

·第二十五届汽车安全技术国际学术会议专题·

基于深度学习的分心驾驶行为检测方法

曹立波¹ 杨洒¹ 艾昌硕¹ 颜京才² 李旭升²

(1.湖南大学,汽车车身先进设计制造国家重点实验室,长沙 410082;2.毫末智行科技有限公司保定分公司,保定 071000)

【摘要】针对现有分心驾驶行为检测方法存在的检测精度低、实时性差等问题,利用基于深度学习的目标检测方法进行了驾驶员分心驾驶行为检测,首先构建分心驾驶行为数据集,包括驾驶员使用手机、饮水和吸烟3种行为的图像,并进行目标物的标注,然后选用轻量化目标检测模型 NanoDet 进行训练验证,结果表明,该方法可以准确并快速地识别出驾驶员在驾驶过程中使用手机、饮水和吸烟的行为。

关键词:分心驾驶 目标检测 数据集标注 轻量化模型

中图分类号:U492.8;TP391.41 **文献标识码:**A **DOI:** 10.19620/j.cnki.1000-3703.20230262

Distracted Driving Behavior Detection Based on Deep Learning

Cao Libo¹, Yang Sa¹, Ai Changshuo¹, Yan Jingcai², Li Xusheng²

(1. State Key Laboratory of Advanced Design and Manufacturing for Vehicle Body, Hunan University, Changsha 410082; 2. Baoding Branch of Haomo Technology Co., Ltd., Baoding 071000)

【Abstract】To address some of the problems in existing distracted driving behavior detection methods, such as low detection accuracy and poor real-time performance, a deep learning-based target detection method was used for driver distracted driving behavior detection. Firstly, a distracted driving behavior dataset was constructed, including images of drivers using mobile phones, drinking water and smoking, and the targets were annotated, secondly a lightweight target detection model NanoDet was selected for training and validation. The results show that the method can accurately and quickly identify driver behaviors including using mobile phones, drinking water and smoking while driving.

Key words: Distracted driving, Target detection, Dataset annotation, Lightweight model

【引用格式】曹立波,杨洒,艾昌硕,等.基于深度学习的分心驾驶行为检测方法[J].汽车技术,2023(6):49-54.

CAO L B, YANG S, AI C S, et al. Distracted Driving Behavior Detection Based on Deep Learning[J]. Automobile Technology, 2023(6): 49-56.

1 前言

在“人-车-路”复杂道路交通系统中,驾驶员行为是影响交通安全的重要因素之一。中国社科院与某保险机构的相关调查结果表明,34%的受访者在驾驶过程中存在分心驾驶的行为^[1]。此外,美国高速公路安全管理局的报道显示,美国的交通事故中25%~30%的案例与分心驾驶行为有关^[2]。因此,如果能够实时监测驾驶员的分心驾驶行为并进行预警,将对提升道路交通安全具有重要意义。

目前,分心驾驶检测方法主要分为间接检测法和直接检测法。间接检测法即通过传感器检测车辆的行驶状态特征,如加速度、运行轨迹和转向盘转角变化等,从而间接获取驾驶员的分心状态。间接检测法的成本相对较低,且不影响驾驶员的操作行为,但检测精度易受驾驶条件和驾

驶员驾驶习惯等影响,研究方法和准确性均需改进^[3-4]。直接监测法即利用驾驶员生理参数或视觉特征直接获取其状态数据。基于生理参数的监测方法通过穿戴医疗设备获取驾驶员的心电信号、脑电信号和肌电信号,提取分析信号信息,获取驾驶员状态^[5],该方法准确率较高,但相关设备成本较高,且可能影响正常驾驶,实际应用难度较高。

