

基于YOLOv8的交通信号灯识别

赵恩兴¹, 王超^{1,2}

¹宿州学院, 信息工程学院, 安徽 宿州

²同济大学, 电子与信息工程学院, 上海

收稿日期: 2023年7月14日; 录用日期: 2023年8月17日; 发布日期: 2023年8月28日

摘要

交通信号灯的识别对于辅助驾驶系统是至关重要的, 它可以帮助减少事故和提高行车安全。本文提出了基于YOLOv8的交通信号灯标志识别方法, 该方法包括数据集的构建、模型的训练、自然场景测试三个主要部分。首先, 通过网络公开的交通信号灯数据集进行标注, 使用YOLOv8算法框架对数据集进行训练, 得出最优模型。最后, 在真实道路场景中对训练好的模型进行了测试, 得到了较为准确的结果。通过实验对比, 我们发现YOLOv8训练后的模型性能优异, 在保证精度的情况下提高检测速度, 还可以解决目标部分遮挡和小尺寸目标检测等问题, 从而提高了识别的准确性和效率。在辅助驾驶系统中应用该方法可以更加精确地判断箭头指向性信号灯和全屏型信号灯, 帮助提高车辆在路面上的运动安全性和稳定性。目前的大多方法仅仅针对于交通信号灯的颜色以及整体交通信号灯位置进行判断识别, 本文会更细化交通灯上各式各样的方向标志颜色做出分类识别, 通过YOLOv8算法在减少参数的情况下还能够大幅度减少计算资源, 通过实验结果表明, 迭代200轮后的模型mAP50-95便达到了82.6%, FPS达到了27.2帧/毫秒。

关键词

YOLOv8, 交通信号灯, 模型训练, 目标检测

Traffic Signals Recognition Based on YOLOv8

Enxing Zhao¹, Chao Wang^{1,2}

¹School of Informatics and Engineering, Suzhou University, Suzhou Anhui

²School of Electronics and Information Engineering, Tongji University, Shanghai

Received: Jul. 14th, 2023; accepted: Aug. 17th, 2023; published: Aug. 28th, 2023

Abstract

The recognition of traffic lights is crucial for driver assistance systems, which can help reduce acci-

文章引用: 赵恩兴, 王超. 基于YOLOv8的交通信号灯识别[J]. 人工智能与机器人研究, 2023, 12(3): 246-254.

DOI: 10.12677/airr.2023.123028

dents and improve driving safety. This paper proposes a traffic light sign recognition method based on YOLOv8, which includes three main parts: data set construction, model training and natural scene testing. First, the traffic light data set disclosed through the network was annotated, the YOLOv8 algorithm framework was used to train the data set, and the optimal model was obtained. Finally, the trained model is tested in a real road scene, and the results are more accurate. Through experimental comparison, we found that the model trained by YOLOv8 has excellent performance, improves detection speed while ensuring accuracy, and can also solve problems such as partial occlusion of the target and small-size target detection, thus improving the accuracy and efficiency of recognition. The application of this method in the auxiliary driving system can judge the arrow directional signal light and the full-screen signal light more accurately, and help to improve the safety and stability of the vehicle on the road. Most of the current methods only judge and identify the color of traffic lights and the overall position of traffic lights. This paper will further refine the classification and recognition of various direction sign colors on traffic lights. The YOLOv8 algorithm can greatly reduce computing resources while reducing parameters. Experimental results show that after 200 iterations, the model achieves a mAP50-95 of 82.6% and an FPS of 27.2 frames per millisecond.

Keywords

YOLOv8, Traffic Signals, Model Training, Pattern Recognition

Copyright © 2023 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

随着汽车工业的快速发展,自动驾驶技术逐渐成为了未来汽车行业的发展方向和趋势。在实现自动驾驶的道路上,交通灯信号灯标识是非常重要的交通设施,在保障交通安全和顺畅通行方面起到至关重要的作用。对于自动驾驶车辆来说,识别路面交通标志是必不可少的一项任务。

