

[自控·检测]

DOI:10.3969/j.issn.1005-2895.2023.03.010

基于改进 YOLOv5 的陶瓷环缺陷检测算法

王旭, 管声启*, 刘通, 张理博, 王静国, 于资江

(西安工程大学机电工程学院, 陕西西安 710048)

摘要:针对陶瓷环缺陷信息微弱且特征各异、种类繁多导致特征难以提取的问题,课题组提出了一种基于改进 YOLOv5 的陶瓷环缺陷检测算法。首先,在 YOLOv5 中的 Backbone 网络添加 CBAM 注意力机制模块,通过学习的方式在空间和通道上对特征图像进行权重分配,有效地提升模型对于不同类型缺陷通道特征和空间特征信息的提取能力;然后,采用 CARAFE 算子替换原 YOLOv5 中的最近邻上采样算子,该模块基于输入特征自适应生成上采样内核,有效的增加模型的感受域;最后,添加一个新的特征融合层,提取较低的空间特征并将其与深层的特征进行融合生成新的特征图,提升模型对小目标缺陷的检测能力。实验结果表明课题组提出的陶瓷环缺陷检测算法检测所有缺陷种类平均精度均值可以达到 90.7%,能够实现陶瓷环缺陷的检测。

关键词:陶瓷环;缺陷检测;改进 YOLOv5;注意力机制;CARAFE 上采样

中图分类号:TP391.4;TH145.1 文献标志码:A 文章编号:1005-2895(2023)03-0066-06

Ceramic Ring Defect Detection Algorithm Based on Improved YOLOv5

WANG Xu, GUAN Shengqi*, LIU Tong, ZHANG Libo, WANG Jingguo, YU Zijiang

(School of Mechanical and Electrical Engineering, Xi'an Polytechnic University, Xi'an 710048, China)

Abstract: Aiming at the problem of weak defect information and diverse features of ceramic ring defects, which makes it difficult to extract detection features, a ceramic ring defect detection algorithm based on improved YOLOv5 was proposed. First, a CBAM attention mechanism module was added to Backbone in YOLOv5 to assign weights to feature images in space and channels by learning, which effectively improved the ability of model to extract channel features and spatial feature information for different types of defects; Then the nearest neighbor upsampling operator in original YOLOv5 was replaced by the CARAFE operator. This module adaptively generated an upsampling kernel based on input features, effectively increased the perceptual domain of the model; Finally, a new feature fusion layer was added to extract the lower spatial features and fuse them with the deeper features to generate a new feature map, and improved the ability of model to detect small target defects. The experimental results show that the mean average precision of the proposed ceramic ring defect detection algorithm for detecting all defect types can reach 90.7%, which can effectively achieve the detection of ceramic ring defects.

Keywords: ceramic ring; defect detection; improved YOLOv5; attention mechanism; CARAFE upsampling

当前我国的陶瓷制造业快速发展,陶瓷产业规模已居世界首位^[1]。特别是21世纪以来,随着工业机械化迅速的发展,已经开始大规模工业化流水线生产,在我国已经出现了很多主要陶瓷生产基地,如广东等几大陶瓷生产区。目前我国已成为全世界陶瓷出口大国

之一^[2]。陶瓷环零件制作作为陶瓷产业的一个重要环节,在工业生产中已经进入批量生产的阶段。对于陶瓷环零件的缺陷检测,目前主要通过人工检测^[3],而人工检测存在效率低、成本高等问题,难以满足现在高速发展的工业需求。

收稿日期:2022-10-25;修回日期:2023-03-10

基金项目:西安市创新能力强基计划:人工智能技术攻关项目(21RGZN0021)。

第一作者简介:王旭(1996),男,陕西武功人,硕士研究生,主要研究方向为机器人视觉。通信作者:管声启(1971),男,陕西西安人,博士,教授,主要研究方向为纺织品表面质量在线检测、机器人视觉等。E-mail:sina1300841@163.com

