

SAD-YOLOv5: 基于 YOLOv5 的铝合金表面缺陷检测方法

袁俊森, 凌六一

安徽理工大学 人工智能学院, 安徽 淮南 232000

摘要:目的 铝合金铸件表面缺陷检测是工业中的一个重要应用, 正确且快速地检测出铸件表面的缺陷可以大大提高产量和质量。针对图像中缺陷目标较小, 缺陷类别易混淆, 定位不精准等问题, 提出了一种在一级检测器基础上改进的 SAD-YOLOv5 模型。方法 针对一般卷积神经网络中由于跨步卷积和池化层导致网络训练过程中信息丢失的问题, 通过引入空间到深度(space-to-depth, SPD) 模块避免细粒度信息的丢失, 提高对小目标的特征学习能力; 为进一步提升网络模型精度, 在网络的 Head 中引入自适应空间特征融合(adaptively spatial feature fusion, ASFF) 和 Decoupled Head, 其中 ASFF 通过实现不同特征之间的自适应融合, 抑制了不同尺度特征之间的不一致性, 保留更有鉴别性的信息, 从而提升网络学习能力; 使用 Decoupled Head 替换原先的耦合头, 将分类和回归进行解耦, 使分类更加关注纹理信息, 回归更加关注边缘信息, 二者各司其职, 进一步提升网络判断能力。结果 在自己拍摄的铸件缺陷检测数据集中的测试结果表明, SAD-YOLOv5 的 mAP@0.5 和 mAP@0.5:0.95 分别为 95.1% 和 68%, 较基线模型(YOLOv5) 分别提升了 1% 和 3.3%。结论 SAD-YOLOv5 能更准确地完成铝合金铸件的表面缺陷检测任务。

关键词: 铝合金铸件; 表面缺陷检测; YOLOv5; SPD; ASFF; Decoupled Head

中图分类号: TP183; TP399 **文献标识码:** A **doi:** 10.16055/j.issn.1672-058X.2025.0003.010

SAD-YOLOv5: Aluminum Alloy Surface Defect Detection Method Based on YOLOv5

YUAN Junsen, LING Liuyi

School of Artificial Intelligence, Anhui University of Science and Technology, Anhui Huainan 232000, China

Abstract: Objective Surface defect detection of aluminum alloy castings is a critical application in industry. Accurate and rapid detection of defects on the surface of castings can significantly improve production and quality. In response to challenges such as small defect targets, easily confused defect categories, and imprecise localization in images, an improved SAD-YOLOv5 model based on a primary detector was proposed. **Methods** Addressing the problem of information loss during network training caused by strided convolutions and pooling layers in general convolutional neural networks, space-to-depth (SPD) module was introduced to avoid the loss of fine-grained information and enhance the feature learning capability for small targets. To further improve the model accuracy, adaptively spatial feature fusion (ASFF) and Decoupled Head were introduced in the network's Head. ASFF achieved adaptive fusion between different features, suppressing inconsistencies among features of different scales to retain more discriminative information and enhance network learning capability. Decoupled Head replaced the original coupled head to decouple the classification and regression, allowing classification to focus more on texture information and regression to focus more on edge information. This division of responsibilities further enhances the network's decision-making capability. **Results** Testing on a self-

收稿日期: 2023-10-11 修回日期: 2023-12-18 文章编号: 1672-058X(2025)03-0077-07

基金项目: 安徽省自然科学基金项目(项目标号: 2208085ME128)。

作者简介: 袁俊森(1997—), 男, 安徽合肥人, 硕士研究生, 从事机器视觉和目标检测研究。

通信作者: 凌六一(1980—), 男, 安徽淮南人, 教授, 博士, 从事智能信息处理、光电处理研究。Email: lyling@aust.edu.cn.

引用格式: 袁俊森, 凌六一. SAD-YOLOv5: 基于 YOLOv5 的铝合金表面缺陷检测方法[J]. 重庆工商大学学报(自然科学版), 2025, 42(3): 77-83.

