究

宋长明, 梁朝阳, 肖露, 宋蒙, 彩朔

(中原工学院理学院, 河南 郑州 450007)

✉ cmsongh@163.com; 2020003053@zut.edu.cn; 2020103058@zut.edu.cn;

songmeng3057@163.com; 1049636164@qq.com



摘要:交通安全监控图像中的汽车驾驶员安全带检测,可以协助交通管理部门规范驾驶员的驾驶行为。针对汽车驾驶员安全带检测中目标尺寸较小、特征对齐难度较大、检测速度较慢等问题,在 MobilenetV2 轻量化主干网络的基础上,引入 Oriented RCNN 旋转框目标检测算法,提出注意力特征融合模块(Attention Feature Fusion Module, AFFM),从而构建了一种轻量高效的端到端旋转框安全带检测算法。该算法的平均精度(AP)达到 0.905,查全率(Recall)达到 0.949,参数量(Params)仅需要 18.54 MB,端到端的检测推理速度(FPS)达到每秒 14.6 张图片。实验结果表明,该算法有效提高了监控图像中汽车驾驶员安全带检测性能,在实际应用中具备一定的竞争力。

关键词:安全带检测;旋转目标检测;注意力机制

中图分类号:TP391.41 **文献标识码:**A

Research on Lightweight and Efficient Seat Belt Detection Algorithm Based on Rotating Frame

SONG Changming, LIANG Chaoyang, XIAO Lu, SONG Meng, CAI Shuo

(College of Science, Zhongyuan University of Technology, Zhengzhou 450007, China)

✉ cmsongh@163.com; 2020003053@zut.edu.cn; 2020103058@zut.edu.cn;

songmeng3057@163.com; 1049636164@qq.com

Abstract: Driver's seat belt detection in the traffic safety monitoring image can help the traffic management department to regulate the driver's driving behavior. Aiming at the problems of small target size, difficult feature alignment and slow detection speed in the detection of automobile driver's safety belt, this paper proposes a lightweight and efficient end-to-end rotating frame safety belt detection algorithm, by introducing the Oriented RCNN rotating frame target detection algorithm and proposing the Attention Feature Fusion Module (AFFM), based on the MobilenetV2 lightweight backbone network. The average precision (AP) of the algorithm reaches 0.905, the recall rate (Recall) reaches 0.949, the parameter quantity (Params) only requires 18.54 MB, and the end-to-end detection and inference speed (FPS) reaches 14.6 images per second. The experimental results show that this algorithm effectively improves the detection performance of car driver safety belts in monitoring images, and has certain competitiveness in practical applications.

Keywords: safety belt detection; rotated object detection; attention mechanism

0 引言(Introduction)

基于道路监控的汽车驾驶员安全带检测,是智慧交通建设

中不可或缺的一环,规劝驾驶人员养成佩戴安全带的良好驾驶习惯,保障驾驶人员人身安全,是交通管理部门不可推卸的责

任。现有的安全带检测算法研究中,部分学者使用机器学习的方式,人为地选取某些特征作为安全带检测识别标志。例如,谢腾等^[1]选取方向盘与安全带构成的空间特征作为安全带图像特征,周彬^[2]选取安全带边缘直线特征作为安全带图像特征。这些方法对图片质量要求过高,并且算法使用中对门限阈值的设定受主观因素干扰较大,缺乏鲁棒性。随着深度学习被越来越多地应用在图像识别领域,在交通安全监控检测方向利用卷积神经网络进行安全带检测也成为新的发展方向^[3]。付春芬^[4]首先根据经验阈值粗定位车窗位置,其次利用积分投影定位车窗,最后利用卷积神经网络对车窗位置图片样本进行训练检测。吴天舒^[5]先使用 YOLO 目标检测方法定位驾驶员位置,再用语义分割网络检测安全带。王甄延^[6]提出一种改进的 YOLOv3 车辆检测算法、前挡风玻璃定位检测算法和安全带检测算法,实现了对乘车人员与驾驶人员位置、驾驶员及乘车人员是否佩戴安全带的检测。这些使用深度学习的安全带检测算法,都采用了分两步或三步的策略,通过先检测车、车窗或者驾驶员位置,再通过目标检测网络或图像分割网络检测安全带,不能进行端到端的训练检测,会消耗大量算力资源,并且由于采用的都是水平框检测安全带,会导致模型将目标框中大量非目标特征学习进来,出现严重的特征不对齐问题。

