

基于改进 YOLOv5 的铝型材瑕疵检测算法

刘 柱,董 琴*,杨国宇,陈朝峰
(盐城工学院 信息工程学院,江苏 盐城 224001)

摘 要:基于铝型材表面瑕疵类别多样,对实时检测快速精准的需求,提出一种基于改进 YOLOv5 的瑕疵检测算法。通过在原始骨干网络的基础上增加新检测层并使用 K-means++ 算法改进锚框的生成方式,提升检测尺度,避免忽视低层语义信息。对铝型材瑕疵数据集离线增强,丰富样本容量;在 Backbone 网络结构中融入新的卷积结构和 E-CBAM 注意力机制,提高网络的特征提取能力的同时降低冗余计算,提升模型检测性能;采用 EIoU Loss 作为整个网络结构的损失函数来加快收敛效率,解决难易样本不平衡的问题。实验结果表明,在铝型材瑕疵数据集上将改进后 YOLOv5 检测模型与原始 YOLOv5 模型进行比较,平均精度 mAP 提升 2.9 百分点,召回率 Recall 提升 3.9 百分点,速度 FPS 达至 45.8,将近年来的代表性算法 YOLOv3、YOLOv4、SSD、Faster-rcnn 与改进后的检测算法在铝型材瑕疵数据集上进行性能比较,通过综合对比检测精度、检测速度等重要参数证明改进后的 YOLOv5 检测算法更好地兼顾了检测效率和检测精度。所提方法满足了铝型材工厂生产现场瑕疵检测要求。

关键词:YOLOv5;铝型材;注意力机制;瑕疵检测;损失函数;锚框

中图分类号:TP391

文献标识码:A

文章编号:1673-629X(2023)10-0183-06

doi:10.3969/j.issn.1673-629X.2023.10.028

Aluminum Profile Defect Detection Algorithm Based on Improved YOLOv5

LIU Zhu,DONG Qin*,YANG Guo-yu,CHEN Chao-feng

(School of Mechanical Engineering,Yancheng Institute of Technology,Yancheng 224001,China)

Abstract:In order to meet the requirement of rapid and accurate real-time detection due to various types of defects on aluminum profiles,a defect detection algorithm based on improved YOLOv5 was proposed. By adding a new detection layer on the basis of the original backbone network and using K-means++ algorithm to improve the generation mode of anchor box,the detection scale is increased to avoid ignoring the low-level semantic information. Off-line enhancement of aluminum profile defect data set is conducted to enrich the sample size. New convolution structure and E-CBAM attention mechanism are integrated into Backbone network structure to improve feature extraction capability,reduce redundant calculation,and improve model detection performance. EIoU Loss is adopted as the loss function of the whole network structure to accelerate the convergence efficiency and solve the problem of unbalance of difficult and easy samples. The experimental results show that by comparing the improved YOLOv5 detection model with the original YOLOv5 model in the aluminum profile defect data set,the average precision mAP increases by 2.9 percentage points,the recall rate increases by 3.9 percentage points,and the speed FPS reaches 45.8. Representative algorithms YOLOv3,YOLOv4,SSD,Faster-rcnn proposed in recent years were compared with the improved detection algorithm on the aluminum profile defect data set. Through comprehensive comparison of detection accuracy,detection speed and other important parameters,it is proved that the improved YOLOv5 detection algorithm better takes into account detection efficiency and accuracy. The proposed method meets the requirement of defect detection in aluminum profile factory.

Key words:YOLOv5;aluminum profiles;attention mechanism;defect detection;loss function;anchor

0 引言

伴随“中国制造 2025”的提出,铝型材凭借其独特的质轻,柔韧性好等特点,在机械工业领域有着不可动

摇的地位,其全行业的产量和消费量在世界范围内逐年递增^[1]。在加工速度上涨的同时,由于环境等不可抗力因素,使得产品表面不可避免地产生瑕疵问题,这

收稿日期:2023-01-06

修回日期:2023-05-10

基金项目:江苏省产学研合作项目(BY2022460);盐城工学院高层次人才科研启动项目(XJR2022001)

