

DOI: 10.13382/j.jemi.B2306681

基于自适应多尺度特征融合网络的金属齿轮端面缺陷检测方法研究*

王 宸^{1,2,3} 杨 帅¹ 周 林¹ 华珀玺¹ 王生怀¹ 吕 江⁴

(1. 湖北汽车工业学院机械工程学院 十堰 442000; 2. 中国工程科技十堰产业技术研究院 十堰 442000; 3. 上海大学上海市智能制造与机器人重点实验室 上海 200072; 4. 东风汽车零部件(集团)有限公司刃量具分公司 十堰 442000)

摘 要:针对目前金属齿轮端面结构复杂,导致缺陷的小目标占比度高和尺度变化大引起的检测准确率低,难以满足企业实时在线检测需求等问题。本文基于 YOLOv5s 网络提出了一种基于自适应多尺度特征融合网络的金属齿轮端面缺陷检测方法(YOLO-Gear)。首先,搭建了一个齿轮端面缺陷检测试验台,并制作了齿轮端面缺陷数据集。然后,提出了自适应卷积注意力模块(convolutional block attention module-C3,CBAM-C3),CBAM-C3 通过将通道注意力(channel attention module,CAM)和空间注意力(spatial attention module,SAM)相结合加强了对金属齿轮缺陷小目标缺陷自适应的特征学习与特征提取,及时对模型中的权重参数进行学习和优化,提高了模型对小目标缺陷的检测准确度;最后,提出了重复加权双向特征金字塔网络(bidirectional feature pyramid network,BiFPN),通过自适应控制不同尺度的特征图之间的融合程度,提高了模型对缺陷多尺度检测能力。试验表明,YOLO-Gear 模型在齿轮端面缺陷测试集上的平均精度达到了 99.2%,F1 值为 0.99,FPS 值为 33。相较于其他深度学习模型,本文提出的 YOLO-Gear 模型提高了检测的精度和效率,能够满足企业的实时在线检测需求。

关键词: 齿轮端面;YOLOv5s;CBAM;BiFPN;缺陷检测

中图分类号: TP391.4;TH164 文献标识码: A 国家标准学科分类代码: 520.2060

Research on metal gear end-face defect detection method based on adaptive multi-scale feature fusion network

Wang Chen^{1, 2, 3} Yang Shuai¹ Zhou Lin¹ Hua Boxi¹ Wang Shenghuai¹ Lyu Jiang⁴

(1. Hubei University of Automotive Technology, Shiyan 442000, China; 2. China Academy of Engineering Science and Technology Shiyan Industrial Technology Research Institute, Shiyan 442000, China; 3. Shanghai Key Laboratory of Intelligent Manufacturing and Robotics, Shanghai University, Shanghai 200072, China; 4. Cutting and Measuring Tools Branch of Dongfeng Motor Parts Co., Ltd., Shiyan 442000, China)

Abstract: The high proportion and large-scale variation of small targets with defects caused by the complex structure of metal gear end faces have led to low detection accuracy, making it difficult to meet the real-time online detection needs of enterprises. In this paper, we propose a metal gear end face defect detection method based on an adaptive multi-scale feature fusion network (YOLO-Gear) using the YOLOv5s network. Firstly, we establish a gear end face defect detection test platform and create a gear end face defect dataset. Then, we introduce the adaptive convolutional block attention module (CBAM-C3) which combines channel attention module (CAM) and spatial attention module (SAM) to enhance the adaptive feature learning and extraction for small target defects in metal gears, effectively improving the detection accuracy of the model for small target defects. Finally, we propose the bidirectional feature pyramid network (BiFPN), which repetitively weights and fuses features from different scales, thereby improving the model's ability to detect defects at multiple scales. Experimental results demonstrate that the YOLO-Gear model achieves an average precision of 99.2%, an F1 score of 0.99, and an FPS value of 33 on the gear end face defect test set. Compared to other deep learning models, the proposed YOLO-Gear model in this paper improves both detection accuracy and efficiency, meeting the real-time online detection needs of enterprises.

Keywords: gear end-face; YOLOv5s; CBAM; BiFPN; defect detection

收稿日期: 2023-06-30 Received Date: 2023-06-30

* 基金项目: 国家自然科学基金(51475150)、教育部人文社科项目(20YJCZH150)、湖北省重点研发计划项目(2021BAA056)、湖北省高等学校中青年科技创新团队计划项目(T20200018)、湖北省社科基金(21Q174)、湖北汽车工业学院博士基金(BK201905)项目资助

0 引言

随着工业生产的不断发展,企业更加注重追求更高的质量、稳定的可持续性、更低的浪费和更快的制造,零缺陷制造(zero-defect manufacturing, ZDM)作为最有效的质量提升方式,逐渐成为更加环保、高效的零缺陷生产线的新标准。ZDM 的策略包括检测、预测、修复和预防,其中缺陷检测是最重要的部分。在制造过程中,金属齿轮会出现划痕、凹痕和断齿等缺陷,说明设备存在残留铁屑或夹紧不当、干扰、碰撞等问题。如果不及时有地解决这些问题,可能会出现各种不良现象。基于经验的人工采样和检测方法容易出现错误、遗漏和耗时问题,这对企业的可持续性非常不利^[1-3]。

东风汽车公司某齿轮厂,齿轮每年的产量高达三百多万件。检测金属齿轮的端面是否有缺陷成为了该公司面临的一项重要任务。而金属齿轮是现代工业中用来传递动力和运动的基本元件^[4]。金属齿轮端面常见缺陷有断齿、划痕。如果存在断齿、划痕等缺陷,会导致齿轮在运动时产生噪音、磨损等问题。甚至可能会加速设备的损坏,影响整个机器的运行。所以为了避免对工厂的营业不利,对于金属端面缺陷的检测是非常有必要的^[5-6]。

近些年来图像处理技术的快速发展,越来越多的学者聚焦于将基于机器视觉的自动检测技术应用到表面缺陷检测中。包从望等^[7]通过 Canny 算子、双边滤波和拉式锐化增强算法等传统处理工具对齿轮图像边缘进行处理,提取到了较好的齿轮边缘信息,算法缺陷准确度达到了 93.8%。Vivekananthan 等^[8]通过对图像预处理后,转换为灰度图,再采用 Otsu 法设定阈值范围,以提高裂纹检测的准确性。使用 Sobel 滤波器检测图像像素边缘上的裂纹;Pasrommatis 等^[9]通过采用 Otsu 阈值法对小模数塑料齿轮进行图像分割,把连通域的面积值作为齿轮是否合格的判定标准。赵佩瑶^[10]等针对板式换热器板片产生的缺陷,根据先验知识对缺陷区域进行了解,然后定位缺陷的感兴趣区域,最后采用区域分割算法分割出不同的区域,进而进行缺陷检测,显著提高了缺陷的检测准确率。

