

标签结合现实场景的交通标志分类检测研究*

张成,张瑞宾,王曙道

(桂林航天工业学院 汽车与交通工程学院,广西 桂林 541004)

摘要: 交通标志在车辆的安全行驶和自动驾驶中都有着大量的研究。由于交通标志的种类繁多且受各种因素的影响,交通标志的分类检测也是一个具有挑战的难题。为此,提出了一种标签结合现实道路场景的交通标志分类检测方法,该方法分为数据生成部分和目标检测部分。实验结果表明,利用该方法生成训练数据,能够有效地训练深度卷积神经网络,实现现实场景交通标志的分类检测,并且优化的检测模型相比文中提到的模型具有更小的体积和更快的速度。

关键词: 交通标志;自动驾驶;数据增强;深度卷积神经网络;检测

中图分类号: TP391

文献标识码: A

DOI:10.16157/j.issn.0258-7998.211840

中文引用格式: 张成,张瑞宾,王曙道. 标签结合现实场景的交通标志分类检测研究[J]. 电子技术应用, 2022, 48(3): 27-31, 36.

英文引用格式: Zhang Cheng, Zhang Ruibin, Wang Shudao. Research on classification and detection of traffic signs based on tags combined with real scenes[J]. Application of Electronic Technique, 2022, 48(3): 27-31, 36.

Research on classification and detection of traffic signs based on tags combined with real scenes

Zhang Cheng, Zhang Ruibin, Wang Shudao

(School of Automobile and Traffic Engineering, Guilin University of Aerospace Technology, Guilin 541004, China)

Abstract: There are a lot of researches on traffic signs in the safe driving and automatic driving of vehicles. Due to the wide variety of traffic signs and the influence of various factors, the classification and detection of traffic signs is also a challenging problem. To this end, a traffic sign classification and detection method combining tags with real road scenes is proposed. The method is divided into a data generation part and a target detection part. Experimental results show that the use of this method to generate training data can effectively train deep convolutional neural networks to achieve classification and detection of traffic signs in real scenes, and the optimized detection model has a smaller size and faster speed than the model mentioned in the article.

Key words: traffic signs; automatic driving; data enhancement; DCNN; detection

0 引言

在车辆安全和自动驾驶领域,交通标志的检测有着很大的实用价值。真实的交通场景复杂多变,交通标志易受到光照、雨雾和遮挡等外在因素的影响。传统的检测算法根据交通标志的形状、颜色等特点^[1-6],使用不同尺度大小的滑动窗口对待检测图片进行潜在目标区域提取,之后对潜在区域通过HOG(Histograms Of Oriented Gradient)^[7]、Gabor^[8]、Haar-like^[9]等人工提取特征方法,结合支持向量机、BP(Back Propagation)神经网络、极限学习机和最近邻算法等常用的机器学习算法完成分类的任务。这些方法若要完成细分类检测问题,工作量巨大,且最后的效果也不尽理想。

深度学习方法不同于前面的方法,它利用深度卷积神经网络完成特征提取,实现交通标志的检测任务。目前常用方法可分为候选区域和逻辑回归。候选区域的网络(如RCNN(Region-Convolutional Neural Network)^[10]、Faster R-CNN^[11])先提取出候选的区域特征,之后根据候选区域的特征进行位置和类别的学习,这种方法突出了出色的检测精度,牺牲了计算的时间和存储资源;逻辑回归的网络(如YOLO(You Only Look Once)^[12]、SSD(Single Shot Detector)^[13])直接将预测边界框的坐标和类别设置为回归问题,提升了网络的检测速度,但是针对具体的任务网络模型还需要进一步调整,且完成交通标志检测的研究需要数据庞大的交通标志数据集。

针对交通标志检测研究方法中的问题,本文提出了一种标签结合现实场景的交通标志分类检测方法。首

* 基金项目:广西高校中青年教师基础能力提升项目(2019KY0819)

先,为得到深度学习算法所需要的数据集,采用交通标志标签结合现实场景自动生成;然后,为解决常规算法体积大、速度慢的不足,对深度学习算法进行调整优化;最后,完成交通标志的分类检测。