作者简介:刘 柱(1998-),男,硕士研究生,研究方向为目标检测、深度学习;通讯作者:董 琴(1974-),女,副教授,研究方向为深度学习、人工智能、大数据。

不仅仅是铝型材外观问题,它严重影响到了铝型材在应用时的可靠性和安全性。

通过调研发现,传统的检测方法成本高、操作复杂度高且对环境要求苛刻,无法实现可视化过程检测。伴随工业化的发展,机器视觉检测凭借精确、无接触测量等优点被广泛应用于瑕疵表面检测^[2]。早期利用机器视觉进行表面瑕疵检测,主要是通过对缺陷照片进行图像处理提取瑕疵特征,这类方法能对特定瑕疵较准确地进行检测^[3]。但对于铝型材瑕疵种类较多、尺度不一的情况,存在泛性较差、特征提取复杂、检测精确度不高等问题,因此在实际生产中局限性太大,难以满足生产需求。

近些年,基于深度学习的瑕疵检测算法因具有稳定高效、鲁棒性强的特点,在很多场景中得到了广泛的应用,取得了瞩目的成绩,并逐渐取代传统表面瑕疵检测方法。YOLO(You Only Look Once)算法的检测性能更为突出,在YOLO的基础上涌现出了YOLOv3^[4]、YOLOv4^[5]等众多改进算法。例如,刘浩等人在机床刀具的检测过程中采用了改进的YOLOv3算法,取代了传统的人工目测^[6]。孙永鹏等人采用改进YOLOv4的检测方法提升冲压件的检测效率和检测精度并取得了良好的效果^[7]。乔欢欢等人使用YOLOv5算法进行交通标志检测,取得了良好的检测效果^[8]。

上述研究结果表明,YOLO算法在瑕疵检测方面具有良好的鲁棒性。因此,针对准确快速的检测需求,提出一种基于改进YOLOv5(You Only Look Once Version-5)的铝型材瑕疵检测算法。首先,在head部分通过K-means++算法设计新的检测端,提升小尺度瑕疵的检测效果。其次,改进原Backbone网络结构,加入新的卷积层,并插入改进注意力机制E-CBAM,增强整个模型网络特征提取能力,减少原始信息丢失。最后,使用EIoU Loss(Efficient IoU Loss)作为整个网络结构的损失函数,加快收敛效率,从而解决难易样本不平衡的问题。通过评估不同改进策略对检测模型的影响,及真实产线下数据集实验数据对比不同检测算法,确保改进YOLOv5算法在满足工厂实际需求的同时,兼顾检测精度和检测效率。

1 改进YOLOv5算法模型

1.1 改进YOLOv5网络模型

YOLOv5是一阶段的目标检测算法,主要模型结构可分为Input(输入端)、Backbone(骨干网络)、Neck(颈部网络)及Head(输出网络)。

Input主要用于对输入数据进行规范化调整、数据增强、自适应锚框计算。当数据进入网络进行训练时,输入端会起到保持图片的大小一致作用,当输入数据

量不够时,会自动对图像进行裁剪、反转扩充数据量,在训练阶段在初始锚点框的基础上输出对应的预测框。Backbone主要用于输入数据的特征提取,YOLOv5使用CSPDarknet作为骨干网络,在不同图像细粒度上聚合并形成图像特征的卷积神经网络。Neck网络借鉴CSPnet^[9]研发出CSP2结构实现不同深度的特征图信息融合,增强神经网络对特征信息的表达能力,有效提高网络的特征融合性能。Head以CIoU Loss作为bounding box损失函数,使用三个不同尺度的Detect检测端在特征图上进行瑕疵位置确定及种类分别的预测。

为使得网络模型达到检测更精确、检测尺度更广、速度更快的瑕疵检测效果,针对铝型材瑕疵检测模型结构的改进将通过改进原有Backbone网络结构、加入优化注意力机制(E-CBAM)增强特征信息提取、添加新的检测端扩大识别尺度且锚框使用K-means++算法聚类生成、采用EIoU边界损失函数加快网络收敛的方法实现^[10],改进后的网络结构模型如图1所示。

1.2 改进检测尺度及锚框优化

1.2.1 检测尺度优化

原有的YOLOv5模型结构采用三个检测端,以本实验为例,若输入的图片数据为 640×640 ,通过卷积池化等操作步骤最终得到的检测尺度为 80×80 、 40×40 、 20×20 ,这对于一般的检测任务是较为合理的,但对于铝型材表面瑕疵尺度多样化,且存在尺度差较大的瑕疵类别,为提升检测精度,在Neck网络结构中再一次进行上采样及张量拼接,将浅层特征与深层特征融合增大检测特征尺度 10×10 ,同时在Backbone网络结构中增添一组CSP及CBS模块。