目前有许多现有的对交通信号灯识别的方法,有基于深度学习和 Opencv 的交通灯识别算法研究,利用 Opencv 提取目标为交通灯的图片区域,将各个像素点的颜色由 RGB 空间转换到 HSV 空间[1]。在 HSV 空间内,分别统计红色、绿色和黄色的点数,此方法能够对于交通灯不同颜色进行有效分类与识别;有基于卷积神经网络的交通信号灯实时识别方法[2] [3] [4],这项研究使用 YOLOv3 算法,通过改进特征提取、调整特征融合方式及尺度、优化损失函数来提升交通信号灯的检测效果;有基于深度学习的交通信号灯识别算法[5] [6] [7],提出了 DA-FasterR-CNN [8]算法对候选区域中的交通灯进行检测识别,利用改进的 ResNext [9]网络替代 VGG [10]作为网络骨干以提高神经网络浅层特征层对细节信息提取能力,引入特征融合模块,将从基础网络中输出的两个不同尺度特征层的信息进行融合以产生一个新的高级语义特征层,提高了小型目标的检测性能。笔者在此提出了基于 YOLOv5 [11]算法对格式各样的交通信号灯标志进行识别,并通过最新的 YOLOv8 [12]结构对相同数据集进行训练评估,以实现更快的模型拟合速度以及更优的模型精度,因为模型可以对于国内当前较多场景的情况做出判断,如图 1 所示,最上层的图是原始图,中间层是传统交通信号灯识别,最后一层是本文训练出的模型推理后的图片,所以可将模型部署在应用程序,使得该程序亦可应用于智能导航系统,将交通信号灯标志识别结果上传至导航系统中,进行数据比较,结合地图信息、车辆定位以及实时交通信息,以对车辆行驶决策提供帮助。



Figure 1. Traffic light recognition
图 1. 交通灯识别

2. 算法概述

2.1. YOLOv5 概述

YOLOv5 是 Ultralytics 于 2020 年推出的一种物体检测模型, Ultralytics 是最初的 YOLOv1 和 YOLOv3 的创始人。YOLOv5 在 COCO 基准测试数据集上实现了 SOTA 性能, 同时训练和部署也非常快速和高效。YOLOv5 进行了一些架构上的更改, 最值得注意的是将模型构造成三个组件(the backbone, neck, and head)的标准化实践。YOLOv5 的主干网络是 Darknet53, 这是一种新的网络架构, 专注于特征提取, 其特征是小的过滤器窗口和残留连接。部分连接使体系结构能够实现更丰富的梯度流, 同时减少计算量。在 YOLOv5 中描述的 neck 连接着 backbone 与 head, 其目的是对 backbone 提取的特征进行聚合和细化, 重点是增强跨不同尺度的空间和语义信息。空间金字塔池(Spatial Pyramid Pooling, SPP) [13]模块消除了网络的固定大小限制, 从而消除了扭曲、增大或裁剪图像的需要。然后是 CSP-Path Aggregation Network 模块 [14], 该模块整合了主干中学习到的特性, 缩短了低层和高层之间的信息路径。YOLOv5 的头部由三个分支组成, 每个分支预测不同的特征尺度。在模型的最初发布中, 创建者使用了三个网格单元格大小, 分别为 13×13 、 26×26 和 52×52 。每个 head 生成边界框、类概率和置信分数。最后, 网络使用非最大抑制(Non-maximum Suppression, NMS)过滤掉重叠的边界框。

2.2. YOLOv8 概述

YOLOv8 [14]是 YOLO 对象检测模型的最新版本。这个最新版本的架构与之前的版本相同, 但与 YOLO 的早期版本相比, 它有很多改进, 比如新的神经网络架构, 利用了特征金字塔网络(FPN)和路径聚合网络(PAN), 如图 2 所示, 以及一个新的标签工具, 简化了注释过程。这个标记工具包含一些有用的功能, 比如自动标记、标记快捷方式和可定制热键。这些特性的组合使得为训练模型而对图像进行注释变得更加容易。FPN 的工作原理是逐渐降低输入图像的空间分辨率, 同时增加特征通道的数量。这就产生

了能够在不同的比例尺和分辨率下检测物体的特征地图。另一方面, PAN 架构[15]通过跳过连接来聚合不同层次的网络特性。通过这样做, 该网络可以更好地捕捉多个尺度和分辨率的特征, 这对于准确地检测不同大小和形状的物体至关重要。