虽然传统图像处理算法提出较早且发展比较成熟,在表面缺陷检测领域应用广泛。但是随着制造工业的不断发展,传统算法已难以满足日益提高的检测精度与智能化需求,基于卷积神经网络^[11-12](convolutional neural networks, CNN)的目标检测算法凭借强大的表征能力与出色的可移植性逐渐取代传统算法成为表面缺陷检测领域研究的热点。目前基于 CNN 的缺陷检测网络从结构上可分为单阶段算法和两阶段算法。Singh 等^[13]通过基于 faster R-CNN、SSD(single shot detector)和 YOLOv3 (

you look only once)网络做了对比试验,发现 YOLOv3 在实现机器人焊缝形状检测任务中表现优异,同时相较于其他模型,YOLO 模型更容易实现实时检测;向宽等^[14]使用改进的 Faster RCNN 模型处理铝材表面缺陷检测问题,并利用 ROI 代替感兴趣区域池化层(ROI pooling),提高了小目标的检测能力,为工业铝材质量检测提供一种有效的应用参考;Bakir 等^[15]利用热成像图像和基于卷积神经网络的深度学习分类器,提出了一种用于分类光伏组件中缺陷的缺陷检测系统。检测平均精度均值(mAP)达到了 95.05%;李勇等^[16]提出了一种基于改进 BiFPN 的微电机电枢表面缺陷模型,采用通道注意力机制增强改进的 BiFPN 结构,对提取到的不同维度特征进行融合,并对特征进行筛选,检测准确度高达 98.42%;Appe 等^[17]提出了一种改进的 YOLOv5 番茄检测算法,用于检测番茄是否破损。将 CBAM 注意力机制加入 CAM-YOLO 中,提高了对番茄的检测能力。

综上所述,由于金属齿轮端面结构复杂,导致缺陷小目标占比度高,缺陷尺度变化大等问题,使得传统检测方法高度依赖高水平的专家经验;此外,传统的检测方法更多的关注局部特征,缺乏对缺陷的全局关注,在复杂的工作场景中容易受到干扰,缺乏稳定性,鲁棒性和泛化性也有一定的局限性,难以应用于更具有挑战性的工业场景,因此,传统的检测方法难以对金属齿轮端面的小目标和多尺度缺陷进行分类和定位。而深度学习中各种方法都有自己有的优缺点,Faster RCNN 模型检测精度高,但是计算量大;而 YOLO 系列能实现实时检测,检测速度快,但是检测精度比较低,可以通过改进网络提升检测精度。目前有关齿轮端面缺陷检测的深度学习算法研究较少。

本文在 YOLOv5s 网络中融合了 BiFPN 特征提取策略,并引进了自适应注意力机制 CBAM-C3 模块,最终提出了一种基于自适应多尺度特征融合的金属齿轮端面缺陷检测方法(YOLO-Gear)用于检测金属齿轮端面缺陷。得到的 YOLO-Gear 模型可以对齿轮端面缺陷样本图像进行特征融合,进一步提高特征的辨识度,并且突出特征图的关键信息,丰富全局信息,较为显著地提高了金属齿轮端面缺陷检测的精度。

1 试验台的搭建和数据集的建立

1.1 齿轮端面缺陷智能检测系统

搭建的金属齿轮端面缺陷智能检测平台,如图 1 所示。金属齿轮端面主要有划痕和断齿两种缺陷,如图 2 所示。由于缺陷与金属齿轮端面的色差不是很大,因此为了确保采集到的图像质量,需要特定的光源环境,本文采用点阵光源对齿轮上、下端面照射,保证了齿轮上、下

端面的亮度,CCD 工业相机穿过点阵光源对放在 PLC 控制的传送带上的齿轮上端面进行图像采集;两个传送带之间装配有电磁翻转装置,齿轮经过此装置完成上、下端面的翻转,从而完成齿轮下端面图像的采集。两次采集到的金属齿轮上、下端面图像输入到 YOLO-Gear 模型进行端面缺陷检测。若检测到有缺陷,则通过 PLC 控制剔除装置,将有缺陷的金属齿轮剔除到回料仓,如图 3 为金属齿轮端面缺陷检测流程图,图 4 为检测装置结构设计简图。

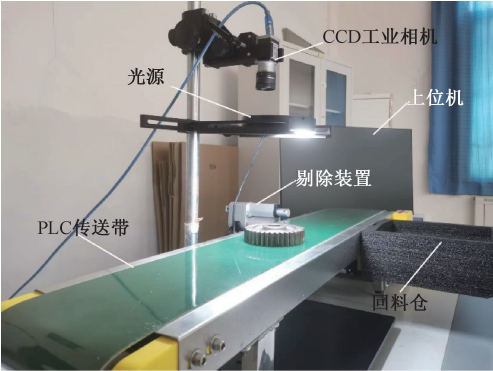


图 1 齿轮端面缺陷智能检测平台

Fig. 1 Intelligent inspection platform for gear surface defects

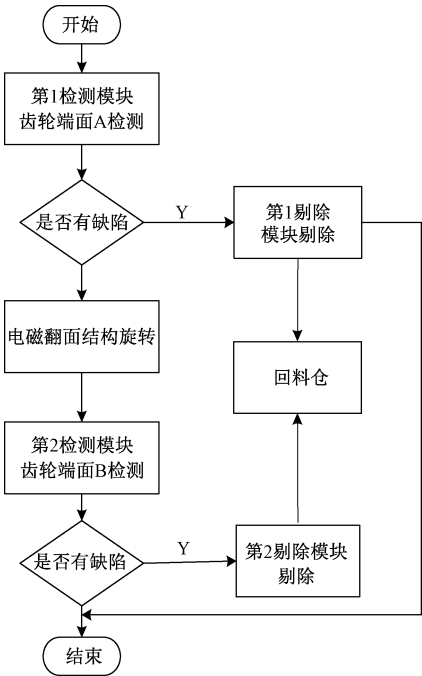


图 3 齿轮端面缺陷检测试验流程

Fig. 3 Flow chart for gear surface defect detection test

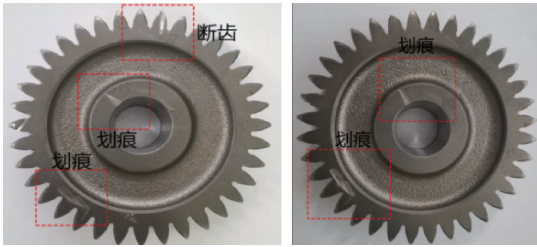


图 2 金属齿轮端面缺陷

Fig. 2 Gear surface defects

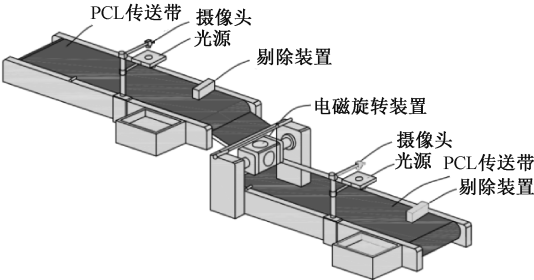


图 4 齿轮端面缺陷智能检测装置结构设计简图

Fig. 4 Sketch of the structure design of the gear surface defect inspection platform