YUAN Junsen, LING Liuyi. SAD-YOLOv5: Aluminum alloy surface defect detection method based on YOLOv5[J]. Journal of Chongqing Technology and Business University (Natural Science Edition), 2025, 42(3): 77-83.

captured dataset for casting defect detection showed that SAD-YOLOv5 achieved $mAP@0.5$ and $mAP@0.5:0.95$ of 95.1% and 68%, respectively. This represented a 1% and 3.3% improvement over the baseline model (YOLOv5).

Conclusion SAD-YOLOv5 demonstrates the ability to more accurately perform surface defect detection tasks on aluminum alloy castings.

Keywords: aluminum alloy casting; surface defect detection; YOLOv5; SPD; ASFF; Decoupled Head

1 引言

铝合金铸件具有强度高、密度小等特点,作为轻量化用材在制造业中发挥着重要作用。而铝合金铸件在生产过程中,总会因为种种原因产生气孔、裂纹、划痕等缺陷,这些缺陷不仅影响铸件的外观,更对铸件的使用寿命、物理性能和表面质量产生重大影响,甚至可能会导致严重事故的发生,所以铸件的缺陷检测至关重要。

现如今的缺陷检测主要是人工检测和基于单一机制的检测方法。人工检测受限于人的主观影响,容易出现漏检、误检的情况,且人工检测准确率不高。基于单一机制的检测方法,例如 X 光检测、涡流检测、漏磁检测等,这些方法主要通过缺陷的一些物理特征进行检测,存在很大局限性,检测条件也较为苛刻,且检测成本高昂。随着人工智能技术的不断发展和普及,设计一种准确、快速、高效的自动化检测方法已成为工业生产中的迫切任务。

近年来,随着人工智能技术的不断发展,已有学者将目标检测算法应用于铸件缺陷检测。奥地利科技研究所^[1]早在 2014 年就已使用卷积神经网络(CNN)来实现对轨道表面空洞缺陷进行分类;Liu 等^[2]将 Faster-RCNN 应用在检测研磨和抛光水龙头的缺陷上,实验证明该架构可以有效抑制异常检测;Xu 等^[3]在 YOLOv3 中使用 K-means++ 分析数据集上的锚框大小信息,实现了对金属表面细小缺陷更多特征的提取;Duan 等^[4]用双密度卷积层替换 YOLOv3 中的单密度卷积层,并添加一个检测头,从而获得更丰富的语义信息,提高对铸件的缺陷检测能力;Li 等^[5]使用 MobileNet 结构优化 SSD 的骨干网络,以便简化检测模型,从而实现对罐装生产线容器密封表面缺陷的检测。

上述算法虽然都取得了不错的效果,但普遍缺少对于铸件缺陷较小这一特性的深入探讨和研究。为此,提出了一种改进的 YOLOv5 模型用于铝合金表面缺陷检测——SAD-YOLOv5。该模型引入 SPD(Space-to-depth)^[6]来避免一些细节信息在提取特征过程中的丢失,加强模型对小目标的特征学习能力;使用 ASFF^[7]对不同尺度特征进行融合,提升模型对小目标的检测能力;替换原先的检测头,使用 Decoupled Head 将分类和回归任务进行解耦,从而进一步提升模型的精度。实验证明,SAD-YOLOv5 可以在铝合金表面缺陷数据集中达到 95.1% mAP ,较之其他模型,有着更高

的精度。

综上所述,本文的主要贡献包括:自主拍摄铝合金表面缺陷数据集用作模型特征学习与验证;在原模型的骨干网络和颈部网络添加 SPD 层,实现特征的下采样,保留更多的特征信息,避免一些细节信息的丢失;使用 ASFF 实现对输入的不同尺度特征融合,抑制不同尺度特征信息之间的不一致性;使用 Decoupled Head 避免原先耦合检测头带来的精度损失。

2 相关工作

2.1 目标检测

目标检测是计算机视觉的最基本问题之一。2014 年是目标检测发展的一道分水岭,在此之前,目标检测依靠传统算法解决问题,在此之后,目标检测开始趋于基于深度学习的算法。

