doi:10.3969/j.issn.1672-4348.2020.01.011

车前小型障碍物图像检测与分类方法

陈炳煌[1,2]

(1. 福建工程学院 信息科学与工程学院, 福建 福州 350118;
2. 福建省汽车电子与电驱动技术重点实验室, 福建 福州 350118)

摘要: 针对车辆辅助驾驶系统中遇到的障碍物小的特点和对实时性的高要求,提出一种基于卷积神经网络 YOLO 图像检测算法优化并增加分类计数的方法。通过对小石子和道路坑洞这 2 种极易引发车辆事故的典型小型障碍物图像建立数据库,针对数据库利用 k-Means+优化 k 值并配置新的锚定值,对取自车载视频的图像进行检测识别。新增的分类和计数算法可快速、直观地获得检测结果,实现驾驶员快速决策的目标。实验结果表明,该方法可对小石子和道路坑洞等小型障碍物有效地检测识别和分类计数,检测速度也满足系统的实时性要求。

关键词: 目标检测; 小型障碍物; YOLO; 分类
中图分类号: TP274 **文献标志码**: A **文章编号**: 1672-4348(2020)01-0057-06

Image detection and classification of small obstacles in front of vehicles

CHEN Binghuang [1,2]

(1. School of Information Science and Engineering, Fujian University of Technology, Fuzhou 350118, China;
2. Fujian Key Laboratory of Automotive Electronics and Electric Drive, Fuzhou 350118, China)

Abstract: In view of the small obstacles encountered in the vehicle's assisted driving system and the high real-time requirements, an optimized method based on convolutional neural network YOLO image detection algorithm was proposed and the classification counting was added. By establishing a database for images of such typical small obstacles as small stones and road potholes, which tend to cause vehicle accidents, the images from the vehicle video were detected and identified by using k-means+ to optimize k value and configure new anchors for the database. The newly added classification and counting algorithm can obtain test results quickly and intuitively and realize the goal of rapid driver decision-making. Experimental results show that this method can effectively detect, identify, classify and count small obstacles such as small stones and road potholes, and the detection speed also meets the real-time requirements of the system.

Keywords: object detection; small obstacles; YOLO; classification

感知和理解环境的能力是汽车智能系统的基础和前提[1]。智能车辆通过感知周围环境并做出分析后,将信息提供给控制系统,引导车辆避开障碍物行驶。例如谷歌和德国的无人驾驶汽车配备先进的智能识别、自动驾驶和导航系统,可以探测到行人和其他车辆[2-5]。然而,对于高度明显低于车辆高度的障碍物或者目标较小的障碍物,如小石子和道路坑洞等小型障碍物,车载系统在图像检测和识别上仍有一些不足之处[6]。这些小障碍物体积和几何尺寸都较小,导致车辆在快速行驶过程中很难检测到,从而造成车辆受损甚至人员受伤[7]。而车载视频检测小障碍物的低

收稿日期: 2019-10-28
作者简介: 陈炳煌(1978-),男,福建仙游人,讲师,博士研究生,研究方向:人工智能在图像目标检测的研究和智能系统。

检出率和实时性低的问题将不能保障车辆安全行驶。对于车辆前方障碍物的视频图像检测,最重要的是图像检测算法的实时性和鲁棒性[8]。

在图像目标检测领域,Girshick、Ren 等[9-10]分别提出了快速区域卷积神经网络(fast regional convolutional neural network, Fast R-CNN)和更快速区域卷积神经网络(Faster R-CNN),可以提高图像检测的准确率。但这些算法检测速度较慢,不适合实时检测车前障碍物。Redmon 等[11]提出 YOLO (you only look once)可以满足检测速度的要求,其视频检测速度达到 45fps。而且 YOLO 对于小型目标的识别精度也较高,这对于实时的车前小型障碍物的图像检测与识别是个优势。

为了使驾驶辅助系统能快速地并且实时地识别小型障碍物,提出一种优化车载视频图像检测车辆前方小障碍物和新增分类计数的方法,将优化后的 YOLO(optimized YOLO,O-YOLO)应用于道路坑洞和小石子等小障碍物的图像检测。

