

基于深度学习的汽车油箱盖区域检测

李 锁¹, 杨凯悦¹, 由金池¹, 廖 伟², 李悦宁³

(1. 沈阳理工大学机械工程学院, 辽宁 沈阳 110159, 2. 沈阳航空航天大学工程训练中心, 辽宁 沈阳 110136, 3. 中国科学院大学, 北京 100049)

摘要:针对传统的目标检测方法油箱盖形状识别准确率低的问题, 提出融合注意力机制和深度学习的汽车油箱盖区域检测方法。首先, 通过添加注意力模块对 YOLOv5 算法进行改进, 加强提取检测对象的特征, 提升 YOLOv5 算法对小目标对象的检测准确率; 然后, 在 C3 模块以及主干网络中最后一层分别添加 SE 模块以及 CBAM 模块; 最后, 通过 Loss 以及 mAP 训练结果对比得出在哪个位置添加何种模块对目标检测准确率提升效果更好。实验结果表明, 在 C3 部分融合了 SE 模块之后, 损失更加小接近 0, 提升了检测能力, mAP 提高了 0.5%, 基于改进算法后进行目标检测, 置信度提升了 2% 左右。

关键词: 目标检测; YOLOv5 算法; 注意力机制; 深度学习

中图分类号: TP24

文献标识码: A

文章编号: 2096-9759(2023)06-0013-04

Area Detection of Automobile Fuel Tank Cap Based on Deep Learning

LI Suo¹, YANG Kaiyue¹, YOU Jinchi¹, LIAO Wei², LI Yuening³

(1. College of Mechanical Engineering, Shenyang Ligong University, Shenyang 110159, China;

2. Engineering Training Center of DUT, Shenyang Aerospace University, Shenyang 110136, China;

3. University of Chinese Academy of Sciences, Beijing 100049, China)

Abstract: Aiming at the low accuracy of the traditional target detection method for the shape recognition of the fuel tank cap, a new method for the detection of the fuel tank cap region based on attention mechanism and deep learning is proposed. First of all, YOLOv5 algorithm is improved by adding attention module to enhance the detection accuracy of YOLOv5 algorithm for small target objects; Then, SE module and CBAM module are added to C3 module and the last layer of backbone network respectively; Finally, through the comparison of the training results of Loss and mAP, we can find out which module is added at which position to improve the accuracy of target detection better. The experimental results show that after the SE module is partially fused in C3, the loss is smaller and closer to 0, and the detection ability is improved. The mAP is improved by 0.5%. The target detection based on the improved algorithm increases the confidence by about 2%.

Key words: target detection, YOLOv5 algorithm, attention mechanism, deep learning

1 引言

随着人工智能技术、物联网技术、语音技术等迅猛发展, 人们对于生活品质的追求也越来越高, 传统人工式加油服务已经无法满足当今互联网时代用户的体感, 以科技化、现代化、创新化为核心的无人智能加油站已经成为新方向。由于传统的目标检测方法油箱盖形状识别准确率低, 为了解决上述问题, 本文基于引入注意力机制深度学习方法, 提出了一种对于汽车油箱盖区域进行检测的方法。

基于深度学习的目标检测优于传统目标检测, 深度学习是一种卷积神经网络, 它由卷积层、池化层和全连接层组成^[1], 通过对卷积层进行网络训练, 研究数据的特性并根据特征点得出结论。其中视觉注意力机制在图像领域有着广泛的应用, 很多显著性检测方法也通过加入注意力模块抑制背景噪声, 突出检测目标。陈维境等提出将改进的通道注意力机制与空间注意力机制并行结合, 输出的两个注意力特征加权融合再反馈至通道-空间联合注意力机制中, 从而得到细粒度更高的显著图; 王凯诚等提出了注意力精炼模块, 对通道单独整合的注意力进行非局部的空间操作, 并且用显著真值监督空间操作后的特征, 增强预测位置的准确性^[2-4]。

