

DOI:10.19651/j.cnki.emt.2417525

改进 YOLOv5s 的轴承座缺陷检测算法^{*}

王浪¹ 胥云¹ 李琦² 高亮¹ 张佳骏¹

(1. 四川轻化工大学机械工程学院 宜宾 644000; 2. 四川长征机床集团有限公司 自贡 643000)

摘要: 目前数控机床轴承座缺陷检测主要依赖人工目检,无法满足工业生产高精度、高效率 and 低错误率的要求。针对以上问题,提出基于改进 YOLOv5s 的数控机床轴承座缺陷检测算法。首先以 HardSwish 激活函数替换 ConvNeXt2 中的 GELU,并结合 CSC 模块提出全新 CSCConvNeXt2-HS 结构,用以替换 backbone 网络 C3 模块,在降低计算复杂度的同时提升关键信息的特征提取能力;在 Neck 网络中引入尺度序列特征融合模块,提升模型对多通道信息的提取能力;最后采用 Focal-Inner Loss 损失函数,在提高训练收敛速度的同时,降低了类别分布不平衡带来的影响。实验表明,改进模型的准确率为 91.09%,召回率为 81.97%,平均精度均值为 84.40%,处理速度为 61.73 fps,各项评估指标较原始模型 YOLOv5s 分别提升 2.52%、4.47%、6.7% 和 1.12 fps,能满足工业生产需求。

关键词: YOLOv5s; HardSwish 激活函数; CSCConvNeXt2-HS; 尺度序列特征融合; Focal-Inner Loss 损失函数
中图分类号: TN919.2 **文献标识码:** A **国家标准学科分类代码:** 520.2099

Improved YOLOv5s algorithm for bearing seat defect detection

Wang Lang¹ Xu Yun¹ Li Qi² Gao Liang¹ Zhang Jiajun¹(1. School of Mechanical Engineering, Sichuan University of Science & Engineering, Yibin 644000, China;
2. Sichuan Changzheng Machine Tool Group Co., Ltd., Zigong 643000, China)

Abstract: Currently, defect detection of CNC machine tool bearing seats primarily relies on manual visual inspection, which cannot meet the demands for high precision, high efficiency, and low error rates in industrial production. To address these issues, a defect detection algorithm based on an improved YOLOv5s is proposed for CNC machine tool bearing seats. Firstly, the HardSwish activation function is used to replace the GELU in ConvNeXt2, and a novel CSCConvNeXt2-HS structure is introduced, incorporating the CSC module to replace the C3 module in the backbone network. This modification reduces computational complexity while enhancing the feature extraction capability of key information. In the Neck network, a Scale Sequence Feature Fusion module is introduced to improve the model's ability to extract multi-channel information. Finally, The final proposal of the Focal-Inner Loss loss function not only improves the speed of training convergence but also reduces the impact brought about by the imbalance in class distribution. Experimental results show that the improved model achieves an accuracy of 91.09%, a recall rate of 81.97%, and a mean Average Precision of 84.40%, with a processing speed of 61.73 fps. All evaluation metrics show improvements of 2.52%, 4.47%, 6.7%, and 1.12 fps compared to the original YOLOv5s model, thereby satisfying industrial production requirements.

Keywords: YOLOv5s; HardSwish activation function; CSCConvNeXt2-HS; scale sequence feature fusion; Focal-Inner Loss function

0 引言

随着现代制造业的快速发展,数控机床作为精密加工的核心设备,其性能的稳定性和可靠性对生产效率和产品质量都有着至关重要的影响。其中数控机床轴承座零件是

连接轴和机床的主要部件之一,它装上轴承后是机床进给系统的支撑与承载系统,降低运动部件的磨损量,提高机床的稳定性和工作精度。但在生产过程中,由于材料、测量技术和工艺等原因,导致生产的轴承座零件存在缺陷,致使进给系统精度超差,甚至还会导致轴承或丝杆损坏。因此,为

收稿日期:2024-12-03

^{*} 基金项目:自贡市科技局校地合作项目(2022CDZG-19)资助

解决不良轴承座造成的影响,实现对轴承座缺陷的快速、准确检测,对于保障数控机床的正常运行和提高工作效率具有重要意义。

提到传统的数控机床轴承座零件缺陷检测主要依赖于人工目检,这种方法不仅效率低下,而且受限于检测人员的经验和技能,难以保证检测的一致性和准确性。文献[1]提出随着计算机视觉、机器视觉和深度学习技术的发展,自动化缺陷检测方法逐渐成为研究热点。然而,现有的自动化检测方法在处理复杂工业场景时,仍面临着检测速度慢、漏检误检率高等问题。