在传统算法阶段,检测器依靠人工设计特征,为使检测器具有更好的表达能力并降低对计算资源的消耗,人们使用了诸多技巧,但是在 HOG^[8]和 DPM^[9]检测器之后,人工检测器没有进一步的发展,甚至一度处于停滞状态。后来随着卷积神经网络(CNN)^[10]在图像分类任务取得巨大进步并应用到目标检测任务中,人们开始逐步关注深度学习在目标检测任务中的运用。现阶段基于深度学习的目标检测方法主要分为两类:两阶段的目标检测算法和单阶段的目标检测算法。两阶段算法先生成一系列候选框,然后对候选框进行筛选,判断候选框是否包含检测目标,并对目标的位置进行修正,这一类算法的典型代表有 R-CNN^[11]、Fast R-CNN^[12]、Faster R-CNN^[13]等;而单阶段算法并不会产生候选框,直接将边框定位问题转换为回归问题,直接产生目标的位置坐标和类别概率,这一类算法的典型代表有 YOLOX^[14]、SSD^[15]、FCOS^[16]等。

2.2 多尺度特征融合

特征融合是目标检测中的一个重要问题。卷积神经网络逐层提取目标特征,不同层次特征图的特征表达能力不同。浅层特征感受野较小,几何细节信息表征能力强,能够较好地反映明暗、边缘等细节信息;深层特征感受野较大,语义信息表征能力强,能够更好地反映目标的整体结构。但是传统的目标检测模型通常只采用特征提取网络的最后一层输出进行目标的分类和定位,这导致小目标的有效信息获取较少,降低了小

标大小适中时并不会暴露出来,反而会筛选出一些冗余信息,使模型可以更好地学习目标物特征。但是在检测铸件缺陷这类小目标时,冗余信息的过度假设不再成立,这些缺点就会被进一步放大并暴露出来。Sunkara 等^[6]提出了一种适用于小目标和低分辨率图像的构建块:SPD-Conv。在本次试验中使用 SPD-Conv 中的 SPD 层,SPD 层实际上是对提取到的特征图进行一次下采样,并且保留通道维度中的信息,不造成任何信息损失。

输入任意大小 $S \times S \times C_1$ 的特征图映射,SPD 层会将其划分为式(1)中的子特征图映射:

$$f_{i,j} = X[i:S:scale, j:S:scale] \quad (1)$$

$$i \in [0, scale-1], j \in [0, scale-1]$$

式(1)中的 X 为输入特征图映射, f 为子特征图映射, $scale$ 是划分子特征图映射的比例因子。得到这些新的子特征图映射后,将其按照通道维度进行拼接从而使其空间维度降低,通道维度增加,即输入大小为 (S, S, C_1) 的特征图输出大小为 $(\frac{S}{scale}, \frac{S}{scale}, scale^2 * C_1)$ 。图 2 给出了当 $scale=2$ 时的例子,从图中可以看出,SPD 层相当于进行了一次因子为 2 的下采样,从而获得了 4 个尺寸大小为 $(\frac{S}{2}, \frac{S}{2}, C_1)$ 的 $f_{0,0}, f_{1,0}, f_{0,1}, f_{1,1}$ 子特征图映射,随后将这些子特征图映射进行级联得到 $(\frac{S}{2}, \frac{S}{2}, 4 * C_1)$ 。

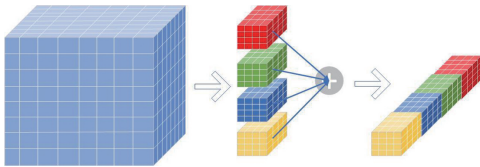


图 2 $scale=2$ 时的 SPD 过程图

Fig. 2 SPD process diagram when $scale=2$

3.2 ASFF

特征金字塔结构虽然是现今解决多尺度问题的常用方法,但是该方法对不同尺度之间存在的不一致性没有得到很好的解决,这种不一致性干扰了训练过程中梯度计算,并降低了特征金字塔的有效性,极大地限制了特征金字塔的精度。为此 Wang^[7]使用 Adaptively Spatial Feature Fusion (ASFF) 方法来解决该问题,该方法借鉴空间过滤冲突信息来抑制这种不一致性,只保留了有用的信息进行组合。