1 相关工作

完成交通标志检测的深度学习算法研究,一是需要数据庞大的交通标志数据集,二是需要针对交通标志检测的算法。

1.1 交通标志数据集

目前,用于交通标志检测的LISA^[14]搜集的是美国交通标志,有47种标志类型。德国交通标志检测数据集GTSDB^[15]应用也非常广泛,但是数据集只有900张图片,数据规模较小,且没有对交通标志进行细分。BTSD^[16]是比利时的交通标志数据集,数据集包含约7300个图像数据,包含不同场景下的交通标志样本,数据样本的分布不够均匀。针对中国交通标志的TT100K^[17],从10万个腾讯街景全景图中创建了一个大型的交通标志数据集,也是迄今中国公开的最大交通标志数据集。TT100K虽然拥有众多的数据,但也无法对中国的交通标志实现细分,且数据集内的不同交通标志的样本也不够均匀,部分常见交通标志数据较多,其他不常见标志数据较少,导致无法很好地实现最终的检测效果。因此,本文提出了交通标志标签结合现实场景的方式自动生成数据集,可以很好地解决数据集问题。

1.2 深度学习的交通标志检测

针对交通标志的检测算法的研究有很多,Rajendran等^[18]通过加深ResNet^[19]网络,采用反卷积模块方法提升了网络的检测效果,但是这加大了网络的规模。Aghdam等^[20]采用膨胀卷积方式完成滑动窗口,并且利用数据集统计的信息加快前向传播,得到了较高的精确度,但该方法的结果受限于使用的数据集。YOLO算法在目标检测领域有着广泛的研究,该算法在保证检测精度的同时,具有较快的检测速度。

1.2.1 YOLO 算法基本思想

YOLO算法的检测流程如图1所示。

首先,将输入的图像划分成 $S \times S$ 个网格,然后每一个网格预测 B 个边界框。对每个预测的边界框都用置信

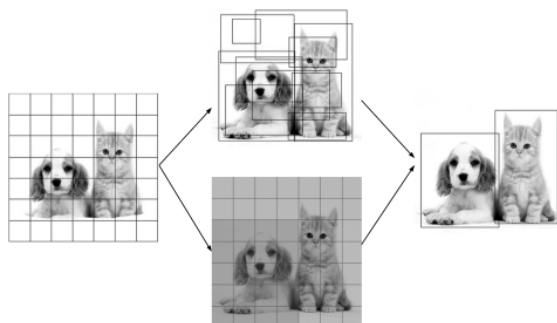


图1 YOLO 预测流程

度进行评分。置信度分数反应了该网格包含待检测物体的信心以及预测的准确程度,定义为:

$$\text{confidence} = \text{Pr}(\text{Object}) \cdot \text{IOU}_{\text{pred}}^{\text{truth}} \quad (1)$$

式中,confidence为置信度值,Pr(Object)为预测的目标概率,pred表示预测边界框的面积,truth为实际边界框的面积,IOU是实际边界框与预测边界框的交集和并集的比值。

在检测时,每一个预测种类的置信度(classConfidence)为条件类别概率(Pr(class_i|Object))与目标置信度的乘积:

$$\text{classConfidence} = \text{Pr}(\text{class}_i | \text{Object}) \times \text{confidence} \quad (2)$$

1.2.2 目标框位置预测

YOLO中使用锚框来预测待检测物体的边界框,如图2所示,每个固定框预测4个边界坐标(t_x, t_y, t_w, t_h)。通过相对网格坐标来预测目标框的中心位置,假设网格距离图像左上角的距离为(c_x, c_y),该网格预设的固定框宽高为(p_w, p_h),则预测的表达式如下:

$$\begin{cases} b_x = \sigma(t_x) + c_x \\ b_y = \sigma(t_y) + c_y \\ b_w = p_w e^{t_w} \\ b_h = p_h e^{t_h} \end{cases} \quad (3)$$

式中, σ 为逻辑函数,可以将边界框坐标约束在[0,1],用来预测相对于网络中心的偏移; b_x, b_y 为归一化后相对于单元格的数值, b_w, b_h 为归一化后相对于预设固定框的值。

