

基于 YOLO v3 的交通标志牌检测识别*

潘卫国^{1,2}, 刘 博³, 陈英昊³, 石洪丽³

(1. 北京联合大学 北京市信息服务工程重点实验室 北京 100101;

2. 北京联合大学 机器人学院 北京 100027; 3. 北京联合大学 应用科技学院 北京 100101)

摘 要: 在无人驾驶和辅助驾驶领域,交通标志牌检测识别是重要的。针对目前基于 YOLO 的检测方法能够达到实时的检测效果,但在准确率方面有所降低的问题,提出了基于感兴趣区域(ROI)的交通标志牌检测方法。首先根据交通标志牌的颜色特性得到候选区域;再利用交通场景图像规则确定交通标志牌的 ROI;最后在交通标志牌的 ROI 基于 YOLO v3 实现对交通标志牌的检测识别。实验结果表明:由于本文提出的方法去除了图像中部分干扰因素,使得算法在检测精度上得到了提升,也能满足实时性的需求,并在无人驾驶车辆上进行了验证。

关键词: 目标检测; 感兴趣区域; 深度学习; 交通标志牌

中图分类号: TP33 **文献标识码:** A **文章编号:** 1000-9787(2019)11-0147-04

Traffic sign detection and recognition based on YOLO v3*

PAN Weiguo^{1,2}, LIU Bo³, CHEN Yinghao³, SHI Hongli³

(1. Beijing Key Laboratory of Information Service Engineering, Beijing Union University, Beijing 100101, China;

2. College of Robotics, Beijing Union University, Beijing 100027, China;

3. College of Applied Science and Technology, Beijing Union University, Beijing 100101, China)

Abstract: In the field of self-driving and driver assistance system, traffic sign detection and recognition is very important. Aiming at that current detection method based on YOLO can achieve real-time detection effect, but the accuracy is decreased, a problem, a traffic sign detection method based on region of interest(ROI) is presented. color features of traffic sign are firstly used to get the candidate regions. Secondly, ascertain the ROI of traffic sign based on the rules of traffic scene image. Finally, detect traffic sign in ROI based on YOLO v3. The experimental result show that self-driving or intelligent vehicle show that the proposed approach eliminate partial interference which makes the detection precision of algorithm is improved and can meet the needs of real-time performance. It is verified in remotely piloted vehicle.

Keywords: object detection; region of interest(ROI); Deep learning; traffic sign

0 引 言

交通标志牌检测是无人驾驶和辅助驾驶的重要组成部分,也是实现安全驾驶的一个重要环节。车辆需要检测到前方道路的交通标志才能获取路况信息,这样才能增加自动驾驶的安全性和减少事故发生的可能性。

目前已有的相关研究包括 Xiong C Z^[1]提出的一种基于 Faster R-CNN 检测交通标志牌的方法。Guo J K^[2]提出的算法不仅可以识别出交通标志牌,还可以识别道路信息、设施服务和交通信号灯标识。Lee H S 等人^[3]提出了一种基于 SSD 实时的交通标志牌识别方法,同时可以提供精确的位置估计及标志牌精确的边界信息。Li C P^[4]提出了基于 channel-wise 显著图和深度学习的端到端的交通标志检

测方法。基于 channel-wise 的特征提取、层次性细化和特征融合等方法。Rehman Y W 等人^[5]提出了一种基于 d-patches 的方法解决交通标志牌检测过程中遮挡的问题。Wang D D 等人^[6]提出了一种基于级联的方法检测交通标志,同时使用了基于中层语义特征的显著图提升检测的速度。Yuan Y 等人^[7]提出了一种针对视频中交通标志检测、追踪和识别的增量式的框架。Zhu Z 等人^[8]提出了一种改进的交通标志定位方法,通过交通标志牌的颜色形状信息,使交通标志牌的边界框更加紧致。Zhu Y G 等人^[9]利用级联网络实现了一种针对文本类标志牌的检测方法。Liu Z G 等人^[10]提出了一种基于区域的多尺度卷积神经网络方法,能够检测识别出交通场景中较小的交通标志牌。目前

收稿日期: 2019-09-03

* 基金项目: 国家自然科学基金资助项目(61802019, 61871039); 北京市教育委员会科技计划资助项目(KM201711417005, KM201911417001)