梁海波等^[2]提出天然气钢制管道的缺陷检测,提出优化的最小二乘支持向量机(least square support vector machines, LSSVM),可有效识别天然气钢制管道腐蚀缺陷。符前坤等^[3]提出对聚乙烯管道(polyethylene pipe, PE)进行缺陷检测时通过图像预处理后在YOLOv5模型中引入坐标注意力机制(coordinate attention, CA),提升目标检测的特征提取能力。王龙达等^[4]提出通过灰度拼接构建缺陷数据集后,基于SqueezeNet神经网络对医用透明方杯进行缺陷检测。严舒等^[5]提出一种基于特征增强和多尺度融合的印刷电路板(printed circuit board, PCB)缺陷检测方法FCM-YOLO,可增强对小目标和背景复杂情况的特征提取能力。文献[6]提出一种新型部分卷积PConv应用于FasterNet模型,在同时减少冗余计算和内存访问来更有效的提取空间特征。张波等^[7]提出金字塔池化模块,将空洞卷积加入到YOLOv5s网络的快速空间金字塔池化(spatial pyramid pooling-fast, SPPF)模块,有效提高了无人机航拍图像对小目标的感知能力。董耿耿等^[8]提出新梅目标检测模型SFFYOLOv5s,解决新梅在树干树叶遮挡、果实重叠情况下难以准确检测的问题。

目前,单阶段目标检测(you only look once, YOLO)系列算法以高效的目标检测性能在工业视觉检测领域展现出巨大的潜力,但其针对数控机床轴承座缺陷检测这一特定任务,面对轴承座表面存在铁屑、小目标和遮挡目标以及复杂光照等恶劣的工作环境时,仍存在一定局限性,例如目标特征提取能力不足,对背景复杂的鲁棒性有待提高等。

本文针对轴承座缺陷检测的复杂工作环境以及特征提取能力弱等问题提出一种基于改进YOLOv5s的数控机床轴承座缺陷检测算法。

1 改进的YOLOv5s算法

1.1 改进算法的整体框架

本文是在YOLOv5s算法的基础上进行改进,算法整体架构如图1所示。在backbone网络上,采用全新的CSCConvNeXtv2-HS结构替换C3结构,在减少模型参数的同时提高模型特征提取能力;在Neck网络上,采用尺度序列特征融合(scale sequence feature fusion, SSFF)模块来增强网络的多尺度信息提取能力;最后引入Focal-Inner

Loss损失函数替换完整交并比(complete intersection over union, CIoU)损失函数,提高对小目标检测精度,简化计算过程并且使模型在样本类别分布不平衡时亦能显示出更好的性能。

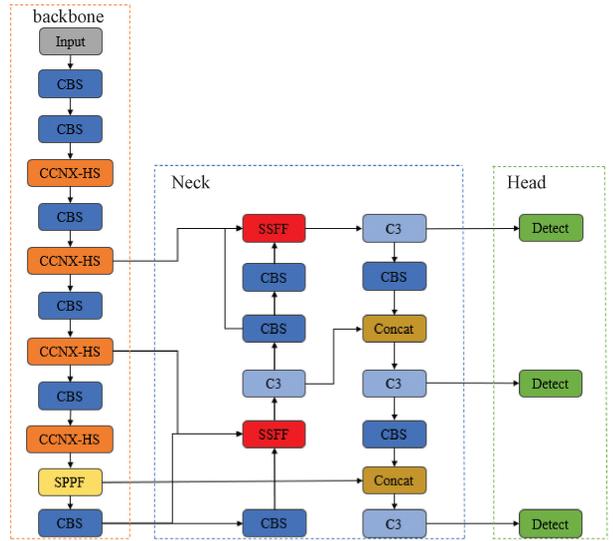


图1 改进模型的整体架构

Fig. 1 Improving the overall framework of the model

1.2 CSCConvNeXtv2-HS模块

文献[9]提出ConvNeXt是一种融合了卷积神经网络(convolutional neural network, CNN)与递归神经网络(recurrent neural network, RNN)的混合模型,利用CNN提取局部特征和RNN捕捉长程依赖关系的优势,实现更加有效的文本分类,而且相较于YOLOv5s中的C3结构更加轻量,但ConvNeXt的特征提取不够精细,并且由于没有引入注意力机制,导致模型难以准确聚焦重要信息部分。因此Woo等^[10]提出将一个全卷积掩码自编码器框架和一个新的全局响应归一化(global response normalization, GRN)层添加到ConvNeXt架构中,首次提出ConvNeXtv2模型。相比ConvNeXt模型,文献[11-12]提出的ConvNeXtv2模型采用了深度可分离卷积和组卷积,减少了参数和计算量,并且ConvNeXtv2引入了自注意力机制,使得模型在长距离依赖关系的捕捉上更灵活且在多模态特征融合和正则化方法等方面也有显著的提升和优化。激活函数则采用高斯误差线性单元(Gaussian error linear unit, GELU),虽然减少了梯度消失的现象,但是GELU激活函数计算复杂度较高,在计算资源有限的情况下,本文以文献[13]提出的HardSwish激活函数替换ConvNeXtv2中GELU激活函数,其中HardSwish激活函数定义为:

$$\text{HardSwish}(x) = \begin{cases} 0, & x \leq -3 \\ x, & x \geq 3 \\ \frac{x \times \text{ReLU6}(x+3)}{6}, & \text{其他} \end{cases} \quad (1)$$

