

# 基于改进 RT-DETR 的车门内拉手表面缺陷检测方法<sup>\*</sup>

徐仟祥 曾勇 卢倩 南玉龙

(盐城工学院机械工程学院 盐城 224000)

**摘要:** 针对车门内拉手表面的缺陷目标小、多尺度、易反光等问题。首先,通过使用碗状光源和降低图像采集表面夹角的方法,解决内拉手表面图像采集时因表面弯曲和镜面反射导致的缺陷特征被覆盖问题。然后,针对传统的 RT-DETR 模型存在缺陷检测精度差,速度慢等问题,提出一种改进的 RT-DETR 目标检测方法。该方法首先以 RT-DETR 为基础架构,在主干网络中采用并行的膨胀卷积与 CA 注意力机制并结合卷积重参数化的方式,以增加网络感受野和建立长距离的语义信息的同时提高网络推理速度。其次,通过添加额外的检测层来增加网络对小目标检测的特征提取能力。紧接着,在多尺度特征融合阶段使用了改进的 BIFPN 结构以提高模型信息交互的能力。最后,消融实验表明,相较于传统的基于 RT-DETR 的检测方法,本文提出的改进 RT-DETR 的检测方法,平均精度提升了 6.5%,检测速度为传统模型的 1.6 倍,同时模型的参数量仅为原网络的 76.5%。验证了本文所提方法的有效性。

**关键词:** 车门内拉手;缺陷检测;RT-DETR;膨胀卷积;卷积重参数化;BIFPN

中图分类号: TP391;TP29;TN0 文献标识码: A 国家标准学科分类代码: 510.4050

## Surface defect detection method for inner handle of car door based on improved RT-DETR

Xu Qianxiang Zeng Yong Lu Qian Nan Yulong

(College of Mechanical Engineering, Yancheng Institute of Technology, Yancheng 224000, China)

**Abstract:** To address the challenges of small defect targets, multi-scale issues, and high reflectivity on the surface of the inner car door handle, we first tackle the problem of defect features being obscured during image acquisition due to surface curvature and mirror reflection by using a bowl-shaped light source and reducing the angle of the image acquisition surface. Then, recognizing the limitations of traditional RT-DETR models, such as poor detection accuracy and slow speed, we propose an improved RT-DETR object detection method. This method builds upon the RT-DETR framework, utilizing parallel dilated convolutions and the CA attention mechanism combined with convolutional re-parameterization in the backbone network to increase the receptive field and establish long-distance semantic information while improving the network inference speed. Additionally, extra detection layers are added to improve the network's feature extraction capability for small object detection. In the multi-scale feature fusion stage, we use an improved BIFPN structure to enhance the model's information interaction capability. Finally, ablation experiments show that, compared to traditional RT-DETR-based detection methods, our proposed improved RT-DETR method increases the mean Average Precision by 6.5%, achieves a detection speed 1.6 times that of the traditional model, and reduces the model's parameter count to only 76.5% of the original network, validating the effectiveness of our proposed method.

**Keywords:** door handle; defect detection; RT-DETR; dilated-Conv; convolutional re-parameterization; BIFPN

## 0 引言

车门内拉手作为汽车车门中的重要部件,由于受到制

造工艺和物料输送环境的影响,其表面易出现凹坑、浅划痕、深划痕等质量缺陷。因此车门内拉手在出厂前必须进行表面缺陷检测,以确保出厂质量。当前,传统的人眼检测

因检出率低、劳动强度大等问题,已无法适应当今企业对高质、高效、低成本质检的要求。

由于车门内拉手表面存在弯曲和高反光特性,导致局部高亮饱和淹没缺陷特征,降低了表面缺陷检测的成功率。针对高反光金属表面,目前通常采用红外检测、超声扫描和激光探伤等无损检测方法<sup>[1-2]</sup>。然而,由于这些方法的高昂成本,导致它们仅能在部分高精密零件中进行有限的抽查。与其他检测方法相比,基于机器视觉的表面缺陷检测方法具有成本低、适用范围广、检测效率高,因此广泛应用于诸多领域<sup>[3]</sup>。近年来,众多学者基于机器视觉针对高反光金属工件表面缺陷检测越来越深入。冯维等<sup>[4]</sup>通过改变图像的曝光率,将不同曝光时间下的图片进动态合并,并运用自适应多阈值与 Haar-like 算法来提取缺陷。姜美华等<sup>[5]</sup>利用穹顶光源,在灯罩内部上放置与铁球同心的图案,获得了钢球表面的反射图像,并运用自适应阈值与 Canny 算法提取反射图案的缺陷信息。龚肖杰<sup>[6]</sup>通过采用俩条与板状钢材之间成的 45°夹角的光源来降低金属表面反光的问题,从而获得钢材表面图片,并采用改进的 YOLO 算法进行了缺陷的识别。以上研究方法,在仅存在单个曲面且表面缺陷不复杂的情况下获得了良好的成绩。然而,由于车门内拉手表面存在多个高曲率表面,且表面存在多尺度、小目标缺陷,因此传统的检测方式并不适用。

当前,大多数的目标检测都是使用卷积神经网络(CNN),并在目标检测领域中获得了良好的成绩,如单阶段检测器 SSD<sup>[7]</sup>和 YOLO<sup>[8]</sup>系列、二阶段检测器 R-CNN<sup>[9]</sup>和 Faster R-CNN<sup>[10]</sup>。但无论是单阶段系列还是二阶段系列,都依赖非极大值抑制(non-maximum suppression, NMS)这种后处理方式,以选取置信度最高的候选框进行定位。相较于 YOLO 和 R-CNN,RT-DETR 网络不再利用 NMS 的方式,而是使用查询向量作为软锚框来定位目标<sup>[11]</sup>,减少了 NMS 所需要的时间。但 RT-DETR 原本创作的目的是服务于自然图像处理,在小目标和多尺度目标检查上精度较低。针对 RT-DETR 网络在小目标、多尺度缺陷检测上精度较差,周良钰等<sup>[12]</sup>通过在特征提取网络增加了 SeNet,并结合 FPN 特征金字塔融合提高小目标的检测精度,但并未考虑在多尺度缺陷的情况下感受野不足的问题<sup>[13]</sup>。金睿蛟等<sup>[14]</sup>通过引入多尺度旋转可变形注意力机制和旋转交并比,提高了旋转遥感领域的小目标多尺度检测的精度,但此方法并不适用于旋转特征较少的车门内拉手表面缺陷问题。张柏圆等<sup>[15]</sup>通过将 deformable 卷积的最佳稀疏空间采样方法和 Transformer 的关系建模能力相结合以提高网络在密集场景下的检测能力,但其忽略了 DETR 采用 Transform 结构,具有较大的参数量,在处理复杂任务时实时性较差。