的交通标志牌检测算法以深度学习为主,但是如何快速准确的检测出交通场景中的交通标志牌是一个难题。

在交通信号识别的过程中,交通信号的检测是关键的一步,直接决定了后续分类的效果。近几年,涌现出的基于深度卷积网络的目标检测可以分为 single-stage^[11,12] 和 two-stage^[13-15] 两种,在目标检测方面展现了显著的优势;基于 single-stage 的方法具有较快的检测速度,而基于 two-stage 的方法需要产生许多候选区域,检测速度较慢,但具有很高的准确率。目前研究的难点是如何达到这两类算法之间的平衡,既能够达到高的准确度又具有实时的检测效果。

1 交通标志牌检测

根据国家交管部门的规定,交通标志分为警告标志、禁令标志、指示标志、指路标志、旅游区标志和其他标志。本文的方法主要检测警告、禁令和指示三类交通标志牌,如图1所示。



图1 交通标志牌类别

本文提出的交通标志牌检测流程:

首先将图像进行颜色空间的转换,根据交通标志牌的颜色特性提取出 ROI,再基于 YOLO v3 实现对场景中交通标志牌的检测。

ROI 获取: 将输入图像的颜色空间由 RGB 转换为 HSV

$$H = \begin{cases} 0^\circ \Delta = 0 \\ 60^\circ \times \left(\frac{G' - B'}{\Delta} + 0 \right) & \mathcal{C}_{\max} = R' \\ 60^\circ \times \left(\frac{B' - R'}{\Delta} + 2 \right) & \mathcal{C}_{\max} = G' \\ 60^\circ \times \left(\frac{R' - G'}{\Delta} + 4 \right) & \mathcal{C}_{\max} = B' \end{cases}$$

$$S = \begin{cases} 0 & \mathcal{C}_{\max} = 0 \\ \frac{\Delta}{\mathcal{C}_{\max}} & \mathcal{C}_{\max} \neq 0 \end{cases}$$

$$V = \mathcal{C}_{\max}$$

式中 $R' = R/255$, $G' = G/255$, $B' = B/255$, $\mathcal{C}_{\max} = \max(R', G', B')$, $\mathcal{C}_{\min} = \min(R', G', B')$, $\Delta = \mathcal{C}_{\max} - \mathcal{C}_{\min}$ 。

在 HSV 颜色空间提取指示、警示、禁止类交通标志牌的主颜色,分别是蓝色、黄色和红色,色彩的范围如下所示:

H, S, V 三分量都归一化到 $[0, 1]$ 范围,具体地

$$\text{blue: } \begin{cases} 0.56 < H < 0.7 \\ 0.17 < S < 1.0 \\ 0.19 < V < 1.0 \end{cases}$$

$$\text{yellow: } \begin{cases} 0.06 < H < 0.19 \\ 0.17 < S < 1.0 \\ 0.19 < V < 1.0 \end{cases}$$

$$\text{red: } \begin{cases} 0.0 < H < 0.04, 0.87 < H < 1.0 \\ 0.17 < S < 1.0 \\ 0.19 < V < 1.0 \end{cases}$$

在提取完三大类标志牌主颜色后,对获得的图像进行二值化处理,通过计算二值图像中连通区域(8 连通),去除掉连通区域较小的区域(350~550),获得图像 $Image$,同时考虑图像中目标由远及近的规律,对整幅图像的底边部分($w \times h/5$)不处理。

感兴趣区域的左上点坐标 $U_L = (x_1, y_1)$,其中 $x_1 =$

$$\begin{cases} x_{\min} & x_{\min} < 4h/5 \\ 0 & \text{otherwise} \end{cases}, \quad y_1 = y_{\min}$$

右下点坐标: $D_R = (x_2, y_2)$,其中 $x_2 =$

$$\begin{cases} x_{\max} & x_{\max} < 4h/5 \\ h & \text{otherwise} \end{cases}, \quad y_2 = y_{\max}$$

式中 x_{\min} 为 $Image$ 中所有像素 x 方向上最小值, y_{\min} 为 $Image$ 中所有像素 y 方向上最小值; x_{\max} 为 $Image$ 中所有像素 x 方向的最大值, y_{\max} 为 $Image$ 中所有像素 y 方向上最大值。 h 为图像 x 方向的长度,最后获得的 ROI。