1 小型障碍物图像检测与分类

为了能实时并有效地检测小型障碍物,并给车辆驾驶员发出相应警示信息,需要对车前小型障碍物的图像和视频进行检测识别和分类计数。整体步骤如下:

(1)建立小型障碍物图像数据库,并标注。

(2)采用 k-Means+对数据库进行聚类,找到最优 k 值。

(3)针对最优 k 值和小型障碍物图像数据库进行针对性分析计算,得到训练所需的 Anchor 值。

(4)通过 O-YOLO 对图像数据库的训练集进行训练,得到最优训练权重。

(5)将该训练权重应用到测试集,并根据图像特点设定感兴趣区域(region of interest,ROI),对目标进行检测识别和分类计数,输出结果。

1.1 YOLO 算法原理

YOLO 是一种使用深度卷积神经网络学习的特征来检测对象的对象检测器,有 3 个版本。YOLO v3[12]使用了新的网络 Darknet-53,其结构如图 1 所示(以 416×416 为例)。该模型使用了大量性能良好的 3×3 和 1×1 卷积层,有 53 个卷积层。YOLO v3 从输入图像中提取出特征,然后将输入图像分割成 13×13 的网格单元。如果 ground truth 中的对象的中心坐标位于网格单元中,则将对该对象进行预测。每个网格单元预测 3 个大小不同的边界框(13×13、26×26、52×52)。目标检测预测了 ground truth 与所提出的边界框之间的 IOU(Intersection over Union)。而类别检测预测的是该类别是目标的概率。输出特征图中有两个维度(宽、高),如 13×13。另一个维度(深度)是

$$\text{Box} \times (5 + \text{Class}) \tag{1}$$

式中,Box 表示每个网格单元预测的边界框数,Class 表示目标类别数,5 表示为 4 个坐标和 1 个客观分值。

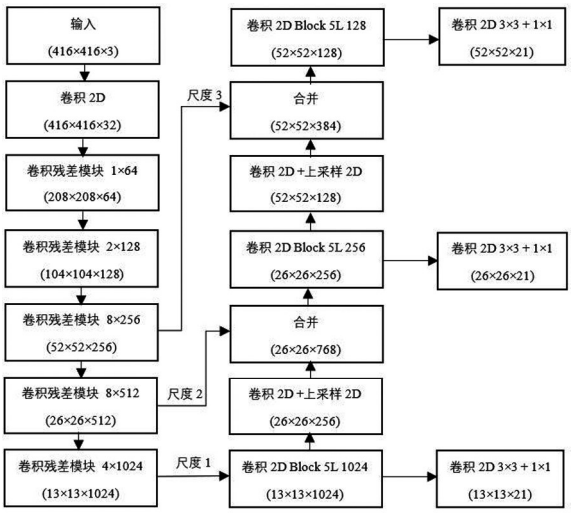


图 1 YOLO v3 的网络框架

Fig.1 Network of YOLO v3

YOLOv3 采用类似 SSD(single shot multiBox detector)的多尺度策略,以类似于特征金字塔网络的方式从不同尺度提取特征,使用 13×13、26×26、52×52 三个尺度的特征图进行预测,使得预测鲁棒性更强。

1.2 k 值与 Anchor 值的优化

通过 YOLO v3 可实现高速的特征提取和高速的对象检测,但是车前障碍物需要实时准确地检测出来,对检测精度的要求也很高。为解决车前的小型障碍物检测精度较低的问题,研究对原始的 YOLO v3 进行优化。

YOLO 借鉴区域卷积神经网络(regional convolutional neural network,R-CNN)的思想,引入 Anchor 值,它是一组 k 个簇中心框的宽高值,影响目标检测速度和分类识别精度[13]。通过k-Means+对数据集中标记的目标框进行维度聚类可以确定 k 值和 Anchor 值,找到目标框的统计规律。原始