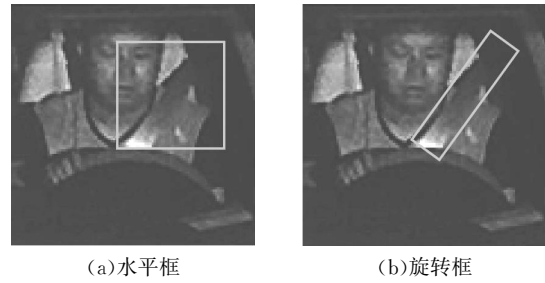
本文所构建的轻量高效的端到端的旋转框安全带检测算法,使用旋转框对齐倾斜安全带的特征区域,很好地改善了特征不对齐的问题;结合轻量化主干网络,降低了模型参数量和算力消耗,提升了模型推理速度;构建了注意力特征融合模块(AFFM),增强了安全带特征表达,提升了模型检测精度。实验结果表明,本文算法的各项指标都达到较高水平。

1 相关工作(Related work)

1.1 旋转框检测算法

普通的检测算法,使用的是水平锚框,包含较多背景区域,如图 1(a)所示,这会导致最终分类置信度与定位精度之间存在较大偏差。因此,降低锚框中背景区域的占比,可以有效改善特征不对齐的问题,更好的解决方案是使用旋转的锚框,如图 1(b)所示。近年来,旋转的锚框在交通检测领域和遥感检测领域已经有了长足的发展,如旋转的区域候选网络 RRPN,基于区域候选网络 RPN(Region Proposal Network)架构引入旋转候选框实现任意方向的场景文本检测^[7]。基于旋转的锚框得到旋转的感兴趣区域 RoI(Region of Interest),然后提取相应特征。区别于 RRPN 设置很多旋转锚框,RoI Transform 是在 RPN 阶段通过全连接从水平锚框学习得到旋转 RoI,基于旋转 RoI 提取特征,然后进行定位和分类,这种方式大大提高了基于旋转框的检测精度^[8]。为了解决小型、杂乱的旋转目标带来的检测问题,YANG 等^[9]在 Faster RCNN 的通用对象检测框架上构建了一种旋转框目标检测方法。XU 等^[10]提出了一种称为顶点滑动的旋转框检测方式,它先检测水平框,再通过学习回归水平框上的 4 个顶点滑动偏移量得到旋转框。XIE 等^[11]提出了 Oriented RCNN 旋转框目标检测算法,首先使用旋转的 RPN,以低计算成本的方式生成高质量的旋

转候选框,其次通过旋转 RoI 对其提取固定尺寸大小的特征,最后将这些特征作为旋转检测头的输入,执行分类和回归任务。



(a) 水平框 (b) 旋转框

图 1 旋转框和水平框检测效果示意图

Fig. 1 Schematic diagram of the detection effect of rotating and horizontal frames

1.2 注意力机制和特征融合网络

目前,解决目标检测中物体的尺度变化问题有两种方法。一种是直接在不同尺度的图像特征上识别目标,并使用非最大抑制组合预测结果^[12];另一种是融合卷积神经网络从不同层次图像中提取的不同尺度的特征用于近似多尺度图像特征,获得高分辨率的语义特征,从而提升预测效果^[13]。近年来,注意力机制被广泛应用在目标检测和图像分割等解决计算机视觉相关问题的网络设计上。注意力机制网络中压缩和激励网络(Squeeze and Excitation Networks, SENet)将全局空间信息挤压到通道权重信息中,捕获通道相关性^[14]。卷积注意力模块(Convolutional Block Attention Module, CBAM),可以在通道和空间维度上添加注意力机制,获得空间和通道上的重要特征^[15]。

2 旋转框安全带检测算法(Rotate frame seat belt detection algorithm)

本文针对安全带检测问题,引入旋转框目标检测算法 Oriented RCNN,提出了 AFFM,结合 MobilenetV2 主干特征提取网络,构建了一种轻量高效的端到端的旋转框安全带检测算法^[16]。基于旋转框的安全带检测算法网络结构如图 2 所示,左侧是 MobilenetV2 主干网络,用于提取图像不同层次的特征;中间部分是 AFFM,结合注意力机制融合不同尺度之间的特征;右侧部分是旋转的 RPN 模块和旋转的 RCNN 检测头模块。

