

DOI:10.19651/j.cnki.emt.2416025

改进 YOLOv8 的烟盒缺陷检测算法*

王震洲 杨榕 宿景芳

(河北科技大学信息科学与工程学院 石家庄 050018)

摘要:近年来,人们对于烟盒包装质量的要求越来越高,在现代化的生产中,烟盒的生产速度大幅度提升,生产设备也都实现了智能化。但是,对于烟盒的表面质量检测仍然采用人工的方式进行。针对人工烟盒表面缺陷检测容易发生漏检、错检等问题,提出了一种基于改进 YOLOv8 的烟盒缺陷检测算法。首先在 YOLOv8 的颈部网络引入了 GD 机制,提高了模型对于不同层级间的信息融合能力;其次加入了尺度序列特征融合模块,增强了网络对不同尺度信息的提取能力;最后使用 RT-DETR 的 Decoder 替换 YOLOv8 的头部网络,这使得网络模型无需依赖于复杂的 NMS 后处理步骤,大大简化了检测流程并提高了效率。实验结果表明:改进后的算法模型在自制的烟盒缺陷数据集上与 YOLOv8 算法相比检测精度提升到了 94.6%,检测速度达到了 121.4 FPS。并且与其他目标检测算法相比,改进后的算法在检测精度和检测速度方面有一定的优势,更适合应用在烟厂对烟盒表面质量的检测。

关键词: 烟盒缺陷检测;YOLOv8;GD 机制;SSFF;RT-DETR Decoder

中图分类号: TN9 **文献标识码:** A **国家标准学科分类代码:** 520.6040

Improved YOLOv8 cigarette box defect detection algorithm

Wang Zhenzhou Yang Rong Su Jingfang

(School of Information Science and Engineering, Hebei University of Science and Technology, Shijiazhuang 050018, China)

Abstract: In recent years, there has been an increasing demand for higher quality cigarette pack packaging. While modern production has significantly increased the speed of cigarette box production and made production equipment more intelligent, surface quality inspection of cigarette boxes still relies on manual methods. Addressing the issues of human error such as missed or incorrect detections in surface defect inspection, a cigarette box defect detection algorithm based on improved YOLOv8 is proposed. Firstly, a Gather-and-Distribute mechanism is introduced into the neck network of YOLOv8 to enhance the model's fusion capability for information across different hierarchies. Secondly, a scale sequence feature fusion module is incorporated to strengthen the network's ability to extract information from different scales. Finally, the head network of YOLOv8 is replaced with the Decoder of RT-DETR, eliminating the need for complex post-processing steps such as Non-Maximum Suppression, thereby simplifying the detection process and improving efficiency. Experimental results show that the improved algorithm model achieves a detection accuracy of 94.6% and a detection speed of 121.4 FPS on a self-made cigarette box defect dataset compared to YOLOv8. Moreover, compared with other object detection algorithms, the improved algorithm has certain advantages in terms of detection accuracy and speed, making it more suitable for application in cigarette factories for surface quality inspection of cigarette boxes.

Keywords: defect detection of cigarette boxes;YOLOv8;GD mechanism;SSFF;RT-DETR Decoder

0 引言

在烟盒的生产制造过程中,会存在着较多种类型的缺陷,如划痕、胶点和污渍的缺陷^[1]。在印刷烟盒图案和文字时,使用油墨、颜料、溶剂等化学物质。这些物质有时会残

留在烟盒表面,尤其是在印刷设备调试、清洗不彻底或者印刷速度过快的情况下容易产生污渍。由于烟盒面纸是使用胶水粘连成烟盒的,因此,在烟盒制造的过程中,有极大的几率会在烟盒的表面留下胶点,从而制造出不符合产品质量要求的残次品。同时,烟盒在生产线上传送带上运动时,由

收稿日期:2024-05-12

* 基金项目:河北省教育厅青年基金(QN2023185)项目资助

于惯性烟盒之间会发生碰撞,从而产生一些带有划痕缺陷的残次品。由于缺陷检测部分通常是生产流水线中的最后一道工序,其检测结果将直接影响烟盒质量,因此烟厂对于烟盒外包装的检测精确度和速度都有着较高的要求。

目前一些企业仍然采用传统的人工方式进行缺陷检测,这种方法存在一系列问题。首先,它需要大量的人力资源,耗费时间和成本。其次,人工检测容易受到主观因素的影响,导致检测结果的不一致性和不准确性。此外,人工检测还容易出现疲劳和注意力不集中的情况,可能导致漏检或误检,从而影响产品质量和交货周期。为了解决这些问题,许多企业正在积极转向自动化缺陷检测技术。基于机器视觉的缺陷检测技术^[2-3]在取代传统人工检测方面表现出了显著的优势,尤其是在提高检测准确性和效率方面。然而,传统的基于机器视觉的缺陷检测系统往往需要依赖人工设计的特征提取算法,这些算法的鲁棒性和泛化能力有时候存在局限,难以满足实际生产环境中复杂多变的需求。随着深度学习技术的快速发展,特别是在目标检测领域取得的突破性进展,基于深度学习的缺陷检测^[4-6]逐渐成为了研究和应用的热点。深度学习技术通过构建深层神经网络,可以自动学习和提取数据中的特征,从而克服了传统方法中对人工设计特征的依赖性,具有更好的泛化能力和适应性。使用深度学习的缺陷检测技术更多的应用于半导体芯片、太阳能电池、带钢表面、焊缝等领域,但是在烟盒领域进行缺陷检测的研究较少。