式中: $ReLU6$ 是修正线性单元 (rectified linear unit, RELU) 函数的一种变体, 将输出限制在 $0 \sim 6$ 之间。

GELU 激活函数的定义为:

$$f(x) = x \times \phi(x) \quad (2)$$

式中: $\phi(x)$ 为标准正态分布的累积分布函数。GELU 激活函数的计算涉及到概率密度函数的积分, 在计算效率上远不如 HardSwish, 并且在训练时内存占用额外多出 52%; 文献[14]提出 HardSwish 在训练初期损失下降速度可能不如其他激活函数快, 但其后续的波动较小, 稳定性更好。

ConvNeXt 和 ConvNeXtv2-HS 具体结构如图 2 所示。

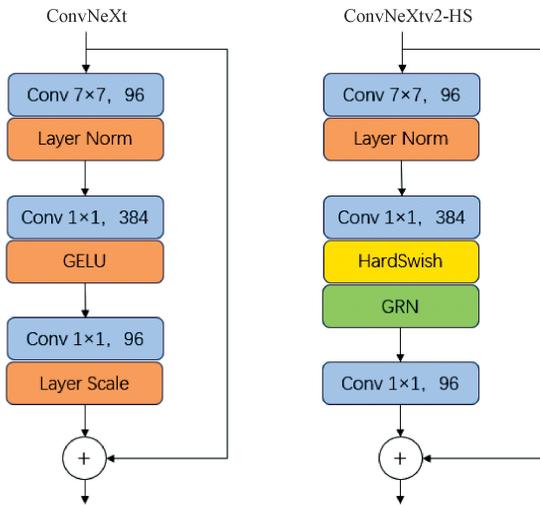


图 2 ConvNeXt 和 ConvNeXtv2-HS 的结构图

Fig. 2 Structure diagram of ConvNeXt and ConvNeXtv2-HS

通道分割卷积 (channel split convolution, CSC) 结构是一种卷积操作的变体, 易于拓展, 其通过分离通道来进行卷积运算, 文献[15-16]提出 CSC 结构可显著减少计算成本并减少模型对某些特征噪音的敏感度。

本文将 ConvNeXtv2-HS 模型与 CSC 结构结合, 提出全新的 CSCConvNeXtv2-HS 模型, 模型结构如图 3 所示。

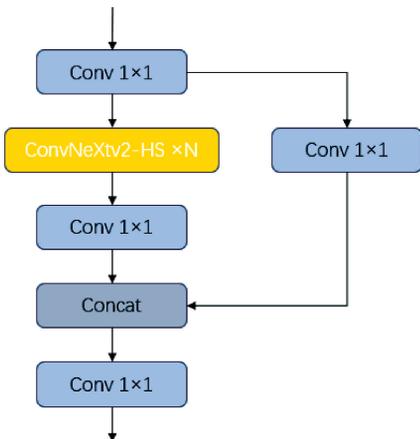


图 3 CSCConvNeXtv2-HS 模型结构图

Fig. 3 CSCConvNeXtv2-HS model structure diagram

1.3 SSFF 模块

Kang 等^[17] 提出一种基于注意力尺度序列融合的框架 ASF-YOLO, 其采用 SSFF 模块来增强提取多尺度信息的能力。本文针对数控机床轴承座多尺度问题, 将 YOLOv5s 模型与 SSFF 模块相结合。该模块相较于仅采用求和或 Concat 来融合金字塔特征的结构, 能够更好的将深层特征图的高维信息与浅层特征图的详细信息相结合, 其中图像大小在降采样过程中发生变化, 但尺度不变的特征不会发生变化。其结构原理如图 4 所示。

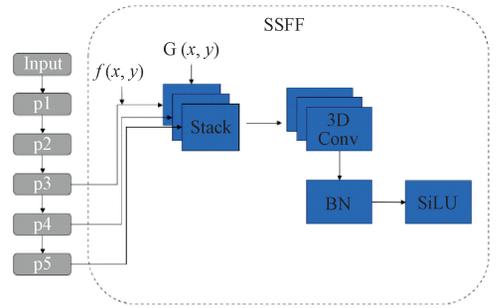


图 4 SSFF 结构原理图

Fig. 4 Schematic diagram of SSFF structure

主干网络的 P3、P4、P5 先与一系列标准差逐渐增加的高斯核进行卷积, 具体如下:

$$F_{\sigma}(i, j) = \sum_u \sum_v f(i-u, j-v) \times G_{\sigma}(u, v) \quad (3)$$

$$G_{\sigma}(x, y) = \frac{1}{2\pi\sigma^2} e^{-\frac{(x^2+y^2)}{2\sigma^2}} \quad (4)$$