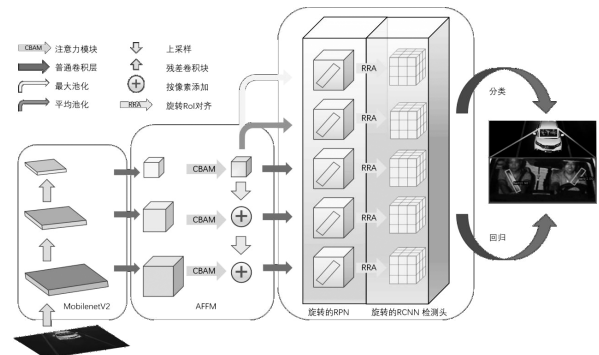


图 2 基于旋转框的安全带检测算法网络结构图

Fig. 2 Network structure diagram of seat belt detection algorithm based on rotated frame

2.1 注意力特征融合模块

针对安全带这一较小目标,普通的特征融合方式容易丢失特征,本文在金字塔特征融合网络 FPN 的基础上,结合 CBAM 卷积注意力模块,提出了一种针对小目标特征融合的注意力特征融合模块。它的主要作用是在空间位置上和特征通道上,增强目标的特征表达,从而达到提高检测精度的目的。其中, CBAM 卷积注意力模块结构如图 3 所示,输入的特征图首先经过通道注意力模块,其次将得到的通道注意力特征图与输入的特征图相乘,给不同通道施加注意力权重;以同样的方式再输入空间注意力模块,最终得到被赋予了空间和通道注意力权重的特征图。

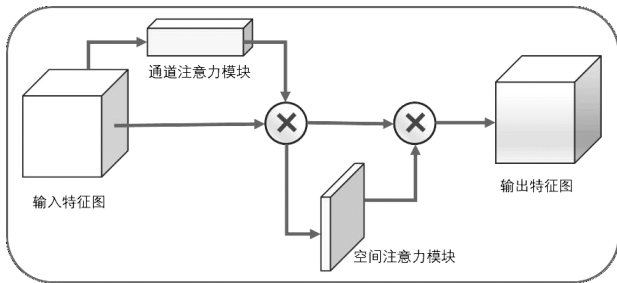


图 3 卷积注意力模块结构图

Fig. 3 Structure diagram of convolutional attention module

2.2 Oriented RCNN 旋转框目标检测

两阶段的旋转目标检测算法 Oriented RCNN,由旋转的 RPN 和旋转的 RCNN 检测头构成。首先通过旋转的 RPN 生成高质量旋转候选框,其次通过旋转 RoI 对齐提取固定大小尺寸的特征,最后将这些特征作为旋转的 RCNN 检测头的输入,执行分类和回归任务。

旋转的 RPN 是在 RPN 网络上构建的,拓展了 RPN 回归分支的输出维度,由原来的 4 个变为 6 个。如图 4(a)所示,对于每个位置的锚框,旋转的 RPN 输出为 $(x, y, w, h, \Delta\alpha, \Delta\beta)$,其中 (x, y) 为旋转候选框的中心坐标, w 和 h 表示旋转候选框的宽和高。 $\Delta\alpha$ 和 $\Delta\beta$ 表示旋转候选框外接矩形顶边和右边的偏移量。

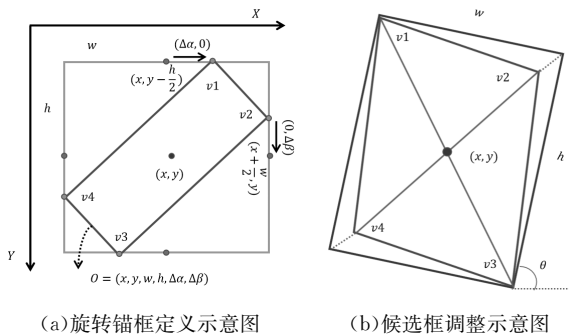


图 4 旋转框定义和候选框调整

Fig. 4 Rotate frame definition and proposal frame adjustment

旋转的 RCNN 检测头首先进行候选框调整操作,如图 4(b) 所示,将上述提到的旋转候选框的顶点坐标集 $V =$

$\{v1, v2, v3, v4\}$ 由平行四边形转换为矩形,将平行四边形较长的对角线作为矩形的对角线。然后对所有调整后的旋转候选框进行旋转 RoI 对齐操作,将得到的特征图作为旋转检测头的输入,执行分类和回归任务。