针对上述问题,同时结合车门内拉手表面的固有高反光特性,本文基于设计的车门内拉手表面缺陷机器人自动检测系统,提出了一种基于改进 RT-DETR 的车门内拉手

表面缺陷检测方法。旨在解决检测密集小目标、多尺度缺陷等场景下实时性和精度难以平衡的问题。同时,本文的创新点如下:

- 1) 针对车门内拉手的表面特征,设计提出了一种集成工业机器人的自动化检测系统,实现无人质检的目的。
- 2) 在特征提取网络中增加了小目标检测层,加强了对小尺度缺陷的关注。
- 3) 采用并联多组不同尺度的膨胀卷积,并结合结构重参数化的方式,设计了一种快速多尺度特征提取模块(F-MRFFE),提高了多尺度缺陷检测的精度并减少了网络的计算量。
- 4) 在多尺度特征融合阶段采用改进的 BiFPN 结构,加强了不同层级图像信息交互的能力。

## 1 系统设计与采集方式

### 1.1 检测系统硬件设计

车门内拉手表面缺陷自动检测系统硬件包括工业计算机、工业机器人、输送线、机器视觉系统和装载容器等多个核心组件,如图 1 所示。其中,工业计算机主要用于对检测线中各设备的集成控制,以及对采集图像的处理与识别;工业机器人用于从装载容器中夹取内拉手零件到相机处采集图像;输送线用于装载内拉手零件容器的传输,实现流水作业;机器视觉系统由穹顶光源和工业相机构成,用于对内拉手零件表面缺陷图像的采集;装载容器中分布有规则排布的凹槽,用于固定车门内拉手,如图 2 所示。

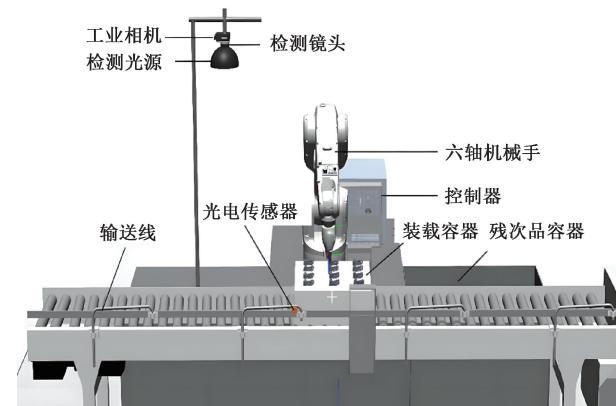


图 1 车门内拉手表面缺陷自动检测系统

Fig. 1 Automatic detection system for surface defects of door inner handle

### 1.2 缺陷检测系统的工作流程

系统控制器通过串口协议控制输送线的启停,运行时输送线将装载容器送至目标位置,并在目标位置,设置光电开关检测装载容器是否达到目标点位。当装载容器通过输送线送到目标位置,系统控制六轴机械手按照事先编写好的机器人运动程序从装载容器中夹取内拉手零件,并将其送至工业相机处进行图像采集,如图 3 所示,采集的图像传

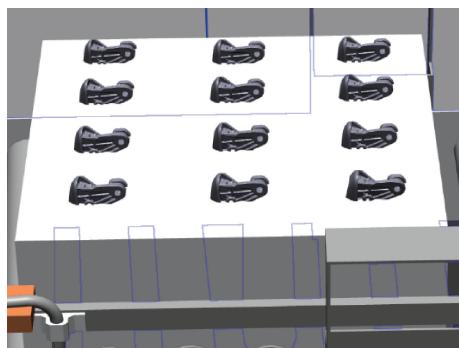


图 2 车门内拉手装载容器

Fig. 2 Handle loading container in the door

输至上位机进行缺陷检测。由于内拉手表面属于复杂形面,需要在机械手的带动下在相机面前呈现多个位姿才能将表面全部采集完整。系统若检测出缺陷,则机械手将有缺陷的内拉手零件送至废品站,否则将其送至原位,直至当前容器中的所有零件全部检测完,此时,再次启动输送线开始下一个装载容器中内拉手表面的检测。系统工作流程如图 4 所示。