其中, f 表示二维特征图, F_{σ} 是通过使用标准差 σ 逐渐增加的二维高斯滤波器进行一系列卷积和平滑处理的特征图。

然后将这些不同尺度的特征图水平堆叠, 并使用 3D 卷积提取它们的尺度序列特征。SSFF 模块 (因为高分辨率的特征图 P3 对小目标的检测和分割至关重要, 因此 SSFF 是基于 P3 设计的) 实现步骤如下:

- 1) 使用 1×1 的卷积将 P4 和 P5 的特征层通道数改为 256。
- 2) 使用最近邻方法调整它们的大小至 P3 级别的尺寸。
- 3) 使用 unsqueeze 方法增加每个特征层的维度, 将每个特征层从三维张量 (高度、宽度、通道) 转变为四维张量 (深度、高度、宽度、通道)。
- 4) 然后沿着深度维度连接 4D 特征图, 形成一个便于后续卷积的 3D 特征图。
- 5) 最后, 使用 3D 卷积、3D 批量标准化和 SiLU 激活函数完成尺度序列特征提取。

1.4 Focal-Inner Loss 损失函数

YOLOv5s 模型采用的是 CIoU 损失函数, 该损失函数主要考虑了边界框的中心点距离和长宽比, 当目标框与预测框的纵横比相同时, CIoU 的性能将大幅降低, 而且对于小目标而言, 可能会造成梯度爆炸。

文献[18]提出 Inner-IoU 损失,即通过辅助边界框计算交并比(intersection over union, IoU)损失。其中真实边界(ground truth, GT)框和锚框分别表示为 b^{gt} 和 b , GT 框和 Inner-GT 框的中心点为 (x_c^{gt}, y_c^{gt}) , 锚框和 Inner-锚框的中心点为 (x_c, y_c) 。GT 框的宽度和高度分别表示为 w^{gt}, h^{gt} , 锚框的宽度和高度分别为 w, h , 缩放因子表示为 $ratio$, 可以控制辅助边界框的尺寸。Inner-IoU 的描述图具体如图 5 所示。

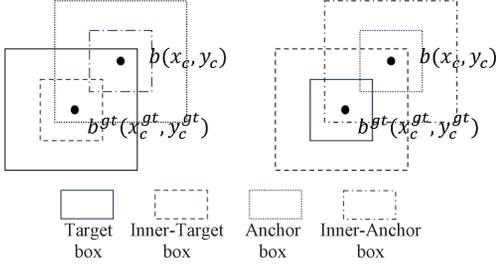


图 5 Inner-IoU 的描述图

Fig. 5 Description of Inner-IoU

公式如下:

$$b_l^{gt} = x_c^{gt} - \frac{w^{gt} \times ratio}{2} \quad (5)$$

$$b_r^{gt} = x_c^{gt} + \frac{w^{gt} \times ratio}{2} \quad (6)$$

$$b_t^{gt} = y_c^{gt} - \frac{h^{gt} \times ratio}{2} \quad (7)$$

$$b_b^{gt} = y_c^{gt} + \frac{h^{gt} \times ratio}{2} \quad (8)$$

$$b_l = x_c - \frac{w \times ratio}{2} \quad (9)$$

$$b_r = x_c + \frac{w \times ratio}{2} \quad (10)$$

$$b_t = x_c - \frac{h \times ratio}{2} \quad (11)$$

$$b_b = x_c + \frac{h \times ratio}{2} \quad (12)$$

$$inter = (\min(b_l^{gt}, b_r) - \max(b_l, b_l^{gt})) \times (\min(b_b^{gt}, b_b) - \max(b_t, b_t^{gt})) \quad (13)$$

$$union = (w^{gt} \times h^{gt})(ratio)^2 + (w + h)(ratio)^2 - inter \quad (14)$$

$$IoU^{inner} = \frac{inter}{union} \quad (15)$$

可推导出:

$$L_{inner} = 1 - \frac{(A \cap B)}{|A| + |B| - (A \cap B)} \quad (16)$$

其中, A 为预测框, B 为真实框; 使用较小尺寸的辅助边界框计算 IoU 损失有助于高 IoU 样本的回归以及加速收敛; 使用较大尺寸的辅助边界框计算 IoU 损失可以加速低 IoU 样本回归, 简化了计算过程。