2.3 MobilenetV2 主干特征提取网络

Mobilenet 是为了将复杂的网络能够应用在移动端和资源受限情境下的轻量级神经网络的解决方案。在 Mobilenet 中提出的方案主要是深度可分离卷积和两个超参数,即宽度缩放因子、分辨率缩放因子。如图 5 所示,在 MobilenetV2 的网络设计中,除了继续使用深度可分离卷积结构,还使用了扩张层和压缩映射层。扩张层是使用 1×1 卷积将低维空间映射到高维空间,即扩大通道数,压缩映射层也是使用 1×1 卷积,它的目的是希望把高维特征映射到低维空间,即减小通道数,其中残差连接是在输入和输出的部分进行连接。

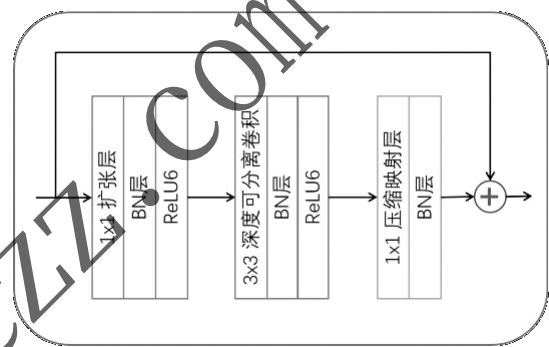


图 5 残差卷积模块结构图

Fig. 5 Structure diagram of residual convolutional module

3 实验与结果分析 (Experiments and results analysis)

3.1 数据集

本数据集包含 557 张道路监控画面,划分为训练集 281 张、验证集 184 张、测试集 92 张。每张图片的像素大小为 $1\ 600 \times 1\ 264$ 。数据集为 DOTA 格式。

3.2 实验平台及模型参数

实验使用的实验设备操作系统为 64 位的 Windows 11, CPU 选用英特尔(R)酷睿(TM) i5-12400F, GPU 选用英伟达 Ge Force GTX1080TI 11 GB, CUDA 版本为 11.3, 所有模型均使用基于 PyTorch 的 MMRotate 旋转框目标检测框架。输入图片尺寸为 $1\ 024 \times 1\ 024$, Batch_size 设置为 2, Epoch 设置为 72, 学习率为 0.005, 权重衰减参数设置为 0.000 1, 动量设置为 0.9, 使用随机梯度下降优化算法 SGD。

3.3 评价指标

通常用真实结果的正反例和预测结果的正反例,作为评价指标的评判基准,其中 TP 是被模型预测为正类的正样本,即真正例。TN 是被模型预测为负类的负样本,即真反例。FP

是被模型预测为正类的负样本,即假正例。 FN 是被模型预测为负类的正样本,即假反例。

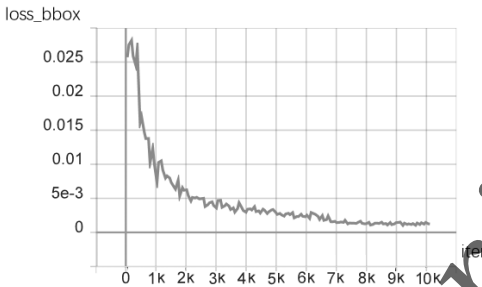
准确率($Precision$)反映了被分类器判定的正例中真正的正例样本的比重,如公式(1)所示,在目标检测场景中,无法全面衡量模型效果。查全率($Recall$)反映了被正确判定的正例占总正例的比重,如公式(2)所示。平均精度(AP)是查准率曲线查全率曲线与坐标轴所围成的面积,表示查准率与查全率的关系,是目标检测中较为常用的评价指标, AP 值越高,模型效果越好。本文采用 $Recall$ 和 AP 作为模型评价指标。

