

基于轻量化YOLOv5s的集装箱锁孔识别算法

李源鑫, 郭忠峰, 杨钧麟

引用本文

李源鑫, 郭忠峰, 杨钧麟. [基于轻量化YOLOv5s的集装箱锁孔识别算法](#)[J]. 计算机科学, 2024, 51(6A): 230900021-6.

LI Yuanxin, GUO Zhongfeng, YANG Junlin. [Container Lock Hole Recognition Algorithm Based on Lightweight YOLOv5s](#) [J]. Computer Science, 2024, 51(6A): 230900021-6.

相似文章推荐 (请使用火狐或 IE 浏览器查看文章)

Similar articles recommended (Please use Firefox or IE to view the article)

[面向产线AI质检的少样本评测方法研究和验证](#)

Study and Verification on Few-shot Evaluation Methods for AI-based Quality Inspection in Production Lines

计算机科学, 2024, 51(6A): 230700086-8. <https://doi.org/10.11896/jsjcx.230700086>

[基于BERT和CNN的药物不良反应个案报道文献分类方法](#)

Literature Classification of Individual Reports of Adverse Drug Reactions Based on BERT and CNN

计算机科学, 2024, 51(6A): 230400049-6. <https://doi.org/10.11896/jsjcx.230400049>

[DUWe:动态未知词嵌入方法在Web异常检测中的应用](#)

DUWe:Dynamic Unknown Word Embedding Approach for Web Anomaly Detection

计算机科学, 2024, 51(6A): 230300191-5. <https://doi.org/10.11896/jsjcx.230300191>

[WiCare:一种非接触式的老人如厕跌倒监测模型](#)

WiCare:Non-contact Fall Monitoring Model for Elderly in Toilet

计算机科学, 2024, 51(6A): 230700044-8. <https://doi.org/10.11896/jsjcx.230700044>

[深度学习驱动下IaaS云运维异常检测算法的研究进展](#)

Research Progress of Anomaly Detection in IaaS Cloud Operation Driven by Deep Learning

计算机科学, 2024, 51(6A): 230400016-8. <https://doi.org/10.11896/jsjcx.230400016>

基于轻量化 YOLOv5s 的集装箱锁孔识别算法

李源鑫 郭忠峰 杨钧麟

沈阳工业大学机械工程学院 沈阳 110870

(lyx775764462@smail.sut.edu.cn)

摘要 为提高现有集装箱的锁孔识别检测效率,减少算法参数量以及减小模型大小,提出了一种基于轻量化 YOLOv5s 的集装箱锁孔识别算法。该算法将 YOLOv5s 的 Backbone 主干特征提取网络部分更换为轻量级神经网络模型 MobileNetV3,并对 neck 部分的特征融合结构进行进一步的优化,减少了模型的参数量和计算量,并提高了检测速度。引入注意力机制 SimAM 层,提高了检测的准确率和效率。使用不同的改进方法对模型进行重构后,在自建的集装箱锁孔数据集上进行训练和测试,并与改进的 YOLOv5s 进行对比实验。结果表明,改进后的模型大小仅为 2.4 MB,每幅图像的平均检测时间仅为 5.1ms,平均检测精度达 97.3%;与原始目标检测模型相比,该模型的大小减小了 82.8%,检测速度提高了 39%,在确保高检测精度的前提下展现出了较强的算法实时性。

关键词: 机器视觉;集装箱锁孔;YOLOv5s;轻量化;深度学习

中图分类号 TP181

Container Lock Hole Recognition Algorithm Based on Lightweight YOLOv5s

LI Yuanxin, GUO Zhongfeng and YANG Junlin

School of Mechanical Engineering, Shenyang University of Technology, Shenyang 110870, China