在 YOLOv5 中会有 3 个不同层次的特征信息,它们有着不同的分辨率和不同的通道数,而 ASFF 对于某一层特征,会先将其他层的特征调到同一分辨率和通道数,然后将特征进行整合训练,以期找到最优融合方案。这些不同层次的特征进行自适应融合,使其中所携带的有冲突的信息被过滤,而那些更具有鉴别性的信息被保留

下来并占据主导地位,ASFF 大致流程如图 3 所示。

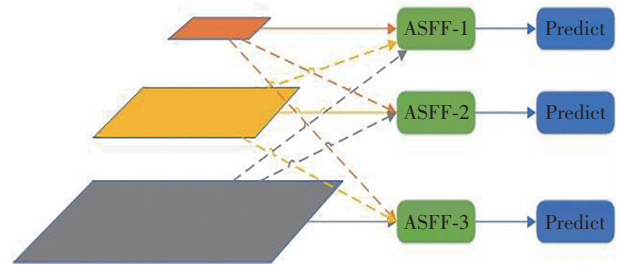


图 3 ASFF 流程图

Fig. 3 Flow chart of ASFF

ASFF 需要将 3 个不同层次的特征信息进行整合。第一步要将它们调整为相同的尺寸,所以需要对每个尺度的上采样和下采样进行更改。对于上采样,首先使用 1×1 卷积将其他层级的通道数转变为当前层级的通道数,然后分别使用插值法提高分辨率。对于 $1/2$ 下采样,使用 stride 为 2 的 3×3 卷积修改通道数即可,对于 $1/4$ 下采样,在卷积之前增加一层 stride=2 的最大池化。调整为统一尺寸后需要将它们进行整合并训练,三个不同层次的特征整合公式如下:

$$y_{ij}^l = \alpha_{ij}^l \cdot x_{ij}^{1 \rightarrow l} + \beta_{ij}^l \cdot x_{ij}^{2 \rightarrow l} + \gamma_{ij}^l \cdot x_{ij}^{3 \rightarrow l}$$

$$\alpha_{ij}^l, \beta_{ij}^l, \gamma_{ij}^l \in [0, 1] \quad (2)$$

式(2)中, $x_{ij}^{n \rightarrow l}$ 表示将 (i, j) 的特征图通道数从 n 转变成 l 。 y_{ij}^l 表示最终输出的特征图。 $\alpha_{ij}^l, \beta_{ij}^l$ 和 γ_{ij}^l 是 3 个不同层次的特征进行整合的权重,且满足 $\alpha_{ij}^l + \beta_{ij}^l + \gamma_{ij}^l = 1$ 。

3.3 Decoupled Head

在目标检测中,分类和回归之间的冲突是一个众所周知的问题。现今,分类和定位的解耦头被广泛应用于一级和二级探测器中,但在 YOLOv5 中仍使用耦合在一起的检测头。YOLOX^[14] 中用实验证明了耦合检测头会损坏检测的精度。这是因为分类和回归任务的关注点并不相同,分类任务更加关注目标的纹理信息,定位任务更加关注目标的边缘信息,因此提出了一个简易的解耦头将这两个任务进行解耦。该 Decoupled Head 先使用一个 1×1 的卷积降低信道维度,然后两个平行的分支都进行两次 3×3 的卷积运算,最后分别使用 1×1 卷积进行计算得到分类和回归的结果。Decoupled Head 结构如图 4 所示。