交通标志检测: YOLO 算法使用一个神经网络直接预测不同目标的类别与位置。其将输入图像分成网格,如果某个目标的中心落在网格中,则该网格负责预测此目标;每个网格单元预测一个概率值和 3 个边界框,每个边界框预测 5 个回归值($x, y, w, h, confidence$), (x, y) 为边界框的中心点坐标, (w, h) 为边界框的宽度和高度, $confidence$ 为置信度,置信度的计算: $P_r(object) \times IOU_{pred}^{truth} = confidence$ 。其中, $P_r(object)$ 为 0 或 1, 0 表示图像中没有目标, 1 表示有目标。 IOU_{pred}^{truth} 表示预测的边界框和实际的边界框之间的交并比,置信度反应是否包含目标以及包含目标情况下预测位置的准确性。置信度阈值设置为 0.5, 则当边界框置信度小于 0.5 时,删除预测的边界框;预测的边界框置信度大于 0.5 时,保留预测的边界框。

边界框预测调整: (t_x, t_y, t_w, t_h) 是网络预测的边界框的 4 个坐标,如果网格偏离图像左上角的量为 (c_x, c_y) , p_w, p_y 是先前边界框的宽度和高度,对应的预测框调整如下

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

类别预测: 在分类阶段, soft-max 有类间互斥的表现,不利于有包含属性关系的数据分类,因此使用多标签分类方法,取而代之的是多个独立的逻辑分类器,在训练阶段对于类别的预测采用的是二元交叉熵函数

$$loss = - \sum_{i=1}^n \hat{y}_i \log y_i + (1 - \hat{y}_i) \log(1 - \hat{y}_i)$$

跨尺度预测:基于特征金字塔网络的思想,获取更丰富的语义信息和更细粒度的特征信息。YOLO v3 在 3 个尺度上预测边界框,最后在每个尺度图上可以得到 $N \times N \times [3 \times (4 + 1 + 3)]$ 个结果, $N \times N$ 个网格,3 种尺度的锚点,边界框的 4 + 1 个值,3 种类别。通过使用 k-means 聚类算法获取边界框的先验信息,通过选择 9 个聚类均分到 3 种尺度上,9 个聚类如下: (10×13) , (16×30) , (33×23) , (30×61) , (62×45) , (59×119) , (116×90) , (156×198) , (373×326) 。

特征提取:YOLO v3 使用 Darknet-53 网络实现特征提取,相比较 Darknet-19 网络添加了残差单元,使用连续的 3×3 和 1×1 卷积层,该网络结果包括 53 个卷积层以及 5 个最大池化层,同时每个卷积层后增加了批量归一化操作和去除 dropout 操作,防止过拟合现象。训练过程损失函数为

$$F(\text{loss}) = \lambda_{\text{coord}} \sum_{i=1}^{S^2} \sum_{j=1}^k I_{ij}^{\text{obj}} ((x_i - \hat{x}_i)^2 + (y_i - \hat{y}_i)^2) + \lambda_{\text{coord}} \sum_{i=1}^{S^2} \sum_{j=1}^k I_{ij}^{\text{obj}} ((\sqrt{\omega_i} - \sqrt{\hat{\omega}_i})^2 + (\sqrt{h_i} - \sqrt{\hat{h}_i})^2) + \sum_{i=1}^{S^2} \sum_{j=1}^k I_{ij}^{\text{obj}} (\sqrt{C_i} - \sqrt{\hat{C}_i})^2 + \sum_{i=1}^{S^2} I_i^{\text{obj}} \sum_{c \in \text{classes}} (p_i(c) - \hat{p}_i(c))^2$$

式中 I_{ij}^{obj} 为网格单元 i 包含交通标志牌的部分, I_{ij}^{obj} 为网格单元 i 中的第 j 个边界框; I_{ij}^{obj} 为网格单元 i 中的第 j 个边界框,其不包含交通标志牌的任何部分。

2 实验

本文实验的计算机配置为: CPU E5-2670v2,显卡 Gtx1080Ti,内存 32G,操作系统为 Ubuntu16.04。为了验证本文提出的方法,同时也在智能车实验平台上进行了实验,智能车实验平台如图 2 所示,车辆配有: GPS,16 线激光雷达,前视摄像头和惯性导航等设备。