本文的使用场景是加油站, 由于监测目标是汽车油箱盖, 采用单目标检测。基于改进的 YOLOv5 算法, 通过“标定-训练-测试”反复迭代的方式得到一个更加适合在加油站检测一体化油枪准确找到油箱盖的深度学习模型。

2 基于传统的汽车油箱盖区域检测

传统的检测方法基于标志性特点进行提取处理, 通过提取算法, 将提取到的部分与模板匹配或者由分类器进行识别检测。当提取的目标对象结构分明具有丰富的纹理属性时, 可以通过特定算法精确提取到目标对象的特征点, 采用点重合的方式对这些特征点进行匹配, 确定目标区域; 反之, 区域的检测主要取决于目标对象的形状或轮廓。

传统目标检测根据形状特征进行匹配检测, 分为两部分: 第一部分是创建提取的特征模板, 第二部分是与创建出的模板进行匹配, 确定检测区域。主要流程为: 首先对输入图像进行灰度处理、均衡化操作、中值滤波等处理, 然后通过选定阈值对图像进行分割使油箱盖周围形成连通区域, 之后提取出符合油箱盖形状特征的区域, 将其作为参考特征模板, 最后通过模板匹配确定目标对象的位置, 具体步骤如图 1 所示。

收稿日期: 2023-03-03

基金项目: 辽宁省教育厅面上青年人才项目(LJKZ0258); 辽宁省科技厅博士科研启动基金计划项目(2022-BS-187)。

作者简介: 李锁(1979-), 男, 辽宁锦州人, 副教授, 博士, 主要研究方向: 机器人控制技术、嵌入式系统; 杨凯悦(2001-), 女, 辽宁锦州人, 研究生, 硕士, 主要研究方向: 智能机器人。

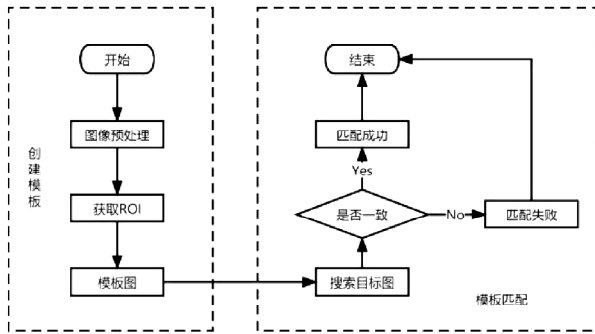


图1 传统目标检测的流程图

3 基于深度学习的汽车油箱盖区域检测

3.1 YOLOv5s 算法介绍

2020年6月, Ultralytics 团队提出了 YOLOv5 算法, 相比于 YOLOv4, v5 更加轻量、快速、准确, 同时在框架上也有改进, v4 使用的是 Darknet, 而 v5 使用更简单、生态更成熟的 Pytorch, 更适用工程应用。在 YOLOv5 算法中包含有 5s、5m、5l、5x 四种训练模型, 他们的不同点体现在网络特定位置的特征提取模块和卷积核的数量, 提取和融合不同特征的能力随着网络的加宽加深也变得越来越强。本文选择最简洁的 YOLOv5s 模型作为检测汽车油箱盖的网络模型, 其结构由输入端 Input、主干网络 Backbone、颈部网络 Neck 以及输出端 Head 四个部分组成^[5], 如图2所示。