YOLO^[7-9]系列算法以其高效的目标检测能力受到广泛关注,为缺陷检测提供了更为高效、精准的技术。张恒等^[10]提出了一种基于 YOLOv5 改进的芯片缺陷检测方法,新增更大尺寸的特征检测器,使用具有高效聚焦学习能力的特征金字塔结构,有效缓解缺陷尺寸差异大造成定位不准问题,提出基于面积的边界框融合算法有效减少冗余框,最终实验结果表明均值平均精度(mean average precision, mAP)只有 79.4%。陈亚芳等^[11]提出了多尺度 YOLOv5 太阳能电池缺陷检测方法,利用可变形卷积和坐标注意力来改进特征提取网络,特征融合网络加入坐标注意力,增强多尺度缺陷特征表达能力,最终该模型表现出了 95.4% 的 mAP,但是检测速度只达到了 51 fps。王春梅等^[12]提出一种轻量级的带钢表面缺陷检测算法,在 YOLOv8 的基础上特征提取网络使用轻量级的 VanillaNet 网络结构,特征融合网络中使用轻量级上采样算子 CARAFE,最终 mAP 达到 80.8%,但是对于与背景相似的缺陷检测效果并不好。Zhang 等^[13]提出了一种基于 YOLOv8 的用于焊缝缺陷检测任务的改进模型。在原有模型结构的基础上进行了优化和改进,包括优化卷积层结构、引入 CARAFE 算子、替换 SPPF 模块、引入基于归一化的注意力模块(normalization-based attention module, NAM)、优化损失函数等,最终 mAP 达到 97.3%,检测速度达到 67.8 fps。目前针对烟盒缺陷的目标检测算法是在 YOLOv3 的基础上进行改进,如

陈同宇等^[14]提出了一种基于改进 YOLOv3 的烟盒缺陷检测方法,在 YOLOv3 的基础网络中添加了四倍的降采样,并且改进损失函数,烟盒缺陷检测 mAP 达到了 97.29%,但是检测速度只有 21 fps。Peng 等^[15]提出了一种改进的 yolov5 算法用于烟盒外观的缺陷检测,通过优化模型结构来增强对不同尺度缺陷特征的表达和融合能力、引入高效聚焦学习机制使模型在复杂背景下也能准确地检测出缺陷区域。该模型 mAP 达到 91.6%,检测速度可达 154 fps。但模型的高计算复杂性、对小缺陷的敏感性有待提升。陈文兵等^[16]提出了一种基于修复对抗网络的烟包表面缺陷检测方法,通过 R-GAN 网络模型对烟包表面缺陷进行检测和修复,修复机制来增强模型的缺陷检测能力,即在生成器生成无缺陷图像的同时,也进行部分缺陷区域的修复。其 mAP 达到 97.5%,但其较高的计算复杂性可能限制了其在实时性要求较高的应用场景中的应用。

基于上述研究,以及烟盒缺陷检测的准确性和效率对于烟草公司的重要性。本文提出了一种基于 YOLOv8 的烟盒缺陷检测改进算法。改进的算法在颈部网络中引入收集-分发机制(gather-and-distribute mechanism, GD 机制)和尺度序列特征融合模块(scale sequence feature fusion, SSFF),增强模型对于跨层之间的信息融合能力以及对于不同尺寸缺陷特征信息的提取能力,最后替换 YOLOv8 头部网络,减少非极大值抑制(non-maximum suppression, NMS)后处理步骤,提升模型缺陷检测速度。

1 YOLOv8 网络介绍

YOLOv8 是 ultralytics 公司在 2023 年初开源的 YOLO 系列最新的版本,目前支持图像分类、目标检测和实例分割任务。YOLOv8 的网络结构主要包括三大部分:特征提取网络(Backbone)、特征融合网络(Neck)和头部网络(Head)。如图 1 所示,特征提取网络包括卷积模块、C2f 模块和 SPPF 模块,卷积模块进行降采样操作来减小特征图的尺寸并增加通道数。C2f 模块首先使用 1×1 的卷积核将输入通道减少到原来的 $1/2$,然后使用 3×3 的卷积核进行卷积操作以提取特征信息,其次使用残差链接将输入直接加到输出中,最后使用 1×1 的卷积核恢复特征图的通道数。SPPF 模块利用三次连续池化,并且结合每一层的输出,保证多尺度融合的同时,降低计算量;特征融合网络采用 PAN-FPN 的思想,特征金字塔网络(feature pyramid network, FPN)用于从不同比例的图像中提取特征,路径聚合网络(path aggregation network, PAN)用于跨网络的不同层聚合这些特征。头部网络负责最终的目标检测和分类任务,包括一个检测头和一个分类头。检测头包含一系列卷积层和反卷积层,用于生成检测结果。分类头则采用全局平均池化来对每个特征图进行分类。