目前公开的交通标志牌数据集由于存在国内外的差异,所以标志牌数据不能直接使用,国内的数据集如 CCTS-DB 只暂时公布了一部分数据, TSD-max 数据集也由于举办“中国智能车未来挑战赛离线比赛”的需要只开放了很少一部分数据; TT100k 是清华和腾讯合作标注的交通标志牌数据,可以公开下载使用,但是该数据也存在着标注样本数量不平衡的问题。目前还没有一个能够完全涵盖所有标志牌类型且每一类数量平衡的数据集,本文实验的数据来源包括 CCTSDB、TSD-max、TT100k 和自采数据标注 4 部分共计 29 000 张标注图像,其中自行标注的数据 9 000 张,采集的地点城市包括:北京、天津、上海、苏州、常熟、深圳、重庆等地;包含不同天气、不同时间、不同季节的交通场景数据,标注的类别如图 1 所示分为三类,标注过程中对于图像中信

号灯目标区域小于的不予标注。随机抽取 23 000 张图像作为训练集,余下的 6 000 张图像作为测试集。

本文所提出方法的效果图如图 2 所示。



图 2 测试平台与识别效果

从准确率、召回率、平均精度均值和 F1 方面对算法进

行评价。准确率计算 $\text{precision} = \frac{tp}{tp + fp}$,召回率计算 $\text{recall} = \frac{tp}{tp + fn}$,其中 tp 为指正样本被正确识别为正样本; tn 为负样本被正确识别为负样本; fp 为假的正样本,即负样本被错误识别为正样本; fn 为假的负样本,即正样本被错误识别为负样本。

$$F_1 \text{ 的计算 } F_1 = 2 \times \frac{\text{precision} \cdot \text{recall}}{\text{precision} + \text{recall}}$$

用来衡量二分类模型精确度的一种指标,值越大说明模型越好。实验效果对比如表 1 所示。

表 1 实验效果对比

方法	召回率/%	查准率/%	F_1 /%	FPS
本文方法	74.80	81.60	78.05	39.2
YOLO v3	73.20	80.50	76.68	38.1
VGG16	88.18	90.24	89.20	3.6
Res50	88.67	95.12	91.78	3.0
Res101	86.21	92.76	89.36	2.9

从表 1 中前两行可以看出,本文的方法在上述各方面都得到了提升,主要得益于所提方法过滤了大部分的干扰(如汽车尾灯等),在帧率方面也得到了提升,但是提升的幅度不大,原因主要是由于在确定 ROI 方面存在一定的经验因素的影响,同时由于在处理过程中默认将图像底部 1/5 高度的区域去除掉,并且根据 ROI 选择的策略,如果 ROI 的点坐标落在底部区域,将在原始图像大小上进行交通标志牌的检测,所以提升幅度不大,本文提出的算法已经可以满足在日程行驶速度下智能车辆对交通标志牌的实时处理。表中的第 3~5 行,分别是基于 Faster RCNN 的方法对交通标志牌的检测结果,分别对应于使用 VGG16、Res50 和 Res101 三种不同的特征提取网络实现的效果,整体上召回率和查准率要优于基于 YOLO 的方法,但在处理的实时性上却远不如基于 YOLO 的方法。

本文中检测的对象为交通标志牌,分布于全路段中,为了保证能够在智能车辆正常行驶的过程中实时检测交通标志牌,从本文的实验结果中可以看出,本文所提出的方法在

准确率和实时性能上都得到了一定程度的提升,并且在实际的无人驾驶车辆平台上得到了验证,能够满足在全路段下实时的处理。

3 结束语

通过实验对比验证了所提方法的有效性,并根据交通标志牌数据集的现状,标注了9000张交通场景图像,扩充了现有的数据集。下一步将在数据的多样性和平衡性方面扩充数据,同时由于光照等因素交通标志牌的颜色也会有所差异,颜色的提取也需要进一步的完善。

参考文献:

- [1] XIONG C Z, WANG C, MA W X, et al. A traffic sign detection algorithm based on deep convolutional neural network [C]//IEEE International Conference on Signal and Image Processing, Beijing, IEEE 2016: 676–679.
- [2] GUO J K, LU J Y, QU Y Y, et al. Traffic-sign spotting in the wild via deep features [C]//IEEE Intelligent Vehicles Symposium, Changshu, IEEE 2018: 120–125.
- [3] LEE H S, KIM K. Simultaneous traffic sign detection and boundary estimation using convolutional neural network [J]. IEEE Transactions on Intelligent Transportation Systems, 2018, 19(5): 1652–1663.
- [4] LI C P, CHEN Z X, WU Q M J. Deep saliency with channel-wise hierarchical feature responses for traffic sign detection [J]. IEEE Transactions on Intelligent Transportation Systems 2018, 20(7): 2497–2509.
- [5] REHMAN Y W, KHAN J A, SHIN H. Efficient coarser-to-fine holistic traffic sign detection for occlusion handling [J]. IET Image Processing 2018, 12(12): 2229–2237.
- [6] WANG D D, HOU X W, XU J W. Traffic sign detection using a cascade method with fast feature extraction and saliency test [J]. IEEE Transactions on Intelligent Transportation Systems, 2017, 18(12): 3290–3302.
- [7] YUAN Y, XIONG Z T, WANG Q. An incremental framework for video-based traffic sign detection, tracking, and recognition [J]. IEEE Transactions on Intelligent Transportation Systems, 2017, 18(7): 1918–1929.
- [8] ZHU Z, LU J M, RALPH R. An optimization approach for localization refinement of candidate traffic signs [J]. IEEE Transactions on Intelligent Transportation Systems 2017, 18(11): 3006–3016.
- [9] ZHU Y G, LIAO M H, YANG M K, et al. Cascaded segmentation-detection networks for text-based traffic sign detection [J]. IEEE Transactions on Intelligent Transportation Systems 2018, 19(1): 209–219.
- [10] LIU Z G, DU J, TIAN F. MR-CNN: A multi-scale region-based convolutional neural network for small traffic sign recognition [J]. IEEE Access 2019, 19(7): 57120–57128.
- [11] GIRSHICK R. Fast R-CNN [C]//IEEE International Conference on Computer Vision, Boston, IEEE 2015: 1440–1448.
- [12] 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.
- [13] REDMON J, DIVVALA S, GIRSHICK R, et al. You only look once: Unified real-time object detection [C]//IEEE International Conference on Computer Vision, Las Vegas, IEEE 2016: 779–788.
- [14] REDMON J, FARHADI A. YOLOv3: An incremental improvement [C]//IEEE Conference on Computer Vision and Pattern Recognition, Salt Lake City, IEEE 2018: 89–95.
- [15] 张汇, 杜煜, 宁淑荣. 基于Faster RCNN的行人检测方法 [J]. 传感器与微系统 2019, 38(2): 147–153.

作者简介:

潘卫国(1985–),男,通讯作者,博士,讲师,研究领域为目标检测、三维可视化, E-mail: ldtweigu@ buu. edu. cn。

刘博(1995–),男,本科,专业方向为数字图像处理,目标检测。

(上接第146页)

参考文献:

- [1] 郭全民, 张海先. 基于图像处理技术的混凝土路面裂缝检测方法 [J]. 传感器与微系统 2013, 32(4): 61–64.
- [2] ABDELQADER I, ABUDAYYEH O, KELLY M E. Analysis of edge-detection techniques for crack identification in bridges [J]. Journal of Computing in Civil Engineering 2003, 17(4): 255–263.
- [3] IYER S, SINHA S K. Segmentation of pipe images for crack detection in buried sewers [J]. Computer-aided Civil & Infrastructure Engineering 2010, 21(6): 395–410.
- [4] FELZENSZWALB P F, HUTTENLOCHER D P. Efficient graph-based image segmentation [J]. International Journal of Computer Vision 2004, 59(2): 167–181.
- [5] FELZENSZWALB P F, HUTTENLOCHER D P. Image segmentation using local variation [C]//IEEE Computer Society Conference on Computer Vision and Pattern Recognition, IEEE Computer Society, 1998: 98.
- [6] 张晶晶, 聂洪玉, 喻强. 基于多尺度输入图像渗透模型的桥梁裂缝检测 [J]. 计算机工程 2017, 43(2): 273–279.
- [7] 李刚, 贺拴海, 巨永锋, 等. 远距离混凝土桥梁结构表面裂缝精确提取算法 [J]. 中国公路学报 2013, 26(4): 102–108.
- [8] 吴赵丽, 梁栋, 赵晋陵, 等. 基于无人机遥感影像茛苕草盖度的图像分割方法 [J]. 传感器与微系统 2018, 37(4): 51–53.
- [9] 马荣贵, 徐琨, 刘飞飞. 高速公路路面裂缝识别算法研究 [J]. 交通信息与安全 2014(2): 90–94.
- [10] 侯伟. 基于透视图的二次曲面重建 [D]. 西安: 西北工业大学 2006.

作者简介:

张晓(1990–),男,硕士研究生,研究方向为数字图像处理, E-mail: 1938729779@qq.com。

高忠文(1961–),男,教授,研究生导师,主要研究领域为计算机信息领域,故障模式分析。