Focal Loss 旨在解决类别不平衡问题, 尤其是在样本间存在显著差异的情况下。公式如下:

$$FL(p_i) = -\alpha_i(1 - p_i)^\gamma \log(p_i) \quad (17)$$

式中: p_i 是模型对真实类别的预测概率; α_i 是与类别相关的加权因子, 用于平衡类别; γ 是一个调节参数, 用于控制分类样本的相对损失降低程度。

虽然 Focal Loss 可以应对类别不平衡问题, 但是在增强对难分类样本关注后易受噪音影响, 出现过拟合现象。文献[19]提出通过 ℓ_n 范数和基于 IoU 损失在目标检测中的缺陷, 设计一种回归版本的 Focal Loss。本文参考其思想将结合 Focal Loss 损失函数和 Inner-IoU 损失函数提出全新的 Focal-Inner Loss 损失函数, 以替换 YOLOv5s 中的 CIoU 损失函数, 其定义为:

$$L^{Focal-Inner} = IoU^\gamma L_{inner} \quad (18)$$

该损失函数结合两者的优点, 不仅关注预测框和目标框的重叠区域以及内部区域与边界框的关系, 而且在面对类别不平衡问题和难易样本分布不均情况, 显示出更好的性能。

2 数据分析

2.1 实验环境及参数设置

本文实验采用 PyTorch1.12.1 作为深度学习模型框架, python3.8 为编程语言。实验的操作系统为 64 位 Win11, GPU 为 NVIDIA GeForce RTX 4050 16 G 显存, 使用 CUDA11.3 对 GPU 加速。模型训练参数 epoch 设置为 300, batch-size 为 16, imgsz 为 640, workers 为 8, 初始学习率为 0.01, 使用 SGD 优化算法, 后续所有实验训练参数设置皆与此一致。

2.2 数据集预处理

数据集图片皆来源于四川长征机床集团有限公司。在不同环境、不同光照条件下采集图像 250 张, 为避免数据不足产生过拟合现象, 增强模型泛化能力, 进行数据增强处理增加至 700 张, 如: 旋转、亮度调节、添加噪音以及遮挡目标等操作。图片像素 640×640 , 数据集中训练集、验证集和测试集分别有 560 张、70 张、70 张, 比例为 8:1:1, 部分采集图片如图 6 所示。



图 6 数据集部分图像

Fig. 6 Partial images of the dataset

3 实验与结果分析

3.1 评价指标

本文采用常用评价指标评估模型性能: 准确率 (precision, P), 召回率 (recall, R), 平均精度均值 (mean average precision, mAP), 参数量 (parameters, paras), 处理速度 (frames per second, FPS) 和复杂度 (giga floating-point operations per second, GFLOPs)。定义如下:

$$Precision = \frac{TP}{TP + FP} \quad (19)$$

$$Recall = \frac{TP}{TP + FN} \quad (20)$$

$$AP = \int_0^1 P(R) dR \quad (21)$$

$$mAP = \frac{1}{n} \sum_{i=1}^n AP(i) \quad (22)$$

$$FPS = \frac{1\ 000}{pre + inf + NMS} \quad (23)$$

其中, TP 表示模型将实际为正类别的样本正确预测为正类别; FP 表示模型将实际为负类别的样本错误预测为正类别; FN 表示模型将实际为正类别的样本错误预测为负类别; pre 表示图像预处理时间; inf 表示推理速度; NMS 表示后处理时间 (即对模型输出结果进行转换)。由于本文检测类别仅一种, 即 $n = 1$, 因此 $mAP = AP$ 。

3.2 损失函数对比实验

为验证 Focal-Inner Loss 损失函数对模型检测效果的有效性和先进性, 本文将在基础模型 YOLOv5s 上分别引

入最新的 Focal-Inner Loss 和文献[20]提出的 Shape-IoU 以及文献[21]提出的 MPDIoU 损失函数, 替换 CIoU 损失函数进行性能对比, 其中 CIoU 损失函数是 YOLOv5s 的初始损失函数。对比结果如表 1 所示。

表 1 损失函数对比实验

损失函数	P/%	R/%	mAP/%	FPS/fps
CIoU	88.57	77.50	77.70	60.61
MPDIoU	91.18	81.58	81.61	45.87
Shape-IoU	87.38	77.59	78.84	68.49
Focal-Inner Loss	91.09	80.00	81.45	80.65

相比原始损失函数 CIoU, Focal-Inner Loss 和 MPDIoU 在的 P、R、mAP 上均有较大的提升, 其中 MPDIoU 比 Focal-Inner Loss 的提升略高一点, 但 MPDIoU 的 FPS 下降 24.32%, 而 Focal-Inner Loss 的 FPS 提升 33.06%, 因此 Focal-Inner Loss 更加符合工业要求的高时效性。