与 YOLOv5、YOLOv6、YOLOv7 相比,YOLOv8 使用了更强大和深层的骨干网络,其骨干网络参考了 YOLOv7

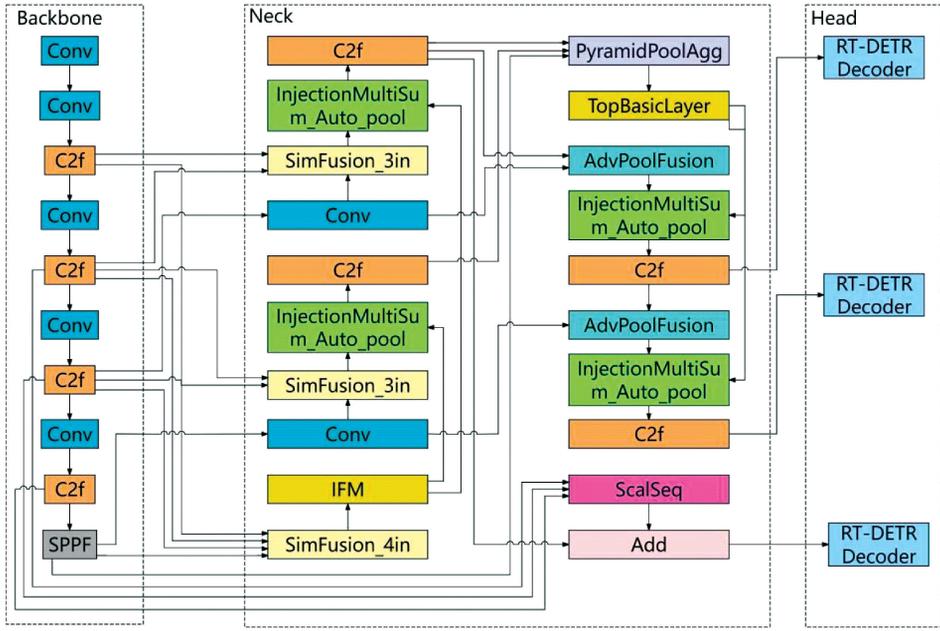


图 2 改进后的 YOLOv8 网络结构

Fig. 2 The improved YOLOv8 network structure

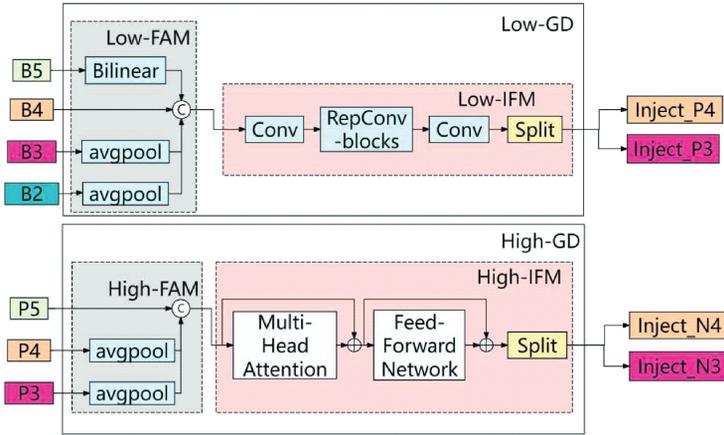


图 3 GD 机制

Fig. 3 Gather-and-Distribute mechanism

Low-GD 模块主要用于融合模型浅层的特征信息。该模块通过特征对齐模块(FAM)将不同层级的特征统一到 B4 层的特征尺寸,如式(1)所示,并通过信息融合模块(IFM)将这些特征进行融合,如式(2)所示。随后,注入模块(Inject)将全局信息注入到浅层特征中,实现了低层级信息的有效交互和融合。

$$F_{align} = Low_FAM([B2, B3, B4, B5]) \quad (1)$$

$$F_{fuse} = RepBlock(F_{align})$$

$$F_{inj_P3}, F_{inj_P4} = Split(F_{fuse}) \quad (2)$$

其中, F_{align} 是 Low-FAM 模块特征对齐的结果, F_{fuse} 是 Low-FAM 的输出经过多层 RepBlock 提取信息得出的结果, RepBlock 生成的特征在通道维度上分为 F_{inj_P3} , F_{inj_P4} , 然后与 P3, P4 层的特征进行融合。

High-GD 模块则专注于融合模型深层的特征信息。在这一模块中, High-FAM 将 P5 层的尺寸作为特征对齐的目标尺寸进行特征对齐, High-IFM 的结构有所变化,由于 High-FAM 输出的特征尺寸变小了,为了充分融合全局信息,将 Low-IFM 中的 RepBlock 替换成了 transformer 模块。涉及到的相应公式如式(3)、(4)所示。Inject 模块将全局信息注入到深层特征中,进一步增强了模型对于烟盒缺陷部位的检测能力。

$$F_{align} = High_FAM([P3, P4, P5]) \quad (3)$$

$$F_{fuse} = Transformer([P3, P4, P5])$$

$$F_{inj_N4}, F_{inj_N5} = Split(Conv1 \times 1(F_{fuse})) \quad (4)$$

其中, F_{align} 是 High-FAM 模块特征对齐的结果, F_{fuse} 是 F_{align} 经过 transformer block 融合得出的结果,通过

Split 操作将 F_{fuse} 沿着通道维数划分为 F_{inj_N4}, F_{inj_N5} , 为后续 Inject 模块做准备。

Inject 模块使用注意力操作融合信息, 可以更有效地将全局信息注入到不同的级别中去。当前级别的特征和由 IFM 生成的特征分别表示为 F_{local} 和 F_{inj} , 使用两种不同的 Conv 和 F_{inj} 进行计算。由于当前级别的特征和 IFM 生成的特征的大小不一, 使用平均池化或者双线性插值来将大小统一, 如式 (5)、(6) 所示。然后通过注意力计算将特征进行融合如式 (7) 所示, 最后经过 RepBlock 进一步提取和融合信息如式 (8) 所示。

$$F_{global_act_Pi} = \text{resize}(\text{Sigmoid}(\text{Conv}_{act}(F_{inj_Pi}))) \quad (5)$$

$$F_{global_embed_Pi} = \text{resize}(\text{Conv}_{global_embed_Pi}(F_{inj_Pi})) \quad (6)$$