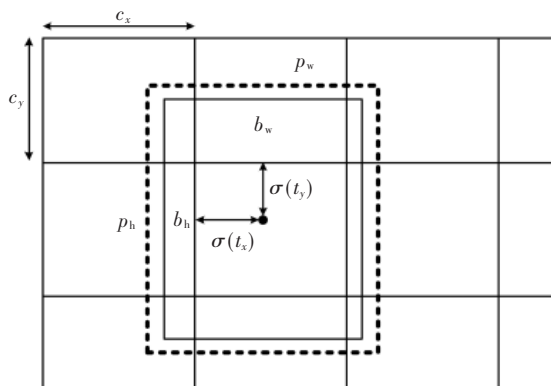


图2 目标框位置预测

2 标签结合现实场景的交通标志分类检测

本文提出的标签结合现实场景的交通标志分类方法。整体结构如图3所示。原始的交通标志标签通过数据增强,之后与现实背景结合得到了大量的数据,用以训练网络;通过训练的网络证明生成数据的有效性,并通过网络的检测结果,进一步调整优化数据和网络;最后实现较优的检测方法。

2.1 标签结合现实场景数据生成

2.1.1 标签数据增强

深度学习网络实现检测任务需要强有力的数据支持。中国的交通标志可分为7大类,包括数百个标志。其中

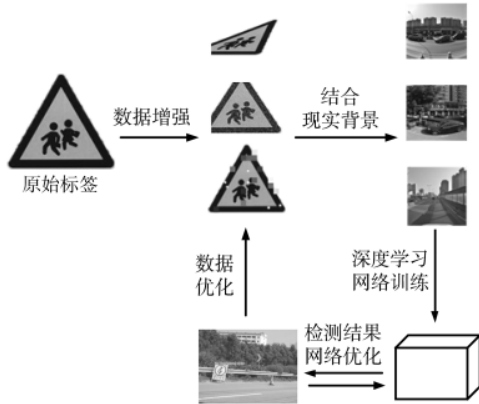


图3 整体结构

常见的3大类为警告标志、禁令标志和指示标志,多数研究也只将交通标志划分为这3大类而没有细分。本文为了便于研究,分别从3大类中抽取3个标签,如图4所示。



图4 部分中国交通标志标签示例

交通标志拥有完整的图案标签,将这些标签进行有效利用,可以降低任务的成本。传统的数据增强多是按批次统一加入噪声、旋转等方法,这种变换方式不能够很有效地增强样本的多样性。本文针对现实交通标志的环境,设置多种数据增强方式,且进行随机组合可以有效地扩张数据的多样性,如图5所示。生成的部分增强图片如图6所示。

2.1.2 结合现实场景

完成交通标志标签的数据增强后,并不能直接用于交通标志检测的研究。深度学习网络在检测时不仅需要识别出物体,还需要准确地找出物体的位置。这意味着网络不仅需要学习检测目标的深层特征,还需要能够学习到背景的相关特征,从而有效定位分类。交通标志因不同的拍摄距离,在照片上的尺寸大小不一。

充分考虑到交通标志背景和尺寸问题,本文选择

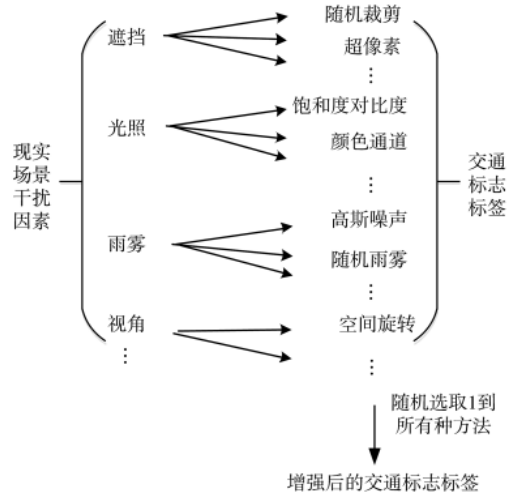


图5 交通标志标签增强流程



图6 增强交通标志标签示例

TT100K中不含有交通标志图片以及搜集的道路图片作为背景图片。通过以下步骤将2.1.1节中增强后的交通标志标签与背景进行融合,如图7所示,并自动生成用于训练的交通标志类别位置文件:

- (1)交通标志标签大小限定,交通标志因不同的拍摄距离,表现在照片上的大小不一,在进行分析后本文选择相对背景图片0.01~0.2进行尺寸限定;
- (2)交通标志标签位置限定,交通标志在照片中的位置也是不固定的,因此本文通过随机的方式将步骤(1)处理的图片融合在背景中;
- (3)交通标志标签类别位置文件生成,交通标志检测任务需要相应的标签位置信息,通过步骤(1)、(2),可以直接读取生成的数值,免除了大量的人工标注成本。