Abstract In order to improve the efficiency of container lock hole recognition and reduce the number of algorithm parameters and model size, a container lock hole recognition algorithm based on lightweight YOLOv5s is proposed. This algorithm replaces the Backbone feature extraction network of YOLOv5s with a lightweight neural network model MobileNetV3, and further optimizes the feature fusion structure of the neck part, which reduces the number of parameters and calculation amount of the model and improves the detection speed. The accuracy and efficiency of detection are improved by introducing the attention mechanism SimAM layer. After the model is reconstructed with different improvement methods, the training and testing are carried out on the self-built container lock hole data set, and the comparison test is carried out with the improved YOLOv5s. The results show that the size of the improved model is only 2.4 MB, the average detection time of each image is 5.1ms, and the average detection accuracy is 97.3%. Compared with the original target detection model, the size of the model is reduced by 82.8%, and the detection speed is increased by 39%, showing strong real-time algorithm on the premise of ensuring high detection accuracy.

Keywords Machine vision, Container lock hole, YOLOv5s, Lightweight, Deep learning

1 引言

随着全球贸易的不断发展,集装箱运输^[1-2]成为一种非常重要的运输方式,在全球范围内被广泛应用。然而随着集装箱数量的增加和运输规模的扩大,对集装箱装卸效率的要求也越来越高。传统的集装箱码放方式通常需要大量的人工干预,效率低且容易出现人为错误,因此,开发一种自动化集装箱码放系统是实现集装箱装卸自动化与智能化的重要环节。其中,集装箱的锁孔识别是集装箱自动码放系统的重要组成部分,如何更加精准、高效地进行锁孔识别具有重要的研究意义。

近年来,计算机视觉和图像处理技术的快速发展为集装箱锁孔识别提供了新的解决方案。Mi 等^[3]采用 HSV 颜色空间和图像二值化定位集装箱箱体位置和角件位置,再通过

Canny 算子和霍夫变换进行边缘检测精确定位锁孔位置。Hao 等^[4]采用 CenterNet 进行关键目标检测,并采用卡尔曼滤波降低视频流中关键点坐标的噪声,实现锁孔关键点的跟踪。Zhang 等^[5]采用支持向量机(SVM)定位集装箱的锁孔区域图像,再使用 Canny 算子提取锁孔的边缘轮廓来定位锁孔。Kunimits 等^[6]将分布式传感网络与扩展模板匹配相结合来计算吊具和集装箱的相对定位。Mi 等^[7]利用 HOG + SVM 的方法对集装箱一边锁孔进行识别,然后通过对称性,采用翻转镜像算法快速识别另一边锁孔。

上述方法虽然都能保证良好的检测精度,但检测速度普遍较慢,很难满足日益扩大的港口集装箱码垛需求。为提升集装箱锁孔检测效率,本文提出了一种基于轻量化 YOLOv5s 的集装箱锁孔识别算法。首先,收集大量的集装箱及集装箱锁孔的图像并对其进行标注,制作相应的数据集。然后,使用

基金项目:辽宁省教育厅 2021 年度科学研究经费项目(面上项目)(LJKZ0114)

This work was supported by the 2021 Scientific Research Funding Project of Liaoning Provincial Department of Education(LJKZ0114).

通信作者:郭忠峰(guozf@sut.edu.cn)

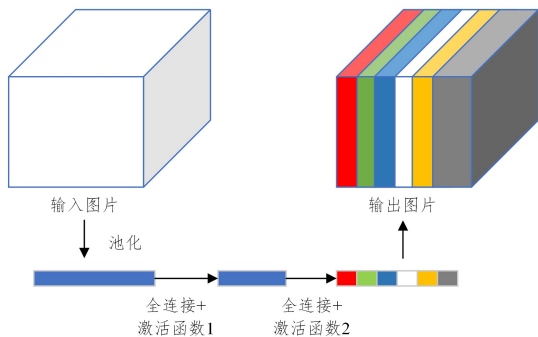


图 4 SE 模块
Fig. 4 SE module

MobileNetV3 中引入了 h-swish 激活函数,它是在 swish 激活函数上进行改进的一种激活函数。swish 函数因具有平滑且非线性的激活特征,在深层网络中表现出了良好的性能。然而,swish 函数的计算复杂度较高,需要消耗大量的计算资源和求导时间。为了应对这一问题,h-swish 函数在 swish 函数的基础上引入了线性修正单元 ReLU6,它将输入最大值限制为 6,以避免过大的输入对梯度造成不利影响。这样可以提高函数的计算效率,并保持一定的非线性特性。h-swish 函数的引入兼顾了计算效率和模型性能,使得模型在保持较高准确性的同时能够更加高效地运行。基本想法是用一个近似函数来逼近 swish 函数,让 swish 函数变得不那么光滑。h-swish 激活函数的具体计算式如下:

$$h(x) = x \frac{R(x+3)}{6} \quad (1)$$

$$R(x) = \min(\max(x, 0), 6) \quad (2)$$

3.2 更换 neck 部分卷积模块

GhostConv 通过舍弃部分常规卷积操作,采用线性变换进行特征提取,以更好地利用特征之间的相关性和冗余性。这种方法在几乎不影响检测精度的前提下,减少了模型的参数量和计算量,并提高了检测精度。具体结构如图 5 所示。

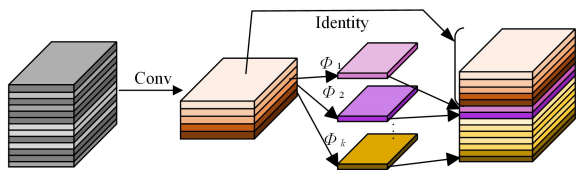


图 5 GhostConv 模块
Fig. 5 GhostConv module

GhostConv 的卷积过程^[15]如下:首先使用步长为 1 的 1×1 卷积核将输入特征图的通道数减半,然后对上一步得到的特征图进行 5×5 卷积核的深度卷积操作,最后将两个卷积结果进行拼接。GhostBottleneck 结构通过第一个 GhostConv 将输入特征图的通道数减半,再经过第二个 GhostConv 将特征图的通道数恢复到原来的大小,最后将经过 3×3 深度卷积的残差边与 GhostBottleneck 的输出相加,以融合特征。将原来的 C3 模块中的 Bottleneck 替换为 GhostBottleneck,得到了 C3Ghost 模块,如图 6 所示。这种改进减少了原结构中大多数 3×3 传统卷积的使用,从而减少了模型的参数量和计算量,并提高了运行速度。

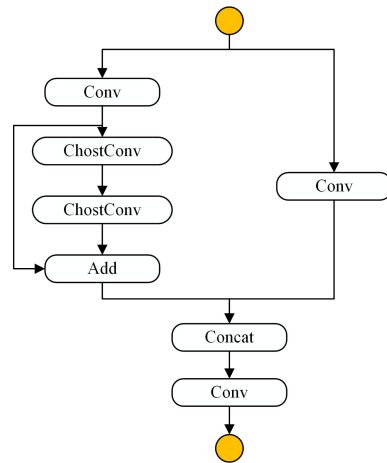


图 6 C3Ghost 模块
Fig. 6 C3Ghost module

3.3 添加注意力机制

在人类的感知过程中,注意力发挥着重要的作用,它使我们能够有选择地关注重要的部分,以获取感兴趣的信息。将这种注意力的机制引入机器学习领域,形成了相应的注意力机制^[16-17]。在机器学习中,注意力机制使得模型在学习过程中能够集中精力在更加重要的信息上,这种机制可以帮助模型更好地理解输入数据,提取关键特征,并减少对无关信息的处理。通过集中注意力于重要的部分,模型可以在减少计算量的同时节约其他成本,从而提高效率和性能。

SimAM 注意力机制^[18]是一种无需向原始网络添加额外参数,即可直接为特征图推导出三维权重,使得网络学习更多有辨别性的神经元以提高网络的特征提取能力的注意力机制。相比现有的 1-D 通道注意力和 2-D 空间注意力方法,SimAM 注意力机制无需额外参数来计算 3-D 注意力权重,只需要通过一个能量函数来计算注意力权重。