1.2 数据集标注与扩充

本文的数据集来源于东风汽车公司某齿轮厂,采集到的齿轮端面图像是在工厂专业的技术人员操作下进行的,图像质量符合工厂的检测标准。本文试验使用开源图像标注工具 LabelImg,标注齿轮缺陷位置。如图 5 所示。左边为划痕标签的样标,右边为断齿标签的样标。

采集到原始的数据集图片一共有 450 张,然后随机按照 8 : 1 : 1 划分为训练集、测试集和验证集。为了保证检测效果,采用水平翻转,变形缩放和随机裁剪等数据增强方法,将数据集扩展为 2 300 张照片。具体每个类型的数据如表 1 所示,数据集增强的效果如图 6 所示。经过数据增强后的数据集不但扩充了图像的数量,还引入了噪声干扰,丰富了金属齿轮端面缺陷姿态和尺度,有

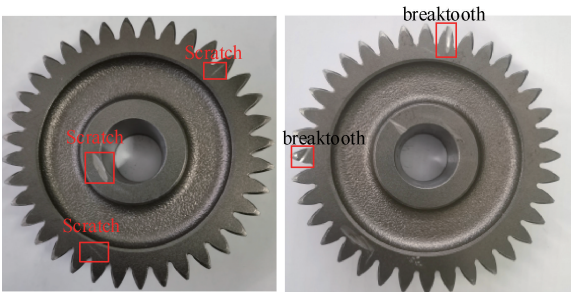


图 5 标注样式图

Fig. 5 Dimension style diagrams

利于后期训练出更稳定的模型。

表 1 缺陷类型数据表
Table 1 Defect Type Data Sheet

标签名称	总数据量	训练数据集	验证数据集	测试数据集
		(80%)	(10%)	(10%)
全部	2 300	1 840	230	230
Break-tooth	1 550	1 226	161	145
Scratch	1 916	1 541	191	184

2 算法原理介绍与改进

针对金属齿轮端面缺陷检测算法图像特征提取能力不足、小目标检测度低等问题。本文对 YOLOv5s 进行改进得到 YOLO-Gear 模型,模型示意图如图 7 所示。改进的 YOLO-Gear 模型结构由“主干”、“颈部”和“头部”3 个部分组成。“主干”部分负责提取图像的低层特征,“颈

部”部分负责将低层特征转换为高层特征,“头部”部分负责对高层特征进行分类和回归,以得到最终的目标检测结果。

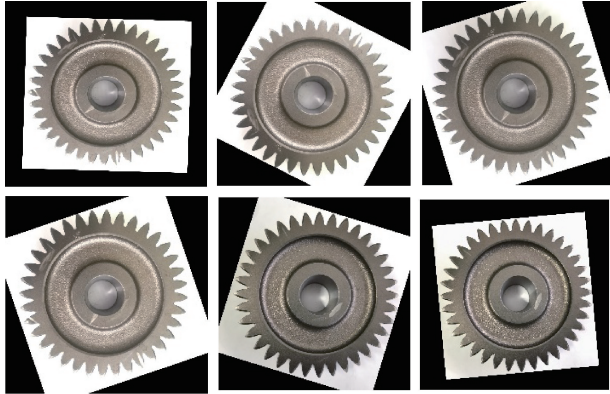


图 6 数据集增强的效果
Fig. 6 Effect of dataset enhancement

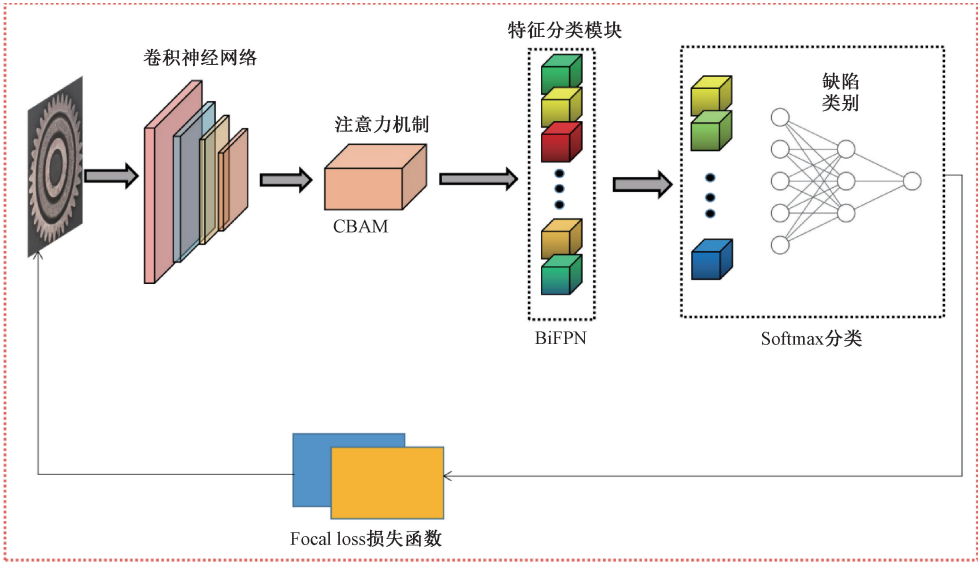


图 7 YOLO-Gear 模型示意图
Fig. 7 Diagram of the YOLO-Gear model

2.1 CBAM-C3 注意力机制

YOLOv5s 使用的是 SE^[18] 注意力机制,SE 注意力机制能够通过自适应地学习通道之间地关系来增强重要特征地表达能力,但是在处理金属齿轮缺陷小目标问题上存在一定的不足。由于小目标的尺寸相对较小,其空间信息对于检测和定位至关重要,SE 注意力机制是将特征图压缩为全局描述向量,这种操作可能导致空间信息的丢失,对于金属齿轮小目标缺陷检测很不友好。因此本文引入 CBAM^[19] 注意力机制,CBAM 通过将通道注意力和空间注意力相结合,对特征图进行加权。通道注意力模块用于建模通道之间的关系,以提取重要的语义特征。