图7 结合现实场景多尺度生成图片

2.2 交通标志分类检测深度卷积网络

YOLOv5具有检测速度快模型轻量化的特点,有4个版本分别为YOLOv5S、YOLOv5M、YOLOv5L和YOLOv5X,网络的深度和广度逐渐增加。考虑到交通标志检测对于速度和文件体积的要求,YOLOv5进行优化,形成本文方法,如图8所示。其中,网络中使用的卷积核个数为YOLOv5S的一半,并且在Focus模块中加入最大池化操作。

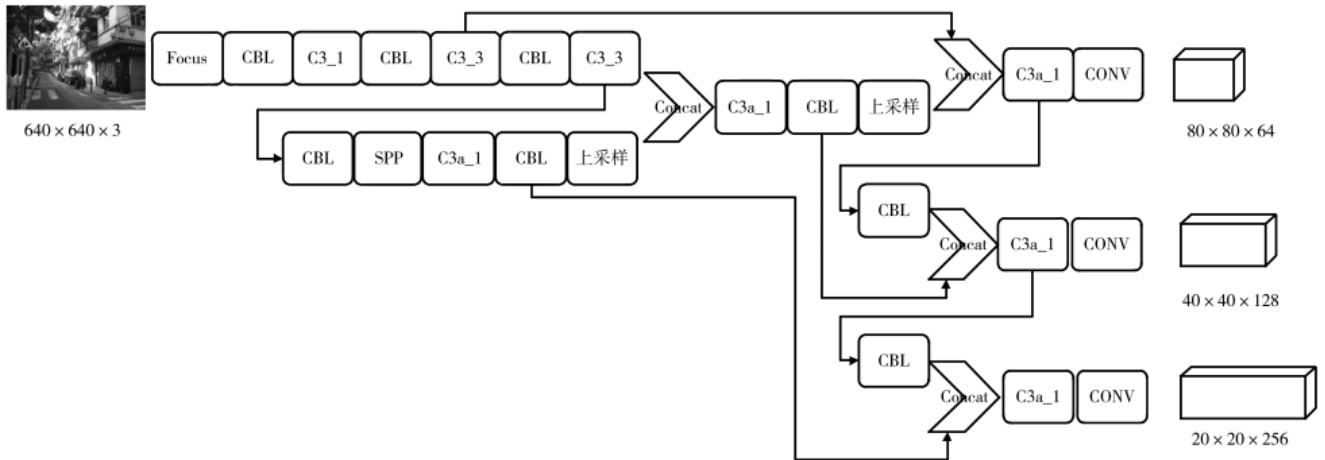


图8 网络结构

其中 CBL 模块如图 9 所示。批量正则化(BN)是避免梯度消失和梯度爆炸的有效手段。网络中,在卷积操作(CONV)和激活函数中间加入 BN 模块,可以提高模型的泛化能力,加快学习收敛速度。



图9 CBL 模块

Focus 模块如图 10 所示。在获得输入图像时,进行滑动取样(Slice)和最大池化(Maxpool)操作,处理后的数据进行组合(Concat),从而将高分辨率图片转换成低分辨率多通道数据,之后进行卷积等操作。