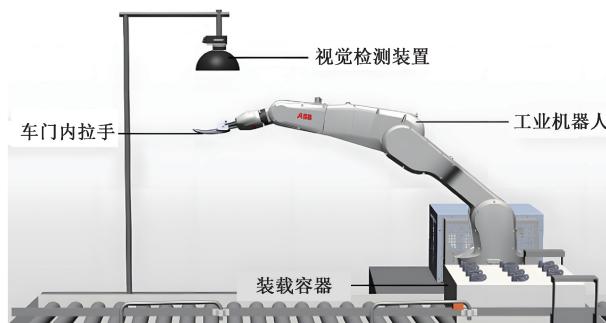


图 3 车门内拉手表面缺陷检测示意图

Fig. 3 Schematic diagram of surface defect detection of door inner handle

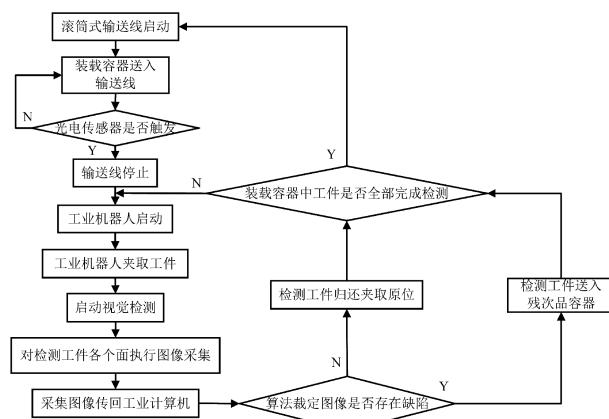


图 4 系统运行流程图

Fig. 4 System operation flow chart

### 1.3 图像数据采集

车门内拉手表面是复杂自由曲面,主要包含 5 个面,其中 2 个是大曲率自由曲面,3 个是近似平面,且具有高反光特性,如图 5 所示。为了避免高反光特性导致的局部亮度饱和淹没特征的问题,图像数据采集首先采用光线更加均匀的碗状 LED 穹顶光源,其次采用分区采集的方法改善局部过曝的问题,如图 6 所示,将曲面法向夹角 30° 范围内的图像采集变为 15° 范围内的图像采集可以消除小范围的过曝。为了确保图像采集质量,对车门内拉手表面进行多角度的 7 次图像采集,即对近似平面进行 1 次采集,对大曲率面分区域 2 次采集,得到如图 7 所示的采集效果。



图 5 车门内拉手

Fig. 5 The handle inside the door

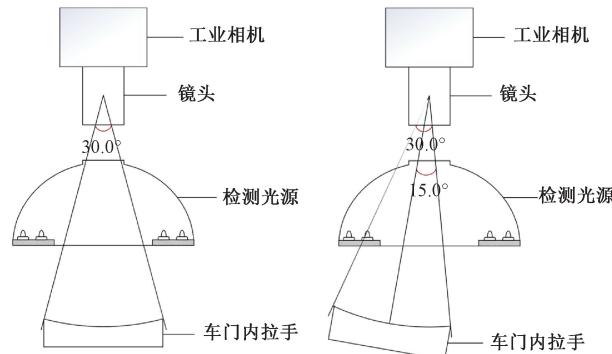


图 6 图像采集示意图

Fig. 6 Schematic image acquisition

## 2 RT-DETR 网络模型

### 2.1 RT-DETR 算法原理

RT-DETR 是一种实时端到端的对象检测网络,利用 Vision Transform(ViT)有效处理输入特征,从而在保证精度的同时提高了实时性能。RT-DETR 整体网络模型主要由三大块组成:首先,是基于 CNN 架构的骨干特征提取网络(本文为了平衡速度和精度采用 ResNet34 作为原始网络的主干网络),用于提取特征。其次,是由基于注意力机制的单尺度内特征交互模块(attention-based intrascale feature interaction, AIFI)和跨尺度特征融合模块(cross-scale feature fusion module, CCFM)构成的高效混合编码器,实现内部尺度交互及跨尺度融合。最后,通过 IOU 感知查询和具有辅助预测头的解码器,实现无后处理(如 NMS

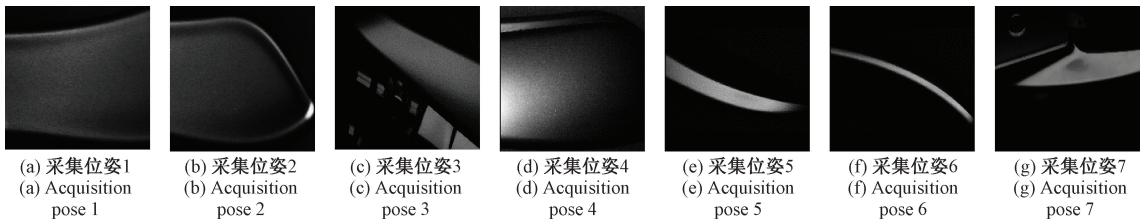


图 7 分区域图像采集

Fig. 7 Sub-regional image acquisition

等)的输出,从而提高了检测效率,减少了推理时间。

## 2.2 改进的 RT-DETR 网络模型

图 8 为本文提出的改进网络模型 PSB-DETR。皆在实时准确的检测小目标与多尺度缺陷。该模型使用快速多尺度感受野特征提取(F-MRFFE)模块堆叠搭建主干网

络,在增大感受野的同时,建立不同尺度和范围特征之间的相互依赖性。针对小目标特征难以提取问题,本文添加了一个新的检测层。此外,在整合多个尺度特征时,为了增强特征信息之间的协同作用,本文对 BIFPN 结构进行了改进,以提高特征信息的全面性。

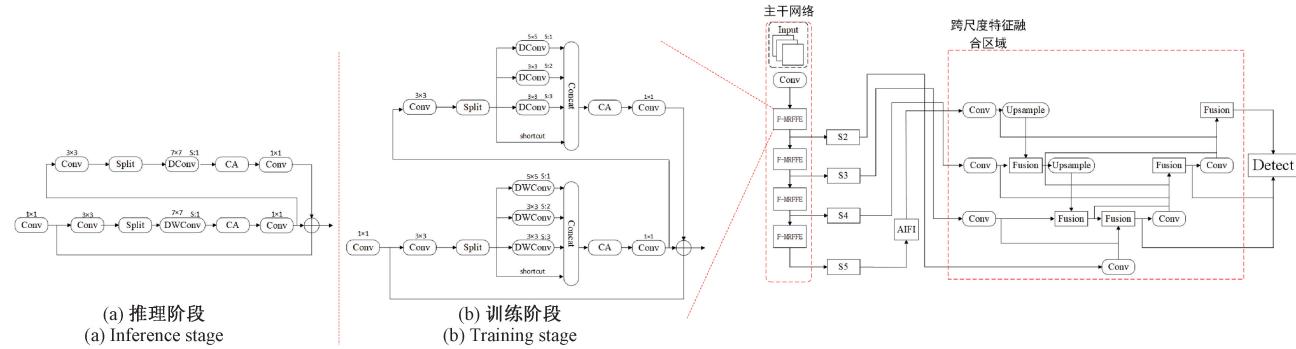


图 8 PSB-DETR

Fig. 8 PSB-DETR

### 1) 小目标检测层

在原始的 RT-DETR 主干网络中,从第 3 次下采样开始,采用基于注意力机制的 AIFI 以及 CCFM 进行特征的融合。由于浅层的特征图具有丰富的轮廓、颜色、纹理等信息。对于大目标的检测,浅层的语义信息对模型最后的输出权重贡献较小。然而,对于一些细小的凹坑缺陷等目标,这些浅层的语义信息却至关重要。在输入含有缺陷的车门内拉手图像后,特征提取网络中第 1 次下采样至第 5 次下采样的通道特征,