空间注意力模块则关注特征图的不同空间位置,保留有助于检测小目标的细节信息。通过结合通道和空间注意力,CBAM 可以更好地适应小目标的特征表示需求。另外,CBAM 相比于 SE 注意力机制,CBAM 的计算复杂度相对较低。CBAM 的模块设计简单且轻量化,不需要引入额外的全连接层。这使得 CBAM 更适合在资源受限的环境下使用,如嵌入式设备或实时目标检测任务。此外,CBAM 的设计是可学习的,模块中的权重参数可以通过训练数据进行学习和优化。这使得 CBAM 能够根据具体任务和自制的数据集的特点进行自适应调整,以更好地适应金属齿轮缺陷小目标的检测需求。通过学习得到的

权重参数,CBAM 可以自动选择和加权特征图中最具有区分性和重要性的信息。如图 8 所示。从图中可以看出,CBAM 由输入、通道注意力模块、空间注意力模块和输出组成。

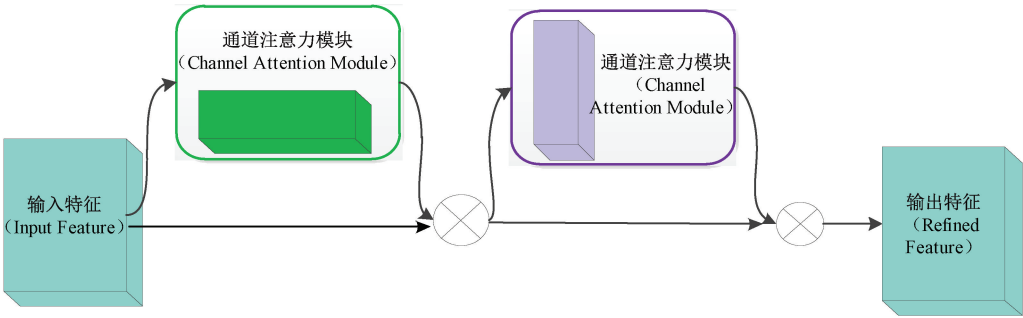


图 8 CBAM 模块
Fig. 8 CBAM module

CAM 模块关注输入图片中有意义的信息。如图 9 所示为 CAM 模块的结构图。将输入的特征图经过两个并行的平均池化层和最大池化层,将特征图从 $C \times H \times W$ 变为 $C \times 1 \times 1$ 的大小,然后经过共享的多层感知模块,在该模块中,它先将通道数压缩为原来的 $1/r$,再扩张到原通道数,经过 ReLU 激活函数得到两个激活后的结果。

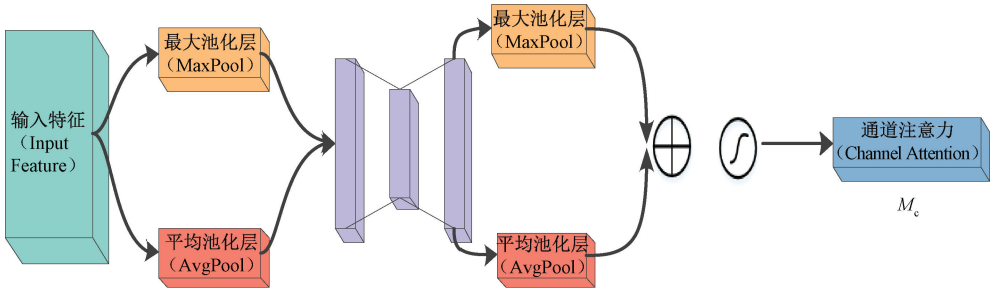


图 9 CAM 模块
Fig. 9 CAM module

SAM 模块关注的是目标的位置信息,将通道注意力模块的输出结果通过平最大池化层和均池化层,得到两个 $1 \times H \times W$ 的特征图,然后经过 Concat 操作对两个特征图进行拼接,通过 7×7 卷积变为 1 通道的特征图,再经过一个 sigmoid 得到空间注意力模块的特征图,最后将输出

将这两个输出结果进行逐元素相加,再通过一个 sigmoid 激活函数得到通道注意力模块的输出结果,再将这个输出结果乘原图,变回 $C \times H \times W$ 的大小。CAM 公式如式 (1) 所示。

$$M_c(F) = \sigma \left(\text{MLP} \left(\text{AvgPool} (F) \right) + \text{MLP} \left(\text{MaxPool} (F) \right) \right) = \sigma \left(W_1 \left(W_0 (F_{\text{avg}}^c) \right) + W_1 \left(W_0 (F_{\text{max}}^c) \right) \right) \quad (1)$$

结果乘原图变回 $C \times H \times W$ 大小,如图 10 为 SAM 模块的结构图。SAM 公式如式 (2) 所示。

$$M_s(F) = \sigma \left(f^{7 \times 7} \left(\left[\text{AvgPool} (F) ; \text{MaxPool} (F) \right] \right) \right) = \sigma \left(f^{7 \times 7} \left(\left[F_{\text{avg}}^s ; F_{\text{max}}^s \right] \right) \right) \quad (2)$$

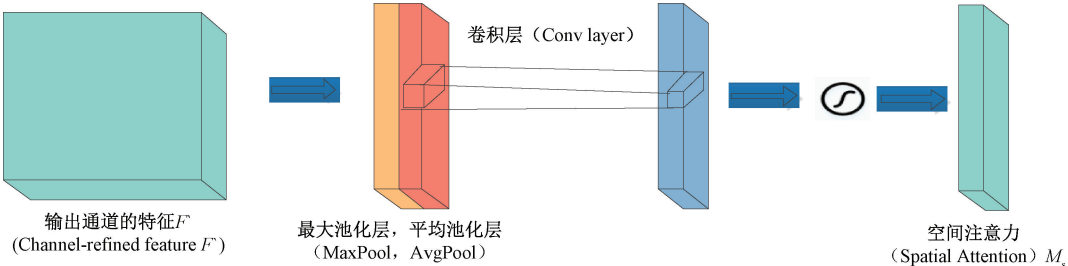


图 10 SAM 模块
Fig. 10 SAM module

本文将 CBAM 嵌入到神经网络 C3 通道中,构建了 CBAM-C3 模块,组合的 CBAM-C3 模块增强了网络对金属齿轮缺陷小目标的关注度。CBAM-C3 模块如图 11 所示。C3 模块中最后一个标准卷积模块 Conv 被 CBAM 模块取代,与直接在网络中插入 CBAM 相比,它减少了一个卷积块的计算。CBAM 模块的输入来自 C3 模块中卷积操作和瓶颈(BottleNeck)操作后的 Concat 的特征图。首先,使用 CAM 子模块将输入特征处理成基于通道的 M_c 特征图,然后将 M_c 与输入特征相乘,得到 SAM 子模块的输入特征。接下来,通过 SAM 操作将输入特征修改为基