相比原有检测结构,新的检测端进一步加强了对特征信息的提取,并将浅层特征进一步融合得到更加丰富、高效的语义信息,拓宽感受野得到更多有用的特征信息,从而有效提升模型检测精度。

1.2.2 检测锚框优化

每个检测端的检测锚框采用K-means++算法聚类生成^[11],原始算法的自适应锚框只适用于常规尺寸的数据集,而本次实验所用的铝型材瑕疵数据集瑕疵复杂程度较大,尺度大小差距较大,采用K-means++算法减小真实框与预设框的差距,避免发生错检、漏检现象,提升检测精度。

相比于传统的K-means^[12]算法,K-means++算法使用IOU(Intersection Over Union)也就是预测框与真实框的交并比来计算样本与聚类中心的距离,从而控制不同大小边框的误差,公式如下:

$$\text{DiStance}(\text{box}, \text{cluster}) = 1 - \text{IOU}(\text{box}, \text{cluster}) \quad (1)$$

式中, box 表示真实框集合, cluster 表示聚类中心集合, $\text{IOU}(\text{box}, \text{cluster})$ 表示真实框与聚类中心交集与并集的比例。

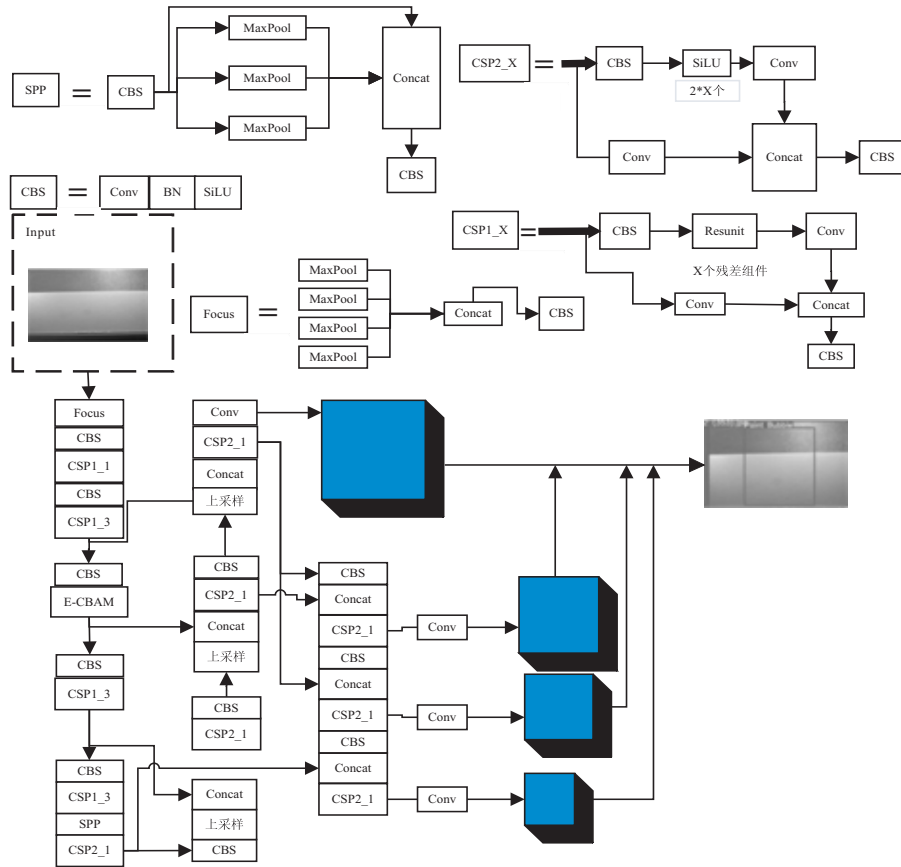


图 1 改进后的 YOLOv5 网络框图

IOU 值越高表示真实框与聚类中心越接近, 相关度也就越高。初始化簇中心的方法为逐个选取 k 个簇中心, 且离其它簇中心越远的样本点越有可能被选为下一个簇中心。计算公式如下:

$$P(X) = \frac{D(X)^2}{\sum_{X \in \alpha} D(X)^2} \quad (2)$$

其中, α 表示数据集, $D(X)$ 表示每个样本与当前已有聚类中心之间的最短距离, $P(X)$ 表示下一个聚类中心的概率大小。

最终利用聚类分析得到最优的 12 个不同比例大小的锚框参数避免收敛速度慢的现象, 其聚类结果和分配策略如表 1 所示。

表 1 聚类结果及分配策略

Detector size	Anchor size			Number of bounding boxes
80×80	(47, 68)	(50, 101)	(62, 235)	3
40×40	(68, 279)	(81, 324)	(179, 249)	3
20×20	(343, 212)	(284, 342)	(426, 345)	3
10×10	(605, 362)	(633, 165)	(623, 295)	3