近年来,广大学者开展了深入的研究。由于陶瓷环具有很好的绝缘性,因此用途较广的电磁无损检测技术就不能应用于陶瓷的质量检测^[4]。文献[5]通过滤除噪声实现对陶瓷零件缺损缺陷的识别,但检测的缺陷类别单一,不适用于多类缺陷的检测;文献[6]对电容器的缺陷采用基于形状的模板匹配算法进行检测;文献[7]通过最大类间方差法与形态学等操作建立形状匹配模板实现对冲压件外形缺陷的检测,但基于模板匹配的检测方法在缺陷大小和形状发生变化时,准确率较低且匹配速度较慢;文献[8]对裂缝缺陷预处理后,建立深度卷积网络,检测准确率较高;文献[9]构建了基于卷积网络的带钢表面缺陷检测模型,自动获得目标的本质特征,从而进行识别分类,准确率达到 98%。但大多数基于卷积网络的缺陷检测方法对于多种类缺陷及小目标检测的效果不太理想。

为了解决陶瓷环缺陷小、种类多难以检测的问题,课题组通过研究卷积神经网络的机理,设计了一种适用于陶瓷环缺陷的目标检测算法,实现了对陶瓷环缺陷准确的检测。

1 YOLOv5 算法

YOLOv5 的网络结构包括 Input 端、Backbone 网络、Neck 部分和 Prediction 端^[10],YOLOv5 网络模型结构如图 1 所示。Input 端包括 Mosaic 数据增强、自适应图片缩放和自适应锚框计算;Backbone 模块主要包含 Focus、跨阶段局部网络(cross stage partial network, CSPNet)^[11]和空间金字塔池化(spatial pyramid pooling, SPP)3 部分;Neck 部分采用特征图金字塔网络(feature pyramid networks, FPN)+路径聚合网络(path aggregation network, PAN)结构;Prediction 端的边界框损失函数使用 GIOU 函数,并采用加权非最大抑制(weighted non maximum suppression, weighted NMS)对非极大值^[12]进行抑制,进而获得最优目标检测框^[13]。图 1 中,输出层有 3 个尺寸的输出通道^[14]。

由于陶瓷环缺陷小、特征信息微弱,针对陶瓷环缺陷检测采用的 YOLOv5 算法对陶瓷环的检测效果并不理想,仍需进一步改进。

2 改进的 YOLOv5 算法

为了解决陶瓷环缺陷种类繁多、缺陷体积小和特

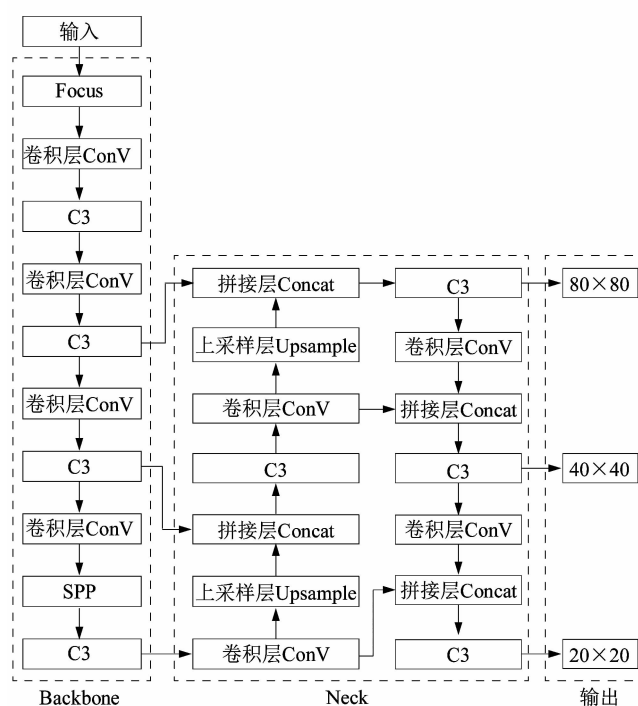


图 1 YOLOv5 网络结构

Figure 1 YOLOv5 network structure

征信息微弱难以提取的问题,课题组对 YOLOv5 原网络结构进行了 3 处改进:①添加 CBAM 注意力机制模块以有效提升模型对于不同类型缺陷通道特征和空间特征信息的提取能力;②采用 CARAFE 算子替换原 YOLOv5 中的最近邻上采样算子,有效的增加模型的感受域;③添加了一个新的微小特征融合层,提升模型对小目标的检测能力。改进后的 YOLOv5 算法检测框架如图 2 所示。