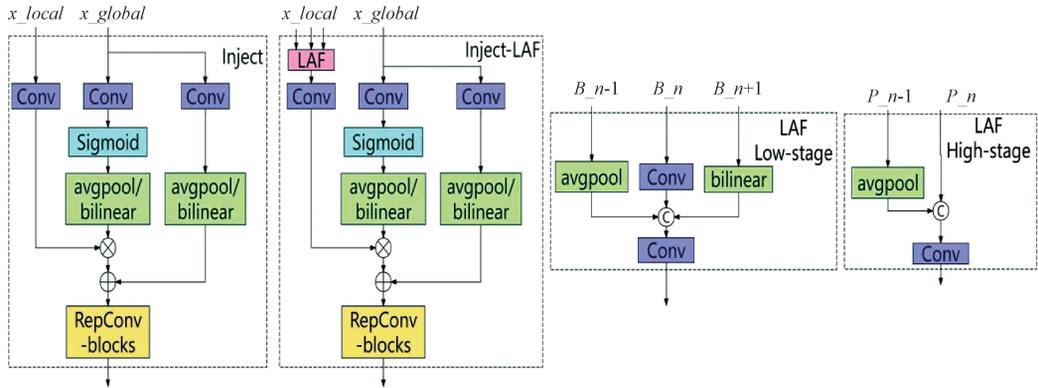


图 4 信息注入模块

Fig. 4 Information injection module

2.3 尺度序列特征融合

本文提出的尺度序列特征融合模块旨在解决烟盒缺陷中划痕的多尺度问题。该模块通过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, 能够更好地处理图像尺度的变化。在下采样过程中, 尺度不变特征可以保持不变, 因此 SSFF 模块利用这一特性来聚合网络不同层的特征。通过融合多个尺度的特征, SSFF 模块能够提供更为丰富和细致的特征表示, 从而改善模型在处理不同大小烟盒划痕时的性能。

SSFF 模块输入的缩放图像可以通过以下方式获得:

$$F_{\sigma}(\omega, h) = G_{\sigma}(\omega, h) \times f(\omega, h) \quad (9)$$

$$G_{\sigma}(\omega, h) = \frac{1}{2\pi\sigma^2} e^{-\frac{(\omega^2+h^2)}{2\sigma^2}} \quad (10)$$

其中, $f(\omega, h)$ 表示具有宽度 w 和高度 h 的二维输入图像, $F_{\sigma}(\omega, h)$ 是通过使用二维高斯滤波器 $G_{\sigma}(\omega, h)$ 对一系列卷积进行平滑而生成的。 σ 是用于卷积的二维高斯滤波器的标准差的缩放参数。

SSFF 模块将不同尺度的特征图水平堆叠, 并使用三维卷积提取它们的尺度序列特征。由于高分辨率特征图 P3 层包含大部分小目标检测至关重要的信息, 因此 SSFF 模块主要应用在 P3 层。如图 5 所示, SSFF 模块首先使用 1×1 的卷积改变 P4 和 P5 层的通道数为 256, 以确保通道

$$F_{att_fuse_Pi} = \text{Conv}_{local_embed_Pi}(F_{local}) \times F_{global_act_Pi} + F_{global_embed_Pi} \quad (7)$$

$$P_i = \text{RepBlock}(F_{att_fuse_Pi}) \quad (8)$$

为进一步增强 Inject 模块的性能, GD 机制引入了轻量级的邻层融合模块 (lightweight adjacent layer fusion, LAF) 如图 4 所示, 对于输入的局部特征 (表示为 B_i 或 P_i), 首先进行与邻域特征的融合操作, 以促进局部特征的上下文理解。随后, 经过该融合操作的局部特征图将进入注入模块, 以接受更多层次的信息传递和交互。这样, 局部特征图不仅能够捕获局部区域的特征, 还能够获得来自周围环境的多层次信息, 使得局部特征具有更多的信息层级, 进而提高了模型的准确性和稳定性。

数一致性。使用最近邻插值方法将其大小调整为 P3 层的大小, 以保持尺度一致性。使用 unsqueeze 方法增加每个特征层的维度, 将其从三维张量 (高度, 宽度, 通道) 更改为四维张量 (深度, 高度, 宽度, 通道)。然后沿深度维度拼接四维特征图, 以形成用于后续卷积的三维特征图。最后使用三维卷积、三维批量归一化和 SiLU 激活函数完成尺度序列特征提取, 通过这些操作, 模型可以有效地提取尺度序列特征, 捕获不同尺度下的目标信息和特征细节, 从而提高目标检测的准确性和鲁棒性。

在整个处理过程中, SSFF 模块通过适当的尺度调整、维度转换和特征融合操作, 有效地整合了不同尺度的特征信息, 并利用三维卷积操作来提取更丰富的尺度序列特征, 从而帮助模型更好地理解 and 处理复杂的烟盒缺陷检测任务。这些步骤的结合能够显著提升模型性能, 特别是在处理具有多尺度特征的烟盒图像数据时表现突出。