YOLO v3 的 Anchor 值是由微软 COCO(common objects in context)图像数据集聚类确定的。因 COCO 图像数据集中有 80 个类别,其 Anchor 值是由这 80 个类别生成的,但这并不适用于前方小障碍物数据集。因此,为了提高图像检测识别和分类的精度,聚类的 k 值和 Anchor 值需要重新计算确定。小障碍物数据集目标框的宽高经过 k-Means+聚类的最优 k 值由样本聚类误差平方和方法确定[14]。

$$SSE = \sum_{i=1}^{k} \sum_{s \in C_i} |s - d_i|^2 \qquad (2)$$

式中,误差平方和(SSE)是样本聚类误差平方和方法的核心指标,s 是样本,d_i 是第 i 个聚类的中心点。k 越大,SSE 越小,说明样本聚合程度越高。通过对小障碍物图像数据集的训练样本进行聚类得到 SSE 曲线,最先趋于平缓的点即为合适的 k 值,如图 2 所示,k 为 3。从表 1 可发现小障碍物数据集的宽高维度值基本表征了小石子和坑洞的几何尺寸。

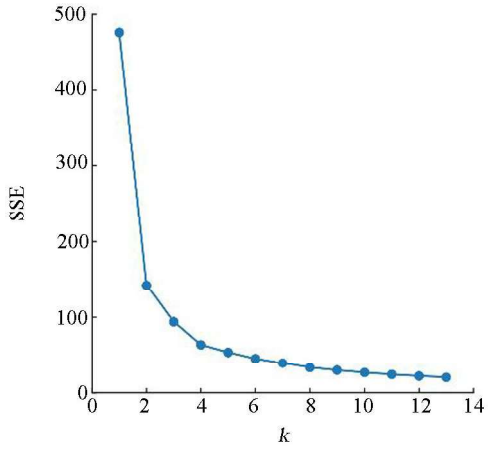


图 2　SSE 聚类曲线

Fig.2　SSE clustering curve

1.3　ROI 的确定

在实际应用中,车前图像的获取是由设置在车内的摄像头完成,摄像头可以记录行车视频,拍摄事故过程中的重要照片。绝大多数情况下图像的上半部分是距离车辆较远的区域,因此将图像的下半部分确定 ROI 有利于图像检测和分类计数、提高检测效果,如图 3 所示。而摄像头的分辨率不同,导致图像的输入源不同,因此根据图像的比例来确定 ROI。

$$0.03 < X\_center < 0.97$$

$$0.46 < Y\_center < 1.00 \qquad (3)$$

表 1　不同数据集的 Anchor 值

Tab.1　Anchor values of different data sets

数据集 k	COCO		小障碍物	
	宽	高	宽	高
1	10	13	15	15
2	16	30	56	40
3	33	23	214	160
4	30	61		
5	62	45		
6	59	119		
7	116	90		
8	156	198		
9	373	326		

式中,X_center 和 Y_center 分别为图像目标框的 x 轴和 y 轴中心的坐标位置。0.03 和 0.97 分别为宽度比例,0.46 和 1.00 分别为高度比例,当图像分辨率是 1 920×1 080 时,按比例其宽度为 70~1 850,高度为 500~1 080,从而匹配任何大小的图像。

图 3　ROI 的确定

Fig.3　Setting ROI

1.4　计数与分类

图像检测目标对象的计数一般比较简单,原始 YOLO v3 可以检测图像中物体的总数。但是每一类对象不能被原始 YOLO v3 单独计数。因此,提出一种分类计数算法,并自动生成报表。

Counting and classification algorithm

STEP 1: The counter and the count cache are initialized: js=0, buf=0

STEP 2: Detect small obstacles byO-YOLO. The total number and class of objects can be detected.

STEP 3: For i=0 to num

If (0.03<X_center<0.97 & 0.46<Y_center<1)

Then js = js + 1, copy class name to buf

End for.

STEP 4: Calculate the number of the same name and output to the TXT document.