如图 9 所示。如图 9(c)所示,在第 2 次下采样中,凹坑轮廓信息依旧丰富;但在第 3 次下采样开始,则提取一些更为抽象的信息。这些信息对不含小目标的网络在目标分类时,有着较大的贡献。因此,传统的 RT-DETR 目标检测网络通常选择从第 3 层开始输出,进行特征融合。然而由于本研究中所含大量的小目标缺陷,因此为了提高网络对小目标的检测能力,本文在原始的 RT-DETR 网络模型中添加了一个小目标检测层,以保留浅层的信息用于后续特征融合。

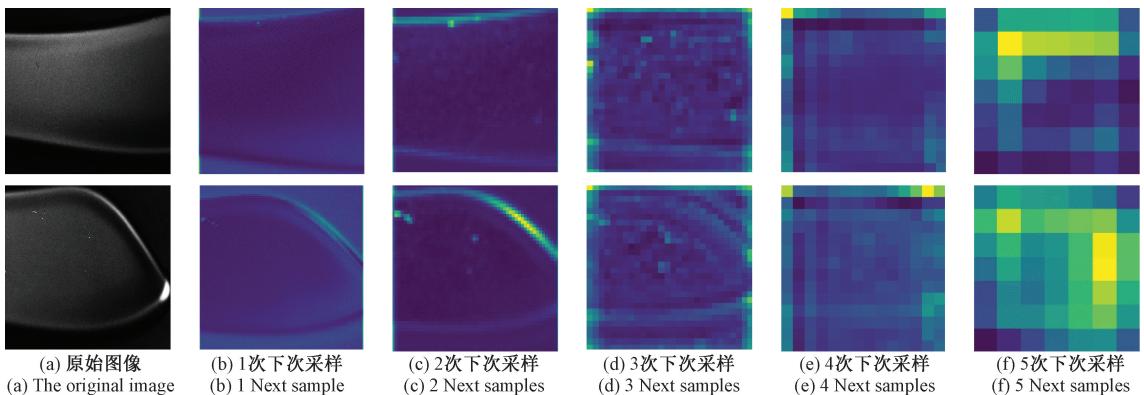


图 9 特征图可视化

Fig. 9 Feature map visualization

## 2) 快速多尺度感受野特征提取模块(F-MRFFE)

RT-DETR 的主干特征提取网络采用的是 ResNet 结构, 其中 ResNet 残差模块仅包含 2 个卷积核大小为  $3 \times 3$  的卷积层, 用于实现图像特征提取。虽然该结构解决了梯度消失问题<sup>[16]</sup>, 但由于感受野的不足以及忽略了语义中的上下文信息, 导致对多尺度缺陷检测的不足。车门内拉手表面缺陷尺寸复杂多样, 划痕长短不一, 位置不固定。为了提高复杂多尺度信息之间的关联性并增大感受野, 以获取更为广泛的语义信息, 同时提升模型在推理时的速度, 本文设计了快速多尺度感受野特征提取模块, 它分为训练和推理俩阶段。在训练阶段, 首先由小内核的普通卷积进行输入图片整体信息的提取, 如式(1)所示。接着, 对输入特征图片进行切片操作, 得到 4 张通道维度相同的子图。然后, 使用一组膨胀因子 S 分别为 1, 2, 3 的膨胀卷积(DConv)并行连接, 用于提取多个尺度的上下文信息, 如式(2)所示。随后, 通过 CA 注意力<sup>[17]</sup>机制建立缺陷之间的短距离的相互依赖关系。接下来利用大小为  $1 \times 1$  的卷积进行通道的整合, 如式(3)所示。最后将得到的特征再次传入上述的结构中, 获取长距离特征, 整体结构如图 10 所示。

$$L_{t-1} = \text{Conv}_{k \times k}(X_t) \quad (1)$$

$$Z_{t-1} = D\text{Conv}_{k^{(m)} \times k^{(m)}}(L_{t-1}), m = 1, \dots, 4 \quad (2)$$

$$P_{t-1} = \text{Conv}_{1 \times 1}\left(\sum_{m=1}^4 Z_{t-1}^{(m)}\right) \quad (3)$$

其中,  $L_{t-1}$  表示通过  $K \times K$  的卷积 Conv 进行局部特征的提取,  $X_t$  表示输入的特征。 $Z_{t-1}$  表示经过并联的 DConv 提取出的多尺度上下文信息,  $m$  表示卷积核的个数。 $P_{t-1}$  表示最后通过  $1 \times 1$  卷积后的输出。

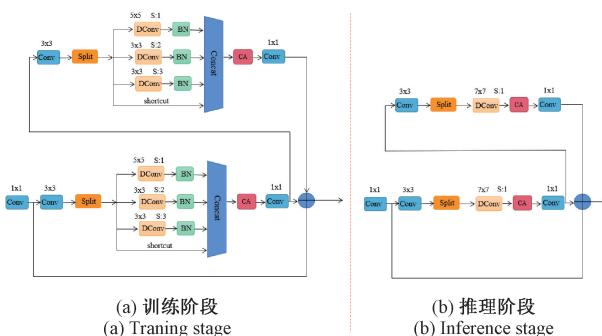


图 10 快速多尺度感受野特征提取模块

Fig. 10 Fast multi-scale receptive field feature extraction module

在推理阶段, 为了提高网络模型的推理速度, 并减少冗余参数的计算。采用卷积重参数化的方式, 将每个并行的膨胀卷积和相应的批归一化层(batch normalization, BN), 利用式(4)进行合并, 得到一个新的卷积。接着, 以实际卷积核尺寸最大的膨胀卷积为基础, 对其他膨胀卷积的实际卷积核大小不足最大卷积的部分进行四周填充零值, 以确保每个膨胀卷积的实际卷积核大小保持一致。最终