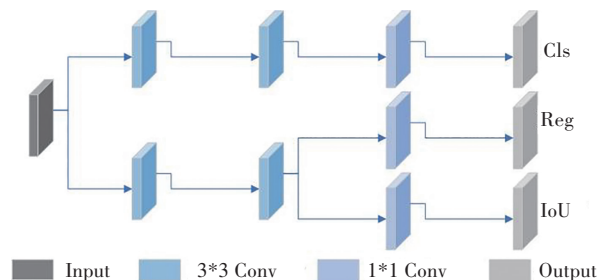


图 4 Decoupled Head 结构图

Fig. 4 Structure diagram of Decoupled Head

4 仿真实验与结果分析

4.1 数据集

本文的数据集是使用双目摄像机自行捕捉采集,从不同角度每间隔 3 s 采集一张图片,共采集 8 000 多张图片,经筛选后保留了 2 665 张图片,标签分配情况如表 2 所示。图片中缺陷被粗略划分为 8 类:sag、hole、shrinkage、surface punch、round punch、piece punch、gritty machines 和 crack。其中,铸件表面出现塌陷被定义为

sag;铸件表面由气体形成的孔洞类缺陷被定义为 hole;铸件表面出现大面积缩陷被定义为 shrinkage;与标准铸件相比,根据未打孔截面形状将缺陷定义为 surface punch、round punch 和 piece punch;铸件在凝固过程中,由于补缩不良产生的孔洞被定义为 gritty machines;铸件表面出现划痕被定义为 crack,具体情况如图 5 所示。在训练过程中,数据集按照 8 : 2 的比例随机分为 2 132 张图像的测试集和 533 张的验证集。

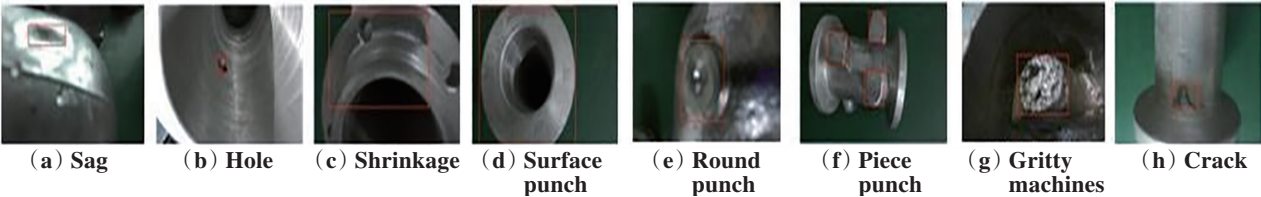


图 5 缺陷样例图

Fig. 5 Sample images of defects

表 2 标签分配情况

Table 2 Assignment of labels

缺陷名称	标签/个
Sag	642
Hole	465
Shrinkage	151
Surface punch	2 408
Round punch	670
Piece punch	2 671
Gritty cavity	640
Crack	182

考虑数据集中的有些样本较少,在实验过程中使用数据增强的方式扩大数据集数量。YOLOv5 中运用的数据增强方式有:随机裁剪、翻转、缩放、旋转、移位和马赛克图像预处理技术。其中马赛克图像预处理技术是将多张图片裁剪到一张图片中,丰富检测物体的背景,将大样本随机裁剪成小样本进行训练,增加了小尺度目标的数量。

4.2 评价指标

采用 precision (P)、recall (R)、mean average precision (mAP)评价改进后的模型性能,评价指标计算如下:

$$P = \frac{T_p}{T_p + F_p} \tag{3}$$

$$R = \frac{T_p}{T_p + F_N} \tag{4}$$

$$A_p = \int_0^1 P dR \tag{5}$$

$$m_{AP} = \frac{\sum_{i=1}^N A_{p_i}}{N} \tag{6}$$

式(3)—式(6)中, N 为类别数目, T_p (True Positive) 是缺陷图像中检测出缺陷的数目; F_p (False Positive) 是无缺陷图中检测出缺陷的数目; T_N (True Negative) 是无缺陷图中未检测出缺陷的数目; F_N (False Negative) 是有缺陷图中未检测出缺陷的数目; A_p 是 P-R 曲线下的面积。mAP 是目标检测算法中的常用评价指标,反映了模型边界框中的目标类别和位置的准确性。

