

# DBF-RTDETR：基于DETR的工业产品表面缺陷检测

李建华，段 勇\*

沈阳工业大学信息科学与工程学院，辽宁 沈阳

收稿日期：2025年12月9日；录用日期：2025年12月30日；发布日期：2026年1月14日

## 摘要

工业产品表面缺陷检测作为智能制造的关键环节，亟需解决复杂纹理干扰、多尺度缺陷共存及实时性要求高等问题。基于此，本文提出一种基于双卷积 - 双向特征金字塔结构RTDETR (DBF-RTDETR)模型的表面缺陷检测的方法，旨在针对工业表面缺陷检测场景进行性能优化。首先，为增强模型对微小缺陷的特征捕获能力，设计了双卷积模块，其采用双分支并行结构，在降低计算量的同时增强了微小缺陷的特征表达能力；其次，构建了双向特征金字塔结构网络，通过双向跨层连接与多尺度特征融合，有效抑制了光照不均与背景纹理干扰；此外，通过引入多样分支块模块，有效提升了推理速度。最后，在工业产品数据集上进行了实验，验证了本文方法具有更好的表面缺陷检测效果，在提取精度、完整性、准确性等方面均有显著提升。

## 关键词

表面缺陷检测，双卷积，双向特征金字塔，DETR，多尺度

# DBF-RTDETR: Industrial Product Surface Defect Detection Based on DETR

Jianhua Li, Yong Duan\*

School of Information Science and Engineering, Shenyang University of Technology, Shenyang Liaoning

Received: December 9, 2025; accepted: December 30, 2025; published: January 14, 2026

## Abstract

Surface defect detection of industrial products, as a key link in intelligent manufacturing, urgently needs to address challenges such as complex texture interference, coexistence of multi-scale defects,

\*通讯作者。

and high real-time requirements. To tackle these issues, this paper proposes a surface defect detection method based on a Dual Convolution-Bidirectional Feature Pyramid RT-DETR model (DBF-RTDETR), aiming to optimize performance for industrial surface defect detection scenarios. First, to enhance the model's feature capture capability for micro-defects, a dual convolution module is designed. This module adopts a dual-branch parallel structure, which reduces computational complexity while strengthening the feature representation ability for tiny defects. Second, a bidirectional feature pyramid network is constructed, which suppresses uneven lighting and background texture interference through bidirectional cross-layer connections and multi-scale feature fusion. Additionally, by introducing the Diverse Branch Block (DBB) module, the inference speed is significantly improved. Finally, experiments on industrial product datasets demonstrate that the proposed method achieves superior surface defect detection performance, with significant improvements in detection precision, integrity, and accuracy.

## Keywords

**Surface Defect Detection, Double Convolution, Bidirectional Feature Pyramid, DETR, Mul-Ti-Scale**

Copyright © 2026 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

## 1. 引言

工业产品表面缺陷检测作为智能制造的核心环节,面临微小缺陷识别(如微米级划痕)、多尺度特征融合(局部凹坑与全局形变)及复杂背景干扰(油污、反光)等严峻挑战,传统人工检测与早期机器视觉方法存在效率低、漏检率高的问题。

在深度学习领域,目标检测方法主要根据其检测流程的差异,划分为两阶段(Two-Stage)检测方法和单阶段(One-Stage)检测方法两大主流模式。单阶段检测算法 YOLO 将目标检测简化为单一回归问题,在实时缺陷检测任务中应用广泛。以钢材表面缺陷检测为例,近年来,许多研究基于 YOLO 系列模型进行改进。Ren 等采用可变形卷积替换 YOLOv5 部分常规卷积[1],提升模型对钢材表面缺陷的检测精度; Teng 等将 YOLOv7 检测头替换为高效解耦检测头[2],增强了模型对钢材表面缺陷的分类和定位能力; Zhu 等提出了一种基于 YOLOv7 的带钢表面缺陷检测算法[3],将 C2f 组件集成到特征金字塔网络,加快了 YOLOv7 训练速度,但模型体积较大; 刘昱等提出的 YOLOv8-RDP 采用 Pconv 检测头减少计算量[4],但对钢材表面缺陷的检测精度提升有限; Jiang 等提出的 YOLOv8n-SDEC 缺陷检测算法用 SPPCSPC 替换 SPPF 模块[5],并用 EIoU 替代 CIoU 函数,提升检测精度,但复杂背景下检测仍有不足。尽管基于 CNN 的检测模型(如 YOLO 系列、FasterR-CNN)通过局部卷积特性提升效率,但其单阶段方法在浅层细节与深层语义的协同利用上存在一定局限性,而两阶段模型因计算冗余往往难以满足实时需求。