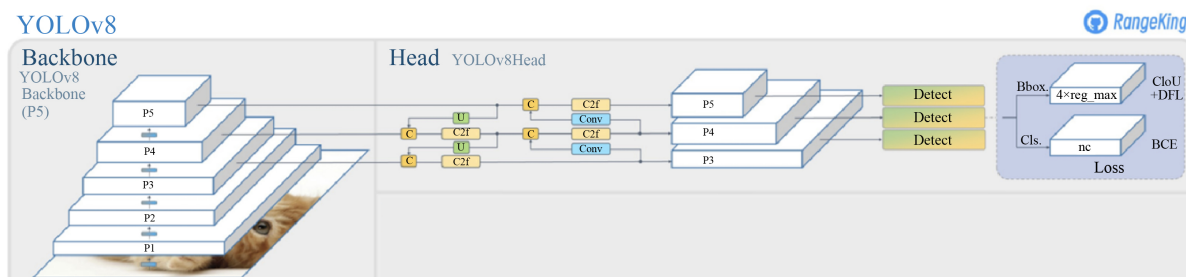


Figure 2. YOLOv8 architecture

图 2. YOLOv8 架构

3. 实验

3.1. 数据集

实验数据集来自国内真实道路场景数据, 共计 7947 张图片数据集示例如图 3 所示:



Figure 3. Dataset example

图 3. 数据集实例

对初始数据使用 YOLOv8n 模型进行初步识别, 再进行人工筛选与标定后得到可用数据共 7947 张图片, 共十八类包括: 红色向前 rup、红色向右 rright、红色向左 rleft、绿色向前 gup、绿色向右 gright、绿色向左 gleft、绿色向下 gdown、红色自行车 rbike、绿色自行车 gbike、红色行人 rperson、绿色行人 gperson、红色圆圈 rcircle、绿色圆圈 gcircle、绿色转弯 gturn、红色转弯 rturn、黄色圆圈 ycircle、黄色向前 yup、黄色向右 yright 黄色对数据集进行分析后得到的可视化结果如图 4 所示。

由可视化结果经分析可以看出 gcircle 类和 rleft 类的数量较多, 反应了国内道路交通状况。从 label 大小分布可以看出 label 的大小主要分布在 0.3~0.7 之间, 其中, x 表示目标框中心点在图像的水平方向上的位置, y 表示目标框中心点在图像的垂直方向上的位置。也反映了道路交通信号灯数据的另一个特点就是整体识别目标而言小目标较多。这将直接影响模型对目标检测的准确性, 也是交通信号灯目标检测的难点。