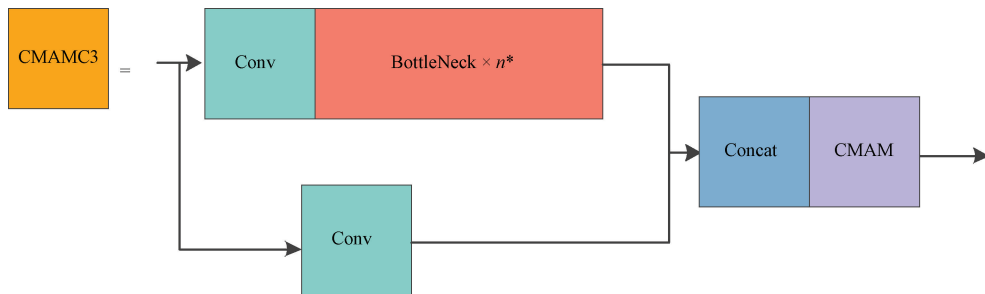


图 11 CBAM-C3 模块

Fig. 11 CBAM-C3 module

2.2 BiFPN 模块

YOLOv5s 中使用了 PANet^[20] 网络结构,由于 PANet 会导致模型训练时间较长、学习率难以调节、特征重叠、难以处理小目标和不适用于单一物体检测等问题。本文提出了将 BiFPN^[21] 网络结构用于齿轮端面缺陷检测。BiFPN 网络结构相比于 PANet 网络结构,可以在不损失速度的情况下提高模型的精度,BiFPN 结构将不同层级的特征图融合在一起,使用多个融合模块进行迭代融合,有效地提取更丰富的特征信息,提高了模型的图像特征表示能力,降低了漏检率。

随着网络层数的加深,提取到的特征信息可能逐渐丢失,因此,为了提高网络模型的检测性能,通常在目标检测网络中广泛使用多尺度特征融合方法。多尺度特征融合通过将来自不同层级的特征图融合在一起从而提升目标检测模型的表现。这种方法有助于减少信息丢失,并提高模型的泛化能力,使其能够在多种情况下更好地进行目标检测。Google 团队 2021 年提出的 BiFPN 是一种加权的双向特征金字塔网络,它能够从两个方向(上下和左右)提取特征,使得 BiFPN 结构能够更有效地提取多尺度特征,并且比 PANet 结构更简单,另外,BiFPN 结构还具有较强的稳定性和适应性,它可以适应多种不同的输入图像,并且能够稳定的处理图像中的目标。这使得 BiFPN 结构在训练过程中更加稳定,并且在实际应用中也能够取得更好的效果。

本文提出的 BiFPN 具体结构如图 12 所示,图 12 中

于空间的特征 M_s , M_s 和输入特征相乘,得到输出特征。这样,CBAMC3 的输出与 C3 的输出在大小上是一样的,但对权重参数的修改可以使网络更好地聚焦于金属齿轮小目标缺陷的部分。此外,骨干网络中的特征被引入颈部网络融合层,从而增加梯度反向传播,避免梯度衰减。由于齿轮端面的缺陷尺寸在毫米级,骨干网络中的 80×80 pixels 和 20×20 pixels 的特征图被引入到颈部网络中,这增加了计算量,但对网络的精度没有很大的提高。因此,本文只选择 40×40 pixels 的特征图,然后引入颈部网络。

CS 为 CSPDarkNet 模块,BiFPN 网络中有自顶向下的通路,传递的是高层特征的语义信息;自底向上的通路,传递的是低层特征的位置信息;同一层在输入节点和输出节点间新加的一条通路。在 BiFPN 网络中,首先通过横向连接和下采样层,实现第 1 次特征融合,BiFPN 网络通过对主 CSPDarkNet 层提取的不同尺度的特征图进行下采样,并通过横向连接将它们融合在一起。然后通过跳跃连接,实现与同一尺度的下采样层和上采样层的第 2 次特征融合。跳跃连接允许 BiFPN 网络在不同尺度的特征图之间跳跃,使得它能够更好地融合不同尺度的信息。最后通过多尺度特征融合得到的特征图输入到最终的检测器中,完成目标检测任务。BiFPN 网络通过多次特征融合,能够更好地捕获目标的特征,从而提高目标检测的准确性。

3 试验结果分析

3.1 试验平台

本试验在 Ubuntu20.0 操作系统下,运用了 Pytorch1.7.1 框架,OpenCV2 和 Pycharm 代码集成开发环境,使用 CUDA11.0 加速 GPU;工作站的硬件是英特尔 3.10 GHz 64 核 CPU,256 GB 内存,两个 NVIDIA Titan XP GPUs。

经过反复训练调参得到最终超参数设置为:初始学习率为 0.001;每次迭代逐次衰减,衰减率为 0.95;优化

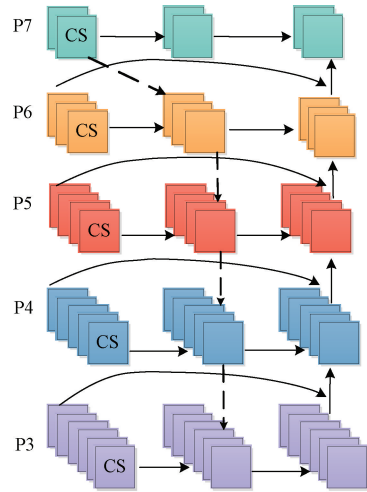


图 12 BiFPN 结构
Fig. 12 BiFPN structure

器为 Adam,指数衰减率为 0.9。网络总共训练 200 次迭代,总共分为两个阶段,先进行 50 轮冻结训练,batch size 设置为 8,然后进行 100 轮的解冻训练,batch size 设置为 4。

3.2 评价指标

本文采用一些评价指标对训练出来的模型进行评价。准确率 (precision) 即模型预测出的正确目标框占比,精度越高说明模型预测出的目标框更准确,能更好的检测到实际目标,如式 (3) 所示;召回率 (Recall) 模型能检测到的所有目标框占有所有预测的目标框的比例,召回率越高说明模型能检测到的目标框越多,能够更全面地覆盖所有的目标,如式 (4) 所示;平均精度 (average precision, AP) 是不同召回率下的准确率的平均值,如式 (5) 所示;mAP 是是一种综合性能指标,衡量了模型在所有类别中的精确率和召回率的平均值,mAP 值越高说明模型在多种精度阈值下的精度都较高,能够在多种情况下表现良好,如式 (6) 所示;其次参数量用来表征模型的大小,F1 值是准确率与召回率的综合值。

$$P = \frac{TP}{TP + FP} \tag{3}$$

$$R = \frac{TP}{TP + FN} \tag{4}$$

式中:TP 为真阳性数量;FP 为假阳性数量;FN 为假阴性数量;TN 为真阴性数量。

$$AP = \int_0^1 P(R) dR \tag{5}$$

$$mAP = \frac{\sum_{i=0}^n AP(i)}{n} \tag{6}$$

3.3 试验结果分析

为了验证本文提出的 YOLO-Gear 网络模型在齿轮

端面缺陷检测的有效性,本文对基础网络 YOLOv5s 进行了消融试验。具体试验内容与结果如表 2 所示。

表 2 消融试验对比结果

Table 2 Comparative results of ablation tests				
方法	第 1 组	第 2 组	第 3 组	第 4 组
YOLOv5s	√	√	√	√
CBAM-C3		√		√
BiFPN			√	√
划痕/%	93.6	95.5	97.7	99.8
断齿/%	91.5	96.3	96.6	99.4
mAP/%	92.8	95.9	96.3	99.2
F1 值	0.95	0.88	0.97	0.99