4.3 实验参数

实验中使用的 SAD-YOLOv5 模型,参考 SPD-Conv 结构,在网络中引入 SPD 层进行下采样,在 Head 部分使用 ASFF 对输入特征图进行自适应融合,用 Decoupled Head 代替原来的检测头。在实验中,使用 SGD 作为优化器,momentum 设置为 0.937,初始学习率为 0.01,使用余弦学习率衰减,weight decay 为 0.000 5,防止数据的过拟合。在模型的初始训练中,首先进行了 3 个 epochs 的 warm-up, warm-up 训练 momentum 设置为 0.8。此外,训练的 batch-size 被设置为 8,总共训练了 300 个 epochs。

4.4 对比实验结果分析

表 3 显示了不同模型在数据集中的评估结果。工业场景下物体检测的精度非常重要,只有保证了精度,才能实现工业场景下的智能检测。因此,在这部分论述中,使用平均精度 (mAP) 对缺陷检测模型进行综合评价,其中 mAP 分为 mAP@0.5 和 mAP@0.5 : 0.95。mAP@0.5 反映了真实框与边界框的耦合度达到 0.5 时的精度,mAP@0.5 : 0.95 表示耦合度从 0.5 到 0.95 的平均 mAP,这在一定程度上反映了预测位置的精度。

表 3 中,方法 A、B、C、D 和 E 分别表示 YOLOv5 原模型、YOLOX、YOLOv7^[21]、SSD 以及 SAD-YOLOv5。

表 3 显示了各种目标检测模型的性能比较,其中

SAD-YOLOv5 模型在 $mAP@0.5$ 指标上表现最优,同时在 $mAP@0.5:0.95$ 上也属次优。这一优势可以归功于该模型采用了多种方法来增强对小目标的检测能力以及边界框的定位精度。具体而言,SAD-YOLOv5 模型通过引入先进的技术和算法,有效地解决了在小目标检测过程中常见的问题。相较之下,方法 D 虽然在 $mAP@0.5:0.95$ 上表现出色,但其 $mAP@0.5$ 却相对较低,与 SAD-YOLOv5 模型相比差距达到了 4.5%。这表明,在目标检测的整体准确度方面,方法 D 明显不如 SAD-YOLOv5 模型。当一个模型在定位精度上表现极高,但如果其目标检测准确度不足,也会极大地影响其实际应用价值。因此,综合考虑各项指标,SAD-YOLOv5 模型在检测精度和定位精度之间取得了良好的平衡,成为当前评估中表现最为出色的模型之一。其综合性能优势使它在实际场景中具有广泛的应用前景,并为目标检测领域的发展提供了有力的支持与启示。

表 3 对比实验结果

Table 3 Comparison of experimental results

方 法	A	B	C	D	E
<i>P</i>	0.963	0.899	0.958	0.624	0.929
<i>R</i>	0.89	0.672	0.893	0.857	0.91
$m_{AP}@0.5$	0.94	0.855	0.941	0.906	0.951
$m_{AP}@0.5:0.95$	0.647	0.558	0.661	0.686	0.68