从 STEP 3 可得到相应的缓存数据如:stone hole stone. STEP 4 中 TXT 文本保存格式为:002345.jpg 图像中有 3 个目标:小石子 2 个,坑洞 1 个。

2 实验

数据集训练和测试的实验环境为:Windows 10 系统,内存 64G,显卡 Nvidia Quadro M4000 (8GB RAM)一张。同时使用 Faster R-CNN 对同一数据集进行训练和测试与本算法进行对比。

2.1 图像数据库

由于没有的小石子和道路坑洞的图像数据集,本文收集了一些不同大小的小型障碍物图片(jpg 格式),并利用车载摄像头采集到小型障碍物 1 080p 高清视频,通过视频转换器获得高清序列图像。针对小石子和坑洞原始图像较少的情况,利用图像剪切、对比度增强、微角度旋转、图像平移、图像翻转、加噪声等方法增大图像数据,共得到 3 575 张图片,形成小型障碍物图像数据库。其中用于 O-YOLO 训练 3 217 张,测试 358 张。设置 2 个类别,按小石子(stone)和坑洞(hole)标注,如图 4 所示。

(a) 小石子 (b) 坑洞

图 4 车前小型障碍物数据集

Fig.4 Dataset of small obstacles in front of a vehicle

2.2 训练

针对小障碍物图像没有预先训练的模型,训练过程较费时,训练前需要根据式(1),设定 O-YOLO 的关键训练参数“filters”。实验中,Box 取 3,Class 取 2,则 filters 是 21。

其他训练参数主要包括:学习率、动量和衰减。在这个实验中,动量是 0.9,衰减是 0.000 5。实验的学习率策略是阶段性的。根据迭代次数,其具体的变更过程如表 2 所示。在迭代次数为 20 000 次时,学习率降低到 0.000 1,可以加快训练速度。O-YOLO 416×416 模型的损失曲线如图 5 所示。最终的平均损失值是 0.367 7。图像输入分辨率的大小影响检测的精度[8],因此使用 5 种不同分辨率图像模型(O-YOLO 288×288、352×352、416×416、480×480、544×544)来训练。

表 2 学习率变化策略

Tab.2 Learning rate's change strategy

迭代次数	学习率/%	倍增因数
0~149	0.000 01	none
150~1 000	$0.10 \times \left(\frac{\text{次数}}{1\,000}\right)^4$	none
1 000~20 000	0.10	1.0
20 000~25 000	0.01	0.1
25 000~32 000	0.001	0.1

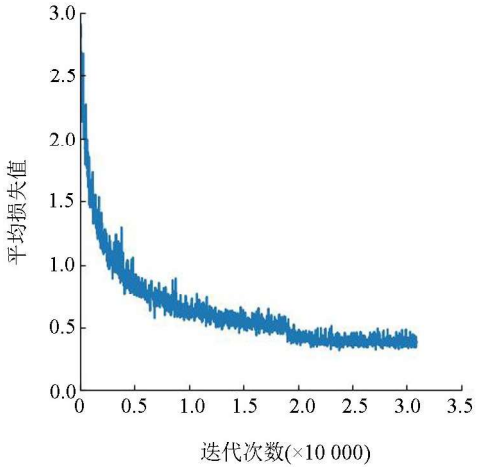


图 5 O-YOLO 416×416 模型的损失曲线

Fig.5 Loss curve of O-YOLO 416×416

2.3 图像检测和对比分析

本实验针对 2 个类别分别计算其精度(precision)和召回率(recall)。图像检测分析中,阈值 t 设置为 0.25,以保证检测精度较高。表 3 显示了 O-YOLO 下 5 个模型的精度、召回率、基于精度与召回率的调和平均值(F1-score)和平均精度均值(mAP)。图 6 显示了精度-召回率(P-R)曲线。