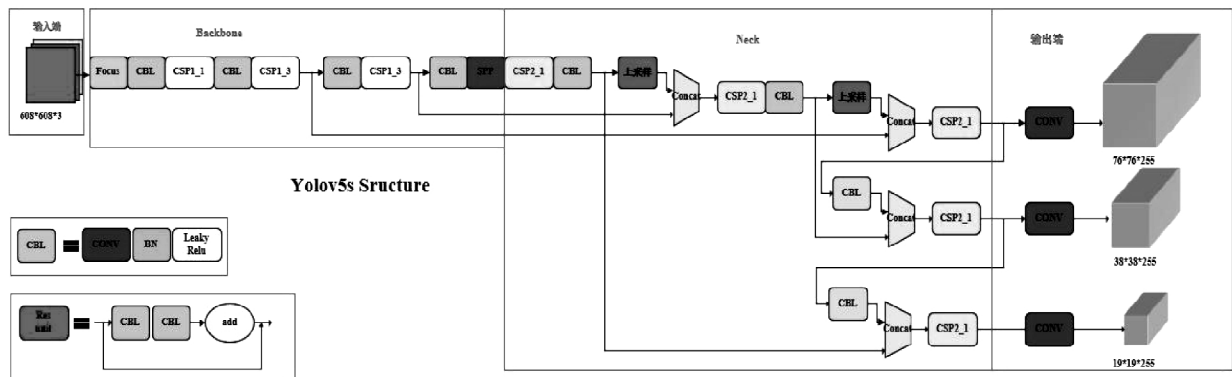


图2 YOLOv5s 模型网络结构

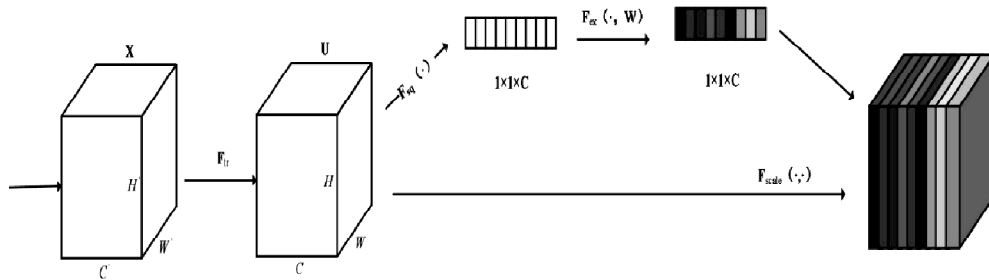


图3 SE 模块的原理和结构

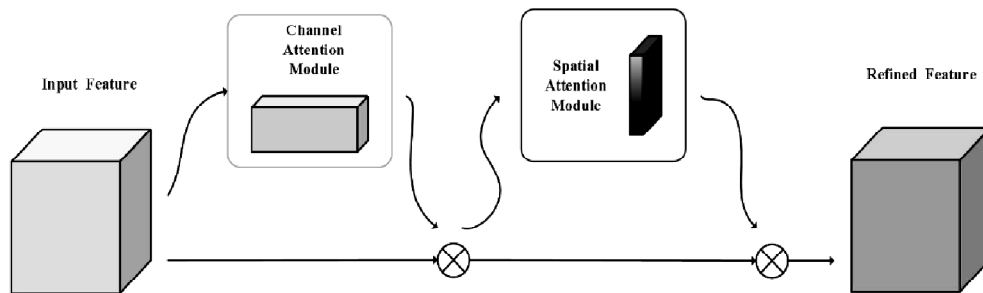


图4 CBAM 模块的原理和结构

Input 端主要对数据进行处理, 通过 Mosaic 算法^[6]进行数据增强, 同时对图片进行自适应缩放^[7]。Backbone 网络主要的目的是对输入的图像进行通用特征提取并传递给下一网络。Yolov5 保留了 v4 的跨阶段局部融合网络 (Cross Stage Partial, CSP) 结构的同时加入了一种新的结构——Focus 结构, 这两种结构的融合加强了网络特征提取的能力。Focus 结构位于 Backbone 的首位, 主要将图像进行切片处理。颈部网络介于主干网络和输出端之间, 用来将主干网络传递过来的图像特征融合聚集在一起, 提高特征的稳定性。输出端网络结构用来对

边界框预测和回归, 主要涉及到边界框的预测以及损失计算, 分别通过 CIOU_Loss^[8]函数以及非极大值抑制 (Non-maximum suppression, NMS)^[9]函数进行计算。