2.4 RT-DETR Decoder

YOLOv8 中的 NMS 算法用于抑制冗余的边界框, 保留最具代表性的目标框。在检测的过程中通常会生成多个候选边界框, 每个边界框都对应着一个可能的目标。由于图像中可能存在多个重叠的边界框, 为了提取出最准确的目标框, 需要使用 NMS 来进行筛选。但是 NMS 会延迟推理速度, 并且不够鲁棒, 需要手动挑选合适的超参数才

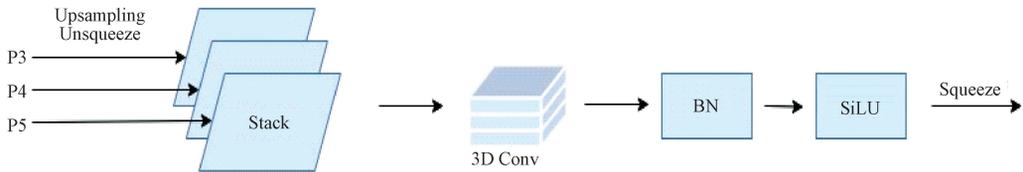


图5 尺度序列特征融合

Fig. 5 Scale sequence feature fusion

能达到最优精度。针对这一缺点,Carion等^[20]提出了不需要NMS后处理的基于Transformer的端到端目标检测器DETR。DETR系列检测器不需要NMS处理,但是与YOLO系列检测器相比速度并未得到提升。随后Lyu等^[19]提出的RT-DETR,从根源上解决NMS对实时检测器带来的速度延迟问题。

RT-DETR中的检测头变换器解码器部分主要负责处理经过混合编码器加工后的特征,并对这些特征进行目标检测。变换器解码器利用Transformer的强大能力来捕捉和理解图像中的复杂模式。它通过自注意力机制来分析 and 加权输入特征,从而能够更准确地定位和识别图像中的不同对象。这种方法使RT-DETR能够在保持高速度的同时,提供比传统方法更精准检测结果。辅助预测头则是变换器解码器的一个关键组件,它被用来进一步提升模型的检测性能。这些预测头直接作用于解码器的输出,负责生成最终的检测结果,包括对象的类别、位置和大小等信息。

RT-DETR Decoder的架构与DETR类似,利用Transformer中的自注意力来捕获不同查询之间的关系,从而建立分数差异来抑制冗余边界框^[21]。在解码器的每个级别中,查询逐渐细化,从而产生一对一的对象集。这极大简化了目标检测过程,并消除了对NMS的依赖。RT-DETR Decoder的结构如图6所示。RT-DETR Decoder主要包括多尺度可变形注意力机制(multi scale deformable attention)、多头自注意力机制(multi-head attention)和前馈神经网络(feed-forward network, FFN)。多尺度可变形注意力机制模块只关注参考周围的一小组关键采样点,该模块可以自然地通过注意力机制聚合多尺度特征图,而无需借助特征金字塔网络,利用多尺度可变形注意模块处理特征映射。多头自注意力机制主要是协助模型捕获不同位置序列之间的依赖关系和相互作用,然后对不同的特征进行关系建模。前馈神经网络是将上一层得到的结果进行非线性变换生成新的表示。

3 实验分析

3.1 数据集制作

本文的数据集来源于荷花香烟盒以及实际生产车间。针对烟盒在制作中容易出现的缺陷类型制作缺陷数据集。本文将烟盒缺陷类型分为4类分别是划痕、胶点、污渍和

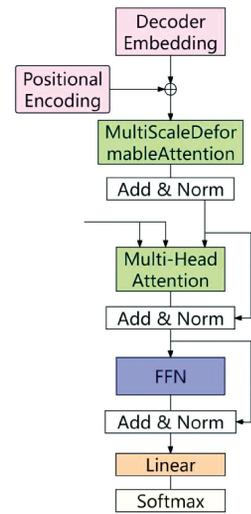


图6 RT-DETR Decoder

Fig. 6 RT-DETR Decoder

侧腰银线未对齐。一共拍摄了450张样本,后续使用图像处理将原始样本进行亮度变换、旋转、平移、翻转、添加噪声等处理,数据增强将数据集扩充到1800张。使用LabelImg工具对样本进行标注,标注类别分为4类:划痕(scratch)、胶点(stain)、污渍(blots)、侧腰银线未对齐(unjustified)。将数据集按6:2:2比例划分为训练集、验证集、测试集,其中训练集1080张、验证集360张、测试集360张。

3.2 训练环境及评价指标

本文实验使用AutoDL服务器,操作系统为ubuntu20.04, GPU为RTX3080Ti, CPU为12 vCPU Intel(R) Xeon(R) Silver 4214R CPU@2.40 GHz,内存90 GB,深度学习框架为Pytorch, Pytorch版本为2.0.0, Python版本为3.8, Cuda版本为11.8。训练时,图片大小为640×640,训练轮次为300轮,批量大小为4,使用AdamW优化器,网络初始学习率为0.0001。

本文为了验证改进后的网络对于烟盒缺陷检测的效果,使用mAP和FPS作为评价指标^[22],其中mAP的计算涉及到准确率(P)、召回率(R)以及平均精度(AP),准确率、召回率的公式如下:

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

其中,TP表示预测正确的正样本数量,FP表示负样

本被预测正确的数量, FN 表示正样本被预测错误的数量。 AP 是以召回率为横坐标, 准确率为纵坐标所绘制的曲线与坐标轴的面积, 公式如下:

$$AP = \int_0^1 P(R) dR \quad (12)$$

mAP 为烟盒各缺陷类别 AP 的平均值, 计算公式如下:

$$mAP = \frac{\sum_{i=0}^n AP(i)}{n} \quad (13)$$

FPS 为模型每秒检测的帧数, 数值越大说明模型检测速度越快。

3.3 实验结果分析

图 7 展示了本文的网络与原始的 YOLOv8 网络在相同的配置下进行 300 轮训练的 mAP 曲线。图中橘色曲线代表使用本文网络训练时的 mAP 曲线, 蓝色代表使用原始 YOLOv8 网络训练时的 mAP 曲线。横坐标表示训练的轮数, 纵坐标表示 mAP 的值。从图中的结果可以看出: 与原始的 YOLOv8 网络相比, 在相同的训练轮数下, 改进后网络模型的 mAP 为 94.6%, YOLOv8 网络模型 mAP 为 93.1%, 提高了 1.5%。

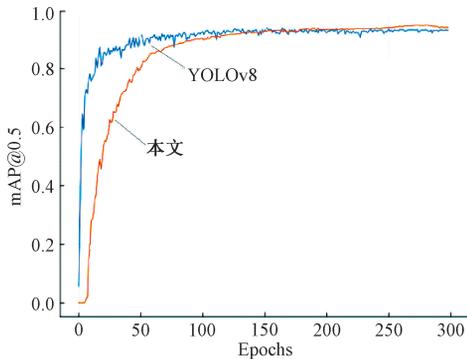


图 7 $mAP@0.5$ 曲线

Fig. 7 $mAP@0.5$ curve

PR 曲线横坐标为召回率, 纵坐标为准确率, 其中横纵坐标与曲线围成的面积表示平均精度均值。图 8 是 YOLOv8 网络模型各缺陷类别的平均精度均值, 图 9 是本文改进后网络模型各类别的平均精度均值。从两图中可以看出, 本文改进后模型缺陷类别的 mAP 有明显提升, 其中划痕提升了 0.7%, 胶点提升了 6.2%, 侧腰银线未对齐提升 1.2%。

检测结果如图 10 所示, 其中图 10(a) 为 YOLOv8 的检测结果, 图 10(b) 为本文算法的检测结果。从图中可以看出, YOLOv8 算法对于污渍缺陷出现了漏检情况, 并且被检测目标的置信度普遍较低。改进之后的算法能够检测出 YOLOv8 算法未检测出来的缺陷, 同时置信度有明显的提升, 显著降低了烟盒缺陷的漏检率。本文算法在烟盒缺陷检测方面表现不错, 相较于原始的 YOLOv8 算法, 大大

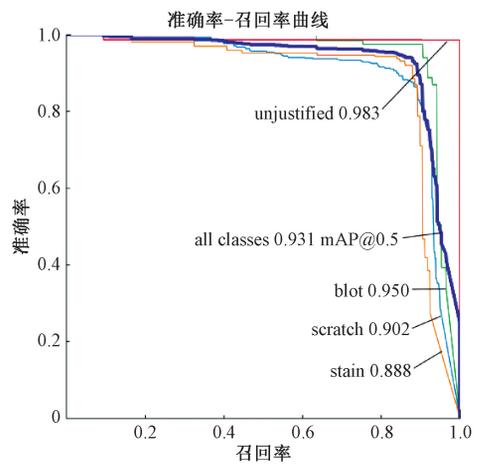


图 8 原始 YOLOv8 模型的 PR 曲线

Fig. 8 PR curve of the original YOLOv8 model

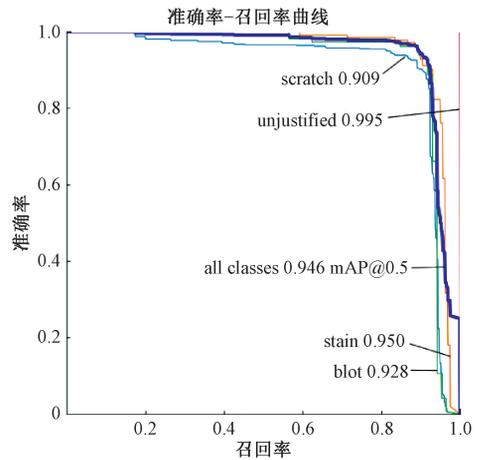


图 9 改进后模型的 PR 曲线

Fig. 9 PR curve of the improved model

减少了出现漏检、错检的情况, 并且提高了检测精度。通过实验综合考虑分析检测精度、检测速度, 本文提出的算法更适合应用于烟盒缺陷检测。

3.4 对比实验

为了验证本文所提出的算法的检测性能, 将本文算法与双阶段目标检测算法 Faster-RCNN、单阶段 SSD 算法以及 YOLO 系列的 YOLOv5 和 YOLOv8 进行了比较, 结果如表 1。从表中可以看出, 双阶段的 Faster-RCNN 算法由于需要生成候选区域, 导致模型检测速度较慢, 不能满足烟盒缺陷检测的实时性要求。同时又由于生成的候选区域较大, 对于胶点或者污渍这种小目标的检测就会出现漏检的情况。SSD 算法的检测精度相比较于 Faster-RCNN 算法提升了 6.7%, 但是它的检测速度比 YOLO 系列的慢。YOLOv5 算法对于在实际生产中最容易出现的划痕这种多尺度的缺陷检测效果并不理想。与 YOLOv5 算法相比 YOLOv8 的 mAP 提升了 3.3%, 在检测速度相差不大的情况下, 对于烟盒缺陷的检测效果有一定的提升。本