随着计算机视觉和深度学习技术的发展,可以通过输入图像进行端到端的分心驾驶行为检测^[6-8]。黄向康^[9]采用改进后的多任务级联卷积神经网络进行驾驶员人脸检测,得到人脸检测框和5个特征点,根据检测结果进行分心行为判别,其中包括通话和吸烟2种行为,检测准确率分别为93.3%和92.4%。庞再统^[10]针对图像分类网络 MobileNetV3 进行改进,提出轻量化模型 SMobileNet,对分心驾驶行为进行识别以实现实时监测。

本文采集分心驾驶数据集,包括驾驶员饮水、使用手机和吸烟3种分心驾驶行为的图像,利用预训练的Yolov5^[11](You Only Look Once)模型得到分心驾驶行为对应目标物的标注文件,最后,采用一种轻量化的目标检测网络NanoDet^[12]进行训练,以便在移动端进行部署。

2 数据集获取及标注

2.1 数据集图像获取

数据集图像来源包括多视角、多模态和多光谱驾驶员行为数据集(Multiview, Multimodal and Multispectral Driver Action Dataset, 3MDAD)和驾驶员监测数据集(Driver Monitoring Dataset, DMD),采集19位驾驶员饮水、使用手机和吸烟的视频,对视频进行裁剪并逐帧输出以便得到图像集合。其中,3MDAD数据集共包含16类驾驶行为,本文提取驾驶员使用右手发短信、使用左手发短信、使用右手通话、使用左手通话、拍照、使用右手饮水、使用左手饮水、吸烟类别的图像。其中,使用右手发短信、使用左手发短信、使用右手通话、使用左手通话、拍照统一为使用手机类行为;使用右手饮水、使用左手饮水统一为饮水类行为。整理后的数据集包括驾驶员不同角度的饮水、使用手机和吸烟3类图像,如表1和图1所示。

表1 数据集图像数量

分心驾驶行为	图像数量/张
饮水	36 738
使用手机	72 648
吸烟	35 039



图1 分心驾驶行为图像示例

2.2 数据集标注

本文采用基于COCO^[13]数据集预训练的Yolov5快速标注方法对数据集图像进行标注。

具有不同网络深度和宽度的5种Yolov5模型在COCO val2017数据集上的性能如表2所示。综合考虑检测精度和推理速度,选取Yolov5l作为推理时的权重文件。

表2 Yolov5的5种网络模型

模型名称	图像尺寸/像素	验证精度 mAP@0.5:0.95/%	验证精度 mAP 0.5/%	推理时间/ms	参数量/ 10 ⁶ 个
Yolov5n	640	28.0	45.7	45	1.9
Yolov5s	640	37.4	58.6	98	7.2
Yolov5m	640	45.4	64.1	224	21.2
Yolov5l	640	49.0	67.3	430	46.5
Yolov5x	640	50.7	68.9	766	86.7

注:mAP(mean Average Precision)为平均精度均值

COCO数据集共包含80个数据类别,本文只需检测瓶子(Bottle)、高脚杯(Wine glass)、茶杯(Cup)和手机(Cellphone)共4类目标物。将瓶子、高脚杯、茶杯统一为水瓶类数据,用于得到驾驶员饮水行为目标检测物的标注文件;手机用于得到驾驶员使用手机行为目标检测物的标注文件。此外,COCO数据集不包含香烟类别,需要对香烟类别进行单独预训练,用于得到驾驶员吸烟行为目标检测物的标注文件。

Yolov5模型推理得到的txt文件包括目标物的种类信息、检测框中心点归一化坐标信息和检测框归一化宽高信息。需要对推理过程中出现的冗余检测、错误检测和遗漏检测等情况进行剔除、修正或重复推理,如图2所示。



图2 数据图像推理结果示例

最终整理得到的正确推理的图像数量如表3所示。将数据集的每一类驾驶行为随机分为训练集和验证集。其中,训练集占数据集总量的96%,验证集占数据集总量的4%。之后对正确的txt文件进行格式转换得到json格式的数据集标注文件。