3.3 消融实验

为了验证本文改进模型各模块的有效性, 在自制数据集上进行消融实验。实验 1 为基础模型 YOLOv5s; 实验 2、3、4 则是在基础模型上分别引入 Focal-Inner Loss 损失函数模块、CSCConvNeXtv2-HS 模块和 SSFF 模块; 实验 5 在 YOLOv5s 上引入 Focal-Inner Loss 损失函数和 CSCConvNeXtv2-HS 结构; 实验 6 则是将 3 个模块都引入基础模型中。具体实验结果如表 2 所示。

表 2 消融实验

Table 2 Ablation experiment

Experiment	Focal-Inner Loss	CSCConvNeXtv2-HS	SSFF	P/%	R/%	mAP/%	Paras	FPS/fps
1	×	×	×	88.57	77.50	77.70	7 022 326	60.61
2	√	×	×	91.09	80.00	81.45	7 022 326	80.65
3	×	√	×	90.29	81.58	82.46	6 911 318	63.69
4	×	×	√	91.92	80.53	81.74	7 350 774	68.49
5	√	√	×	89.72	82.76	83.14	6 911 318	60.61
6	√	√	√	91.09	81.97	84.40	7 239 766	61.73

相比实验 1, 实验 2 在加入 Focal-Inner Loss 损失函数模块后, P、R、mAP 分别上涨了 2.52%、2.50%、3.75%, Paras 不变, 但 FPS 大幅提升了 33.06%, 表明 Focal-Inner Loss 损失函数在提升检测性能的同时也加速了收敛; 实验 3 在 P、R、mAP 上分别上涨了 1.72%、4.08%、4.76%, Paras 降低 1.58%, 表明 CSCConvNeXtv2-HS 模块可有效增强通道间的特征竞争。实验 4 在 P、R、mAP 上分别提升 3.35%、3.03%、4.04%, 虽然 Paras 增加了 4.68%, 但 FPS 却更快, 表明 SSFF 结构有效降低计算复杂度并增强了多尺度特征提取能力。

从实验 6 能够看出, 将 3 个模块都加入到基础模型后, P、R、mAP 分别提升了 2.52%、4.47%、6.70%, Paras 虽有略微增加, 但 FPS 不降反增, 因此将 3 个模块均引入基础模型 YOLOv5s 中时, 模型的识别效果达到最佳, 计算复杂度也更低。

3.4 改进模型对比实验

为验证本文改进算法的有效性和先进性, 选择和目前主流算法 SSD300、YOLOv5s、YOLOv8s、YOLOv10s、Faster-RCNN 以及文献[22]提出的 Rtdetr-R50 进行性能对比实验, 实验结果如表 3 所示。

表3 算法对比实验
Table 3 Algorithm comparison experiment

算法	P/%	R/%	mAP/%	Paras	GFLOPs	FPS/fps
SSD300	—	—	56.15	—	273.2	—
YOLOv5s	88.57	77.50	77.70	7 022 326	15.9	60.61
YOLOv8s	88.12	73.55	75.15	11 135 971	28.6	54.95
YOLOv10s	91.49	74.78	80.42	8 067 110	24.8	4.97
Rtdetr-R50	55.8	44.7	42.8	42 762 787	130.5	37.30
Faster-RCNN	81.83	77.35	60.45	—	370.4	—
Ours	91.09	81.97	84.40	7 239 766	16.4	61.73

实验表明,本文改进算法相较传统算法 SSD 表现出色,在 mAP 提升了 28.25%;相比本文基础模型 YOLOv5s,在 FPS 略微提高的同时,P 提升 2.52%,R 提升 4.47%,mAP 提升 6.7%;相比 YOLOv8s 算法,各项性能指标皆有提升,P、R、mAP 分别提升 2.97%、8.42%、9.25%,FPS 每秒提升 6.78 帧;相比 YOLOv10s,P 下降 0.3%,R 提升 7.19%,mAP 提升 3.98%,相比 FPS,本文算法是 YOLOv10s 的 12.2 倍。而 Paras 和 GFLOPs 虽比改进前有些许增高,但推理速度得到提升;相比 Rtdetr-R50,改进算法在各项指标上皆有更好的表现,其中 P、R、mAP 分别提升 35.29%、37.27%、41.6%,FPS 提升了 65.5%;相比 Faster-RCNN 算法,改进算法在 P、R、mAP 分别提升 9.26%、4.62%、19.95%,GFLOPs 更是仅为 Faster-RCNN 算法的 4.43%。

3.5 可视化分析

本文为直观展示改进模型的性能,在测试集上进行改进模型和基础模型的检测效果可视化分析。考虑到在实际工业生产过程中,难免出现不可预见的情况,例如光照不稳定、相机镜头污染等情况,选取了 3 种不同情况的检测效果进行对比。