4.5 消融实验结果分析

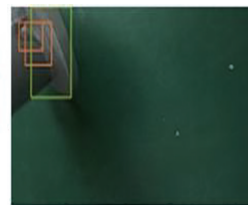
在表 4 中,对比了不同方法(A~E)在目标定位回归方面的性能表现。方法 A 中引入了 SPD,使网络在特征提取过程中保留了更多细节信息,显著减少了细粒度特征的丢失,从而实现了目标的更准确定位。 $mAP@0.5:0.95$ 指标相对于原模型提升了 1.9%,这充分证明了 SPD 的有效性。在方法 C 中,引入了解耦头的概念,将分类和回归任务解耦,使网络能够专注于执行各自的任务,从而得到更精确的结果。实验证明,这一改进使 $mAP@0.5$ 和 $mAP@0.5:0.95$ 分别提升了 0.4% 和 1.4%,为目标检测的精度提升提供了有力支持。在方法 D 中,采用了 ASFF 特征融合方式,通过保留特征信息中的有价值信息,过滤掉不必要的噪声,进一步提升了 2.1% 的 $mAP@0.5:0.95$ 。这一改进策略在铝合金铸件表面缺陷检测中起到了积极的作用,使模型在面对复杂场景时表现更为出色。方法 E 则对前述 3 种改进进行了整合。实验结果显示,这种整合策略取得了显著成果,实现了“1+1>2”的效果。单独采用各种方法时,对 $mAP@0.5$ 的提升并不显著,但通过整合后, $mAP@0.5$ 有了 1.1% 的显著提升,同时 $mAP@0.5:0.95$ 也提升了 3.3%,较之前有了更大的提升。

表 4 消融实验结果

Table 4 Ablation experiment results

方 法	A	B	C	D	E
<i>P</i>	0.963	0.932	0.947	0.962	0.929
<i>R</i>	0.89	0.896	0.899	0.897	0.91
$m_{AP}@0.5$	0.94	0.941	0.944	0.939	0.951
$m_{AP}@0.5:0.95$	0.647	0.666	0.661	0.668	0.68

图 6 与图 7 展示了原模型和 SAD-YOLOv5 检测结果之间的明显差异。SAD-YOLOv5 模型具有更低的误检率,定位更加精准,对小目标的检测效果更为出色。这个结果进一步证实了 SAD-YOLOv5 模型在铝合金铸件表面缺陷检测领域的优越性。



(a) 改进前错检 1



(b) 改进前错检 2



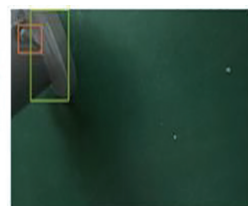
(c) 改进前漏检 1



(d) 改进前漏检 2

图 6 模型改进前效果图

Fig. 6 Images before model improvement



(a) 改进后无错检 1



(b) 改进后无错检 2



(c) 改进后无漏检 1



(d) 改进后无漏检 2

图 7 模型改进后效果图

Fig. 7 Images after model improvement

综上所述,通过引入 SPD、解耦头和 ASFF 特征融合方式,并将其有机整合,SAD-YOLOv5 模型在铝合金铸件表面缺陷检测任务中取得了显著的性能提升,为相关领域的研究和应用提供了有力的支持与启示。

5 结 论

本文针对铝合金铸件的特点,提出了一种改进的YOLOv5目标检测模型——SAD-YOLOv5,通过引入SPD和ASFF加强对小目标的学习和检测能力,再将检测头替换为Decoupled Head,进一步减少耦合头带来的精度损失。通过实验证明,SAD-YOLOv5模型 $m_{AP}@0.5$ 达到了95.1%, $m_{AP}@0.5:0.95$ 达到了68%,比原模型分别提高了1%和3.3%,是更为优异的铝合金表面缺陷检测模型。但由于SPD和ASFF的引入,神经网络的参数量和运算量显著增加,从而使其在实时性方面表现不佳,所以下一步的工作重点是对网络模型进行轻量化,从而降低模型的复杂度和运算量,达到更高的检测速度,实现满足检测精度的同时满足检测速度实时性的要求。

参考文献(References):