3.2 基于改进的 YOLOv5s 算法

随着深度学习的不断改进, 其程度也逐渐复杂, 但检测小目标对象时往往效果不太理想, 因此通过在算法中添加注意力机制模块可以提升小目标对象检测的能力。在网络结构中增添注意力机制其原理跟人们处理重要事情时给与更多的关注十分相似, 都是对关键信息进行强化处理, 在神经网络结构

中配置注意力机制进行模型训练, 可以更有效捕捉到目标的关键信息, 进而使检测的准确率得到改善。

目前, 以 SE^[10](Squeeze and Excitation)为代表的通道注意力机制和以 CBAM^[11](Convolutional Block Attention Module)为代表的空间注意力机制是发展的主流。在卷积神经网络中加入 SE 和 CBAM 注意力机制, 均可以让网络更加关注待检测对象, 提高检测效果, 只不过 SE 模块更注重通道信息, CBAM 模块相比于 SE 模块多了一个空间机制(Spatial Attention), 可以达到两全其美的效果。图 3 为 SE 模块的原理和结构图, 图 4 为 CBAM 模块的原理和结构图。

为了加强提取检测对象的特征, 本文通过在算法中分别添加 SE 模块和 CBAM 模块对原有 YOLOv5s 算法进行改进。但是通过现有的资料发现, 并没有明确表明在具体什么位置添加哪一种注意力机制模块可以提高检测效率。

4 实验结果与分析

4.1 卷积网络训练结果

本文分别在 YOLOv5s 的 C3 模块中以及主干网络中最后一层分别添加 SE 模块以及 CBAM 模块进行融合, 分别对 models 下的 yolov5s.yaml、common.py 以及 yolo.py 进行修改, 将未改进的 YOLOv5s 与融合注意力机制模块的新 YOLOv5s 进行对比, 通过训练的得到 Loss 以及 mAP 训练结果对比图分别为图 5、图 6 所示, 其中模型 A、B、C、D、E 分别代表原有的 YOLOv5s、在骨干网络中最后一层加上 SE 机制、在 C3 结构中融入 SE 机制、在骨干网络中最后一层加上 CBAM 机制以及在 C3 结构中融入 CBAM 机制, 其结果如表 1 所示。

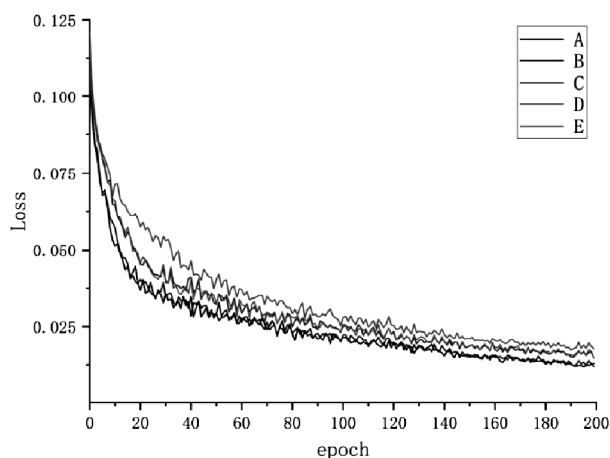


图 5 Loss 训练结果对比

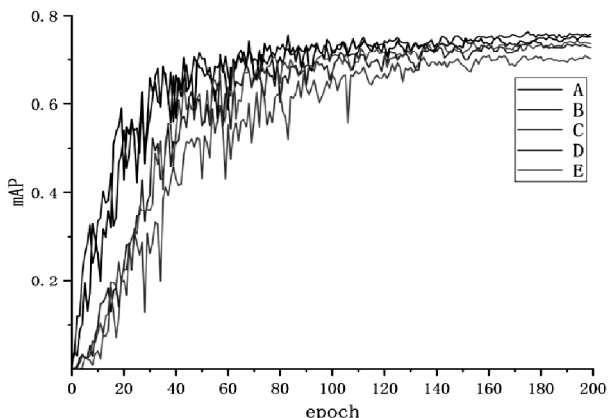


图 6 mAP 训练结果对比

表 1 融合注意力机制模块后的对比表

网络模型	Box_loss	Prediction(%)	Recall(%)	mAP(%)
YOLOv5s	0.025	93.67	91.47	93.30
SE	最后一次	0.029	88.09	87.33
	C3	0.025	92.81	91.47
CBAM	最后一次	0.030	86.89	86.75
	C3	0.034	86.62	81.71