将 3 组经过合并填充处理的并行卷积利用式(5)整合成一个  $7 \times 7$  的普通卷积, 整体转换过程如图 11 所示。

$$BN(\text{Conv}(x)) = \frac{\gamma W(x)}{\sqrt{var}} + \left( \frac{\gamma(b - \text{mean})}{\sqrt{var}} + \beta \right) \quad (4)$$

$$y = (w_1 + w_2 + w_3)x + (b_1 + b_2 + b_3) \quad (5)$$

其中,  $w_1, w_2, w_3$  分别代表 3 组已经通过批归一化(BN)层合并的膨胀卷积的权重, 而  $b_1, b_2, b_3$  则代表这 3 组膨胀卷积合并后的偏置。

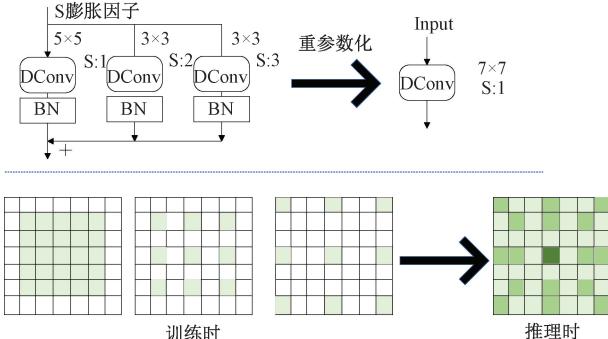


图 11 卷积重参数化示意图

Fig. 11 Schematic structure of convolutional weight parameters

## 3) 改进的 BiFPN 结构

在 RT-DETR 的 CCFM 中, 采用了类似于路径聚合网络(path aggregation network, PANet)网络结构的设计。PANet 网络为了克服单向信息流的限制, 在特征金字塔(feature pyramid network, FPN)的基础上增加了一条额外的自底向顶的路径, 将浅层的语义信息传递到上层, 更多了保留了浅层的特征。然而, 原始 RT-DETR 网络采用在自顶向下和自底向上的结构在传递过程中会损失掉一部分位置和语义的信息, 导致对小目标和多尺度缺陷融合能力不足。加权双向特征金字塔网络(bidirectional feature pyramid network, BiFPN)网络认为如果某个节点只有一条输入边, 则其对于网络的贡献就较小。为了减少计算量, BiFPN 网络移除了第 2 层和第 5 层的特征融合节点。同时, 为加强同层信息的融合, 提出了同层间跨尺度连接的方法, 将同层的输入和输出相连, 在保留深层语义的同时不丢失过多浅层信息。然而, BiFPN 网络在不同层级之间信息融合方面存在不足, 导致对不同尺度大小物体的同时检测能力不足<sup>[18]</sup>。因此, 本研究提出一种改进的 BiFPN 特征融合网络。网络结构如图 12 所示, 其中,  $P_2_{\text{out}}, P_3_{\text{out}}, P_4_{\text{out}}$  和  $P_5_{\text{out}}$  代表了 backbone 网络的 4 个不同层级的输出。 $P_2$  在经过两次下采样后得到  $P_3_{\text{out}}$  和  $P_4_{\text{out}}$ , 而  $P_2$  至  $P_5$  则作为改进 BiFPN 的输入。在自顶向下的特征融合过程中, 网络会获取中间的过渡信息, 计算过程为式(6)、(7), 其中  $P_3^{id}, P_4^{id}$  表示  $P_3$  和  $P_4$  的中间层权重信息,  $\text{Conv}$  表示卷积操作,  $\omega_1, \omega_2$  表示学习到的权重,  $\text{Resize}$  表示上采样操作, 目的是将其形状改变与输出一致,  $\epsilon$  表示偏移量:

$$P_3^{td} = \text{Conv} \left( \frac{\omega_1 \cdot P_3 + \omega_2 \cdot \text{Resize}(P_4)}{\omega_1 + \omega_2 + \epsilon} \right) \quad (6)$$

$$P_4^{td} = \text{Conv} \left( \frac{\omega_1 \cdot P_5 + \omega_2 \cdot \text{Resize}(P_5)}{\omega_1 + \omega_2 + \epsilon} \right) \quad (7)$$

将获取到的中间融合特征重新传输到 P4 和 P5 的输出中,实现了自底向上的特征传递。最终,这些特征被整合并输出。以图中 P4\_out 节点为例,其计算过程如下:

$$P_4^{out} = \text{Conv} \left( \frac{\omega_1 \cdot P_4^{td} + \omega_2 \cdot P_4^{in} + \omega_3 \cdot \text{Resize}(P_3) + \omega_4 \cdot \text{Resize}(P_3^{td})}{\omega_1 + \omega_2 + \omega_3 + \epsilon} \right) \quad (8)$$

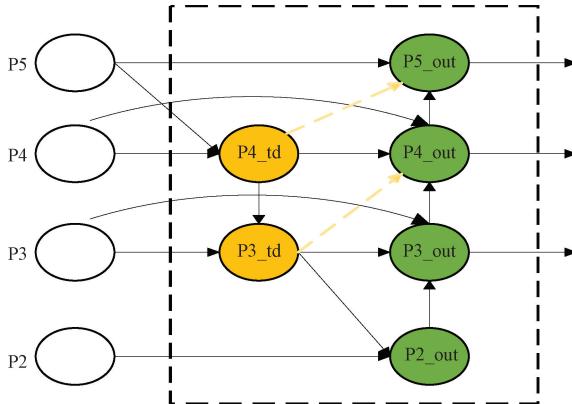


图 12 改进的 BiFPN 网络结构

Fig. 12 Improved BiFPN network architecture

### 3 算法实验与分析

#### 3.1 数据采集

本研究采集了 220 张  $2048 \times 2048$  像素大小的图片。为了提高处理速度和算力,采用了滑动窗口方法进行裁剪。同时为了确保每张图片中存在足够数量的缺陷,且这些缺陷能够覆盖较大的区域,本文将滑动窗口的步长设为 320,窗口设为  $640 \times 640$ 。此外,为了提高算法的鲁棒性,本文对裁剪出来的图片进行了随机的旋转、镜像、扰动缩放等操作,如图 13 所示,最终生成了 3 000 张可供训练的图片,这些缺陷被分成了凹坑(pit)、浅划痕(scratch)和深划痕(Bruise)。并按 3:1:1 的比例划分为训练集、验证集和测试集。