(a) YOLOv8的检测结果
(a) Detection results of YOLOv8



(b) 本文算法的检测结果
(b) Detection results of the proposed algorithm

图 10 检测结果对比

Fig. 10 Comparison of detection results

表 1 本文数据集对比实验结果

Table 1 The data set in this paper compares the experimental results

模型	scratch	stain	blot	unjustified	mAP/%	FPS/fps
Faster-RCNN	82.1	78.6	79.5	85.4	81.7	31.2
SSD	87.7	80.0	83.4	89.7	88.4	98.3
YOLOv5	83.3	92.3	84.2	93.5	89.8	113.6
YOLOv8	90.2	88.8	95.0	98.3	93.1	119.7
Ours	90.9	95.0	92.8	99.5	94.6	121.4

文的算法对于这 4 种烟盒缺陷的检测效果最好,整体的 mAP 达到了 94.6%,并且检测速度也能够满足需求。造成这样的结果的原因是本文算法采用了 GD 机制,增强了不同层级特征之间的交互和信息传递,模型能够更全面地理解图像内容;引入了尺度序列特征融合模块,使得模型更好地捕捉不同尺度目标的特征信息;使用了 RT-DETR Decoder,不需要 NMS 后处理,提高了模型的检测速度。本文算法在缺陷检测方面表现出了不错的效果,相较于 YOLOv8 基础模型提高了 1.5% 的检测精度,同时检测速度也提升到了 121.4 fps。

为了进一步验证本文算法的泛化能力和鲁棒性,本文使用 PUK-Market-Phone 数据集进行改进前后算法的对比验证。该数据集是由北京大学发布,其中包括 3 类手机屏幕表面缺陷:划痕(Scr)、斑点(Sta)和油污(Oil),总共

1 200 张图像,按照 6 : 2 : 2 将数据集划分训练集、验证集、测试集。

实验结果如表 2 所示,从表中可以看出,本文算法在该数据集上的 mAP 要高于 YOLOv8 算法,提升了 1.7%,并且这 3 类手机屏幕表面缺陷的 mAP 也均有提高。表明本文算法具备良好的泛化能力,能够有效地应对不同类型和场景下缺陷检测任务。

表 2 PUK-Market-Phone 数据集对比实验结果

Table 2 Comparative experimental results on PUK-Market-Phone dataset

模型	Scr	Sta	Oil	mAP/%
YOLOv8	91.4	88.1	91.9	90.6
Ours	93.5	90.3	93.0	92.3

3.5 消融实验

为了验证本文算法中各个模块的改进效果,本文使用消融实验对部分改进模块进行消融对比,使用 mAP 值作为目标检测模型检测精度的衡量指标。以 YOLOv8 算法为基础模型,消融实验结果如表 3 所示,“√”代表在算法中使用该改进方法“—”代表未使用该改进方法。由表可知,在 YOLOv8 的基础模型上加入 GD 机制之后,增强了颈部的信息融合能力,提高了模型在不同物体尺寸上的性能,使得 mAP 提高了 0.6%。随后又加入了尺度序列特征融合模块,增强了网络在多尺度信息提取上的能力,提高了烟盒缺陷检测的效率,mAP 提升了 0.3%。最后用 RT-DETR Decoder 替换了 YOLOv8 的头部网络,克服了 NMS 对实时检测器推理速度的延迟和对精度的影响,使得 mAP 提升了 0.6%。

表 3 消融实验结果

Table 3 Results of ablation experiments				%
YOLOv8	GD 机制	尺度序列特征融合	RT-DETR Decoder	mAP
√	—	—	—	93.1
√	√	—	—	93.7
√	√	√	—	94.0
√	√	√	√	94.6

4 结 论

为了避免人工检测烟盒表面质量易发生漏检、错检等问题,本文提出了基于改进 YOLOv8 的烟盒缺陷检测模型。首先改进基础模型的颈部网络,将原来的 FPN 结构替换成 GD 机制,这使得模型能够更好的融合非相邻层之间的特征信息。其次,为了提高模型对不同尺度烟盒缺陷的识别能力,在颈部网络中又加入了尺度序列特征融合模块。最后为了能够满足烟厂对于烟盒表面质量检测的实时性要求,替换了需要 NMS 处理的 YOLOv8 头部网络,模型的检测速度得以提升。为了评估本文所提模型的性能,使用相同的烟盒缺陷数据集在不同算法上进行了对比,对比实验得出的结果证实本文模型在检测精度和速度上都有明显提升。后续需增加样本和场景的多样性,进一步验证模型的泛化能力,以满足实际生产环境中的需求。

参考文献

[1] 吴玉生,李安虎,万亚明,等.基于机器视觉的烟草在线检测技术研究进展[J].激光与光电子学进展,2024,61(8):49-62.
WU Y SH, LI AN H, WAN Y M, et al. Research progress of online tobacco detection technology based on machine vision [J]. Laser & Optoelectronics Progress,2024,61(8):49-62.

[2] 赵朗月,吴一全.基于机器视觉的表面缺陷检测方法研究进展[J].仪器仪表学报,2022,43(1):198-219.
ZHAO L Y, WU Y Q. Surface defect detection method based on machine vision research progress[J]. Journal of Instruments and Meters, 2022, 43(1): 198-219.

[3] CHEN ZH Q, DENG J H, ZHU Q Q, et al. A systematic review of machine-vision-based leather surface defect inspection [J]. Electronics, 2022, 11 (15): 2383.