$$Precision = \frac{TP}{TP + FP} \quad (1)$$

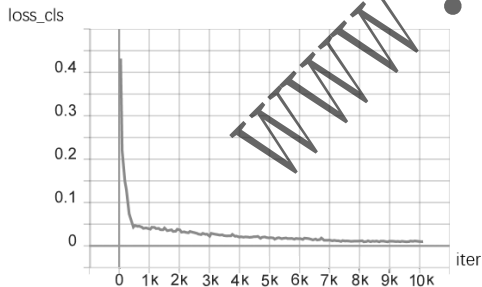
$$Recall = \frac{TP}{TP + FN} \quad (2)$$

3.4 实验结果及对比分析

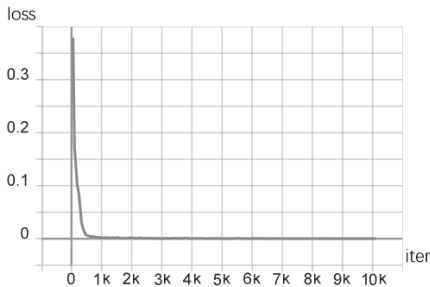
模型训练过程中,随着迭代次数的增加,参数变化如图 6 所示,总的损失函数逐步降低,学习率不断调整,准确度逐渐上升,验证集平均精度逐渐上升,并趋于稳定。以上数据表明模型已经收敛到较好的效果。



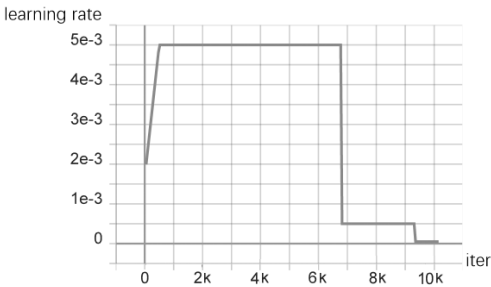
(a)回归损失



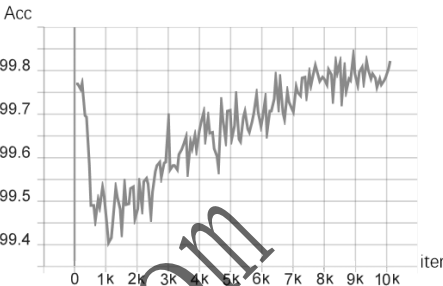
(b)分类损失



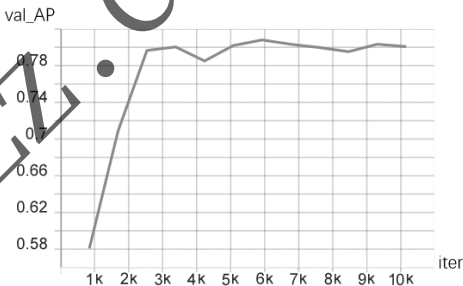
(c)总的损失



(d)学习率



(e)准确率



(f)平均精度

图 6 模型训练过程中各指标变化曲线图

Fig. 6 Change plot of each indicator during model training

本文所构建的旋转框目标检测算法与其他旋转框算法对比实验效果如表 1 所示,本文使用测试数据集和训练验证数据集作为检验数据集,实验结果表明,所建立模型在查全率($Recall$)和平均精度(AP)上均达到了较高水平。

表 1 实验数据效果对比表

Tab. 1 Comparison of experimental data

模型	测试集		训练集 + 验证集	
	$Recall$	AP	$Recall$	AP
Rotated_retinanet ^[17]	0.518	0.355	0.777	0.648
CSL ^[18]	0.579	0.395	0.786	0.662
Rotated_faster_rcnn	0.570	0.454	0.833	0.771
Redet ^[19]	0.558	0.496	0.793	0.694
S ² Anet ^[20]	0.658	0.579	0.861	0.804
Gliding vertex	0.728	0.668	0.889	0.786
R3det_tiny	0.781	0.711	0.926	0.897

模型	续表			
	测试集		训练集 + 验证集	
	Recall	AP	Recall	AP
R3det ^[21]	0.807	0.746	0.901	0.890
RoI_trans	0.816	0.790	0.926	0.899
本文方法	0.860	0.799	0.956	0.906