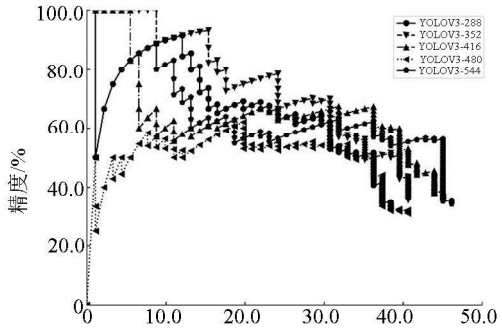
从表 3 可见 O-YOLO 352×352 模型的 mAP 为最好（82.88%）。但是在精度相差不多的情况下，O-YOLO 416×416 模型的精度（86.1%）、召回率（74.2%）和 F1-score 值（79.7%）都优于 O-YOLO 352×352 模型。

表 3　7 种模型的测试结果

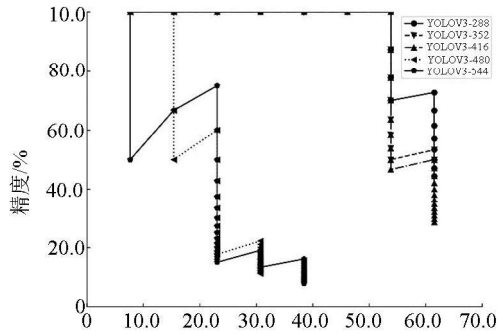
Tab.3　Test results of 7 models（t=0.25）

%

模型	精度	召回率	F1-score	mAP
O-YOLO 288×288	78.2	67.1	72.2	77.38
O-YOLO 352×352	82.0	72.3	76.8	82.88
O-YOLO 416×416	86.1	74.2	79.7	82.30
O-YOLO 480×480	85.2	70.1	76.9	81.06
O-YOLO 544×544	85.1	67.4	75.2	78.69
原始 YOLO 416	78.1	65.2	71.1	67.81
Faster RCNN	83.2	70.2	76.1	80.33



(a) 小石子



(b) 坑洞

图 6　O-YOLO 5 种模型的 P-R 曲线

Fig.6　P-R curves of 5 models of O-YOLO

为进一步说明图像检测识别与分类方法的有效性和实时性，针对同一个小障碍物图像数据集采用原始 YOLO 和 Faster RCNN 算法进行检测分析和对比，如表 3 所示。O-YOLO 416×416 模型的 mAP 值和 F1-score 值分别比 Faster RCNN 算法高 1.97%和 3.60%，而且分别比原始 YOLO 算法高 10.24%和 12.10%，说明经过优化，该算法对于小型障碍物图像检测的精度较为优秀。

针对 O-YOLO 416×416 模型在不同阈值影响下测试的精度和召回率如表 4 所示。随着阈值的增加，精度越来越高，召回率越来越小。当 t = 0.20时，F1-score 指标是最好的，数值为 0.80，表示 416×416 模型的鲁棒性很好。

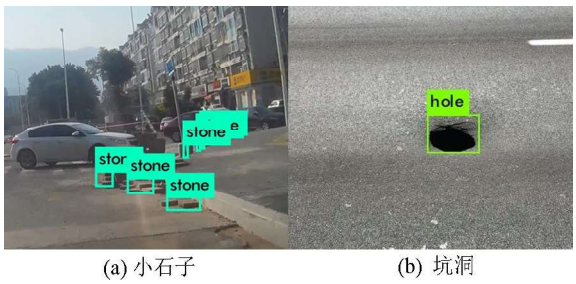
表 4　精度和召回率的测试数值

Tab.4　Test values of precision and recall rate

t	精度/ %	召回率/ %	F1-score/ %
0.10	81	77	79
0.20	85	75	80
0.30	86	73	79
0.40	87	69	77
0.50	88	66	76
0.60	90	62	76
0.70	90	57	70

2.4　分类及计数测试分析

实验使用 O-YOLO 416×416 模型来测试 358 张图像。图 7 为障碍物检测结果图。检测结果可以通过分类计数算法输出到 TXT 文档中。一个文档对应一个识别任务，每个 TXT 文档包含图像名称、障碍物名称和计数结果。