SimAM 注意力模块原理如图 7 所示。通过 3-D 权重,为特征中的每个点分配一个单独的标量值。这种统一的权重赋予了特征网络更具全局性的能力。计算过程如式(3)所示:

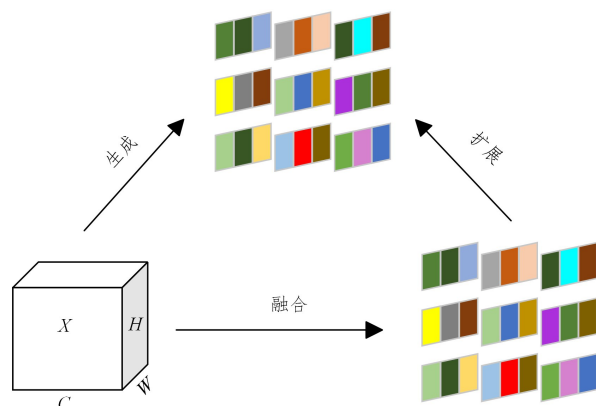


图 7 SimAM 注意力模块
Fig. 7 SimAM attention module

$$\hat{X} = \text{sigmoid}\left(\frac{1}{E}\right) \odot X \quad (3)$$

输出结果为增强后的特征, X 表示输入的特征, \odot 为点积运算,并通过 sigmoid 函数来限制每个通道上的能量函数 E 中可能出现的过大值。E 的计算过程如式(4)所示:

$$E = \frac{4(\sigma^2 + \lambda)}{(t - \mu)^2 + 2\sigma + 2\lambda} \quad (4)$$

其中, t 表示输入特征的值, $t \in X$, λ 为常数 1×10^{-4} , μ 和 σ^2 分别表示 X 中每个通道上的均值与方差, 其计算过程分别如式(5)、式(6)所示:

$$\mu = \frac{1}{M} \sum_{i=1}^M x_i \quad (5)$$

$$\sigma^2 = \frac{1}{M} \sum_{i=1}^M (x_i - \mu)^2 \quad (6)$$

其中, $M = H \times W$ 表示每个通道上值的数量。通过这样的计算, 可以获得每个点的权重, 并利用这些权重来改善网络的识别效果, 同时不会为网络增加额外的参数。改进后的 YOLOv5s-MCS 网络结构如图 8 所示。

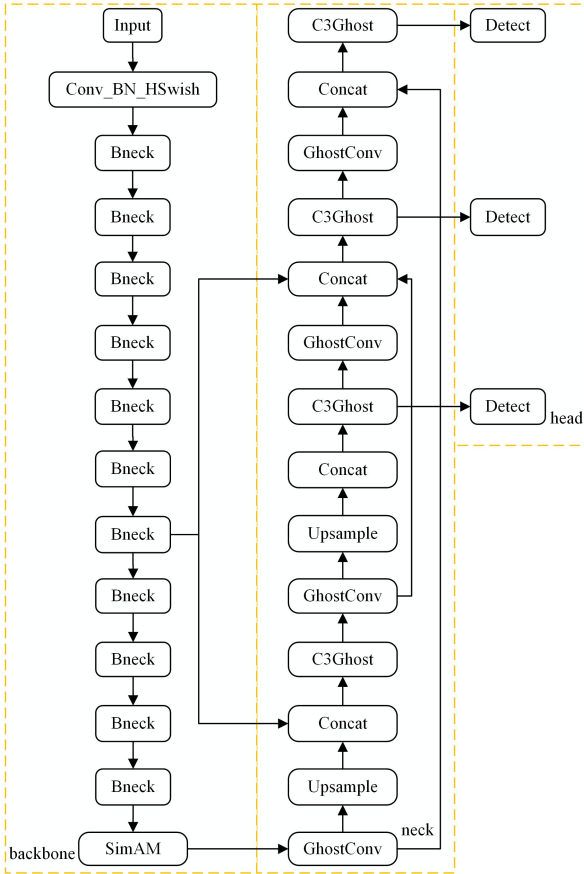


图 8 YOLOv5s-MCS 网络结构

Fig. 8 Network structure of YOLOv5s-MCS

4 实验

4.1 实验配置

本实验采用的操作系统为 Windows10, 采用的深度学习框架为 Pytorch2.0.1, 编程语言采用 Python3.8, CUDA 版本为 12.1, 同时将 cudnn 添加到环境中加速计算机的计算能力, CPU 为 8 核 AMD Ryzen 9 5900HX 3.30 GHz, GPU 为 NVIDIA GeForce RTX 3060 6GB 的计算机来训练模型。初始学习率为 0.01, batch_size 为 16, 总 epoch 为 200。