各个算法实验对比效果图如图 7 所示,左起第一张样例为高光反射区域,左起第二张样例较为昏暗,大部分对比算法均表现不佳;左起第三、第四及第五张样例分别是左肩条形阴影区域干扰、衣领遮挡和手势遮挡,部分对比算法均受到一定影响,锚框位置不准确,以及置信度不高。而本文算法在上述样例中均表现良好,不仅锚框的位置较为接近真实框,而且置信度也达到较高水平。

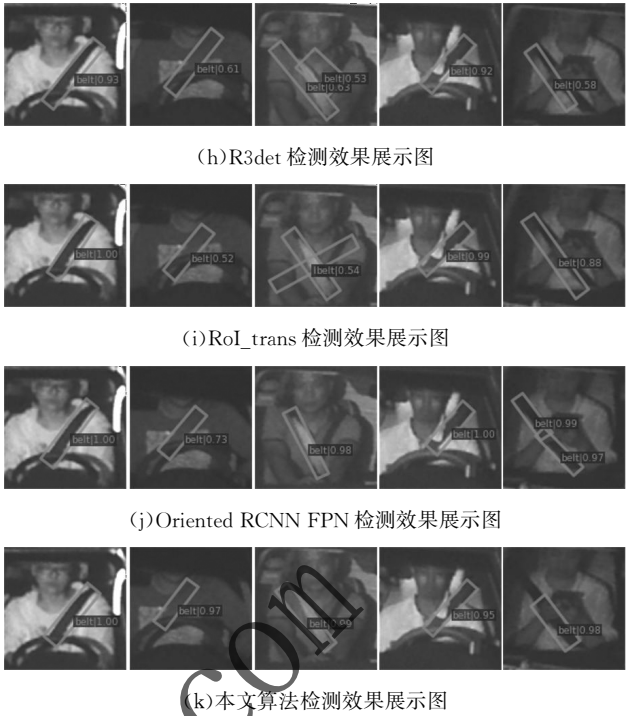


图 7 算法实验效果展示

Fig. 7 Display of the experimental effect of each algorithm

本文所构建的算法与 Oriented RCNN 算法实验效果对比如表 2 所示。本文构建的算法是基于 Oriented RCNN 的改进,第一种改进方案是采用 swin_tiny 作为主干特征提取网络,使用本文提出的 AFFM 作为特征融合模块,第一种方案在保证 Recall 和 AP 的情况下,降低了算力消耗(FLOPs)和模型参数量(Params),不足之处是受限于 swin_tiny 网络的复杂性,检测推理速度(FPS)大打折扣。第二种方案是改用 MobilenetV2 作为主干特征提取网络,该方案在保证查全率(Recall)、平均精度(AP)以及检测推理速度(FPS)的情况下,降低了算力消耗和参数量。

表2 本文方法与 Oriented RCNN 算法对比

Tab. 2 Comparison of the proposed algorithm with Oriented RCNN

模型	主干网络	融合方法	Recall	AP	FLOPs/Params/GFLOPs	MB	FPS/(img/s)
Oriented RCNN	ResNet50	FPN	0.951	0.904	198.53	41.13	14.2
本文方案 1	swin_tiny	AFFM	0.956	0.906	64.59	43.79	8.6
本文方案 2	MobilenetV2	AFFM	0.949	0.905	111.13	18.54	14.6

4 结论(Conclusion)

本文针对道路监控图像的安全带检测问题,构建了一种轻量、高效的端到端的旋转框安全带检测算法,引入旋转框,较好地改善了安全带检测特征不对齐的问题;提出 AFFM,较好地增强了安全带这一小目标的特征融合;结合 MobilenetV2 轻量化主干网络,既降低了参数量和所需算力,又提高了模型检测推理速度。经实验验证,本文构建的基于旋转框的轻量高效的

端到端安全带检测算法,在训练验证数据集上,查全率(*Recall*)达到 0.949,平均精度(*AP*)达到 0.905,参数量(*Params*)仅需要 18.54 MB,检测推理速度(*FPS*)达到每秒 14.6 张图片,满足实际的检测应用需求,具有一定的市场潜力。

参考文献(References)