表3 带有标注的分心驾驶行为数据集

标注编号	分心驾驶行为	检测目标物	图像数量/张
0	饮水	水瓶	34 019
1	通话	手机	47 046
2	吸烟	香烟	32 001

3 模型训练及结果分析

3.1 NanoDet轻量化模型

Yolo、单次多边框检测器(Single Shot Multibox Detector, SSD)、区域卷积神经网络(Region-based Convolutional Neural Network, R-CNN)等目标检测模型可以实现较好的精度和速度,但模型相对较大,不适合在移动端嵌入式设备上部署。而轻量级目标检测模型NanoDet分别对主干网络、特征融合网络和检测头网络进行轻量化,具有检测速度快、模型文件小等优点,并且提供了新卷积神经网络(New Convolutional Neural Network, NCNN)推理框架下的C++代码,可直接在移动端进行部署。NanoDet模型和Yolo系列模型在COCO val2017数据集上的性能对比如表4所示,NanoDet模型具有参数量小、模型尺寸小和推理速度快的优势。因此,本文选用NanoDet作为训练模型。

表4 NanoDet模型和Yolo模型性能比较

模型名称	图像尺寸/像素	验证精度 mAP0.5:0.95/%	推理时间 (4x A76) /ms	参数量 /10 ⁶ 个	模型尺寸 (NCNN bin)/MB
NanoDet-m	320	20.6	10.23	0.95	1.8
NanoDet-m	416	21.7	16.44	0.95	1.8
Yolov4-Tiny	416	21.7	32.81	6.06	23.0
Yolov5n	640	28.4	44.39	1.90	3.8

NanoDet是单阶段的无锚框检测模型,该设计基于全卷积单阶段(Fully Convolutional One-Stage, FCOS)检测算法,NanoDet模型结构如图3所示。

NanoDet的主干网络使用ShuffleNetV2^[14],去掉了最后一层卷积,并且抽取8倍、16倍、32倍下采样的特征输入特征融合网络。ShuffleNetV2通过通道拆分操作,将通道均分为2个分支,一个分支进行卷积操作,并使输入和输出通道数量相等,可以最小化内存访问成本,另

一个分支进行恒等映射,之后将2个分支的通道连接,最后进行通道重组操作,增加通道间的信息交流。

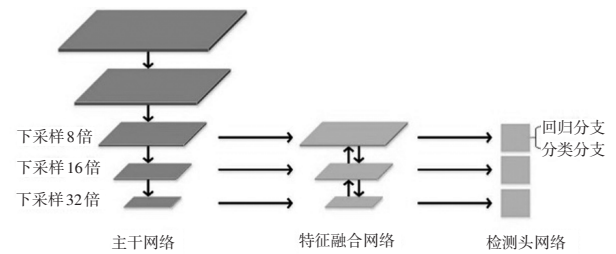


图3 NanoDet模型结构

NanoDet的特征融合网络使用路径聚合网络(Path Aggregation Network, PANet)^[15]进行浅层特征与深层特征的融合。卷积网络中,高层特征感受野大、语义信息丰富,有利于分类;相反,低层特征细节信息丰富、目标位置准确,有利于定位。因此,特征金字塔网络(Feature Pyramid Networks, FPN)^[16]采用自下而上、自上而下以及横向连接来构造特征。自下而上是卷积网络前向传播的过程,选取每个阶段的最后一层特征图构建特征金字塔;自上而下是将上层特征图进行上采样得到新的特征图,新的特征图尺度需要与下层特征图保持一致;横向连接是将新的特征图和原始的下层特征图中每个对应元素相加,实现上层特征和下层特征融合。最后,对融合后的每层特征图进行卷积处理,从而得到融合更加充分的特征图。PANet在FPN后增加了自底向上的通路,从而缩短了低层与顶层特征之间的信息路径。此外,NanoDet为了使模型更加轻量化,使用插值替代卷积进行特征图尺度的缩放。