2.1 注意力机制

计算机视觉中的注意力机制^[15]本质在于模拟人眼的视觉感知^[16],通过训练过后自动调整配置参数,可以像人类视觉一样重视有用信息,减少无效信息的干扰,提高效率与准确率,是目标检测研究中不可或缺的重要组件^[17-19]。课题组通过在 Backbone 网络的最后一个 C3 层后引入注意力机制 CBAM 模块^[20],可以提升模型对于不同类型缺陷通道特征和空间特征信息的提取能力。

CBAM 由通道注意力(channel attention module, CAM)^[21]和空间注意力(spatial attention module, SAM)^[22]2 部分组成,其网络结构如图 3 所示。输入

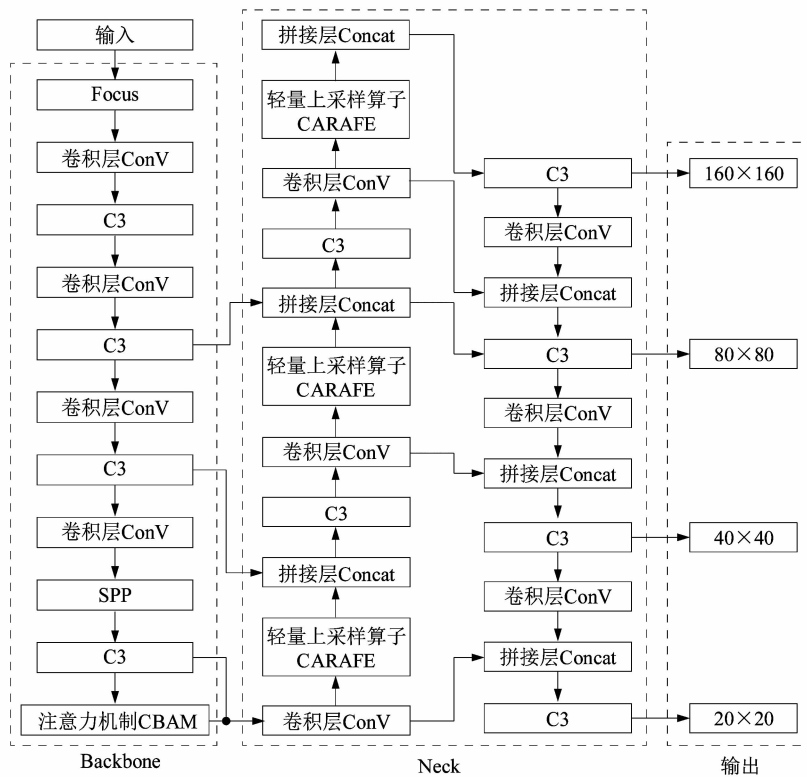


图 2 改进的 YOLOv5 算法框架

Figure 2 Framework of improved YOLOv5 algorithm

的特征依次经过通道注意力模块、空间注意力模块的筛选,最终获得了经过重标定的特征,即强调重要的特征,压缩不重要特征。

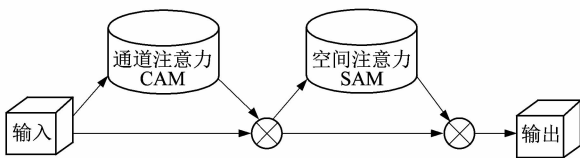


图 3 CBAM 网络结构

Figure 3 CBAM network structure

2.2 CARAFE 上采样

YOLOv5 采用的最近邻上采样只是通过像素点的空间位置来决定上采样核,而且感知域很小,因此课题组采用 CARAFE^[23] 算子替换原 YOLOv5 中的最近邻上采样算子,可以有效地增加模型的感受域,替换后的网络结构如图 2 中 Neck 部分所示。

CARAFE 包含上采样特征模块和特征重组模块,如图 4 所示。假设输入一个形状为 $H \times W \times C$ 的特征图,上采样的倍率为 σ ,采样算子 CARAFE 通过上采