1.3 改进注意力机制 E-CBAM

注意力机制 (Attention Mechanism) 能够根据需求

选择最合适的输入, 能够对齐两个序列之间的 token 关系, 从而实现更好的效果。为了提升 YOLOv5 的网络表达能力, 在 Backbone 网络中对 CBAM^[13] 注意力机制进行改进, 并命名为 E-CBAM (Efficient Convolutional Block Attention Module)。

如图 2 所示, 为选出针对当前数据集最合适的注意力机制, 在 Backbone 网络结构中分别加 CA^[14]、ECA^[15]、SE^[16] 及 E-CBAM 注意力机制后, 通过 pytorch 与 GradCAM^[17] 实现热力图可视化。通过比较特征图层的热力图可以明显看出, E-CBAM 注意力机制针对当前数据集类别检测在精度和准确度要优于其他注意力机制。

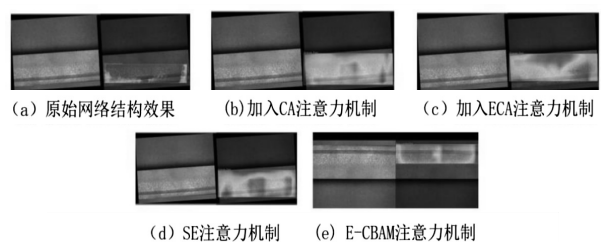


图 2 不同注意力机制热力图效果

通道注意力模块 CAM (Channel Attention Module)^[18] 和经过优化后的空间注意力模块 E-SAM

(Efficient Spatial Attention Module) 两个子模块共同组成 E-CBAM (Efficient Convolutional Block Attention Module), 结构如图 3 所示。

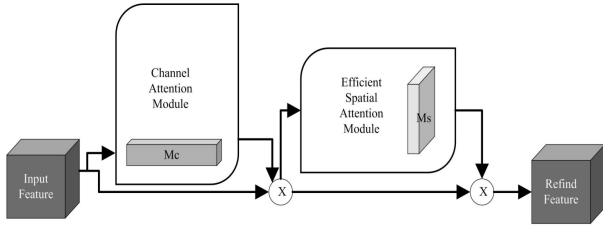


图 3 E-CBAM 结构

从图中可知,输入信息先通过一个通道注意力模块得到加权结果之后,再经过一个空间注意力模块,最终进行加权得到结果。

改进后的空间注意力模块结构如图 4 所示。空间注意力模块的主要工作包括:将 Channel attention 模块输出的特征图信息作为输入数据;将特征信息进行基于 channel 的 global max pooling 和 global average pooling 的处理,然后将这 2 个结果基于 channel 做 concat 操作;经过 3 个卷积核大小为 3×3 的卷积,降维为 1 个 channel;再经过 sigmoid 生成 spatial attention feature;最后将该 feature 和该模块的输入 feature 做乘法,得到最终生成的特征。

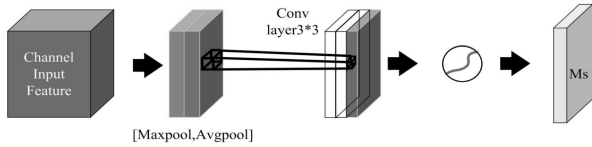


图 4 改进空间注意力模块结构

整体表达公式如下:

$$F_s = M_s(F_c) \otimes F_c \quad (3)$$

$$M_s(F) = \sigma((f^{3 \times 3}([AugPool(F); \\ \text{MaxPool}(F)]))^3) = \\ \sigma((f^{3 \times 3}(F_{avg}^s; F_{max}^s))^3) \quad (4)$$

式中, M_s 分别表示通道及空间注意力模块, F_s 表示通道注意力特征图及空间注意力特征图。通过改进空间注意力的卷积结构,使用连续 3×3 大小的卷积核在原算法的基础上降低了模型的计算量。

1.4 损失函数 EIoU (Efficient IoU Loss)