由上表可知, 将两种注意力机制模块分别在主干网络的不同位置共进行了四次融合, 并不是每次融合所监测的精准度都会发生提升。当 C3 部分融合了 SE 模块之后, 损失更加小接近 0, 小范围提升了检测能力, mAP 提高了 0.5%, 但是在 C3 部分融合 CBAM 模块, 检测能力和精准度都有所下降; 同样在骨干网络最后一层分别加入 SE 模块以及 CBAM 模块, 检测能力和精准度均无法产生提升。综上所述, 本文选择在主干网络中 C3 位置融合 SE 模块。

4.2 检测调整

当车辆驶入可观测的视角范围内并停止运动, 可以通过摄像头识别目标对象, 即汽车油箱盖, 但由于我们训练类别只有一类, 所以在测试过程中, 许多相似形状的物体也会被检测为标定对象, 这会对结果造成影响, 这时要对输出结果的置信度加以判断, 经过多次训练, 发现产生干扰结果的置信度均小于 0.5, 故预定值为 0.5, 当置信值小于预定值则不识别, 反之则可以显示识别对象, 这样可以避免识别过程中出现的干扰物, 提高了识别的准确性, 如图 7、图 8 所示。



图 7 原有检测结果图



图 8 调整后检测结果图

4.3 目标检测

图 9、图 10 分别是基于原本和改进后算法测试后的对比图, 图 11、图 12、图 13 表示在不同光照条件下对改进后的算法进行目标测试过程图, 其结果如表 2 所示。



图 9 基于原有算法的测试



图 10 基于改进算法的测试



图 11 光照良好的测试图



图 12 光照较差的测试图



图 13 光照条件较强的测试图

表 2 基于改进算法的目标测量结果

序号	光照条件	测试置信度
1	良好	0.94
2	较差	0.86
3	较强	0.87

分析: 基于改进算法后进行目标检测, 置信度提升了 2% 左右, 同时光照条件对检测结果影响也很大。

5 结语

本文采集大量汽车油箱盖图片作为数据集, 分别采用传统方法以及深度学习方法对汽车油箱盖进行检测, 通过对比确定本文目标检测的方法, 并在此基础上通过添加注意力机制对 YOLOv5 算法进行改进, 提高检测准确率。当 C3 部分融合了 SE 模块之后, 小范围提升了检测能力, mAP 由 93.30% 提升至 93.80%; 置信度由传统算法的 0.90 提升至 0.92, 在光照良好的情况下, 置信度可达到 0.94。

参考文献:

- [1] 陈国军, 陈巍, 郁汉琪. 基于深度学习的单目视觉水下机器人目标跟踪方法研究 [J]. 机床与液压, 2019, 47(23).
- [2] 李元贞, 赵俊松. 基于深度学习的图像显著性目标检测研究综述 [J]. 软件工程, 2023, 26(01).
- [3] 陈维靖, 周萍, 杨海燕, 等. 通道-空间联合注意力机制的显著性检测模型 [J]. 计算机工程与应用, 2021, 57(19): 214-219.
- [4] 王凯诚, 鲁华祥, 龚国良, 等. 基于注意力机制的显著性目标检测方法 [J]. 智能系统学报, 2020, 15(5): 956-963.
- [5] 仇娇慧, 贝绍铁, 尹明峰, 等. 基于改进 YOLOv5s 的齿轮表面缺陷检测 [J]. 现代制造工程, 2022(03): 104-113.
- [6] Bochkovskiy A, Wang C Y, Liao H. YOLOv4: Optimal Speed and Accuracy of Object Detection [J]. 2020, 45(2): 69-75.
- [7] 王淑青, 顿伟超, 黄剑锋, 等. 基于 YOLOv5 的瓷砖表面缺陷检测 [J]. 包装工程, 2022, 43(09): 217-224.
- [8] 程海博, 熊显名. 基于 GIoU 的 YOLOv3 车辆识别方法 [J]. 桂林电子科技大学学报, 2020, 40(05): 429-433.
- [9] 吴闻起. 基于 YOLOv3 算法的遮挡场景下行人检测 [D]. 哈尔滨工程大学, 2020.
- [10] Jie H, Li S, Gang S, et al. Squeeze-and-Excitation Networks [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2017, 13(10): 213-215.
- [11] Woo S, Park J, Lee J Y, et al. Cbam: Convolutional block attention module [C]// Proceedings of the European Conference on Computer Vision (ECCV). 2018: 3-19.
- [12] Hanss, M. Applied Fuzzy Arithmetic [M]. Springer, Berlin, 2005.
- [13] Andrzej Piegat, Marek Landowski. Horizontal Membership Function and Examples of its Applications [J]. Fuzzy Syst, 2015, 17(1): 22-30.
- [14] S.J.W. Evans, P.C. Waller, S. Dacia. Use of proportional reporting ratios (PRRs) for signal generation from spontaneous adverse drug reaction reports [J]. Pharmacoeconomics Drug Safety, 2001, 10: 483-486.
- [15] Zadeh L A. Fuzzy sets [J]. Information and Control, 1965, 8(3): 338-353.
- [16] Wang GX. Theory and Applications of Fuzzy Numbers (模糊数理论及应用) [M]. Beijing: National Defense Industry Press, 2011.
- [17] Mehran Mazandarani. Granular Differentiability of Fuzzy-Number-Valued Functions [J]. IEEE Transactions on Fuzzy Systems, 2018, 26(1).
- [18] Zadeh L A. The concept of a linguistic variable and its application to approximate reasoning (模糊集合、语言变量及模糊逻辑) [M]. Beijing: China Science Publishing, 1984.
- [19] Li L J, Hu J H, Wang Z, et al. Establishment of a new standard for evaluation of grade of adverse drug reaction severity and application of adverse drug reaction severity index [J]. Pharmacoeconomic Care And Research (药学服务与研究), 2008, 8(1): 9-13.
- [20] Zhang Q, Yan M M, Zhao H, et al. Detection of drug withdrawal signals for drug-induced liver injury based on the database of Shanghai Clinical Drug Quality Control Center [J]. Chinese Journal of Pharmacoeconomics (药物流行病学杂志), 2022, 31(8).
- [21] ZHOU Y J, LIU L. Mining etanercept pharmacovigilance signal based on BCPNN method [J]. Chinese Journal of Hospital Pharmacy (中国医院药学杂志), 2022.
- [22] LIU J, YE G J, LIU W, et al. Detection of ADR reporting signal by fuzzy Bayesian confidence progressive neural network [J]. Chinese Journal of Pharmacovigilance (中国药物警戒), 2022.
- [23] Rajati, M. R., Mendel, J. M. Solving Zadeh's Swedes and Italians challenge problem [J]. Proceedings of Annual Meeting, North American Fuzzy Information Processing Society, pp. 2012, 1-6.
- [24] VietCuong Ngo and Wenchuan Wu, H. Linear Programming Contractor for Interval Distribution State Estimation Using RDM Arithmetic [J]. IEEE Transactions on Power Systems, 2021.
- [25] Chen, T, Y. Remarks on the subtraction and division operations over intuitionistic fuzzy sets and interval-valued fuzzy sets [J]. Int. J. Fuzzy Syst. 2017, 9(3), 169-172.
- [26] Gupta, M. Introduction to fuzzy arithmetic [M]. Van Nostrand Reinhold, New York 1991.

(上接第 8 页)