样特征模块预测上采样核,再经过特征重组模块完成上采样,最终获得一个形状为 $\sigma H \times \sigma W \times \sigma C$ 的输出特征图。

2.2.1 上采样特征模块

1) 特征图通道压缩:假设输入一个形状为 $H \times W \times C$ 的特征图,首先用一个 1×1 卷积将他的通道数压缩到 $H \times W \times C_m$ 。

2) 内容编码及上采样核预测:假设上采样核尺寸为 $K_{up} \times K_{up}$ (上采样核越大意味着更大的感受野和更大的计算量),如果希望对输出特征图的每个位置使用不同的上采样核,那么需要预测的上采样核形状为 $\sigma H \times \sigma W \times K_{up} \times K_{up}$ 。对于 1) 中压缩后的输入特征图,利用一个 $K_{encoder} \times K_{encoder}$ 的卷积层来预测上采样核,输入通道数为 C_m ,输出通道数为 $\sigma^2 K_{up}^2$,然后将通道维在空间维展开,得到形状为 $\sigma H \times \sigma W \times K_{up}^2$ 的上采样核。

3) 上采样核归一化:对 2) 中得到的上采样核利用 softmax 进行归一化,使得卷积核权重和为 1。

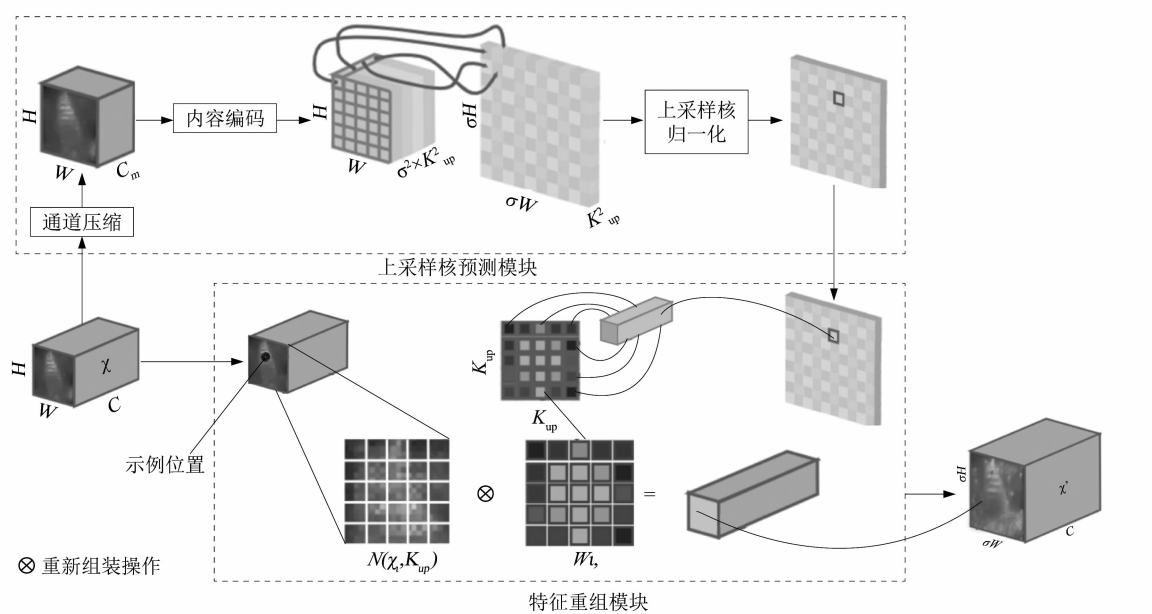


图 4 CARAFE 上采样结构

Figure 4 CARAFE upsampling structure

2.2.2 特征重组模块

将输出特征图中的每个位置映射回输入特征图,取出以之为中心的 $K_{up} \times K_{up}$ 区域,并和预测出的该点的上采样核作点积得到输出值。相同位置的不同通道共享同一个上采样核。

2.3 增加小目标检测层

陶瓷环缺陷尺寸微小,YOLOv5 算法对于小目标检测的效果并不理想,为了提升 YOLOv5 算法的目标检测性能,添加一个新的特征融合层。特征融合层经过 Backbone 网络 4 倍下采样输出,然后再和 8 倍下采样特征图融合,最终生成一个尺寸为 160×160 像素的特征图,如图 2 中输出端所添加的 160×160 像素特征融合层所示。