铝型材瑕疵种类较多,像素数量不均衡,因而容易忽略个别瑕疵的重要性,原始的损失函数纵横比差异不是宽高分别与其置信度的真实差异,因此在很大程度上会降低模型优化相似性。EIoU Loss^[19] 在 CIoU^[10] 的基础上将纵横比拆开计算,使目标与锚盒的宽度和高度之差最小,从而加快收敛效率。针对 BBox 回归过程中样本数据问题,利用梯度下降方法, EIoU Loss 基于 Focal Loss 把高质量的锚框和低质量的锚框分开,通过分类加权的方法提高回归精度,因此

采用 EIoU Loss 作为整个网络结构的损失函数。该函数的惩罚公式如下:

$$L_{IoU} = 1 - IoU \quad (5)$$

$$L_{dis} = \frac{\rho^2(b, b^{gt})}{c^2} \quad (6)$$

$$L_{asp} = \frac{\rho^2(w, w^{dt})}{C_w^2} + \frac{\rho^2(h, h^{dt})}{C_h^2} \quad (7)$$

$$L_{EIoU} = L_{IoU} + L_{dis} + L_{asp} \quad (8)$$

$$L_{F-EIoU} = IoU^\gamma L_{EIoU} \quad (9)$$

式中, C_w 和 C_h 分别表示最小的宽度和高度, L_{IoU} 表示 IoU 损失函数, L_{dis} 表示距离损失函数, L_{asp} 表示宽高损失函数, γ 为控制异常值抑制程度的参数。 b 和 b^{gt} 表示预测框和真实框的中心点, ρ 表示计算两个中心点间的欧氏距离, c 表示能够同时包含预测框和真实框的最小闭包区域的对角线距离。

2 实验结果与分析

2.1 实验环境及参数设定

在实验过程中选用 Windows 10 x64 的操作系统, 处理器为 Intel(R) Xeon(R) Gold 6258R CPU @ 2.70 GHz, 显卡型号是 NVIDIA Tesla V100S-PCIE-32 GB, 采用 Py-Torch1.12 框架及 CUDA11.3 图形加速, 运行内存为 32 G, 显存是 256 G, 采用 SGD 优化器, 编译语言为 Python3.7。

2.2 数据集处理

数据集采用广东某铝型材厂家提供的竞赛专用数据集, 和实际生产中的情况直接相关, 具有公开性、实用性。采用该数据集可以使得训练后的模型更具实用性, 实验所用数据集中包含不导电、角位漏底、桔皮、漏底、杂色、起坑、漆泡共计 7 种常见表面瑕疵, 如图 5 所示。

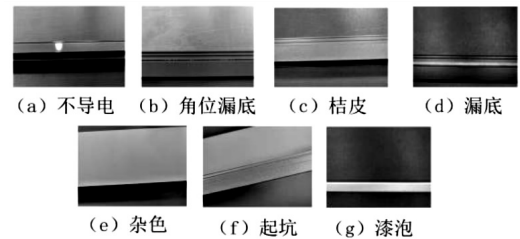


图 5 数据集样本类别

通过对数据集进行离线增强, 如旋转、亮度变化、裁剪等处理方式, 将每种类型的瑕疵数据控制在 700 张, 共计 4 900 张, 并按照 8:2 的策略划分训练集和检测集, 由于原始数据集的标签格式为 json, 因此需通过脚本文件对数据集标签进行格式转换, 调整为 YOLO 使用的 txt 格式数据标签。

2.3 评价指标

实验采用计算机视觉领域通用检测评价标准, AP

(average precision)、mAP (mean average precision) 以及 FPS (Frames Per Second) 作为模型性能评价指标, AP 表示检测精度, mAP 表示平均检测精度, FPS 表示检测速度, 其计算公式如下:

$$P = \frac{TP}{TP + FP} \quad (10)$$

$$R = \frac{TP}{TP + FN} \quad (11)$$

$$AP = \int_0^1 P(r) dr \quad (12)$$

$$mAP = \frac{1}{N} \sum_{i=1}^N AP_i \quad (13)$$

式中, TP (True Positive) 表示检测正确数量, FN (False Negative) 表示漏检数量, FP (False Positive) 表示错检数量, P (Precision) 为查准率, R (Recall) 为查全率, N 为目标检测任务的类别数量。AP 的值越大, 代表这个