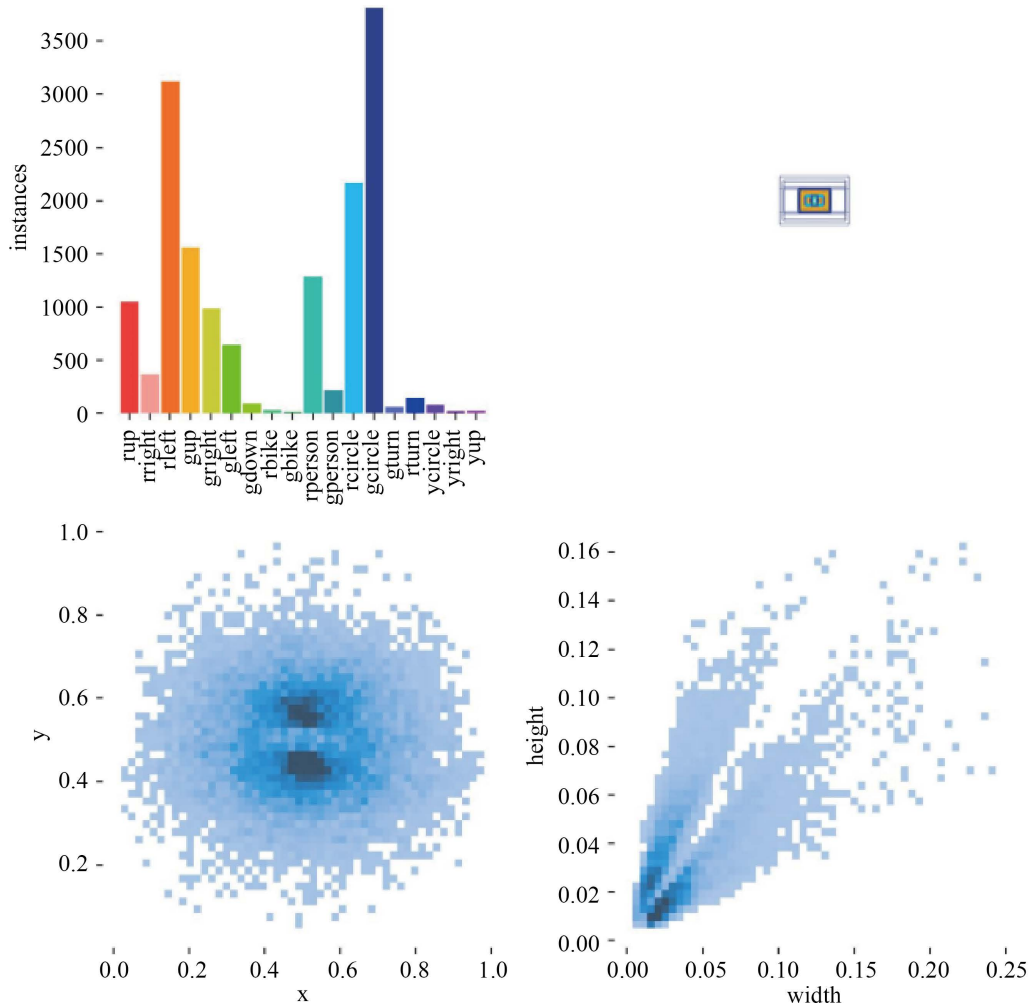


Figure 4. Number, size and center point distribution of labels
图 4. 标签数量、大小及中心点分布

3.2. 实验环境与模型训练

训练参数：预训练权重为 yolov8n.pt，输入图像尺寸为 640×640 ，最大迭代次数 250 轮，batch_size 按 GPU 规格设置为 32，训练线程设置为 2，其余参数默认。本实验环境配置如表 1 所示：

Table 1. System resulting data of standard experiment
表 1. 标准试验系统结果数据

参数	配置
系统环境	Windows11
CPU	12thGenIntelCorei7-12700
GPU	RTX3080
GPU 加速	CUDA10.2
语言	Python3.9
训练框架	Pytorch

3.3. 性能评估

Precision-confidence curve (准确率 - 置信度曲线)是在机器学习和信息检索任务中用于评估模型性能的一种可视化工具。它显示了模型预测结果的准确率与置信度之间的关系。在二分类任务中,模型通常会给出一个置信度或概率来表示对样本属于正类的度量。通过调整阈值,可以将置信度转换为分类预测结果。准确率是指在某个置信度阈值下,模型的预测结果中真正例的比例。准确率 - 置信度曲线通过在不同置信度阈值下计算准确率,然后将其绘制在坐标系中,展示了模型在不同置信度水平下的性能。从图 5 中可以发现随着置信度的提高,所有分类的准确率均可以达到 100%,但是可能存在样本在训练集和验证集上的分布不平衡。