- [1] SOUKUP D, HUBER-MÖRK R. Convolutional neural networks for steel surface defect detection from photometric stereo images[C]//International Symposium on Visual Computing. Cham: Springer International Publishing, 2014: 668–677.
- [2] LIU M W, LIN Y H, LO Y C, et al. Defect detection of grinded and polished workpieces using faster R-CNN[C]//2021 IEEE/ASME International Conference on Advanced Intelligent Mechatronics (AIM). Piscataway, NJ: IEEE, 2021: 1290–1296.
- [3] XU Y, ZHANG K, WANG L. Metal surface defect detection using modified YOLO[J]. Algorithms, 2021, 14(9): 257–271.
- [4] DUAN L, YANG K, RUAN L. Research on automatic recognition of casting defects based on deep learning[J]. IEEE Access, 2020, (9): 12209–12216.
- [5] LI Y, HUANG H, XIE Q, et al. Research on a surface defect detection algorithm based on MobileNet-SSD[J]. Applied Sciences, 2018, 8(9): 1678–1695.
- [6] SUNKARA R, LUO T. No more strided convolutions or pooling: A new CNN building block for low-resolution images and small objects[C]//Joint European Conference on Machine Learning and Knowledge Discovery in Databases. Cham: Springer Nature Switzerland, 2022: 443–459.
- [7] WANG C Y, BOCHKOVSKIY A, LIAO H Y M. Scaled-YOLOv4: Scaling cross stage partial network[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. Piscataway, NJ: IEEE, 2021: 13029–13038.
- [8] BAKHEET S, AL-HAMADI A. A framework for instantaneous driver drowsiness detection based on improved HOG features and naïve Bayesian classification[J]. Brain Sciences, 2021, 11(2): 240–255.
- [9] FELZENSZWALB P, MCALLESTER D, RAMANAN D. A discriminatively trained, multiscale, deformable part model[C]//2008 IEEE Conference on Computer Vision and Pattern Recognition. Piscataway, NJ: IEEE, 2008: 1–8.
- [10] KRIZHEVSKY A, SUTSKEVER I, HINTON G E. ImageNet classification with deep convolutional neural networks[J]. Communications of the ACM, 2017, 60(6): 84–90.
- [11] SAMANTARAY S, DEOTALE R, CHOWDHARY C L. Lane detection using sliding window for intelligent ground vehicle challenge[C]//Innovative Data Communication Technologies and Application: Proceedings of ICIDCA 2020. Singapore: Springer International Publishing, 2021: 871–881.
- [12] GIRSHICK R. Fast R-CNN[C]//Proceedings of the IEEE International Conference on Computer Vision. Piscataway, NJ: IEEE, 2015: 1440–1448.
- [13] REN S, HE K, GIRSHICK R, et al. Faster R-CNN: Towards real-time object detection with region proposal networks[J]. Advances in Neural Information Processing Systems, 2015, 28: 91–99.
- [14] AHMAD H M, RAHIMI A. Deep learning methods for object detection in smart manufacturing: A survey[J]. Journal of Manufacturing Systems, 2022, 64: 181–196.
- [15] LIU W, ANGUELOV D, ERHAN D, et al. SSD: Single shot multibox detector[C]//Computer Vision-ECCV 2016: 14th European Conference. Amsterdam: Springer International Publishing, 2016: 21–37.
- [16] TIAN Z, SHEN C, CHEN H, et al. FCOS: Fully convolutional one-stage object detection[C]//Proceedings of the IEEE/CVF International Conference on Computer Vision. Piscataway, NJ: IEEE, 2019: 9627–9636.
- [17] LIN T Y, DOLLAR P, GIRSHICK R, et al. Feature pyramid networks for object detection[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Piscataway, NJ: IEEE, 2017: 2117–2125.
- [18] HE K, ZHANG X, REN S, et al. Spatial pyramid pooling in deep convolutional networks for visual recognition[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2015, 37(9): 1904–1916.
- [19] LIU S, QI L, QIN H, et al. Path aggregation network for instance segmentation[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Piscataway, NJ: IEEE, 2018: 8759–8768.
- [20] ZAIDI S A, ANSARI M S, ASLAM A, et al. A survey of modern deep learning based object detection models[J]. Digital Signal Processing, 2022, 126: 103514–103533.
- [21] WANG C Y, BOCHKOVSKIY A, LIAO H Y M. YOLOv7: Trainable bag of freebies sets new state of the art for real-time object detectors[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. Piscataway, NJ: IEEE, 2023: 7464–7465.

责任编辑:吴 倩