Nanodet基于FCOS系列的共享权重检测头进行优化,由于移动端模型推理由CPU进行计算,共享权重并不会加速推理过程,而且在检测头非常轻量化的情况下,共享权重使得其检测能力进一步下降,因此每一层特征使用一组卷积。并且使用批归一化(Batch Normalization, BN)替代组归一化(Group Normalization, GN)作为归一化的方式,BN在推理时能够将其归一化的参数直接融合进卷积中,可以省去这一步计算。此外,为了将其轻量化,选择使用深度可分离卷积^[17]替换普通卷积减少参数量,并将卷积堆叠的数量减少到2组。

损失函数分为2个部分。一部分是广义交并比损失(Generalized Intersection over Union Loss, GIoULoss)^[18]对应目标预测框的输出,GIoULoss相对于交并比损失(Intersection over Union Loss, IoULoss),在解决预测框和真实框不重合的问题时,能够更好地反映预测框和真实框的关系,可在目标检测任务中取得更好的结果:

$$L_{\text{IoU}} = 1 - \left(\frac{|A \cap B|}{|A \cup B|} - \frac{C - (A \cup B)}{C} \right) \quad (1)$$

式中, A 为预测框; B 为真实框; C 为框 A 和框 B 的最小包围框。

另一部分是广义焦点损失 (Generalized Focal Loss, GFL)^[19], 可细分为分布焦点损失 (Distribution Focal Loss, DFL) 和质量焦点损失 (Quality Focal Loss, QFL)。DFL 对应目标预测框的输出, 考虑到真实的分布通常与标注位置接近, 使网络能够快速地聚焦到标注位置附近的数值。QFL 对应目标类别输出, QFL 对焦点损失 (Focal Loss)^[20] 进行改进, 联合表示类别和定位质量, 即将焦点损失中的离散类别标签转换为 0~1 范围内的连续质量标签:

$$L_Q(\sigma) = -|y - \sigma|^\beta ((1 - y) \log(1 - \sigma) + y \log(\sigma)) \quad (2)$$

式中, y 为 0~1 范围内的质量标签值; σ 为预测值; β 为缩放因子的超参数, 用于控制降权的速率。

3.2 训练环境及超参数

训练过程中图像的尺寸为 320 像素×320 像素; 设置

训练轮次数 (Epochs) 为 300; 批训练大小 (Batch-Size) 为 128; 优化器使用带动量的随机梯度下降法, 如果当前收敛效果好, 即可加速收敛, 如果收敛差, 则减慢其步伐, 其中动量设为 0.9, 权值衰减设为 0.000 1; 使用多步长调整学习率, 对于不同的训练阶段使用不同的学习率, 一方面可以加快训练的过程, 另一方面可以加快网络收敛, 其中初始学习率为 0.14, 衰减系数为 0.1。此外, 训练环境设置如表 5 所示。

表 5 训练环境设置

操作系统	Ubuntu 18.04
编程语言	Python 3.7
深度学习框架	Pytorch 1.4.0
CPU	Intel®(R) Core™ i9-7900x CPU @ 3.30 GHz
GPU	NVIDIA TITAN Xp
Cuda	10.0