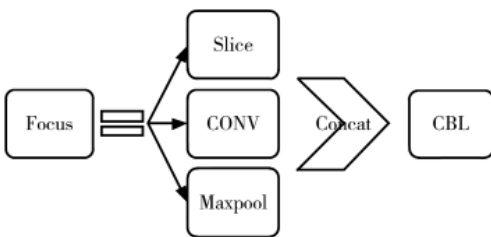


图10 Focus 模块

C3_X 模块如图 11 所示。网络在深度增加时会出现退化问题,残差模块(Resunit)在加深网络的同时,有效地

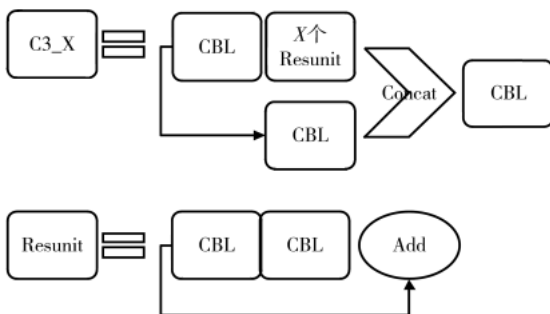


图11 C3_X 模块

解决了这一问题。将残差模块的输出与卷积后特征进行组合,进一步增加网络特征的获取。

SPP 模块如图 12 所示。该模块可以有效提升网络的感受野,将特征中的重要部分分离出来,同时也不会给网络增加硬件消耗。

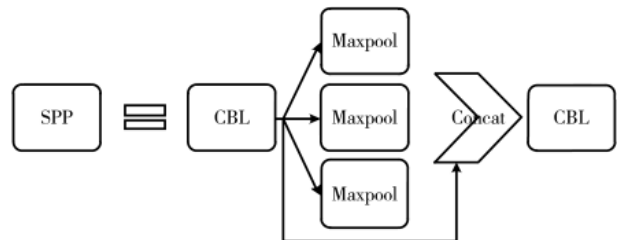


图12 SPP 模块

3 实验与结果分析

3.1 实验环境及数据

本文研究的硬件环境为 i5 处理器,内存为 16 GB,显卡为 NVIDIA GeForce 1660,操作系统为 Ubuntu 18.04。软件环境为 PyTorch 2.0, Python 3.8。

研究中,对交通标志中的 9 类标签数据进行检测。训练数据使用本文提出的标签结合现实方法生成,测试数据集选自 TT100K 以及采集的图片,每类 30 张共 270 张用于检测网络的效果。

网络训练时,初始学习率设置为 0.001,权重衰减系数为 0.000 5,迭代次数为 500。

3.2 评价指标

目标检测算法常用的两个指标为 mAP 和检测速度(f/s)。其中,mAP 代表多类别中 AP 的平均值。AP 是查全率(Recall)和查准率(Precision)曲线围成的面积。Recall 和 Precision 计算时,需要根据真实标签将检测结果划分为真正例(t_p)、真反例(t_n)、假正例(f_p)、假反例(f_n)。

Recall 表示正确检测到的样本数占真实样本数的比例,计算公式如下:

$$\text{Recall} = \frac{t_p}{t_p + f_n} \quad (4)$$

Precision 表示正确检测到样本数占总检出样本数的比例,计算公式如下:

$$\text{Precision} = \frac{t_p}{t_p + f_p} \quad (5)$$

3.3 实验结果与分析

实验中,为验证标签结合现实场景生成数据集的有效性,以及交通标志标签和背景对于检测结果的影响,首先,将9类交通标志标签每类分别增强10张、20张以及30张,分别与10张背景进行融合。使用本文方法与YOLOv3tiny、YOLOv5S分别训练后得到的mAP值进行对比,如图13所示。

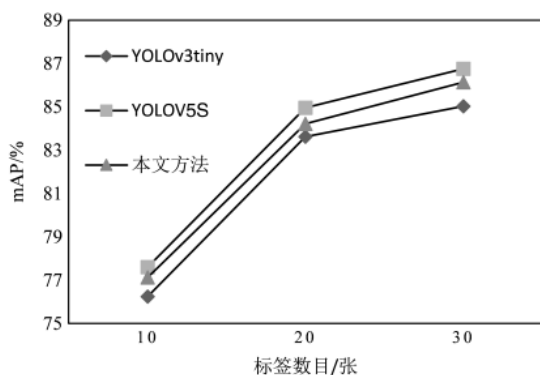


图13 不同标签增强数目检测效果对比

结果表明,本文提出的数据生成方式可以有效地训练3个网络,并且随着交通标志标签增强数目的增加,网络的检测效果也随之改善。

在9类交通标志标签每类增强30张的基础上,增加背景图片为20张、30张,使用本文方法与YOLOv3tiny、YOLOv5S分别训练后得到的mAP值进行对比,如图14所示。