在镜头模糊的情况下进行缺陷检测,其中左侧为 YOLOv5s 检测图,右侧为改进模型检测图,可以看出 YOLOv5s 存在漏检现象,改进模型在此情况下表现良好。对比效果如图 7 所示。

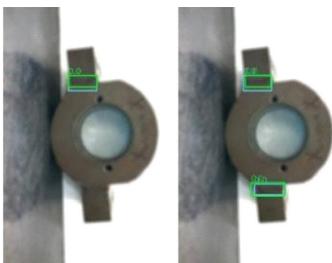


图7 模糊检测效果对比图

Fig.7 Comparison chart of blurriness detection effects

强光条件下检测效果对比如图 8 所示,上方为 YOLOv5s 检测图,下方为改进模型检测图,可以发现

YOLOv5s 存在误检和漏检,改进模型不存在误检,并且大大降低了漏检现象。

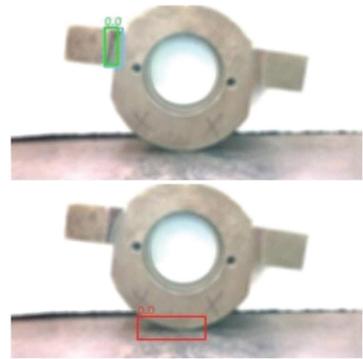


图8 强光检测效果对比图

Fig.8 Comparison chart of detection effects in high-light

在正常情况下采集图像,左侧为 YOLOv5s 检测图,右侧为改进模型检测图,YOLOv5s 存在漏检现象,改进模型检测出 YOLOv5s 未检测出的缺陷。对比效果如图 9 所示。

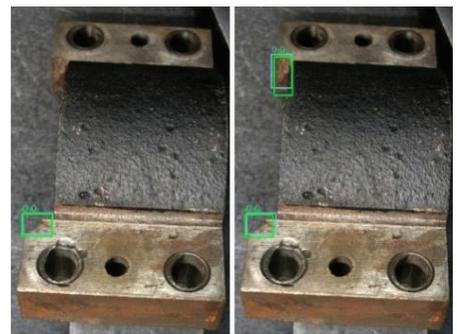


图9 一般检测效果对比图

Fig.9 Comparison chart of general detection effects

综上所述,改进后的模型检测效果优于原始 YOLOv5s 算法,在实际的生产应用中改进模型相比原始模型也有较好的鲁棒性。

4 结 论

针对目前数控机床关键零件轴承座缺陷检测依赖目

检,无法满足工业生产的高精度、高效率 and 低错误率等要求,本文提出基于改进 YOLOv5s 的数控机床轴承座缺陷检测算法。本文在 ConvNeXtv2 结构中引入 HardSwish 激活函数后,结合 CSC 结构,提出 CSCConvNeXtv2-HS 结构,替换原始 YOLOv5s 算法中 backbone 部分的 C3 结构,在提高模型特征提取能力的同时降低计算成本,增强模型稳定性;在 Neck 网络中引入 SSFF 模块来增强网络的多尺度信息提取能力;最后使用 Focal-InnerIoU Loss 损失函数替换原始 YOLOv5s 算法中的 CIoU 损失函数,以加速模型收敛并进一步降低计算复杂度和分类不平衡对性能的影响。实验表明,改进模型相比基础模型 YOLOv5s 在各项指标上分别提升 2.52%、4.47%、6.7% 和 1.12 fps,准确率 P 为 91.09%,召回率 R 为 80.97%,平均精度均值 mAP 为 84.40%,处理速度 FPS 为 61.73 fps,与其他先进算法相比,本改进算法亦有良好的表现。相比人工目检,更加贴合工业生产的高精度、高效率 and 低错误率要求,对于将来工业生产的应用具有重大意义。