3.3 试验结果分析

对训练过程进行可视化分析, 在训练集和验证集上的损失曲线如图 4 所示。

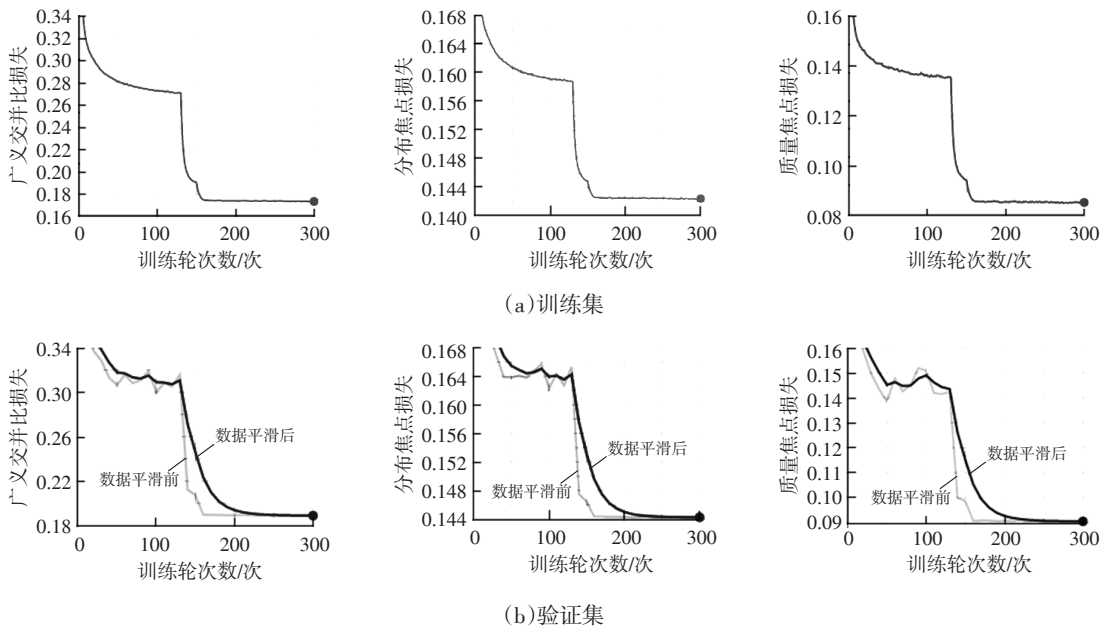


图 4 模型在训练集和验证集上的训练损失曲线

为了评价模型的目标检测性能, 本文采用 mAP 作为评价指标, mAP 是对每类的平均精度 (Average Precision, AP) 求均值:

$$E_{\text{mAP}} = \sum_{i=1}^C E_{\text{AP}_i} / C \quad (3)$$

其中, AP 是精准率-召回率 (PR) 曲线下的面积, PR 曲线以不同置信度阈值下的召回率 (Recall) R 为横坐标, 精准率 (Precision) P 为纵坐标得到:

$$E_{\text{AP}} = \int_0^1 p(r) dr \quad (4)$$

$$R = \frac{N_{\text{TP}}}{N_{\text{TP}} + N_{\text{FN}}} \quad (5)$$

$$P = \frac{N_{\text{TP}}}{N_{\text{TP}} + N_{\text{FP}}} \quad (6)$$

式中, N_{TP} 为真正例数量; N_{FP} 为假正例数量; N_{FN} 为假反例数量; $p(r)$ 为 PR 曲线; C 为目标类别数量。

NanoDet 模型在验证集上的 mAP 可达到 0.833 9, 如图 5 所示。

此外, 对模型的检测效果和推理速度进行测试, 检

测效果如图6所示,测试速度如表6所示。

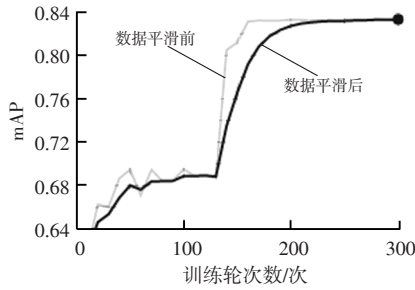


图5 模型验证集mAP曲线



图6 分心驾驶检测效果示例

表6 模型推理速度测试

模型名称	图像尺寸 /像素×像素	图形处理单元(GPU)	推理时间/ms
NanoDet	1 080×1 920	NVIDIA GeForce RTX 3050	7.63

4 结束语

本文采集驾驶员分心驾驶行为的数据集,包括饮水、使用手机和吸烟,利用在COCO数据集上预训练的Yolov5模型得到分心驾驶行为对应目标物的标注文件代替人工标注,采用轻量化的目标检测网络NanoDet进行训练,在验证集上的平均精度均值达到0.833 9,且推理速度和测试准确率皆能达到较好的效果。