4.2 建立数据集

为提高网络模型的鲁棒性和泛化能力, 本文采集了多种场景下不同时间的集装箱及集装箱锁孔图片及视频, 样本数据丰富, 为建立多样化数据集奠定了良好基础。部分数据集图片如图 9 所示。



图 9 数据集示例

Fig. 9 Sample dataset

数据集使用 LabelImg 标注工具进行标注。将每张图片上的集装箱及集装箱锁孔用最小的矩形框标注, 并使用其英文名称作为目标类别。数据集共 6323 张图片, 以 8:1:1 的比例划分为训练集、验证集和测试集。为保证实验的公平性, 在之后进行的各种实验中统一使用此数据集。数据集标签数量及分布如图 10 所示。

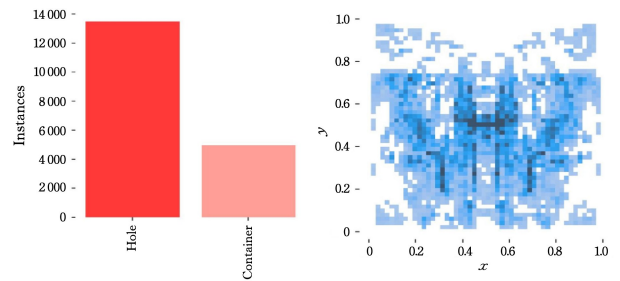


图 10 数据集标签数量及分布示例

Fig. 10 Example of the number and distribution of labels in the dataset

4.3 实验结果与分析

4.3.1 模型评估

使用精度 P (Precision)、召回率 R (Recall)、平均精度均值 mAP (mean Average Precision)、参数量 (parameters)、计算量 (GFLOPs)、模型大小、平均检测时间等作为评价指标对模型进行评估。

相关计算式如下:

$$P = \frac{TP}{TP + FP} \quad (7)$$

$$R = \frac{TP}{TP + FN} \quad (8)$$

$$AP = \int_0^1 P dR \quad (9)$$

$$mAP = \frac{\sum_{i=1}^n AP_i}{n} \quad (10)$$

精度 P 表示在预测为正样本的数据中, 实际为正样本的概率; 召回率 R 表示在实际为正样本的数据中, 被预测为正样本的概率; TP 表示预测为正样本且实际为正样本的数量; FP 表示预测为正样本但实际为负样本的数量; FN 表示预测为负样本但实际为正样本的数量; 平均精度均值 mAP 表示所有类别平均精度 (AP) 的均值。模型大小表示训练结束后所得到的权重文件的大小, 平均检测时间表示模型检测一张图片所需的平均时间。

4.3.2 结果分析

通过消融实验验证本文的改进对集装箱锁孔识别性能的影响。针对 YOLOv5s 算法进行了多种轻量化改进策略, 包

括将主干网络替换为 GhostNet 的 YOLOv5s-G、ShuffleNetV2 的 YOLOv5s-S,以及 MobileNetV3 的 YOLOv5s-M。在这些改进的基础上,对 neck 部分的特征融合结构进行了进一步的优化,采用 GhostConv 模块和 C3Ghost 模块替代普通卷积 Conv 模块和 C3 模块,得到了 YOLOv5s-MC 模型。