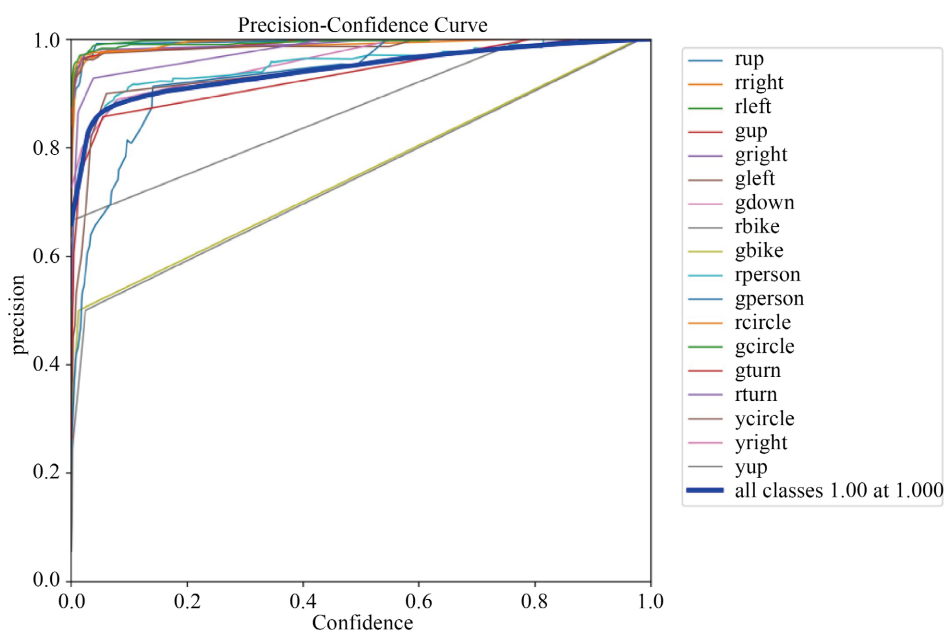


Figure 5. Precision-confidence curve
图 5. 精度置信曲线

Recall-confidence curve (召回率 - 置信度曲线)是用于评估机器学习模型性能的一种可视化工具,特别适用于二分类任务。该曲线显示了模型在不同置信度阈值下的召回率与置信度之间的关系。在二分类任务中,模型通常会输出一个置信度或概率来表示样本属于正类的度量。通过调整阈值,可以将置信度转换为分类预测结果。召回率是指在某个置信度阈值下,模型正确预测为正类的比例。在绘制召回率 - 置信度曲线时,我们根据不同的置信度阈值计算召回率,并将其绘制在坐标系中。曲线上的每个点表示了一个置信度阈值下的召回率。通常情况下,随着置信度阈值的增加,模型的召回率会降低。这是因为较高的置信度阈值会将更多的样本划为负类,导致一些正类样本被错误分类为负类。从图 6 中可以看出所有分类均能到达 98%。说明模型在预测所有类别时,能够正确地找到正例的比例非常高,只有很少一部分正例被模型错误地标记为负例或其他类别。

3.4. 实验结果

混淆矩阵(confusion matrix)是一种可以直观反映算法性能的可视化效果图,通过对混淆矩阵的分析可以直观的看到模型对于每一类目标的检测准确率,以及识别错误的主要原因。以下将通过分析混淆矩阵得到本文训练的模型的准确率以及影响准确率的主要因素,它以矩阵形式呈现了模型在不同类别上的预

测结果与实际标签之间的对应关系。预测为正类 TP (真正例) FP (假正例) 预测为负类 FN (假反例) TN (真反例) 其中, TP 表示模型将正类正确预测为正类的样本数量, FP 表示模型将负类错误预测为正类的样本数量, FN 表示模型将正类错误预测为负类的样本数量, TN 表示模型将负类正确预测为负类的样本数量。