- [1] 谢腾,穆平安,戴曙光,等.一种改进的安全带图像特征提取方法的应用[J].信息技术,2015,39(1):154-158.
- [2] 周彬.面向卡口监控的安全带图像检测的应用与研究[D].武汉:武汉科技大学,2018.
- [3] 吴法.图像处理与机器学习在未系安全带驾车检测中的应用[D].杭州:浙江大学,2013.
- [4] 付春芬.基于深度学习的安全带检测方法研究[D].武汉:华中科技大学,2015.
- [5] 吴天舒.基于深度学习的驾驶员安全带检测[D].沈阳:沈阳工业大学,2019.
- [6] 王甄廷.基于深度学习的汽车安全带检测[D].兰州:西北师范大学,2021.
- [7] MA J Q, SHAO W Y, YE H, et al. Arbitrary-oriented scene text detection via rotation proposals [J]. IEEE Transactions on Multimedia, 2018, 20(11): 3111-3122.
- [8] DING J, XUE N, LONG Y, et al. Learning ROI transformer for oriented object detection in aerial images [C]//IEEE/CVF. 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), Piscataway, NJ: IEEE/CVF, 2019: 2849-2858.
- [9] YANG X, YANG J R, YAN J C, et al. SCRDet: Towards more robust detection for small, cluttered and rotated objects [C]//IEEE/CVF. 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), Piscataway, NJ: IEEE/CVF, 2019: 2849-2858.
- [10] XU Y C, FU M T, WANG Q M, et al. Gliding vertex on the horizontal bounding box for multi-oriented object detection [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2021, 43(4): 1452-1459.
- [11] XIE X X, CHENG G, WANG J B, et al. Oriented R-CNN for object detection [C]//IEEE/CVF. 2021 Conference on International Conference on Computer Vision (ICCV), Piscataway, NJ: IEEE/CVF, 2021: 3500-3509.
- [12] REN S Q, HE K M, GIRSHICK R, et al. Object detection networks on convolutional feature maps [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2017, 39(7): 1476-1481.
- [13] LIN T Y, DOLLAR P, GIRSHICK R, et al. Feature pyramid networks for object detection [C]//IEEE. 2017 Conference on Computer Vision and Pattern Recognition (CVPR), Piscataway, NJ: IEEE/CVF, 2017: 936-944.
- [14] HU J, SHEN L, ALBANIE S, et al. Squeeze and excitation networks [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2020, 42(8): 2011-2023.
- [15] WOO S, PARK J, LEE J Y, et al. CBAM: Convolutional block attention module [C]//Springer. European Conference on Computer Vision, Munich, Germany: Springer, 2018: 3-19.
- [16] SANDLER M, HOWARD A, ZHU M, et al. MobileNetV2: Inverted residuals and linear bottlenecks [C]//IEEE/CVF. 2018 Conference on Computer Vision and Pattern Recognition (CVPR), Piscataway, NJ: IEEE/CVF, 2018: 4510-4520.
- [17] LIN T Y, GOYAL P, GIRSHICK R, et al. Focal loss for dense object detection [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2020, 42(2): 318-327.
- [18] YANG X, YAN J C. Arbitrary-oriented object detection with circular smooth label [C]//Springer. European Conference on Computer Vision, Glasgow, US: Springer, 2020: 677-694.
- [19] HAN J M, DING J, XUE N, et al. Redet: A rotation-equivariant detector for aerial object detection [C]//IEEE/CVF. 2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), Piscataway, NJ: IEEE, 2021: 2786-2795.
- [20] HAN J M, DING J, LI J, et al. Align deep features for oriented object detection [J]. IEEE Transactions on Geoscience and Remote Sensing, 2018, 20(11): 3111-3122.
- [21] YANG X, LIU Q Q, YAN J C, et al. R3Det: Refined single-stage detector with feature refinement for rotating object [C]//AAAI. AAAI Conference on Artificial Intelligence, Menlo Park: AAAI, 2021: 3163-3171.

作者简介:

宋长明(1965-),男,博士,教授。研究领域:偏微分方程,图像处理。

梁朝阳(1996-),男,硕士生。研究领域:图像识别。

肖露(1997-),女,硕士生。研究领域:医学图像分割。

宋蒙(1997-),女,硕士生。研究领域:医学图像分割。

彩朔(1995-),男,硕士生。研究领域:图像识别。