近年来,Transformer 架构凭借其强大的全局上下文信息捕捉能力和并行化优势,在自然语言处理领域取得巨大成功后,迅速被引入计算机视觉领域。以 Detection Transformer (DETR)为代表的检测器,通过将目标检测视为一个端到端的集合预测问题,消除了传统方法中对人工设计组件(如非极大值抑制 NMS 和锚框 Anchor)的依赖,展现了巨大的潜力[6]。然而,尽管前景广阔,将 Transformer 模型直接应用于工业缺陷检测仍面临严峻挑战。首先,工业现场对检测的实时性要求极高,而标准 Transformer 的自注意力机制计算复杂度随图像尺寸平方增长,难以满足高吞吐量的在线检测需求。其次,工业缺陷(如细微划痕、低对比度斑点)通常仅占图像的极小部分,如何让 Transformer 在关注全局信息的同时,不丢失对微小缺

陷的敏感度是一个关键问题。此外, 现有的视觉大模型虽泛化性强, 但在特定工业场景下要达到专家级的检测精度, 仍需针对性的结构优化。而现有的一些改进方案(如 RT-DETR [7]的混合编码器等)仍对多尺度缺陷的高精度定位存在一定不足。

针对 RT-DETR 在工业缺陷检测场景下的优化, 本文提出了一种基于 DBF-RTDETR 的缺陷检测算法。该模型通过系统性的结构优化, 显著提升了模型在复杂背景干扰下的鲁棒性、对小目标的敏感度以及整体计算效率。模型中的 RT-DETR 框架融合了双卷积模块(Dual ConvBlock)与双向特征金字塔网络(BiFPN), 通过并行的卷积分支, 在减少参数量的同时增强了微小缺陷的细节捕获能力。此外, 设计了跨层级双向特征交互路径与动态权重分配机制, 提升了多尺度缺陷的融合效率。实验表明, 研究方法在工业数据集上的表现良好, 为复杂工业场景提供了高精度、高效率的缺陷检测解决方案。

## 2. 本文方法

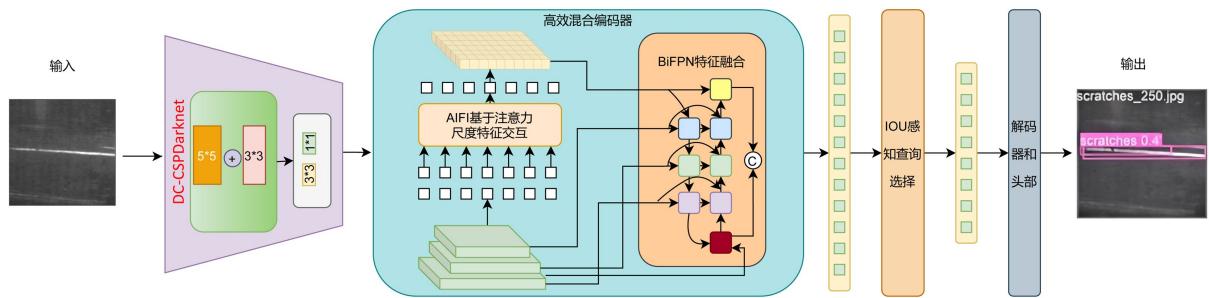
RT-DETR 模型虽通过端到端 Transformer 架构避免了手工锚框设计, 但其在工业产品缺陷检测中存在微小目标特征丢失、多尺度缺陷定位偏差及计算冗余导致实时性不足等瓶颈[8][9]。基于此, 本文进行了三重针对性的研究: (1) 设计了分支特征增强模块, 引入了双卷积模块; (2) 构建了双向特征金字塔网络(BiFPN); (3) 采用了新型多样分支块(Diverse Branch Block, DBB)。本模型从特征提取网络、卷积模块设计、多尺度特征融合机制以及损失函数优化等多个维度对基准模型进行了系统性的结构改进, 提出了一种名为 DBF-RTDETR 的创新检测方法。该方法通过引入深度可变形卷积模块增强了模型对不规则缺陷形态的空间自适应表征能力; 采用双向特征金字塔结构强化了多尺度特征的融合效率, 尤其提升了小目标缺陷的检出率; 并优化了边界框回归损失函数, 加速了模型收敛过程并提高了定位精度。这些协同改进显著增强了模型在复杂工业场景下的特征表示能力、目标检测精度与推理速度, 从而整体提升了模型对于钢材、PCB 等产品表面缺陷的检测性能。