通过分析表 1 的每项改进策略对本文网络的贡献发现,每个模块对模型的整体提升都有不同程度的提升。

试验以 YOLOv5s 模型为基线,表 1 中第 1 组和第 2 组试验对比显示,在 YOLOv5s 网络加入 CBAM-C3 注意力机制之后,mAP 从 92.8%提高到了 95.9%。F1 的值从 0.95 降到了 0.88。模型中引入 CBAM-C3 注意力机制对精度的提高有了明显的进步,但是检测指标 F1 下降,分析可能是因为 CBAM 的空间注意力模块会捕捉不同尺度的特征,但如果网络的设计不适合处理小目标,这种尺度敏感性可能会导致这种问题。第 1 组和第 3 组试验对比显示,在 YOLOv5s 网络中的特征金字塔换成 BiFPN 之后,mAP 从 92.8%提高到了 96.3%,F1 的值从 0.95 提高到了 0.97,精度得以提升主要原因通过分析得出,在齿轮端面检测中,由于缺陷大小不一,需要支持多尺度输入,而 BiFPN 可以自适应地对不同尺度的特征进行融合,以获得更加精细的特征表达,可以更好的捕捉缺陷的细节特征,提高模型对缺陷的检测精度,同时降低误检率。第 1 组和第 4 组试验对比显示,在网络中同时引入 CBAM 模块和 BiFPN,mAP 从 92.8%提高到了 99.20%,F1 指标从 0.95 提高到了 0.99,在网络中同时引入 CBAM 和 BiFPN 不但提高了检测准确度,还弥补了 CBAM 导致的 F1 下降的问题。在 CBAM 注意力机制之后引入 BiFPN,不但可以在处理小目标时更好地捕捉空间和语义信息,减少信息丢失,还提高小目标的检测性能。通过结合 CBAM 和 BiFPN,可以更全面地利用特征的上下文信息和通道关系,提高 F1 指标。二者结合不但可以确保不同层级的特征得到有效融合和交互,还增强小目标的表征能力,同时可以加快模型的收敛速度。如图 13 所示,为 YOLOv5s 模型改进前后收敛曲线变化对比图。结果表明,在网络中同时引入本文提出的改进点,不但提高了算法模型的检测准确率,还可以加快模型的收敛速度。

通过消融试验发现,加入每个改进模块后的模型,都显著的提高了齿轮端面缺陷的检测准确度,本文在缺陷

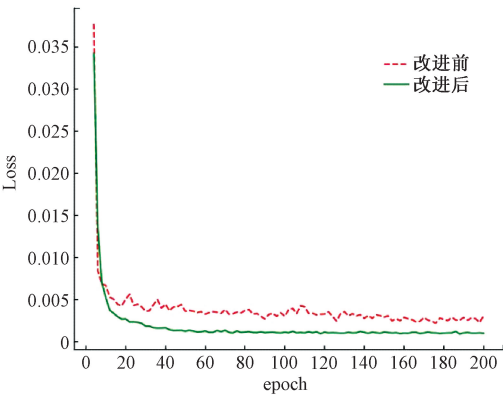


图 13 损失函数变化曲线
Fig. 13 Plot of change in loss function

检测中遇到的检测准确率低,小目标检测难和缺陷尺度变化大的问题得到了有效的解决,相比原始的 YOLOv5s 模型,最终平均精度提升了 6.4%,验证了本文提出 YOLO-Gear 模型在齿轮缺陷检测的可靠性。

PR 曲线的形状可以反映分类器的性能,如果分类器的精度和召回率都很高,则 PR 曲线会接近右上角,即曲线与 x 轴之间的面积越大,表示性能越好。在图 14 的 3 种模型 PR 曲线中,对比发现可知,改进之后模型 YOLO-Gear 是试验中的最优模型。

3.4 对比试验

为了证明 YOLO-Gear 模型在齿轮端面缺陷检测算法的优势,将其与 Faster-RCNN^[22], CenterNet^[23],

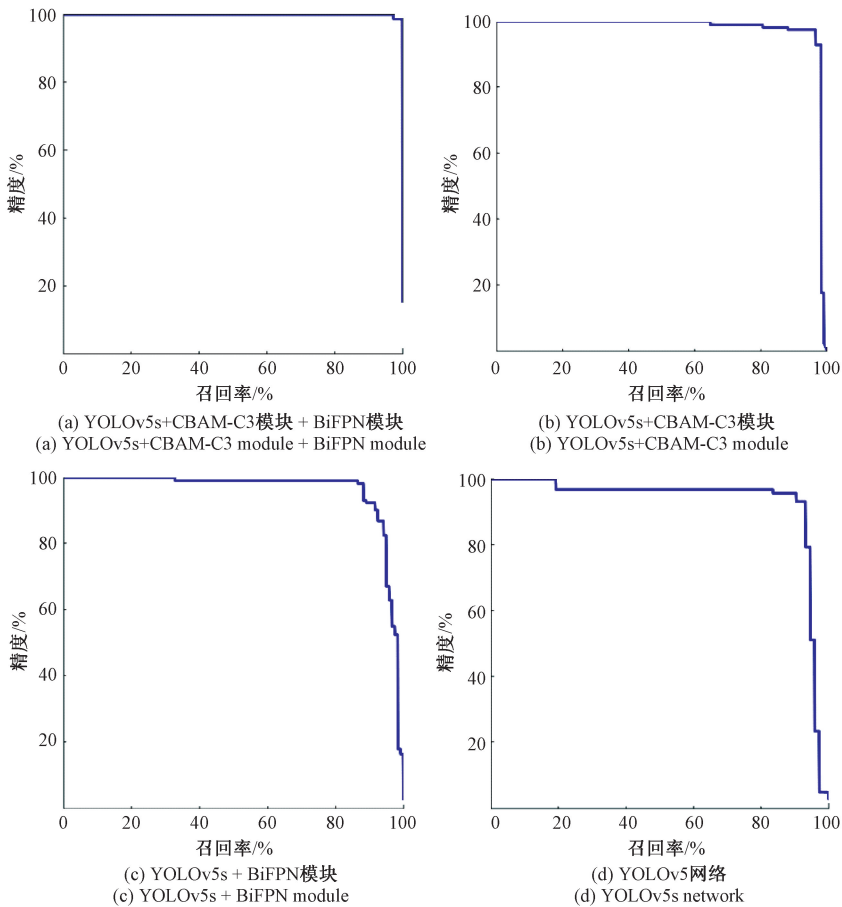


图 14 3 种模型的 PR 曲线
Fig. 14 PR curves for the three models

YOLOv4^[24], SSD^[25] 和原始的 YOLOv5 模型在制作的缺陷数据集进行测试,图 15 为 5 种模型评价指标的散点图,并采用 mAP、F1 值、召回率和 FPS 4 项指标进行评估,试验结果如表 3 所示。表 3 为 5 种试验模型的结果对比, YOLO-Gear 模型在所有的指标中取得了 4 个最优的结