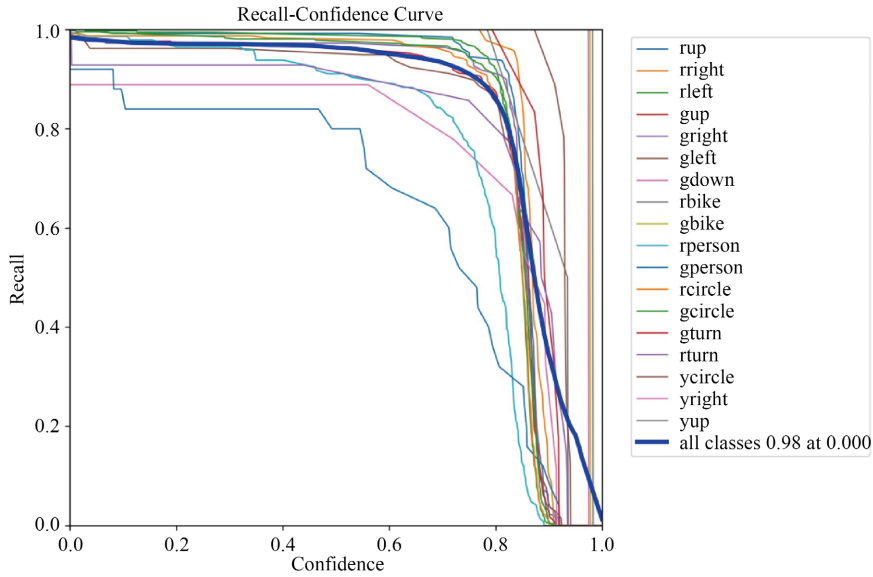


Figure 6. Recall-confidence curve
图 6. 召回置信曲线

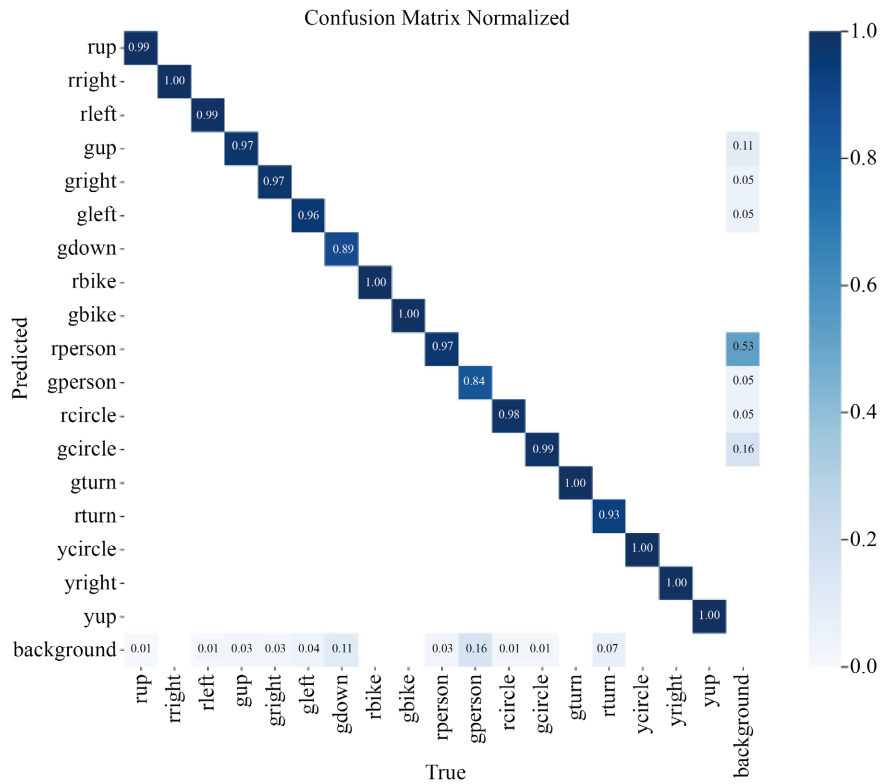


Figure 7. Confusion matrix
图 7. 混淆矩阵

从图7可以看出本实验训练的目标检测框架对于交通信号灯检测有0.826 mAP50-95,精确率为0.98。除 gperson 类外的其他类 mAP 值和精确率都相对较高,主要是因为交通信号灯对于绿色行人标志是一个动态标志,该标志在实际环境中相对较小并且其特征不一,而且模型训练时会进一步压缩图片,导致小目标的特征提取较困难。

部分目标检测效果图如图8所示:



Figure 8. Test result example

图8. 检测结果实例

4. 模型部署

YOLOv8 训练出的模型可以部署在 GUI 客户端程序、Web 网页程序也可部署在算力足够的硬件开发板上,本文暂以 Web 网页程序做实例,Web 网页可以方便部署在大多服务器上,调用训练模型进行部署推理,需要调用对应的 torch.hub.load ()函数,其次定义需要推理的视频或图片放在输入端,最后定义输出端推理结果输出,Web 端如图9所示,其中 conf_thres 和 iou_thres 是在目标检测任务中常用的两个参数,conf_thres (置信度阈值)是一个用于过滤检测结果的阈值。在目标检测中,每个检测结果都会伴随一个置信度(confidences core),表示算法对该结果的预测可信程度。通过设置 conf_thres,我们可以筛选出具有高置信度的检测结果,过滤掉低置信度的结果。iou_thres (交并比阈值)用于判断两个检测结果之间是否重叠。在目标检测中,通常使用交并比(IntersectionoverUnion, IoU)来衡量两个边界框之间的重叠程度。当两个边界框的 IoU 值大于等于 iou_thres 时,表示它们之间存在重叠。通过设置 iou_thres,我们可以控制在多个相似目标中选择哪个作为最终检测结果,用户可以根据自己需求添加更多参数进行应用。