3 实验结果及分析

3.1 实验环境和实验数据集

实验环境的配置:CPU 型号 Intel x299/i7-7820x,采用的双 GPU 型号均是 NVIDIA GeForce RTX 2080Ti,内存均为 16 GiB,实验系统均基于 Ubuntu 操作系统,以 pytorch 作为深度学习的框架。

为了验证改进 YOLOv5 算法是否能准确的识别出陶瓷环缺陷,实验使用工业相机拍摄了 780 张陶瓷环零件缺陷照片。缺陷总共 6 类,分别为:缺损、污点、裂纹、划伤、阴影和变形。每类缺陷 130 张,并采用缩放、平移等数据集扩充方法将陶瓷环缺陷数据集扩充到 1 560 张。实验数据集中的原始图片大小为 $1\ 600 \times 1\ 200$ 像素。陶瓷环缺陷数据集的部分样本如图 5 所示。

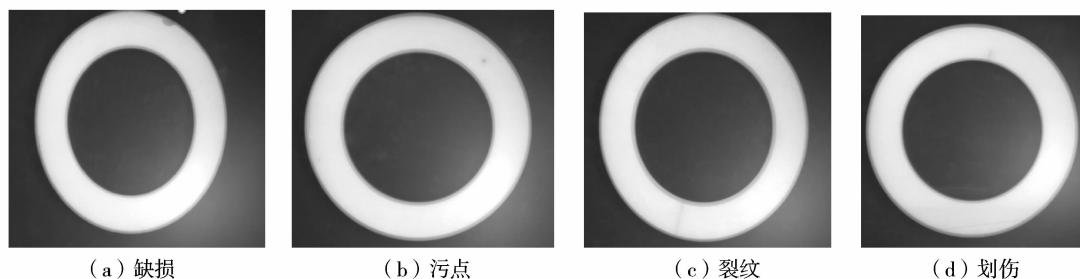


图 5 数据集部分样本

Figure 5 Partial samples of dataset

3.2 网络模型评估指标

课题组选用平均精度均值作为评估指标,以平均精度 (average precision, AP) 和平均精度均值 (mean average precision, mAP) 作为依据。各类缺陷的平均精度 P_A 等于图 6 和图 7 中各类缺陷的 $P-R$ 曲线围成的面积;所有缺陷种类平均精度的均值为 $P_{A,m}$, $P_{A,m}$ 越大,表示缺陷检测的效果越好。

3.3 实验结果与分析

YOLOv5 和改进 YOLOv5 在对缺陷检测数据集进行训练过程中,由于硬件设备的限制,批次大小最大可以设置为 16。因此,课题组将实验部分所有模型的批次大小设置为 16,迭代次数设置为 200。

Yolov5 算法经过 200 轮训练,实验结果如图 6 所示。

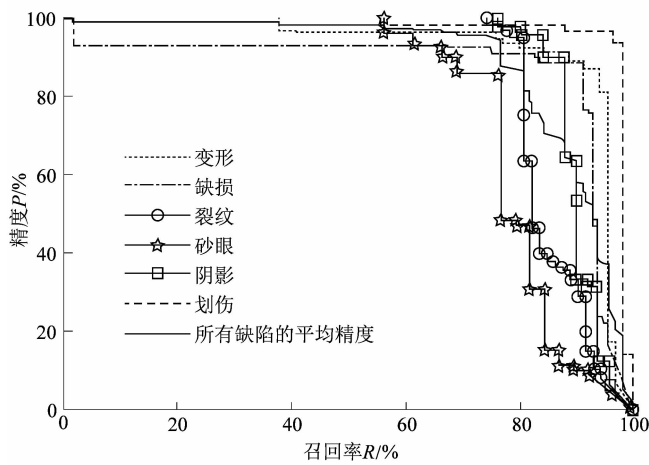


图 6 YOLOv5 的 $P-R$ 曲线

Figure 6 Curves of precision and recall rate of YOLOv5

改进 YOLOv5 在 YOLOv5 中加入 CBAM 注意力机制模块,并采用 CARAFE 算子替换原 YOLOv5 中的最近邻上采样算子,再添加一个新的微小特征融合层,提升模型对小目标的检测能力。改进后的 YOLOv5 算法经过 200 轮训练后, $P-R$ 曲线如图 7 所示。