果,图 16 为 YOLO-Gear 模型的检测结果图,图 16(a) 为预测结果的真实标签,图 16(b) 为预测标签的结果。YOLO-Gear 相比于 Faster-RCNN, CenterNet, YOLOv4, YOLOv5s 的 mAP 分别提高了 5.88%、7.74%、14.4%、6.40%。

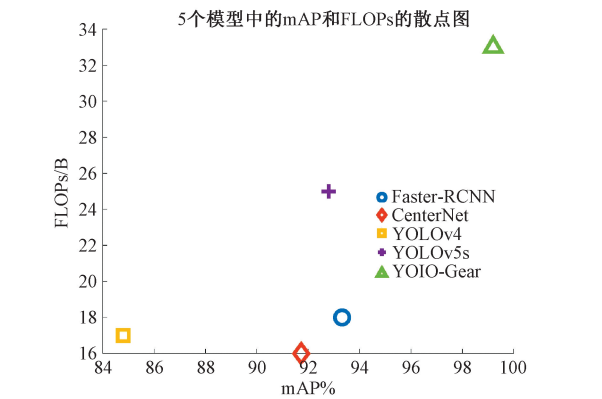
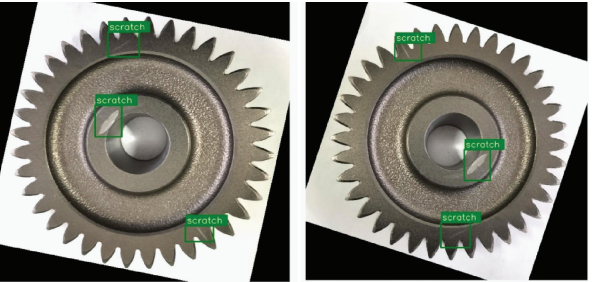


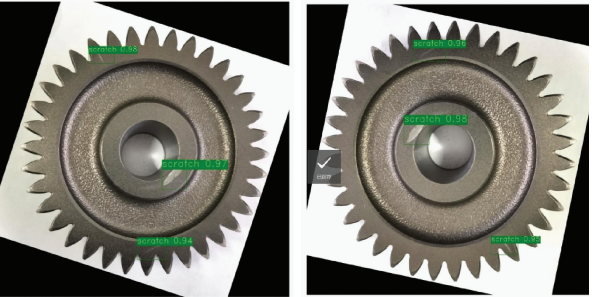
图 15 图像检测效果
Fig. 15 Image detection effect

表 3 5 种模型试验结果对比

模型	mAP/%	FPS	AP/%	
			断齿	划痕
Faster-RCNN	93.32	18	92.6	94.3
CenterNet	91.73	16	87.5	85.2
YOLOv4	84.80	17	76.4	89.2
YOLOv5s	92.80	25	95.4	89.4
YOLO-Gear	99.20	33	99.8	99.4



(a) 预测结果的真实标签
(a) The real label of the predicted result



(b) 预测标签的结果
(b) Predict the result of the label

图 16 YOLO-Gear 模型图像检测结果

Fig. 16 YOLO-Gear model image detection result map

经过改进的 YOLO-Gear 模型实现了 33 fps 的处理速度,平均以 0.13 s 的时间对一张齿轮图像进行检测。齿

轮在传送带上完成端面翻转的时间实际测得为 30 s。因此,平均每个完整齿轮的检测时间为 30.26 s。这种检测速度能够满足企业对齿轮端面缺陷进行实时在线检测的需求,并且能够确保齿轮生产的节拍不低于 1 min/件。

4 结 论

本文针对目前金属齿轮端面结构复杂,导致缺陷的小目标占比度高和尺度变化大引起的检测准确度低,难以满足企业实时在线检测需求等问题,提出了一种基于自适应多尺度特征融合的金属齿轮端面缺陷检测方法。在 C3 模块中融入了自适应注意力模块 CBAM-C3, CBAM-C3 通过将通道注意力 CAM 和空间注意力 SAM 相结合加强了对金属齿轮缺陷小目标缺陷自适应的特征学习与特征提取,及时对模型中的权重参数进行学习和优化;提出了重复加权双向特征金字塔网络 BiFPN,通过自适应控制不同尺度的特征图之间的融合程度,提高了模型对缺陷多尺度检测能力。试验结果表明,在自制的金属齿轮端面缺陷数据集上,改进的 YOLO-Gear 模型相比原始的 YOLOv5s 模型,准确度提高了 6.4%,召回率提高了 0.4%,平均每个完整齿轮的检测时间为 30.26 s,满足企业实时在线检测不低于 1 min/件的要求。

下一步将针对齿轮缺陷样本稀缺的问题,尝试使用对抗半监督和迁移学习与小样本结合的方法,以实现在数据集较少的情况下,也能达到较优的检测效果。

参考文献

[1] KUMAR A, GANDHI C, ZHOU Y Q, et al. Latest developments in gear defect diagnosis and prognosis[J]. Measurement, 2020, 158:107735.

[2] GADEMAWLA E. Development of computer vision algorithms for measurement and inspection of spur gears [J]. Mansoura Engineering, 2020, 7(23): 27-38.

[3] GUERRINI G, FAVIAL L, ALESSANDRO F, et al. The effect of radial infeed on surface integrity in dry generating gear grinding for industrial production of automotive transmission gears [J]. Journal of Manufacturing Processes, 2019, 45: 234-241.

[4] ASHANTA R, HINNAMI S, ESTRADA F, et al. Frequency-tuned salient region detection. CVPR 2009 IEEE Conf [J]. Compute, 2009, 7(20):1597-1604.

[5] KUMAR A, GANDHI C P, ZHOU Y, et al. Latest developments in gear defect diagnosis and prognosis [J]. Measurement, 2020, 158: 107735.

[6] 吕禾丰,陆华才. 基于 YOLOv5 算法的交通标志识别技术研究[J]. 电子测量与仪器学报, 2021, 35 (10): 137-144.