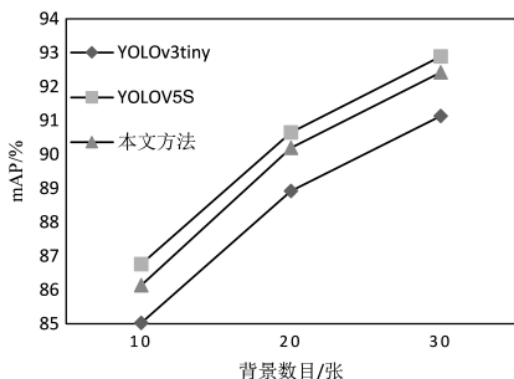


图14 不同背景图片数目检测效果对比

结果表明,增加背景图片可以有效地提升网络的检测效果。最终,通过将交通标志标签每类增强30张,背景图片30张生成训练集,网络检测效果mAP、检测速度和权重大小如表1所示。

实验结果充分证明,本文提出的标签结合现实场景

表1 最终检测结果

检测方法	mAP/%	检测速度/(f/s)	权重大小/MB
YOLOv3tiny	91.13	27	17.8
YOLOv5S	92.89	19	14.4
本文方法	92.25	29	3.8

生成交通标志训练集能够有效地训练检测网络,并且可以通过增加交通标志标签和背景改善网络的检测效果。本文优化的网络在保证检测效果基础上,具有更小的体积和更快的检测速度。部分检测结果如图15所示。



图15 检测结果示例

4 结论

本文针对交通标志的分类检测,提出交通标志标签结合现实场景方式,生成深度学习网络所需要的训练数据。该方法只需要对交通标志标签进行适当的数据增强,之后与背景图片进行结合。训练用的位置类别文件可以自动生成,节省了大量的人力物力。针对交通标志检测优化的网络虽在mAP上低于YOLOv5S,但在检测速度和权重大小上更优,有利于降低硬件成本。

参考文献

- [1] 朱双东,刘兰兰,陆晓峰.一种用于道路交通标志识别的颜色-几何模型[J].仪器仪表学报,2007,28(5):956-960.
- [2] 申中鸿,付梦印,杨毅,等.背景像素突变检测的交通标志图像分割[J].计算机应用研究,2012,29(9):3531-3535.
- [3] PAULO C F,CORREIA P L.Automatic detection and classification of traffic signs[C]//Eight International Workshop on Image Analysis for Multimedia Interactive Services. IEEE Computer Society, 2007.
- [4] GAVRILA D.Traffic sign recognition revisited[C]//Mustererkennung, 21 Dagm-symposium.Springer-Verlag, 1999.
- [5] 许少秋.户外交通标志检测和形状识别[J].中国图象图形学报,2009,14(4):707-711.
- [6] 谷明琴,蔡自兴.无参数形状检测子和DT-CWT的交通标志识别[J].计算机研究与发展,2013(9):99-107.
- [7] DALAL N,TRIGGS B.Histograms of oriented gradients for human detection[C]//2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition.Piscataway: IEEE Press, 2005:886-893.
- [8] LEE T S.Image representation using 2D Gabor wavelet[J].IEEE Transactions on Pattern Analysis and Machine Intellig-

(下转第36页)

表 2 3 种算法实验结果对比

算法	准确率/%	均方误差/(°)
CBF	53.45	68.8
MLP	73.53	15.43
Mask-DOA	93.32	5.89

4 结论

面向船载远程会议系统语音前端需求,本文提出一种适用于船舶舱室环境的小阵列高精度 DOA 估计算法,该方法在 DOA 估计神经网络的基础上引入 Mask 算法降低噪声和混响干扰,从而用较为纯净的语音信号提高 DOA 估计性能。在直径为 7 cm 的六元圆阵上进行的仿真与实验表明,DOA 估计精度达到 5°,且算法模型有着较强的泛化能力,证明了本文算法在船舶舱室内使用的有效性。

同时,需要指出:本文采用深度学习方式构建了信号预处理与 DOA 估计模型系统,虽通过模拟舱室环境实验验证了性能改善,但是考虑到深度学习类算法的效

果与鲁棒性受到数据集大小与数据类型适配的影响,算法的评估验证尚不够充分。特别是,由于船舶舱室属于带行业背景的特殊类型场景,噪声及混响特性复杂、差异性大,无法直接利用通用语音领域的丰富开源数据集、混响特性资源,本文方法训练出的模型不可避免存在数据失配条件下性能下降的不足。