种类在该网络模型上检测效果越好, mAP 值则反映该网络模型对目标检测的效果优劣。

2.4 实验对比

为验证改进算法检测性能的优越性, 将改进后的 YOLOv5 模型与 YOLOv3、YOLO v4、SSD^[20] 等主流算法进行对比, 并以 Precision、Recall、mAP、FPS 作为评估指标, 实验结果如表 2 所示。

通过表 2 可知, 改进后的 YOLOv5 在 mAP 达到了 85.8%, 相较于 YOLOv3 精度提升 10.64 百分点, 召回率提升 8.8 百分点, mAP 提升 7.45 百分点, 各项指标均有提升, 相较于原始算法, 精度提升 2.2 百分点, 召回率提升 3.9 百分点, mAP 提升 2.9 百分点, FPS 提升 4.7 百分点, 同其他算法相比, 检测性能及检测速度均具有优势。

表 2 不同网络结构性能对比

模型	Backbone	Precision/%	Recall/%	mAP/%	FPS
YOLOv3	DarkNet-53	82.56	71.70	78.35	26.13
YOLOv4	CSPDarknet53	89.96	79.40	83.50	33.60
SSD	VGG16	79.88	75.80	76.63	77.17
FasterR-CNN	VGG16	87.14	79.88	82.17	17.21
YOLOv5	CSPDarknet53	91.00	76.60	82.90	41.10
改进后 YOLOv5	CSPDarknet53	93.20	80.50	85.80	45.80

2.5 消融实验

为进一步验证改进策略的有效性, 在原有 YOLOv5 的网络结构接触上设计消融实验, 主要改进策略包括: anchor、检测层、注意力机制 E-CBAM、

EIOU, 在实验过程中为确保实验数据的可靠性, 保证超参数、数据集等条件一致。消融实验结果如表 3 所示。

表 3 YOLOv5 消融实验

模型改进策略	Precision/%	Recall/%	mAP@0.5/%	mAP@0.5:0.95/%
YOLOv5	91.00	76.60	82.90	62.00
YOLOv5+K-means++	88.30	83.00	83.60	63.60
YOLOv5+检测层	96.00	78.00	83.70	63.00
YOLOv5+ECBAM	93.10	80.70	84.10	71.80
YOLOv5+EIOU	90.20	79.70	83.20	68.50
YOLOv5+K-means+++EIOU	95.60	77.60	83.80	69.40
YOLOv5+K-means+++EIOU+检测层	93.10	80.60	84.70	69.60
YOLOv5+K-means+++EIOU+检测层+ECBAM	93.20	80.50	85.80	65.20

通过表 3 可知, 在原始算法的基础上, 使用 K-means++ 锚框生成方式后, 算法的召回率提升了 6.4 百分点, 增加检测层后, 将浅层特征进一步融合得到更加丰富, 高效的语义信息, 从而精确度提升了 5 百分点且召回率略有提升, 在原有的 YOLOv5 基础上加入改进后的注意力模块 E-CBAM 发现, 其精确度、召回率

皆有提升, 其中 mAP@0.5:0.95 更是提升了 9.8 百分点, 提升了边框回归效果, 对于 IOU 效果要求较高的任务是一项有效的改进措施。

通过表 2、表 3 及图 6 数据展示可知, 最终改进算法相比于原始 YOLOv5 算法, 在各项精确度指标皆有提升, 证明了算法改进的有效性, 同时数据集数量满足

了工业生产要求,模型的性能满足了企业需求。

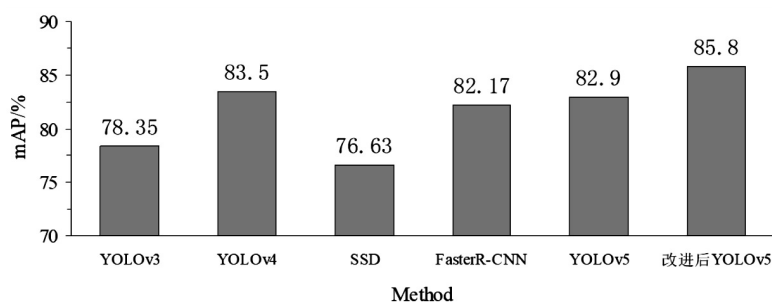


图 6 不同检测模型的 mAP 比较

3 结束语

针对铝型材的不同瑕疵类别,提出了一种基于改进 YOLOv5 的铝型材瑕疵检测算法。通过融入改进后的 E-CBAM 注意力模块,减少冗余计算,弱化背景因素影响,加强铝型材瑕疵的特征提取能力;丰富检测尺度的同时,加入新的卷积核,提升算法对不同尺度目标的检测能力,加强对铝型材表面瑕疵的区域特征学习;采用 K-means++ 锚框生成方式与 EIOU 损失函数相结合,使得算法回归更加专注于高质量锚框,提升了收敛效率,以及模型的精度和鲁棒性。