参考文献

- [1] 郭渊,周俊. 基于机器视觉的轴承缺陷检测研究进展[J]. 机电工程,2024,41(5):761-774.
- GUO Y, ZHOU J. Research progress of bearing defect detection based on machine vision [J]. Mechatronic Engineering, 2024, 41(5): 761-774.
- [2] 梁海波,王怡,贾武升. 基于机器学习的天然气钢质管道缺陷检测方法研究[J]. 安全与环境学报,2023,23(10):3528-3537.
- LIANG H B, WANG Y, JIA W SH. Research on defect detection methods for natural gas steel pipelines based on machine learning[J]. Journal of Safety and Environment, 2023, 23(10): 3528-3537.
- [3] 符前坤,李强,冉文桑,等. 基于图像处理和机器学习的 PE 管道缺陷检测[J]. 现代电子技术,2024,47(21):59-66.
- FU Q K, LI Q, RAN W SH, et al. Research on defect detection methods for PE pipelines based on image processing and machine learning [J]. Modern Electronic Technology, 2024, 47(21): 59-66.
- [4] 王龙达,刘强,任志贵,等. 基于机器视觉的透明方杯位姿与缺陷检测方法[J]. 电子测量技术,2024,47(22):169-180.
- WANG L D, LIU Q, REN ZH G, et al. Research on pose and defect detection methods for transparent square cups based on machine vision [J]. Electronic Measurement Technology, 2024, 47(22): 169-180.
- [5] 严舒,郭颖,黄骏. FCM-YOLO: 一种基于特征增强和多尺度融合的 PCB 缺陷检测方法[J]. 控制与决策,2024,39(10):3181-3189.
- YAN SH, GUO Y, HUANG J. FCM-YOLO: A PCB defect detection method based on feature enhancement and multi-scale fusion [J]. Control and Decision, 2024, 39(10): 3181-3189.
- [6] CHEN J, KAO S, HE H, et al. Run, don't walk: Chasing higher FLOPS for faster neural networks [C]. IEEE Conference on Computer Vision and Pattern Recognition, 2023: 12021-12031.
- [7] 张波,刘隽. 改进 YOLOv5s 的轻量化无人机小目标检测[J]. 电光与控制,2024,31(11):83-89,114.
- ZHANG B, LIU J. Improved lightweight UAV small-target detection based on YOLOv5s [J]. Electronics Optics & Control, 2024, 31(11): 83-89,114.
- [8] 董耿耿,陈小康,樊湘鹏,等. 基于改进 YOLOv5s 的复杂环境下新梅检测方法[J]. 农业工程学报,2024,40(14):118-125.
- DONG G G, CHEN X K, FAN X P, et al. A detection method for xinmei in complex environments based on improved YOLOv5s [J]. Transactions of the Chinese Society of Agricultural Engineering, 2024, 40(14): 118-125.
- [9] CABRERA J J, CESPEDES O J, CEBOLLADA S, et al. An evaluation of CNN models and data augmentation techniques in hierarchical localization of mobile robots [J]. Evolving Systems, 2024, 15(6): 1991-2003.
- [10] WOO S, DEBNATH S, HU R, et al. Convnext v2: Co-designing and scaling convnets with masked autoencoders [C]. IEEE Conference on Computer Vision and Pattern Recognition, 2023: 16133-16142.
- [11] LIU Z, MAO H, WU C Y, et al. A convnet for the 2020s [C]. IEEE Conference on Computer Vision and Pattern Recognition, 2022: 11976-11986.
- [12] SRIVASTAVA N, HINTON G, KRIZHEVSKY A, et al. Dropout: A simple way to prevent neural networks from overfitting [J]. The Journal of Machine Learning Research, 2014, 15(1): 1929-1958.
- [13] HOWARD A, SANDLER M, CHU G, et al. Searching for mobilenetv3 [C]. IEEE International Conference on Computer Vision, 2019: 1314-1324.
- [14] HENDRYCKS D, GIMPLE K. Gaussian error linear units (gelus) [J]. ArXiv preprint arXiv: 1606.08415,2016.
- [15] ZHANG H, WU C, ZHANG Z, et al. Resnest: Split-attention networks [C]. IEEE Conference on Computer Vision and Pattern Recognition, 2022: 2736-2746.
- [16] SANDLER M, HOWARD A, ZHU M, et al.

- Mobilenetv2: Inverted residuals and linear bottlenecks[C]. IEEE Conference on Computer Vision and Pattern Recognition, 2018: 4510-4520.
- [17] KANG M, TING C M, TING F F, et al. ASF-YOLO: A novel YOLO model with attentional scale sequence fusion for cell instance segmentation [J]. Image and Vision Computing, 2024, 147: 105057.
- [18] ZHANG H, XU C, ZHANG S. Inner-IoU: More effective intersection over union loss with auxiliary bounding box [J]. ArXiv preprint arXiv: 2311.02877, 2023.
- [19] ZHANG Y F, REN W, ZHANG Z, et al. Focal and efficient IOU loss for accurate bounding box regression[J]. Neurocomputing, 2022, 506: 146-157.
- [20] ZHANG H, ZHANG S. Shape-IoU: More accurate metric considering bounding box shape and scale[J]. ArXiv preprint arXiv:2312.17663, 2023.
- [21] MA S, XU Y. MPDIoU: A loss for efficient and accurate bounding box regression[J]. ArXiv preprint arXiv: 2307.07662, 2023.
- [22] ZHAO Y, LYU W, XU S, et al. Detrs beat YOLOs on real-time object detection[C]. IEEE Conference on Computer Vision and Pattern Recognition, 2024: 16965-16974.

作者简介

王浪, 硕士研究生, 主要研究方向为机器视觉。

E-mail: wlydnxzb@163.com

胥云(通信作者), 本科, 教授, 国务院津贴专家, 高级工程师, 主要研究方向为智能制造。

E-mail: xuyunzigong@163.com