表 1 消融实验结果

Table 1 Results of ablation experiment

模型	P/%	R/%	mAP@0.5/%	Params	GFLOPs	模型大小/MB	检测时间/ms
YOLOv5s	96.6	97.5	97.7	7025023	16.0	14.0	7.1
YOLOv5s-G	95.7	97.3	97.9	3687239	8.2	7.6	6.9
YOLOv5s-S	94.4	96.6	97.3	3185175	5.9	6.5	5.3
YOLOv5s-M	93.5	96.8	96.8	1390141	2.5	3.0	5.5
YOLOv5s-MC	92.9	96.8	96.6	1051379	1.8	2.4	5.2
YOLOv5s-MCC	93.6	96.8	96.8	1147302	1.9	2.6	5.4
YOLOv5s-MCS	93.4	97.0	97.3	1051379	1.8	2.4	5.1

相比于原始 YOLOv5s 算法,将主干网络替换为 GhostNet,ShuffleNetV2 和 MobileNetV3 都能够有效减少模型参数量和计算量,减小模型大小,其中 GhostNet 和 ShuffleNetV2 的模型分别为 7.6 MB 和 6.5 MB,减小了 45% 和 53%,MobileNetV3 的模型最小,为 3.0 MB,减小了 78%。在此基础上用 GhostConv 模块和 C3Ghost 模块替换 neck 部分的 Conv 模块和 C3 模块,进一步减小模型大小到 2.4 MB。引入 CBAM 注意力机制和 SimAM 注意力机制并进行对比,结果显示,引入 SimAM 注意力机制训练出的模型更小且平均精度更高,达到了 97.3%。

改进的 YOLOv5s-MCS 与原始 YOLOv5s 的 mAP 曲线对比如图 11 所示,YOLOv5s-MCS 的损失函数曲线如图 12 所示。

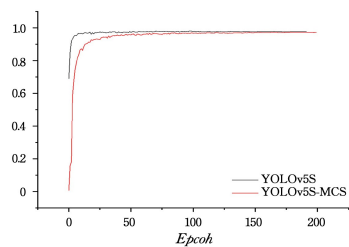


图 11 mAP 曲线图

Fig. 11 mAP curve

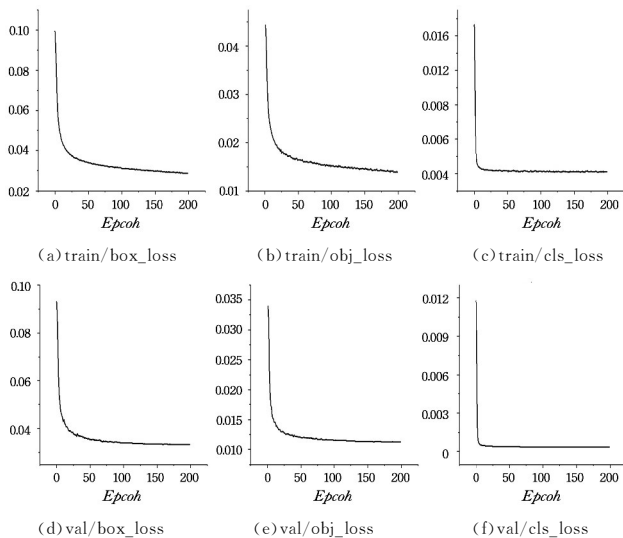


图 12 YOLOv5s-MCS Loss 曲线图

Fig. 12 YOLOv5s-MCS Loss curve

另外,还引入了注意力机制 CBAM 层和 SimAM 层,分别得到 YOLOv5s-MCC 和 YOLOv5s-MCS 模型。通过使用上述 7 种模型作为对照模型,以验证 YOLOv5s-MCS 模型设计的合理性,同时为保证模型测试的有效性,采用了统一的模型训练标准,最终训练所得的模型参数和测试结果如表 1 所列。

相较于原始的 YOLOv5s 算法,改进后的 YOLOv5s-MCS 算法在多个方面取得了显著的提升,其参数量减少了 85%,计算量减少了 88.7%,模型大小减小了 82.8%。这种优化使得改进的算法具有更轻量级的模型结构,降低了计算负担。同时,改进的算法在检测速度方面提高了 39%,精度仅损失 0.4%,能够在保证检测精度的同时更轻量级且有更高的算法实时性,在检测集装箱及其锁孔任务中展现出良好的性能。