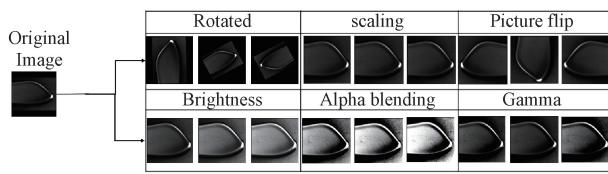


图 13 图像扩充

Fig. 13 Image expansion

#### 3.2 实验环境

本文的硬件实验环境为: 英特尔 Xeon (R) 8350C CPU、英伟达 RTX4070 GPU。软件实验环境为: Windows10 操作系统、Python 语言、Pytorch 学习框架、

CUDA12.0。

#### 3.3 评估指标

在本文中采用平均精度(mAP)来检验模型的分类的准确性和物体的定位性能。检测速度(FPS)来检测模型速度。mAP 的计算公式如下:

$$mAP = \frac{1}{\text{classes}} \sum_{i=1}^{\text{classes}} \int_0^1 P(R) d(R) \quad (9)$$

其中,上式中的  $P$  表示精确度,  $R$  表示召回率。计算公式如下:

$$P = \frac{TP}{TP + FP} \quad (10)$$

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

式中:  $TP$  表示在检测结果中将正确的物体判断为正确的个数,  $FP$  表示将错误的物体判断为正确的个数,  $FN$  表示将正确的物体判断为错误的个数。

#### 3.4 消融实验

为了验证本文所提的 PSB-DETR 各个模块的有效性,设计了一组消融实验,以 RT-DETR 为基准网络,将本文所提出的模块与 RT-DETR 模型组合进行对比实验,实验结果如表 1 所示,○表示模型含有该模块。

表 1 消融实验

Table 1 Ablation experiment

小目标 检测层	F-MRFFE	改进的 BiFPN	mAP/ %	参数/ ( $10^7$ )	FPS
			79.9	3.28	34.1
O			81.8	3.43	32.3
	O		82.3	2.13	58.5
		O	82.0	3.40	34.9
O	O		83.1	2.39	56.3
O	O		85.4	2.45	55.7
O	O	O	86.4	2.51	55.1

由表 1 可知,本文改进的各模块在车门内拉手缺陷检测任务上表现出良好的性能。特别是 F-MRFFE 模块贡献最大,这主要归功于采用卷积重参数化方式并行连接的膨胀卷积,以及 CA 注意力机制,使得多尺度缺陷的特征在充分提取的同时推理速度也大幅增加了,相较于传统 RT-DETR 精度和速度分别提高了 2.4%、24.4 fps。虽然对原始网络添加额外的检测层会在一定程度上降低速度,但却提高了对小尺度目标的识别精度。改进的 BiFPN 模块增强了模型在不同层级间信息融合的能力,进一步提高了多尺度缺陷的识别能力。从表 1 中可以看出,相比基准算法,PSB-DETR 在精度、速度等方面表现更优,能够有效地检测到小目标和多尺寸缺陷,具有良好的车门内拉手表面缺陷检测性能。

### 3.5 对比实验

为了验证本文所提出的方法在车门内拉手上缺陷检测上的优势,本文在相同的车门内拉手表面缺陷数据集的

基础上使用本文的算法 PSB-DETR 对比了现目前主流的算法,结果如表 2。

表 2 性能对比

Table 2 Performance comparison

算法	AP <sub>pit</sub> / %	AP <sub>Scratch</sub> / %	AP <sub>Bruise</sub> / %	mAP@0.5 / %	FPS
SDD	68.8	70.3	71.5	70.2	23.0
Faster R-CNN	75.8	75.6	73.3	74.9	4.3
YOLOV5	81.8	77.4	77.4	78.9	32.6
Multi-CR Yolo <sup>[19]</sup>	83.1	79.5	81.3	80.7	63.0
YOLO-SE <sup>[20]</sup>	82.7	80.2	81.2	81.4	35.2
RT-DETR	81.3	77.2	81.2	79.9	34.1
RT-DETR-Faster <sup>[21]</sup>	83.0	76.5	80.8	80.1	52.0
Ours	<b>86.2</b>	<b>85.4</b>	<b>87.6</b>	<b>86.4</b>	<b>55.1</b>

由表 2 可知,PSB-DETR 在自建的车门内拉手缺陷检测数据集上的精度可以达到 86.4%,较主流的二阶段和一阶段算法都有较大的提升,尤其是在目标尺寸小、缺陷尺度变化大的场合。虽然在速度方面略慢于 Multi-CR Yolo,但在小目标及多尺度缺陷的检测上效果更佳,更适合车门内拉手表明缺陷检测。

图 14 是 PSB-DETR 在车门内拉手表面缺陷检测数据集的部分可视化结果,并对比了基准算法 RT-DETR 的检测结果。从图 14 可以看出,基准算法 RT-DETR 检测精度低,存在漏检错检问题,而 PSB-DETR 能够更好地检测到车门内拉手表面缺陷目标,有效的提高检测的精度,并减少了 RT-DETR 网络的错检误检问题。

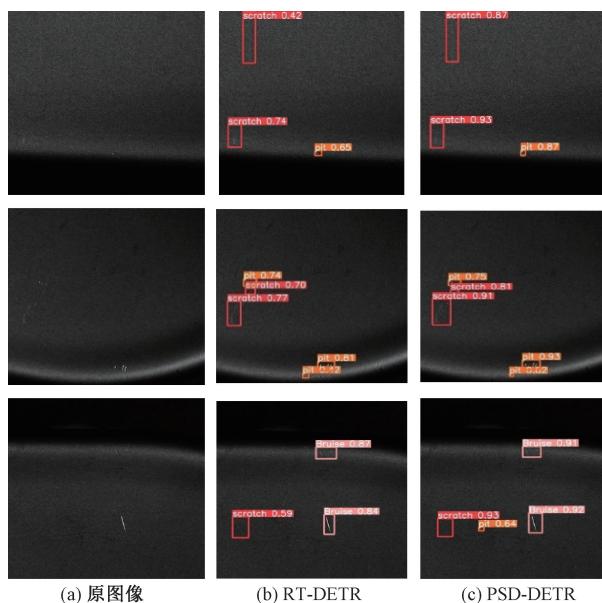


图 14 检测结果图

Fig. 14 Detection result graph

## 4 检测系统实际应用与测试

### 4.1 检测系统软件设计

本系统使用 Winform 框架作为上位机开发平台,采用 C# 语言进行开发和部署改进的 RT-DETR 网络,通过 RobotStudio 软件实现系统仿真与数据交互。在 RobotStudio 软件中设计机器人 7 个采集的位姿,进行图片采集,并将采集到的图片送入上位机中进行缺陷检测,上位机界面如图 15 所示,采集位姿如图 16 所示。