参考文献

- [1] 王晓峰,陆明光,徐莹莹.分心驾驶成为交通出行潜在“杀手”[N].宁波日报,2022-04-19(A03).
WANG X F, LU M G, XU Y Y. Distracted Driving as A Potential “Killer” on the Road[N]. Ningbo Daily, 2022-04-19(A03).
2023年 第6期

- [2] 马艳丽,顾高峰,高月娥,等.基于驾驶绩效的车载信息系统操作分心判定模型[J].中国公路学报,2016,29(4):123-129.
MA Y L, GU G F, GAO Y E, et al. A Distraction Determination Model for In-Vehicle Information System Operation Based on Driving Performance[J]. Chinese Journal of Roads, 2016, 29(4): 123-129.
- [3] 郭城,姚保峰,蔡绍峰.分心驾驶监测技术研究进展[J].安阳工学院学报,2022,21(2):47-49.
GUO C, YAO B F, CAI S F. Advances in Distracted Driving Monitoring Technology[J]. Journal of Anyang Institute of Technology, 2022, 21(2): 47-49.
- [4] 陈仁祥,胡超超,胡小林,等.基于改进YOLOv5的驾驶员分心驾驶检测[J/OL].吉林大学学报(工学版):1-10(2022-09-06)[2022-10-08].
<https://www.cnki.net/KCMS/detail/detail.aspx?dbname=CJFDTotal&filename=JL-GY20220904008&v=MzI0MzlsVjdyQkpGaz1MeUhNZDdH-NEhOUE1wbzlcWk9zSF13azd2QkFTNmpoNFRBemxxM-kEwZkxUN113cWRaT1pxRnkV.>
CHEN R X, HU C C, HU X L, et al. Driver Distracted Driving Detection Based on Improved YOLOv5[J/OL]. Journal of Jilin University (Engineering Edition) :1-10(2022-09-06)[2022-10-08].
<https://www.cnki.net/KCMS/detail/detail.aspx?dbname=CJFDTotal&filename=JL-GY20220904008&v=MzI0MzlsVjdyQkpGaz1MeUhNZDdH-NEhOUE1wbzlcWk9zSF13azd2QkFTNmpoNFRBemxxM-kEwZkxUN113cWRaT1pxRnkV.>
- [5] 关伟,杨柳,江世雄,等.脑电在交通驾驶行为中的应用研究综述[J].交通运输系统工程与信息,2016,16(3):35-44.
GUAN W, YANG L, JIANG S X, et al. A Review of Research on the Use of EEG in Traffic Driving Behaviour[J]. Transport Systems Engineering and Information, 2016, 16(3): 35-44.
- [6] 罗冠泰,陈新伟,肖旺辉,等.基于卷积神经网络的分心驾驶行为识别[J].机电技术,2022(1):33-37.
LUO G T, CHEN X W, XIAO W H, et al. Distracted Driving Behaviour Recognition Based on Convolutional Neural Networks[J]. Electromechanical Technology, 2022(1): 33-37.
- [7] 秦斌斌,彭良康,卢向明,等.司机分心驾驶检测研究进展[J].计算机应用,2021,41(8):2330-2337.
QIN B B, PENG L K, LU X M, et al. Advances in Driver Distracted Driving Detection Research[J]. Computer Applications, 2021, 41(8): 2330-2337.
- [8] 徐毅.基于轻量化卷积神经网络的分心驾驶行为检测与嵌入式实现[D].徐州:中国矿业大学,2021.
XU Y. Lightweight Convolutional Neural Network-Based Distracted Driving Behaviour Detection with Embedded Implementation[D]. Xuzhou: China University of Mining and Technology, 2021.