LYU H F, LU H C. Research on traffic sign recognition

- technology based on YOLOv5 algorithm [J]. Journal of Electronic Measurement and Instrumentation, 2021, 35(10):137-144.
- [7] 包从望,胡才梦,张彩红,等. 基于改进 Canny 算子的齿轮缺陷检测边缘检测算法[J]. 组合机床与自动化加工技术,2023(1):83-86,91.
BAO C W, HU C M, ZHANG C H, et al. Edge detection algorithm for gear defect detection based on improved canny operator [J]. Combined Machine Tools and Automatic Machining Technology, 2023 (1): 83-86,91.
- [8] VIVEKANANTHAN V, VIGNESH R, VASANTHASEELAN S. Concrete bridge crack detection by image processing technique using the improved OTSU method [J]. Materials Today: Proceedings, 2023, 74 (P4):1002-1007.
- [9] PSAROMMATIS F, MAY G, DREYFUS P A, et al. Zero-defect manufacturing: State-of the-art review, shortcomings and future directions in research [J]. ProdRes,2020,58:1-17.
- [10] 赵佩瑶. 板式换热器板片图像 ROI 区域定位及缺陷检测方法研究[D]. 沈阳:沈阳工业大学,2021.
ZHAO P Y. Research on ROI region localization and defect detection method for plate heat exchanger plate images [D]. Shenyang: Shenyang University of Technology,2021.
- [11] BHARATI P, PRAMANIK A. Deep learning technique R-CNN to mask R-CNN: A survey [J]. Computational Intelligence in Pattern Recognition: Proceedings of CIPR 2019, 2020: 657-668.
- [12] 王宸,唐禹,张秀峰,等. 基于改进 EfficientNet 的锻件磁粉探伤智能检测方法研究[J]. 仪器仪表学报, 2021,42(9):89-96.
WANG CH, TANG Y, ZHANG X F, et al. Research on the intelligent detection method of magnetic particle flaw detection for forgings based on improved EfficientNet[J]. Chinese Journal of Scientific Instrument, 2021, 42(9): 89-96.
- [13] SINGH A, KALAICHELV V, KARTHIKEYAN R. Performance analysis of object detection algorithms for robotic welding applications in planar environment[J]. International Journal of Computer Integrated Manufacturing, 2023, 36:1083-1108.
- [14] 向宽,李松松,栾明慧,等. 基于改进 Faster RCNN 的铝材表面缺陷检测方法[J]. 仪器仪表学报,2021, 42(1):191-198.
XIANG K, LI S S, LUAN M H, et al. An improved Faster RCNN-based method for detecting surface defects in aluminum [J]. Chinese Journal of Scientific Instrument, 2021, 42(1):191-198.
- [15] BAKIR H, KUZHIPALLIL F A, MERABET A. Automatic detection of deteriorated photovoltaic modules using IRT images and deep learning (CNN, LSTM) strategies[J]. Engineering Failure Analysis, 2023, 146: 107132.
- [16] 李勇,王杨,方夏,等. 基于改进 BiFPN 的微特电机电机枢表面缺陷检测方法[J]. 机床与液压,2022,50(6): 1-8.
LI Y, WANG Y, FANG X, et al. An improved BiFPN-based method for detecting surface defects in microturbine armatures [J]. Machine Tools and Hydraulics, 2022, 50(6):1-8.
- [17] APPE S N, ARULSEVI G, BALAJI G N. CAM-YOLO: Tomato detection and classification based on improved YOLOv5 using combining attention mechanism[J]. PeerJ Computer Science, 2023, 9: e1463.
- [18] 肖鹏程,徐文广,张妍,等. 基于 SE 注意力机制的废钢分类评级方法[J]. 工程科学学报, 2023, 45(8): 1342-1352.
XIAO P CH, XU W G, ZHANG Y, et al. A classification and rating method for steel scrap based on SE attention mechanism [J]. Journal of Engineering Science,2023,45(8):1342-1352.
- [19] WOO S, PARK J, LEE J Y. CBAM:Convolutional block attention module[C]. European Conference on Computer Vision,2018.
- [20] CHEN J, MAI H S, LUO L, et al. Effective feature fusion network in BIFPN for small object detection[C]. 2021 IEEE International Conference on Image Processing (ICIP). IEEE, 2021:699-703.
- [21] LI X, WANG W, WU L, et al. Generalized focal loss: Learning qualified and distributed bounding boxes for dense object detection [J]. Advances in Neural Information Processing Systems, 2020, 33: 21002-21012.
- [22] 伊欣同,单亚峰. 基于改进 Faster R-CNN 的光伏电池内部缺陷检测[J]. 电子测量与仪器学报,2021, 35(1):40-47.
YI X T, SHAN Y F. Internal defect detection of photovoltaic cells based on improved Faster R-CNN[J]. Journal of Electronic Measurement and Instrumentation, 2021,35(1):40-47.
- [23] 万伟彤,李长峰,朱华波,等. 轻量化 CenterNet 网络的二维码定位算法[J]. 电子测量与仪器学报,2022, 36(5):128-135.
WAN W T, LI CH F, ZHU H B, et al. Lightweight

CenterNet network for 2D barcode localization algorithm[J]. Journal of Electronic Measurement and Instrumentation, 2022, 36(5):128-135.

[24] 王宸,张秀峰,刘超,等. 改进 YOLOv3 的轮毂焊缝缺陷检测[J]. 光学 精密工程,2021,29(8):1942-1954.
WANG CH,ZHANG X F, LIU CH, et al. Improved wheel hub weld defect detection of YOLOv3 [J]. Optics and Precision Engineering, 2021, 29(8):1942-1954.

[25] 孙备,左震,吴鹏,等. 面向无人艇环境感知的改进型 SSD 目标检测方法[J]. 仪器仪表学报,2021,42(9):52-61.
SUN B, ZUO ZH, WU P, et al. Improved SSD target detection method for unmanned vehicle environment perception [J]. Chinese Journal of Scientific Instrument, 2021, 42(9):52-61.

作者简介



王宸(通信作者),2009 年于武汉科技大学获得硕士学位,2019 年于上海大学获得博士学位,现为湖北汽车工业学院副教授,主要研究方向为智能制造及机器视觉。
E-mail:20090011@ huat. edu. cn

Wang Chen (Corresponding author) received his M. Sc. degree form Wuhan University of Science and Technology in 2009 and received his Ph. D. degree from Shanghai University in 2019. He is currently an associate professor at Hubei University of Automotive Technology. His main research interests include intelligent manufacturing and machine vision research.



杨帅,2022 年于南阳师范学院获得学士学位,主要研究方向为机器视觉研究。
E-mail:793132180@ qq. com

research.



周林,2023 年于湖北汽车工业学院获得学士学位,主要研究方向为机器视觉研究。
E-mail: 2294829984@ qq. com

Zhou Lin received his B. Sc. degree from Hubei University of Automotive Technology in 2023. His main research interest includes machine vision research.