图 15 上位机界面

Fig. 15 UI interface

### 4.2 检测系统实验验证

为了进一步验证本文设计的系统在实际检测中的可靠性和稳定性,本研究搭建了实物平台,如图 17 所示。同时将 600 个已经被人工检测过的车门内拉手作为测试样本,分别通过本研究搭建的系统对比原始算法,与改进的算法。为了数据的可靠性,实验分为 3 组,每组实验检测 200 个车门内拉手,检测结果如表 3 和 4 所示,其中为了更加贴合工程实际需求,本文的检测时间的计算方式为:六

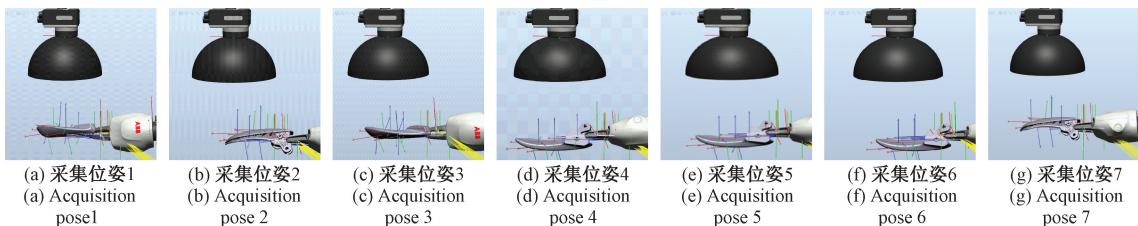


图 16 采集位姿示意图

Fig. 16 Schematic diagram of acquisition pose

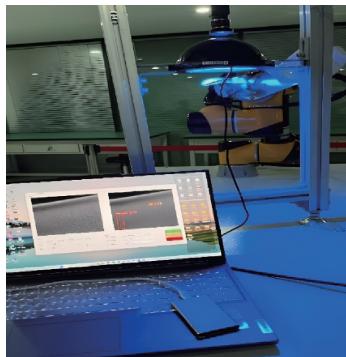


图 17 实物平台

Fig. 17 Physical platform

轴机械手移动时间加采集了 7 次图片并检测的时间。

从表 3、4 中可以得出,本文提出的 PSD-DETR 算法在实际场景中的准确率为 85.83%,远高于原始算法的 80.33%。此外,PSD-DETR 算法在单个车门内拉手的一次性检测时间为 1.65 s/件,优于原始算法的 1.95 s/件。实验证明了系统的可行与可靠性。

表 3 原始算法的实验结果

Table 3 Experimental results of the original algorithm

批次	检测数量	检测不合格数	漏检数	过检数	漏检率/%	误检率/%	准确率/%	检测时间/s
1	200	43	21	17	10.50	8.5	81.00	1.94
2	200	49	18	23	9.00	11.5	79.50	2.02
3	200	58	22	14	11.00	7.0	82.00	1.88
总计	600	150	61	54	10.17	9.0	80.33	1.95

表 4 改进算法的实验结果

Table 4 Experimental results of the improved algorithm

批次	检测数量	检测不合格数	漏检数	过检数	漏检率/%	误检率/%	准确率/%	检测时间/s
1	200	43	15	14	7.5	7.00	85.5	1.64
2	200	49	17	11	8.5	5.50	86.0	1.72
3	200	58	13	15	6.5	7.50	86.0	1.59
总计	600	150	45	40	7.5	6.67	85.83	1.65

## 5 结 论

为了解决车门内拉手表面缺陷检测任务中缺陷目标尺寸较小、存在多尺度变化同时存在易反光等问题,本文首先,基于设计的车门内拉手自动识别系统,考虑车门内拉手表面特征,通过设计合理的相机镜头与工件间的位姿,以减少零件表面反光的问题。其次,基于 RT-DETR 模型进行改进,提出了一种快速多尺度融合和小目标增强的车门内拉手表面缺陷检测方法 PSB-DETR。通过添加

小目标检测层,增强小目标的特征表示,缓解误检、漏检现象,使用快速多尺度感受野特征提取模块,提升了复杂多尺度信息之间的关联性并增大感受野,同时加快了推理速度。在特征融合阶段采用改进的 BIFPN 连接方式,使得特征像素之间进行有效联系,提升了复杂多尺度缺陷目标特征的表达能力。将 PSB-DETR 模型在车门内拉手表面缺陷数据集和实际环境上进行实验验证,结果均表明了该算法在缺陷检测任务上的有效性。