Figure 9. Example of web-side inference

图9. Web 端推理示例

5. 总结

本文提出一种基于 YOLOv8 实现的交通信号灯标志识别, 实现不同环境下图像或者视频中交通信号灯标志的识别。在训练数据集的过程中需要注意迭代次数不能过少, 在实验过程中就出现了因迭代次数过少而过欠拟合从而导致训练模型精度下降的现象, 后来经过调整迭代次数, 识别精度又回归正常。下一步笔者将继续扩大数据集中交通信号灯标志数量, 主要增添不常见的、地域不同的, 数据集中实例较少的交通信号灯标志种类实例, 以及增加其他极端情况下的图像, 从而进一步提高模型分类以及训练精度。另外对已有的基于深度学习的各模型与各优化算法进行优化研究, 并探寻两者之间更好的融合方式, 以达到更高的识别准确率, 能够更好适应现实道路情况。

致谢

文章受到项目支持: 安徽大学信息材料与智能感知安徽省实验室开放课题(IMIS202114)基于深度学习自注意力机制和尺度自适应行人重识别研究。宿州学院重点科研项目(2021yzd01)基于深度学习的注意力和多尺度行人重识别研究。

参考文献

- [1] 余楚礼, 朱强. 基于深度学习和 Opencv 的交通灯识别算法研究[J]. 上海汽车, 2019(7): 19-22.
- [2] 董晓玉. 基于卷积神经网络的交通信号灯实时识别方法研究[D]: [硕士学位论文]. 合肥: 安徽建筑大学, 2020.
- [3] 罗雯. 基于深度学习的交通信号灯识别算法研究[D]: [硕士学位论文]. 南昌: 南昌大学, 2020.
- [4] 王莹, 丁鹏. 基于深度学习的交通信号灯检测及分类方法[J]. 汽车实用技术, 2018(17): 89-91.
- [5] 谭思奇. 基于深度学习的道路交通信号灯检测与识别方法研究[D]: [硕士学位论文]. 重庆: 重庆交通大学, 2022.
- [6] 邓天民, 谭思奇, 蒲龙忠. 基于改进 YOLOv5s 的交通信号灯识别方法[J]. 计算机工程, 2022, 48(9): 55-62.
- [7] 崔帅华. 基于深度学习的交通信号灯检测与识别方法研究[D]: [硕士学位论文]. 焦作: 河南理工大学, 2021.
- [8] Chen, Y.H., Li, W., Sakaridis, C., Dai, D.X. and Van Gool, L. (2018) Domain Adaptive Faster R-CNN for Object Detection in the Wild. 2018 *IEEE/CVF Conference on Computer Vision and Pattern Recognition*, Salt Lake City, 18-23 June 2018, 3339-3348. <https://doi.org/10.1109/CVPR.2018.00352>
- [9] Qi, C., Gao, J., Pearson, S., et al. (2022) Tea Chrysanthemum Detection under Unstructured Environments Using the TC-YOLO Model. *Expert Systems with Applications*, **193**, Article ID: 116473. <https://doi.org/10.1016/j.eswa.2021.116473>
- [10] Wu, J., Zhang, H., Chamola, V., et al. (2022) Brain-Computer Interface-Based Target Recognition System Using Transfer Learning: A Deep Learning Approach. *Computational Intelligence*, **38**, 139-155. <https://doi.org/10.1111/coin.12451>
- [11] 郭磊, 薛伟, 王邱龙, 等. 基于改进 YOLOv5 的小目标检测算法[J]. 电子科技大学学报, 2022, 51(2): 8.
- [12] Solawetz, J. and Francesco (2023) What is YOLOv8? The Ultimate Guide. <https://roboflow.com/>
- [13] He, K.M., Zhang, X.Y., Ren, S.Q. and Sun, J. (2015) Spatial Pyramid Pooling in Deep Convolutional Networks for Visual Recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **37**, 1904-1916.
- [14] Treven, J.R. and Cordova-Esparaza, D.M. (2023) A Comprehensive Review of YOLO: From YOLOv1 to YOLOv8 and beyond. <https://arxiv.org/pdf/2304.00501.pdf>
- [15] Wang, C.Y., Liao, H.Y.M., Yeh, I.H., Wu, Y.H., Chen, P.Y. and Hsieh, J.W. (2019) CSPNet: A New Backbone That Can Enhance Learning Capability of CNN. 2020 *IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*, Seattle, 14-19 June 2020, 1571-1580.