目前,铝型材相关厂家现场配备的摄像头像素较低,无法对相关瑕疵进行精确捕捉,针对该问题,可通过后期人工标注,及对采集的数据进行高分辨率重建,提升数据精度。因此后续研究方向可从瑕疵的尺度、分布及算法的边框回归函数进一步考虑改进。

参考文献:

- [1] 周 济. 智能制造是“中国制造 2025”主攻方向[J]. 企业观察家,2019(11):54-55.
- [2] 王 宇,吴智恒,邓志文,等. 基于机器视觉的金属零件表面缺陷检测系统[J]. 机械工程与自动化,2018,3(4):210-211.
- [3] 余厚云,张 辉. 汽车涡轮壳零件表面质量视觉检测[J]. 自动化仪表,2020,41(11):6-10.
- [4] ZHAO L Q,LI S Y. Object detection algorithm based on improved YOLOv3[J]. Electronics,2020,9(3):537.
- [5] FU H,SONG G,WANG Y. Improved YOLO-v4 marine target detection combined with CBAM[J]. Symmetry,2021,13(4):623.
- [6] 刘 浩,陈再良,王善翔. 基于改进 YOLOv3 算法的刀具表面缺陷检测[J]. 组合机床与自动化加工技术,2021(11):87-90.
- [7] 孙永鹏,钟佩思,刘 梅,等. 基于 YOLOv4 算法的冲压件缺陷检测[J]. 锻压技术,2022,47(1):222-228.
- [8] 乔欢欢,权恒友,邱文利,等. 改进 YOLOv5s 的交通标志识别算法[J]. 计算机系统应用,2022,31(12):273-279.
- [9] 黄志强,李 军,张世义. 基于轻量级神经网络的目标检测研究[J]. 计算机工程与科学,2022,44(7):1265-1272.
- [10] ZHANG Y,REN W,ZHANG Z,et al. Focal and efficient IOU loss for accurate bounding box regression[J]. Neurocomputing,2022,506:146-157.
- [11] LATTANZI S,SOHLER C. A better K-means++ algorithm via local search[C]//International conference on machine learning. Long Beach:IEEE,2019:3662-3671.
- [12] TIAN Weiming,DU Lin,DENG Yunkai,et al. Partition of GB-InSAR deformation map based on dynamic time warping and k-means[J]. Journal of Systems Engineering and Electronics,2022,33(4):907-915.
- [13] WOO S,PARK J,LEE J Y,et al. Cbam:convolutional block attention module[C]//Proceedings of the European conference on computer vision. Munich:IEEE,2018:3-19.
- [14] ZHOU Daquan,HOU Qibin,CHEN Yunpeng,et al. Rethinking bottleneck structure for efficient mobile network design[J]. arXiv:2007.02269,2020.
- [15] WANG Q,WU B,ZHU P,et al. ECA-Net:efficient channel attention for deep convolutional neural networks[J]. arXiv. 1910.03151,2019.
- [16] HU J,SHEN L,SUN G. Squeeze and-excitation networks[C]//2018 IEEE/CVF conference on computer vision and pattern recognition. Salt Lake City:IEEE,2018:7132-7141.
- [17] CZAJKOWSKA J,BADURA P,KORZEKWA Set al. Deep learning based high-frequency ultrasound skin-image classification with multicriteria model evaluation[J]. Sensors,2021,21(17):5846.
- [18] ARK J,WOO S,LEE J Y,et al. A simpleand lightweight attention module for convolutional neural networks[J]. International Journal of Computer Vision,2020,128(9):783-798.
- [19] LI Dongjie,ZHANG Zilei,WANG Baogang,et al. Detection method of timber defects based on target detection algorithm[J]. Measurement,2022,203:111937-1-111937-11.
- [20] LIU Wei,ANGUELOV D,ERHAN D,et al. SSD:single shot multibox detector[C]//Proc of 14th European conference on computer vision. [s. l.]:Springer,2016:21-37.