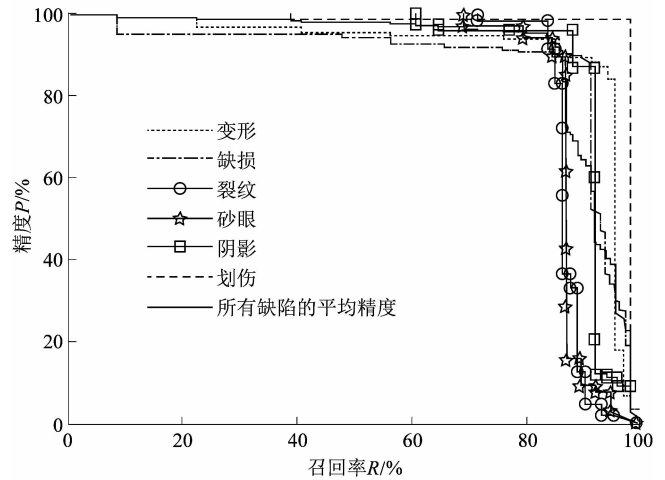


图 7 改进 YOLOv5 的 $P-R$ 曲线

Figure 7 Curves of precision and recall rate of improved YOLOv5

由图 6 和图 7 可以看出,检测陶瓷环各类缺陷的平均精度 P_A 和所有种类缺陷平均精度的均值 $P_{A,m}$ 如表 1 所示。

由表 1 可知 YOLOv5 算法检测的陶瓷环所有种类缺陷平均精度的均值 $P_{A,m}$ 达到 88.9%。但是由于陶瓷环缺陷体积小、种类繁多,且缺陷特征微弱难以提取,原 YOLOv5 算法对于缺陷的识别准确率有待提高。

表 1 改进前后平均精度和平均精度均值

Table 1 Average precision and mean average precision before and after improvement

模型	平均精度 P_A / %						所有缺陷平均精度 均值 $P_{A,m}$ / %
	变形	缺损	裂纹	砂眼	阴影	划伤	
YOLOv5	92.8	87.2	85.9	79.4	90.7	97.6	88.9
改进 YOLOv5	91.8	88.7	87.4	87.4	91.5	97.5	90.7

课题组提出的改进 YOLOv5 算法检测的陶瓷环所有种类缺陷平均精度的均值 $P_{A,m}$ 可以达到 90.7%,和原始的 YOLOv5 模型相比较, $P_{A,m}$ 提高了 1.8%。由此可见,改进后的 YOLOv5 可以有效的提升模型对陶瓷环缺陷的检测精度。

4 结语

为了实现陶瓷环各种缺陷的检测,课题组提出了一种基于改进 YOLOv5 的陶瓷环缺陷检测算法。该算法能够有效增加模型的感受域,获取更多的特征信息,增强模型对于缺陷通道特征和空间特征信息的提取能

力,实验结果显示改进后的 YOLOv5 对陶瓷环缺陷检测的平均精度均值 $P_{A,m}$ 达到 90.7%,为解决小目标检测问题提供了一种有效的方法。但由于现有数据集的数量较少,且有的陶瓷环缺陷数据集质量较差,从而影响检测精度,后期需要加强数据集的采集,以及研究有效的数据集扩充方法。

参考文献:

- [1] 朱天阳,黄用俊,李少杰,等. 精细陶瓷在智能手机上的应用及其制备工艺[J]. 陶瓷, 2018(5):63-70.
- [2] 高青松,李婷.“中国制造 2025”研究进展及评述[J]. 工业技术经济, 2018,37(10):59-66.
- [3] 佚名. 中国智能制造系统解决方案行业现状与发展前景[J]. 电器工业, 2019(11):41-45.
- [4] ZHAO Z K. Review of non-destructive testing methods for defect detection of ceramics [J]. *Ceramics International*, 2021, 47(4):4389-4397.
- [5] 马鼎,吴定祥,唐立军. 环形陶瓷金属涂层缺陷检测方法研究[J]. 电子科技, 2021, 34(6):28-33. DOI: 10.16180/j.cnki.issn1007-7820.2021.06.005.
- [6] 俞洋,陈佐政,陈祝洋,等. 机器视觉在电容器外观缺陷检测中的应用[J]. 电子技术, 2019, 45(9):97-100. DOI:10.16157/j.issn.0258-7998.190134.
- [7] 李永敬,朱萍玉,孙孝鹏,等. 基于形状模板匹配的冲压件外形缺陷检测算法研究[J]. 广州大学学报(自然科学版), 2017, 16(5):62-66.
- [8] 毕松,高峰,陈俊文,等. 基于深度卷积神经网络的柑橘目标识别方法[J]. 农业机械学报, 2019, 50(5):181-186.
- [9] 王立中,管声启. 基于深度学习算法的带钢表面缺陷识别[J]. 西安工程大学学报, 2017, 31(5):669-674. DOI: 10.13338/j.issn.1674-649x.2017.05.012.
- [10] 王莉,何牧天,徐硕,等. 基于 YOLOv5s 网络的垃圾分类和检测[J]. 包装工程, 2021, 42(8):50-56.
- [11] WANG C Y, LIAO H Y M, WU Y H, et al. CSPNet: a new backbone that can enhance learning capability of CNN [C]// Proceedings of the IEEE/CVF conference on computer vision and pattern recognition workshops. Seattle, WA, USA: IEEE, 2020:390-391.
- [12] MISRA D. Mish: a self regularized non-monotonic neural activation function[EB/OL]. (2019-08-23) [2020-11-20]. <https://arxiv.org/abs/1908.08681>.
- [13] 谈世磊,别雄波,卢功林,等. 基于 YOLOv5 网络模型的人员口罩佩戴实时检测[J]. 激光杂志, 2021, 42(2):147-150.
- [14] 林清平,张麒麟,肖蕾. 采用改进 YOLOv5 网络的遥感图像目标识别方法[J]. 空军预警学院学报, 2021, 35(2):117-120.
- [15] LI Y M, TONG S C, LI T S. Composite adaptive fuzzy output-feedback control design for uncertain nonlinear strict-feedback systems with input saturation [J]. *IEEE Transactions on Fuzzy Systems*, 2016, 24(10):841-853.
- [16] 卓天天,桑庆兵. 注意力机制与复合卷积在手写识别中的应用[J]. 计算机科学与探索, 2022, 16(4):888-897.
- [17] 石磊,王毅,成颖,等. 自然语言处理中的注意力机制研究综述[J]. 数据分析与知识发现, 2020, 4(5):1-14.
- [18] 赵广谦,姜培刚,林天然. 基于 CNN-BiLSTM 网络及注意力机制的智能滚动轴承剩余寿命预测方法[J]. 机电工程, 2021, 38(10):1253-1260.
- [19] 崔海朋,姜英昌. 基于注意力机制的海上小目标识别方法[J]. 机电工程技术, 2022, 51(7):100-103.
- [20] WOO S, PARK J, LEE J Y, et al. CBAM: convolutional block attention module [C]// Proceedings of the 2018 European Conference on Computer Vision. Cham:Springer, 2018:3-19.
- [21] LIANG B, LIU Q, XU J, et al. Target-specific sentiment analysis based on multi-attention convolutional neural network[J]. *Computer Research and Development*, 2017, 54(8):1724-1735.
- [22] ZHU X Z, CHENG D Z, ZHANG Z, et al. An empirical study of spatial attention mechanisms in deep networks [C]//2019 IEEE/CVF International Conference on Computer Vision (ICCV). Seoul, Korea (South):IEEE Computer Society, 2019:6687-6696.
- [23] WANG J Q, CHEN K, XU R, et al. Carafe: content-aware reassembly of features [C]//Proceedings of the IEEE International Conference on Computer Vision. Seoul, Korea (South):IEEE, 2019:3007-3016.