改进的 YOLOv5s-MCS 算法对集装箱及其锁孔的检测结果如图 13 所示,计算结果表明,改进的算法能够在保持较高的检测精度的同时更轻量级且具有更高的算法实时性,为提高集装箱及其锁孔检测任务的效率提供了可行且有效的解决方案。

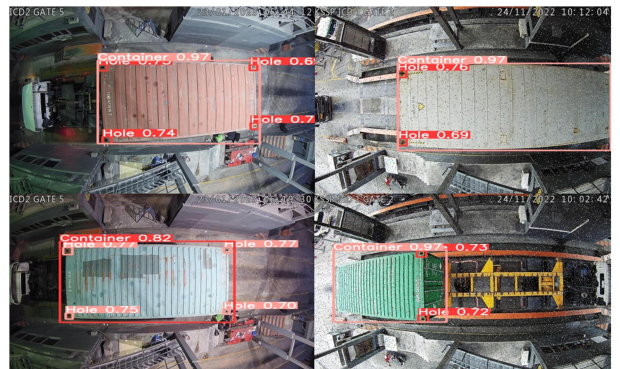


图 13 YOLOv5s-MCS 检测结果示意图

Fig. 13 Schematic diagram of YOLOv5s-MCS detection results

4.3.3 不同网络模型对比实验

为了进一步验证改进的轻量级模型 YOLOv5s-MCS 在集装箱锁孔识别检测任务中的性能,将其与 YOLO 系列中的其他轻量级目标检测算法进行对比实验,包括 YOLOv3-tiny 和 YOLOv4-tiny,对比实验的结果如表 2 所列。由表 2 可知,本研究的改进算法 YOLOv5s-MCS 相比同一系列的目标检测算法,在 mAP、模型参数量、计算量、模型大小和检测时间上均取得了更好的结果。

以上结果进一步验证了算法的有效性和优越性,表明 YOLOv5s-MCS 相对于同系列下的目标检测算法具有更高的检测精度、更少的模型参数量、更低的计算复杂度、更小的模型大小以及更快的检测时间。这些优势为集装箱及其锁孔检测任务提供了一种更具竞争力和可行性的解决方案。

表 2 不同模型对比

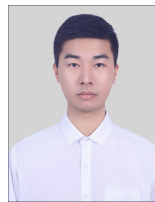
Table 2 Comparison of different models

模型	P/%	R/%	mAP@0.5/%	Params	GFLOPs	模型大小/MB	检测时间/ms
YOLOv3-tiny	93.4	95.1	96.2	8672186	13.0	17.0	6.8
YOLOv4-tiny	93.7	95.3	96.5	6267590	17.2	12.2	6.5
YOLOv5s-MCS	93.4	97.0	97.3	1051379	1.8	2.4	5.1

结束语 本文采用轻量级神经网络模型 MobileNetV3 作为该算法的主干网络,并采用轻量化剪枝的 GhostConv 卷积和 C3Ghost 卷积作为该算法的卷积层,使得模型权重文件大小由 14 MB 减小到 2.4 MB,参数量由 702 万减少到 105 万,大幅提升了算法的识别检测效率;在主干特征提取网络中引入自注意力机制 SimAM 模块,改善网络的识别效果,识别部分的精度为 93.4%,召回率为 97.0%,平均精度为 97.3%,满足集装箱识别检测的精准度需求。对比其他先进算法,本研究提出的 YOLOv5s-MCS 算法在具有一定速度和体量优势的同时保持了较高的检测精度,在集装箱锁孔检测中具有一定的可行性与优越性,为集装箱锁孔识别检测提供了理论依据。但考虑到港口场景的复杂多样性,集装箱运输有时会在恶劣天气的影响下作业,以及模型精度可能受监控设备的分辨率和承载吊具稳定性等外部因素影响。后续将在此基础上对输入视频和模型精度的影响关系进行探究,对模型的训练增加更多的干扰项,使得模型鲁棒性更强,泛用性更高。