[4] 史浩琛,金致远,唐文婧,等.基于深度学习的高精度晶圆缺陷检测方法研究[J].电子测量与仪器学报,2022,36(11):79-90.
SHI H CH, JIN ZH Y, TANG W J, et al. High precision wafer defect detection method based on the deep learning study [J]. Journal of Electronic Measurement and Instrument, 2022, 36(11): 79-90.

[5] ZHANG H M, ZHAO Y, HUANG SH CH, et al. A deep learning-based surface defects detection and color classification method for solar cells [J]. Journal of Circuits, Systems and Computers, 2023, 32(9): 2350156.

[6] 张关应,束云豪,常宸铠,等.基于深度学习的电力杆塔缺陷检测[J].电子测量技术,2024,47(3):116-126.
ZHANG G Y, SHU Y H, CHANG CH K, et al. Power tower defect detection based on depth study[J]. Journal of Electronic Measurement Technology, 2024, 47(3): 116-126.

[7] REDMON J, DIVVALA S, GIRSHICK R, et al. You only look once: Unified, real-time object detection [C]. Las Vegas Computer Vision & Pattern Recognition. IEEE, 2016.

[8] REDMON J, FARHADI A. YOLOv3: An incremental improvement [J]. ArXiv preprint arXiv: 1804.02767, 2018.

[9] TERVEN J, CORDOVA-ESPARZA D. A Comprehensive Review of YOLO: From YOLOv1 to YOLOv8 and Beyond [J]. Arxiv preprint arxiv: 2304.00501, 2023.

[10] 张恒,程成,袁彪,等.基于 YOLOv5-EA-FPNs 的芯片缺陷检测方法研究[J].电子测量与仪器学报,2023,37(5):36-45.
ZHANG H, CHENG CH, YUAN B, et al. Research on chip defect detection method based on YOLOv5-EA-FPNs[J]. Journal of Electronic Measurement and Instrument, 2023,37(5): 36-45.

[11] 陈亚芳,廖飞,黄新宇,等.多尺度 YOLOv5 的太阳能电池缺陷检测[J].光学精密工程,2023,31(12):1804-1815.

- CHEN Y F, LIAO F, HUANG X Y, et al. Multi-scale YOLOv5 solar cell defect detection[J]. *Optical and Precision Engineering*, 2023, 31(12):1804-1815.
- [12] 王春梅,刘欢. YOLOv8-VSC:一种轻量级的带钢表面缺陷检测算法[J]. *计算机科学与探索*, 2024, 18(1): 151-160.
- WANG CH M, LIU H. YOLOv8-VSC: A lightweight steel strip surface defect detection algorithm [J]. *Journal of Computer Science and Exploration*, 2012, 18(1):151-160.
- [13] ZHANG Y, NI Q J. A novel weld-seam defect detection algorithm based on the s-yolo model[J]. *Axioms*, 2023, 12(7): 697.
- [14] 陈同宇,陈洋,范剑伟. 改进 YOLOv3 算法的烟盒缺陷检测[J]. *物联网技术*, 2021, 11(4): 51-52, 55.
- CHEN T Y, CHEN Y, FAN J W. Improved YOLOv3 algorithm for cigarette case defect detection[J]. *Journal of Internet Technology*, 2021, 11(4): 51-52, 55.
- [15] PENG Y, JIANG D, LYU X ZH, et al. Efficient and high-performance cigarette appearance detection based on YOLOv5[C]. Xi'an: 2023 International Conference on Intelligent Perception and Computer Vision (CIPCV). IEEE, 2023: 7-12.
- [16] 陈文兵,车文刚,蔡小尧,等. 基于修复对抗网络的烟包表面缺陷检测方法[J]. *包装与食品机械*, 2023, 41(2): 58-62, 68.
- CHEN W B, CHE W G, CAI X Y, et al. Cigarette packet surface defect detection method based on repair adversarial network [J]. *Packaging & Food Machinery*, 2023, 41(2): 58-62, 68.
- [17] WANG CH CH, HE W, NIE Y, et al. Gold-YOLO: Efficient Object Detector via Gather-and-Distribute Mechanism [J]. *ArXiv preprint arXiv: 2309. 11331*, 2023.
- [18] 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.
- [19] LYU W, XU S, ZHAO Y, et al. DETRs beat YOLOs on real-time object detection [J]. *ArXiv preprint arXiv:2304. 08069*, 2023.
- [20] CARION N, MASSA F, SYNNAEVE G, et al. End-to-end object detection with transformers[C]. *European Conference on Computer Vision*. Cham: Springer International Publishing, 2020: 213-229.
- [21] ZHANG H, LI F, LIU SH L, et al. Dino: Detr with improved denoising anchor boxes for end-to-end object detection[J]. *ArXiv preprint arXiv:2203. 03605*, 2022.
- [22] 杨永跃,夏远超. PCB 缺陷检测深度学习算法的精度改进[J]. *电子测量与仪器学报*, 2023, 37(5): 11-19.
- YANG Y Y, XIA Y CH. PCB defect detection precision of deep learning algorithm to improve [J]. *Journal of Electronic Measurement and Instrument*, 2023, 5(5): 11-19.

作者简介

王震洲, 博士, 教授, 主要研究方向为图像处理、模式识别、测控技术。

杨榕, 硕士研究生, 主要研究方向为图像处理、计算机视觉。

宿景芳(通信作者), 博士, 讲师, 主要研究方向为机器视觉、深度学习和机器人等。

E-mail: 18272565@qq.com