DBF-RTDETR 模型的整体框架如图 1 所示, 其主要由三个核心组件构成: 改进的 DC-CSPDarknet 骨干网络、高效混合编码器以及配备辅助预测头的 Transformer 解码器。该设计旨在解决传统实时检测器在复杂工业场景下面临的挑战, 如复杂背景干扰、小目标检测精度低以及模型部署效率问题。双卷积模块是本文为增强特征表达而设计的结构性创新, 而多样分支块(DBB)模块是一种提升模型性能且不增加推理成本的训练与优化策略。它们共同协作, 分别从特征提取能力和模型效率两个方面提升 DBF-RTDETR 模型的整体性能。AIFI (Attention-Based Intra-Scale Feature Interaction, 基于注意力的尺度内特征交互)模块源自 RT-DETR 模型的高效混合编码器设计。其核心思想是仅对主干网络输出的最高级特征图应用 Transformer 自注意力机制。这是因为高级特征图包含了丰富的语义信息, 自注意力操作能有效捕捉图像中概念实体(物体)之间的全局关联, 有助于后续的定位与识别。同时, 通过避免在低级特征上计算昂贵的注意力, 显著降低了计算成本。模型的工作流程如下: 输入图像首先通过改进的 DC-CSPDarknet 骨干网络提取多尺度特征图(P3, P4, P5)。这些特征随后被送入高效混合编码器进行深度处理, 最后通过 Transformer 解码器生成端到端的检测结果, 无需依赖非极大值抑制(NMS)后处理[10]。

### 2.1. DC-CSPDarknet 骨干网络

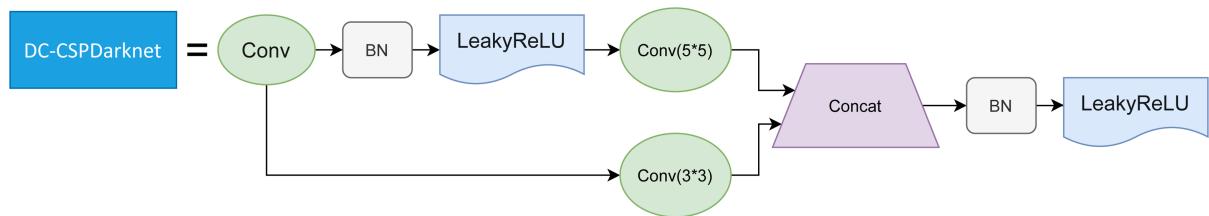
本文对骨干网络架构进行了针对性优化设计, 采用了一种改进的 DC-CSPDarknet 设计。该设计的核心是引入了交叉阶段部分网络(Cross Stage Partial, CSP)结构。相较于基础模型, 该设计能够显著降低模型的整体计算复杂度, 同时增强梯度流的传播效率, 从而有效提升网络的特征学习能力。原始的 RT-DETR 模型所采用的 HGNetV2 骨干网络, 在多尺度特征融合能力方面存在一定的局限性。特别是在处理小尺寸目标检测任务时, 该网络在深度和宽度的设计上可能受到约束, 导致其特征提取的丰富性以及最终的检