- [9] 黄向康. 基于深度学习的分心驾驶行为检测研究[D]. 重庆: 重庆邮电大学, 2021.
HUANG X K. Research on Distracted Driving Behaviour Detection Based on Deep Learning[D]. Chongqing: Chongqing University of Posts and Telecommunications, 2021.
- [10] 庞再统. 基于深度学习的分心驾驶行为识别研究[D]. 成都: 西南交通大学, 2021.
PANG Z T. Research on Distracted Driving Behaviour Recognition Based on Deep Learning[D]. Chengdu: Southwest Jiaotong University, 2021.
- [11] Ultralytics. Yolov5[EB/OL]. (2020.10.21)[2022.10.8]. <https://github.com/ultralytics/yolov5/tree/master>.
- [12] RANGILYU. Nanodet[EB/OL]. (2021.12.27) [2022.10.8]. <https://github.com/RangiLyu/nanodet>.
- [13] LIN T Y, MAIRE M, BELONGIE S, et al. Microsoft COCO: Common Objects in Context[C]// Computer Vision—ECCV 2014, Cham: Springer International, 2014: 740–755.
- [13] MA N N, ZHANG X Y, ZHENG H T, et al. ShuffleNet V2: Practical Guidelines for Efficient CNN Architecture Design [C]// Computer Vision – ECCV 2018, Cham: Springer International, 2018: 122–138.
- [15] LIU S, QI L, QIN H F, et al. Path Aggregation Network for Instance Segmentation[C]// 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), Salt Lake City, UT, USA: IEEE, 2018.
- [16] LIN T Y, DOLLÁR P, GIRSHICK R B, et al. Feature Pyramid Networks for Object Detection[C]// 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Honolulu, HI, USA: IEEE, 2017.
- [17] HOWARD A G, ZHU M L, CHEN B, et al. MobileNets: Efficient Convolutional Neural Networks for Mobile Vision Applications[EB/OL]. (2017–04–17)[2023–04–27]. <https://arxiv.org/abs/1704.04861>.
- [18] REZATOFIGHI H, TSOI N, GWAK J Y, et al. Generalized Intersection over Union: A Metric and A Loss for Bounding Box Regression[C]// 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), Long Beach, CA, USA: IEEE, 2019.
- [19] LI X, WANG W H, WU L J, et al. Generalized Focal Loss: Learning Qualified and Distributed Bounding Boxes for Dense Object Detection[EB/OL]. (2020–06–08)[2023–04–27]. <https://arxiv.org/abs/2006.04388>.
- [20] LIN T Y, GOYAL P, GIRSHICK R, et al. Focal Loss for Dense Object Detection[J]. IEEE Transactions on Pattern Analysis & Machine Intelligence, 2020, 42(2): 318–317.

(责任编辑 王一)

修改稿收到日期为2023年4月4日。

《汽车工程师》征稿启事

《汽车工程师》于1974年创刊,是由中国第一汽车集团有限公司主管、中国第一汽车股份有限公司主办的、国内外公开发行的汽车应用技术类学术期刊,国家新闻出版广电总局第一批认定的学术期刊。

《汽车工程师》以报道汽车产品开发领域的成果为主,专注于技术开发与应用,提供汽车设计创新解决方案。《汽车工程师》将把握电动化、智能化、网联化、共享化的汽车技术主流发展趋势,努力在电池技术、电驱技术、电控技术、高压技术、补能技术、燃料电池技术、智慧控制、智能驾驶、智享座舱、智能悬架、线控转向、线控制动、NVH、功能安全、预期功能安全、信息安全、数据安全、被动安全、高效动力、高效传动、智能管理、低风阻、低滚阻、轻量化领域吸收优质稿源,为我国汽车工程技术创新能力提升贡献力量。

热忱欢迎汽车行业的专家学者不吝赐稿,反映国家重点扶持项目、自然科学基金项目和其他重点项目等研究成果的稿件将优先发表,我们期待与您共同践行“把论文写在祖国大地上”的指示精神,为强大中国汽车工业作贡献!

《汽车工程师》编辑部