(a) 小石子　　(b) 坑洞

图 7　障碍物检测结果

Fig.7　Obstacle detection results

表 5 显示了实际数据为 376 个坑洞和 759 个小石子时不同阈值下的测试数据和相对误差。由表 5 可知，当 t = 0.10 时，O-YOLO 对道路坑洞的检测效果较好；而 t = 0.20 时，O-YOLO 对小石子的检测效果较好；但是总体统计，当阈值 t = 0.20 时 O-YOLO 的检测效果较好。这一组 358 张图像

在 O-YOLO 的测试用时为 35 s,平均一张图像检测用时为 0.097 s;而在同一测试平台上使用 Faster RCNN 来测试用时为 125.3 s,平均一张图像检测用时为 0.350 s,远远低于 O-YOLO 的检测速度,因此该方法在检测速度和精度上均有上佳表现,能够胜任车载实时检测。

表 5 测试数据与实际数据的对比

Tab.5 Comparison between test data and real data

t	坑洞数/个	小石子数/个
0.25	356(5.32%)	745(1.84%)
0.20	368(2.13%)	763(−0.53%)
0.10	375(0.27%)	781(−2.90%)

注:括号内数据为误差百分数

小型障碍物图像检测还存在一些误差,如图 7(a)的右下角没有发现一些小石子,主要是由于目标物过小,O-YOLO 的检测精度还不够高。在此方法的研究基础上增加更多的数据集,并且修改 O-YOLO 的网络结构以提高小障碍物的检测精度。

3 结论

针对车辆前方的小型障碍物检测识别困难、无有效分类和实时性差的问题,本文提出了一种基于 YOLO 的车辆前方小障碍物图像检测优化与分类计数的方法。为提高检测精度,O-YOLO 的 k 值使用 k−Means+方法进行优化,并对 Anchor 值进行针对性的计算。为了对小石子和道路坑洞这 2 类小型障碍物进行准确的分类计数,提出了一种特殊的计数分类方法。实验分析和对比,该算法的检测精度为 0.86,检测速度优于 Faster RCNN 算法,验证了方法的有效性。

参考文献:

[1] 石金进. 基于视觉的智能车辆道路识别与障碍物检测方法研究[D]. 哈尔滨: 哈尔滨工业大学, 2017.

[2] LUO J, YAN B, WOOD K. InnoGPS for data-driven exploration of design opportunities and directions: the case of google driverless car project[J]. Journal of Mechanical Design, 2017, 139(11): 111416−111429.

[3] 郭晶赛. 基于视频流检测的无人驾驶车辆行驶范围生成研究[D]. 北京: 北京交通大学, 2018.

[4] PAWAR S, MUKANE S. A driver assistance system using ARM processor for lane and obstacle detection[M]//Techno-Societal 2016. Cham: Springer International Publishing, 2017: 303−313.

[5] GODHA S. On-road obstacle detection system for driver assistance[J]. Asia Pacific Journal of Engineering Science and Technology, 2017, 3(1): 16−21.

[6] GOEL N, SHARMA R, NIKHIL N, et al. A crowd-sourced adaptive safe navigation for smart cities[C]//2017 IEEE International Symposium on Multimedia (ISM). New York: IEEE, 2017: 382−387.

[7] FLEMING B. Smarter and safer vehicles [automotive electronics] [J]. IEEE Vehicular Technology Magazine, 2012, 7(2): 4−9.

[8] 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. New York: IEEE, 2014: 580−587.

[9] GIRSHICK R. Fast R-CNN[C]. Proceedings of the IEEE International Conference on Computer Vision, 2015: 1440−1448.

[10] REN S Q, HE K M, 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.

[11] 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 (CVPR). New York: IEEE, 2016: 779−788.

[12] REDMON J, FARHADI A. YOLOv3: an incremental improvement[J]. arXiv preprint arXiv:1804.02767, 2018.

[13] 魏湧明, 全吉成, 侯宇青阳. 基于 YOLO v2 的无人机航拍图像定位研究[J]. 激光与光电子学进展, 2017(11): 101−110.

[14] CHEN B H, MIAO X R. Distribution line pole detection and counting based on YOLO using UAV inspection line video[J]. Journal of Electrical Engineering & Technology, 2020, 15(1): 441−448.

(责任编辑: 方素华)