## 参考文献

- [1] 张辉,宋雅男,王耀南,等.钢轨缺陷无损检测与评估技术综述[J].仪器仪表学报,2019,40(2):11-25.
- ZHANG H, SONG Y N, WANG Y N, et al. Review of rail defect non-destructive testing and evaluation [J]. Chinese Journal of Scientific Instrument, 2019, 40(2): 11-25.
- [2] 陈毅.基于数字图像的机械加工表面质量检测技术研究[D].兰州:兰州理工大学,2017.
- CHEN Y. Research on surface quality detection technology of machining based on digital image[D]. Lanzhou: Lanzhou University of Technology, 2017.
- [3] 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. Research progress of surface defect detection methods based on machine vision[J]. Chinese Journal of Scientific Instrument, 2022, 43(1): 198-219.
- [4] 冯维,刘红帝,汤少靖,等.基于HDRI的高反光金属表面缺陷检测方法研究[J].仪表技术与传感器,2019(8):112-116.
- FENG W, LIU H D, TANG SH J, et al. Research on defect detection method for high-reflective-metal surface based on HDRI [J]. Instrumentation Technology and Sensor, 2019(8): 112-116.
- [5] 姜美华,付鲁华,王仲,等.基于图案完整性的高反射曲面缺陷检测新方法(英文)[J].Journal of Measurement Science and Instrumentation, 2016, 7(3):221-228.
- GONG X J. Research on surface defect detection of highly reflective metal workpieces based on deep learning[D]. Tangshan: North China University of Technology, 2023.
- [6] 龚肖杰.基于深度学习的高反光金属工件表面缺陷检测研究[D].唐山:华北理工大学,2023.
- WANG X, FENG Y N. Helmet wearing detection for cyclists based on improved SSD [J]. Electronic Measurement Technology, 2022, 45(21): 90-97.
- [7] 王新,冯艺楠.基于改进SSD的骑行人员佩戴头盔检测[J].电子测量技术,2022,45(21):90-97.
- WANG X, FENG Y N. Helmet wearing detection for cyclists based on improved SSD [J]. Electronic Measurement Technology, 2022, 45(21): 90-97.
- [8] MOHAMMED S, NIMALI T M. A new paradigm for waste classification based on YOLOv5 [J]. Instrumentation, 2021, 8(4):9-17.
- 张胜华,龙嘉健,陆继东.基于Mask R-CNN的锅炉火焰图像核心高温区域提取[J].发电设备,2024,38(4):205-210.
- ZHANG SH H, LONG J J, LU J D. Extraction of core high temperature region of boiler flame image based on mask R-CNN[J]. Power Equipment, 2024, 38(4): 205-210.
- [10] 周诗捷,王玉槐,沈思橙,等.基于改进型Faster R-CNN的仓储环境物体识别技术研究[J].计算技术与自动化,2024,43(2):187-191.
- ZHOU SH J, WANG Y H, SHEN S CH, et al. Research on object recognition in warehouse environment based on improved faster R-CNN [J]. Computing Technology and Automation, 2024, 43(2): 187-191.
- [11] 王文杰,陈伟,路锦通,等.基于RT-DETR-Faster的苹果采摘机器人实时目标检测算法[J].自动化与仪表,2024,39(7):57-62.
- WANG W J, CHEN W, LU J T, et al. Real-time object detection algorithm for apple picking robot based on RT-DETR-Faster [J]. Automation and Instrumentation, 2024, 39(7): 57-62.
- [12] 周良钰,杨硕.基于改进的DETR的目标检测与测距实现[J].信息技术与信息化,2023,(6):75-78.
- ZHOU L Y, YANG SH. Implementation of object detection and ranging based on improved DETR[J]. Information Technology and Informatization, 2023, (6): 75-78.
- [13] CAI X H, LAI Q X, WANG Y W, et al. Poly kernel inception network for remote sensing detection[C]. IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2024: 27706-27716.
- [14] 金睿蛟,王堃,刘敏豪,等.基于DETR和改进去噪训练的光学遥感图像多尺度旋转目标检测(特邀)[J].激光与光电子学进展,2024,61(2):356-366.
- JIN R J, WANG K, LIU M H, et al. DETR with improved denoising training for multi-scale oriented object detection in optical remote sensing images (invited) [J]. Laser & Optoelectronics Progress, 2024, 61(2): 356-366.
- [15] 张柏园,王艺潼,周慧.基于Deformable DETR的密集行人检测[J].电子元器件与信息技术,2023,7(6):60-64.
- ZHANG B Y, WANG Y T, ZHOU H. Dense pedestrian detection based on deformable DETR [J]. Electronic Components and Information Technology, 2023, 7(6): 60-64.
- [16] 王莉,董鹏豪,王瞧,等.基于改进ResNet18的干香菇等级识别[J].国外电子测量技术,2024,43(1):117-125.
- WANG L, DONG P H, WANG Q, et al. Dried shiitake mushroom grade recognition based on improved ResNet18 [J]. Foreign Electronic Measurement Technology, 2024, 43(1): 117-125.

- [17] HOU Q B, ZHOU D Q, FENG J SH. Coordinate attention for efficient mobile network design [C]. IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2021: 13713-13722.
- [18] 张润梅,贾振楠,李佳祥,等.基于多感受野特征增强的改进 EfficientDet 遥感目标检测算法[J].电光与控制,2024,31(7):53-60,96.
- ZHANG R M, JIA ZH N, LI J X, et al. An improved EfficientDet remote sensing target detection algorithm based on multi-sensory field feature enhancement[J]. Electronics Optics & Control, 2024, 31 ( 7 ): 53-60,96.
- [19] 范金豪,崔立志.基于 Yolo\_ES 的垃圾分类目标检测模型[J].电子测量技术,2023,46(1):160-166.
- FAN J H, CUI L ZH. Garbage classification target detection model based on Yolo\_ES [J]. Electronic Measurement Technology, 2023, 46(1): 160-166.
- [20] 姜媛媛,蔡梦南.轻量化的印刷电路板缺陷检测网络 Multi-CR YOLO [J].电子测量与仪器学报,2023,37(11):217-224.
- JIANG Y Y, CAI M N. Lightweight PCB defect detection network Multi-CR YOLO [J]. Journal of Electronic Measurement and Instrumentation, 2023, 37(11):217-224.
- [21] 王文杰,陈伟,路锦通,等.基于 RT-DETR-Faster 的苹果采摘机器人实时目标检测算法[J].自动化与仪表,2024,39(7):57-62.
- WANG W J, CHEN W, LU J T, et al. Real-time object detection algorithm for apple picking robots based on RT-DETR-Faster [J]. Automation and Instrumentation, 2024, 39(7): 57-62.

### 作者简介

徐仟祥,硕士研究生,主要研究方向为深度学习、工业机器人集成。

曾勇(通信作者),教授,主要研究方向为机器视觉、工业机器人集成。

E-mail: zengyong@ycit.edu.cn

卢倩,教授,主要研究方向为智能制造、机器人装配集成。

南玉龙,博士研究生,主要研究方向为机器视觉,深度学习。