参考文献

- [1] CHANG Y M, ZHU X N, WANG L. Review on integrated scheduling of container terminal[J]. Journal of Transportation Engineering, 2019, 12(1): 136-146.
- [2] DONG M W, HUANG L F, GU Y. Review on green job scheduling optimization of container terminals [J]. Journal of Chongqing Jiaotong University(Natural Science Edition), 2022, 41(11): 7-14.
- [3] MI W J, ZHANG Z W, MI C. Study on container hanging holes recognition algorithm based on machine vision[J]. Chinese Journal of Construction Machinery, 2016, 14(5): 399-402.
- [4] HAO Y S, LU B, LIU F, et al. Key point smoothing tracking of container keyhole based on centernet[J]. Journal of Control Engineering, 2021, 28(11): 2108-2113.
- [5] ZHANG J, DIAO Y F, CHENG W M, et al. Container keyhole tracking and center positioning based on video stream[J]. Journal of Computer Applications, 2019, 39(S2): 216-220.
- [6] KUNIMITSU S, ASAMA H, KAWABATA K, et al. Development of crane vision for positioning container[J]. Journal of Robotics and Mechatronics, 2004(2): 16.
- [7] MI C, WANG J, MI W, et al. Research on regional clustering and two-stage SVM method for container truck recognition[J]. Discrete & Continuous Dynamical Systems, 2018, 12(4): 1117-1133.
- [8] JOCHER G, STOKEN A, BOROVEC J, et al. Ultralytics/yolov5:v3.1-Bug fixes and performance improvements(Version v3.1) [EB/OL]. (2020-10-29) [2021-01-19]. <https://gitlab.com/ultralytics/yolov5/-/releases/v3.1>.
- [9] REDMON J, FARHADI A. YOLOv3: An incremental improvement[J]. arXiv: 1804.02767, 2018.
- [10] BOCHKOVSKIY A, WANG C Y, LIAO H Y M. YOLOv4: Optimal speed and accuracy of object detection[J]. arXiv: 2004.10934, 2020.
- [11] HOWARD A, SANDLER M, CHEN B, et al. Searching for mobilenetv3[C] // 2019 IEEE/CVF International Conference on Computer Vision (ICCV). Seoul, Korea (South). New York: IEEE Press, 2019: 1314-1324.
- [12] SIFRE L, MALLAT, STEPHANE. Rigid-Motion scattering for texture classification[J]. arXiv: 1403.1687, 2014.
- [13] HU J, SHEN L, ALBANIE S, et al. Squeeze-and-excitation networks[J]. IEEE Transactions on Pattern Analysis And Machine Intelligence, 2020, 42(8): 2011-2023.
- [14] GLOTOT X, BORDES A, BENGIO Y. Deep sparse rectifier neural networks [J]. Journal of Machine Learning Research, 2011, 15: 315-323.
- [15] ZI B, YIN Z Q, LI Y C, et al. Fast mobile component location method for cable-driven parallel robots based on yolo model[J]. Chinese Journal of Mechanical Engineering, 2019, 55(3): 64-72.
- [16] YANG Y, DENG H. GC-YOLOv3: you only look once with global context block[J]. Electronics, 2020, 9(8): 1235.
- [17] WOO S, PARK J, LEE J Y, et al. CBAM: convolutional block attention module[C] // 2018 15th European Conference on Computer Vision. Munich, Germany, 2018: 3-19.
- [18] YANG L, ZHANG R Y, LI L, et al. SimAM: A simple parameter-free attention module for convolutional neural networks [C] // International Conference on Machine Learning. PMLR, 2021.



LI Yuanxin, born in 1995, postgraduate. His main research interests include machine vision and deep learning.



GUO Zhongfeng, born in 1978, Ph.D, associate professor. His main research interests include robot mechanism design, artificial intelligence machine vision inspection, etc.