因此,面向智能船舶、智慧海洋发展的大趋势,下一步计划结合实际场景下多样化的船舶舱室特性、背景噪声进一步开展所提方法的性能评估、验证。

参考文献

- [1] 长航局.长江航道船载视频会议及远程视频监控系统通过验收[EB/OL].(2011-08-12)[2021-08-30].http://news.hexun.com/2011-08-12/132386171.html.
- [2] 苑立娟,秦芹.云平台船舶状态实时监控[J].舰船科学技术,2021,43(10):178-180.
- [3] 王栋.科考船岸一体化网络设计及发展趋势[J].船舶,(下转第 77 页)

(上接第 31 页)

- ence, 1996, 18(10): 959-971.
- [9] VIOLA P A, JONES M J. Rapid object detection using a boosted cascade of simple features[C]//IEEE Computer Society Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE Press, 2001: 511-518.
 - [10] GIRSHICK R, DONAHUE J, DARRELL T, et al. Rich feature hierarchies for accurate object detection and semantic segmentation[C]//2014 IEEE Conference on Computer Vision and Pattern Recognition, 2014: 580-587.
 - [11] 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 and Machine Intelligence, 2017, 39(6): 1137-1149.
 - [12] REDMON J, DIVVALA S, GIRSHICK R, et al. You only look once: unified, real-time object detection[C]//2016 IEEE Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE Press, 2016: 779-788.
 - [13] LIU W, ANGUELOV D, ERHAN D, et al. SSD: single shot multibox detector[C]//European Conference on Computer Vision. Berlin: Springer, 2016: 21-37.
 - [14] MOGELMOSE A, TRIVEDI M M, MOESLUND T B. Vision-based traffic sign detection and analysis for intelligent driver assistance systems: Perspectives and survey[J]. IEEE Transactions on Intelligent Transportation Systems, 2012, 13(4): 1484-1497.
 - [15] HOUBEN S, STALKAMP J, SALMEN J, et al. Detection of traffic signs in real-world images: the german traffic sign detection benchmark [C]//The 2013 International Joint

- Conference on Neural Networks (IJCNN). IEEE, 2013: 1-8.
- [16] TIMOFTE R, ZIMMERMANN K, VAN GOOL L. Multi-view traffic sign detection, recognition, and 3D localisation[J]. Machine Vision and Applications, 2014, 25(3): 633-647.
- [17] ZHU Z, LIANG D, ZHANG S, et al. Traffic-sign detection and classification in the wild[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2016: 2110-2118.
- [18] RAJENDRAN S P, SHINE L, PRADEEP R, et al. Fast and accurate traffic sign recognition for self driving cars using RetinaNet based detector[C]//2019 International Conference on Communication and Electronics Systems. Piscataway: IEEE Press, 2019: 784-790.
- [19] HE K, ZHANG X, REN S, et al. Deep residual learning for image recognition[C]//2016 IEEE Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE Press, 2016: 770-778.
- [20] AGHDAM H H, HERAVI E J, PUIG D. A practical approach for detection and classification of traffic signs using convolutional neural networks[J]. Robotics and Autonomous Systems, 2016, 84: 97-112.

(收稿日期: 2021-06-07)

作者简介:

张成(1990-),男,硕士,助教,主要研究方向:人工智能技术、自动化驾驶。

张瑞宾(1984-),男,博士,副教授,主要研究方向:人工智能技术、自动化驾驶。

王曙道(1973-),男,硕士,高级工程师,主要研究方向:汽车结构分析。



扫码下载电子文档

版权声明

经作者授权，本论文版权和信息网络传播权归属于《电子技术应用》杂志，凡未经本刊书面同意任何机构、组织和个人不得擅自复印、汇编、翻译和进行信息网络传播。未经本刊书面同意，禁止一切互联网论文资源平台非法上传、收录本论文。

截至目前，本论文已经授权被中国期刊全文数据库（CNKI）、万方数据知识服务平台、中文科技期刊数据库（维普网）、DOAJ、美国《乌利希期刊指南》、JST 日本科技技术振兴机构数据库等数据库全文收录。

对于违反上述禁止行为并违法使用本论文的机构、组织和个人，本刊将采取一切必要法律行动来维护正当权益。

特此声明！

《电子技术应用》编辑部

中国电子信息产业集团有限公司第六研究所