测精度有所妥协。本文所提出的改进型网络结构, 旨在强化模型的特征提取效能, 图 2 具体展示了 DC-CSPDarknet 骨干网络的详细架构。



**Figure 1.** Architecture of the DBF-RTDETR model

**图 1.** DBF-RTDETR 模型结构图



**Figure 2.** DC-CSPDarknet backbone network structure

**图 2.** DC-CSPDarknet 骨干网络结构

传统 CSPDarknet 骨干网络依赖堆叠标准卷积层提取特征, 单路径卷积结构难以兼顾微小缺陷的局部细节与大面积缺陷的全局上下文关联。在本研究中, 双分支架构所产生的特征输出通过通道注意力机制实现了动态融合, 该过程如公式 1 所示。

$$F_{\text{out}} = \sigma \left( \text{GAP} \left( F_{\text{local}} + F_{\text{global}} \right) \right) \odot W_1 F_{\text{local}} + W_2 F_{\text{global}} \quad (1)$$

其中  $\sigma$  为 Sigmoid 函数,  $\odot$  表示通道拼接, GAP 为全局平均池化, 动态权重分配使模型自适应增强了关键缺陷特征的响应强度。

## 2.2. 双卷积模块(Dual ConvBlock)

在特征融合网络中, 我们采用了双卷积模块进行局部特征增强。该模块由两个连续的卷积层构成, 每个卷积层后均包含 BN 层和激活函数。传统卷积操作受限于固定感受野与单一特征提取路径, 难以同时捕捉微米级缺陷的精细纹理(如微小裂纹)和跨区域缺陷的上下文关联(如大面积腐蚀斑块)。为此, 本文设计了双分支并行卷积结构(Dual ConvBlock)。

(1) 采用  $3 \times 3$  深度可分离卷积, 通过分离空间与通道维度卷积操作, 在保持特征分辨率的同时减少计算量, 其输出特征  $F_{\text{local}}$  聚焦于缺陷边缘的梯度变化与局部纹理细节;

(2) 引入  $1 \times 1$  空洞卷积, 输出特征  $F_{\text{global}}$  捕获跨区域的上下文语义信息(如短路区域与邻近线路的异常关联)。首先对  $F_{\text{local}}$  与  $F_{\text{global}}$  进行通道维度拼接, 经全局平均池化(GAP)生成通道描述向量, 再通过两层全连接网络与 Sigmoid 函数生成自适应权重系数  $\alpha$ , 最终融合公式(2)为

$$F_{\text{fused}} = \alpha F_{\text{local}} + (1 - \alpha) F_{\text{global}} \quad (2)$$

该策略使模型能够依据缺陷尺度动态调整特征贡献权重。

### 2.3. 多样分支块(Diverse Branch Block, DBB)

为平衡训练时的表征能力与推理时的效率, 我们在主干网络中采用了多样分支块(DBB)模块。针对工业场景中微小缺陷检测的精度与效率瓶颈, 通过异构卷积操作协同优化与结构重参数化技术, 在不增加推理耗时的前提下提升了模型容量, 显著增强了 RT-DETR 模型在复杂背景下的多尺度特征表达能力[11]。DBB 模块在训练阶段采用多分支异构操作, 包含并行  $5 \times 5$  卷积、 $3 \times 3$  深度可分离卷积及平均池化层, 通过不同尺度的感受野捕捉局部细节特征与全局上下文关联。这种多分支设计模拟了 Inception 架构的特征空间多样性优势, 既保留特征提取的丰富性, 又避免计算量的线性增长。

在模型架构中, DBB 模块主要嵌入了 RT-DETR 的下采样层与特征金字塔网络: 一方面, 将原始步长为 2 的  $3 \times 3$  卷积替换为 DBB 多分支下采样模块, 通过并行平均池化与深度可分离卷积减少信息丢失; 另一方面, 在特征金字塔的跨尺度连接处引入 DBB 模块, 利用空洞卷积与  $5 \times 5$  卷积的序列操作增强多层级特征的语义一致性。

### 2.4. BiFPN 结构

针对工业视觉任务中多尺度目标检测的鲁棒性与效率需求, 本文在模型中引入双向特征金字塔网络(BiFPN), 如图 3 所示, 显著提升了模型对微小缺陷、复杂背景及尺度差异的适应能力[12]。BiFPN 的核心设计思想源于传统特征金字塔网络(FPN)的局限性——单向特征传递导致高层语义信息与低层细节特征的交互不足, 尤其在高密度工业场景中易出现微小目标漏检或跨区域误判问题[13]。

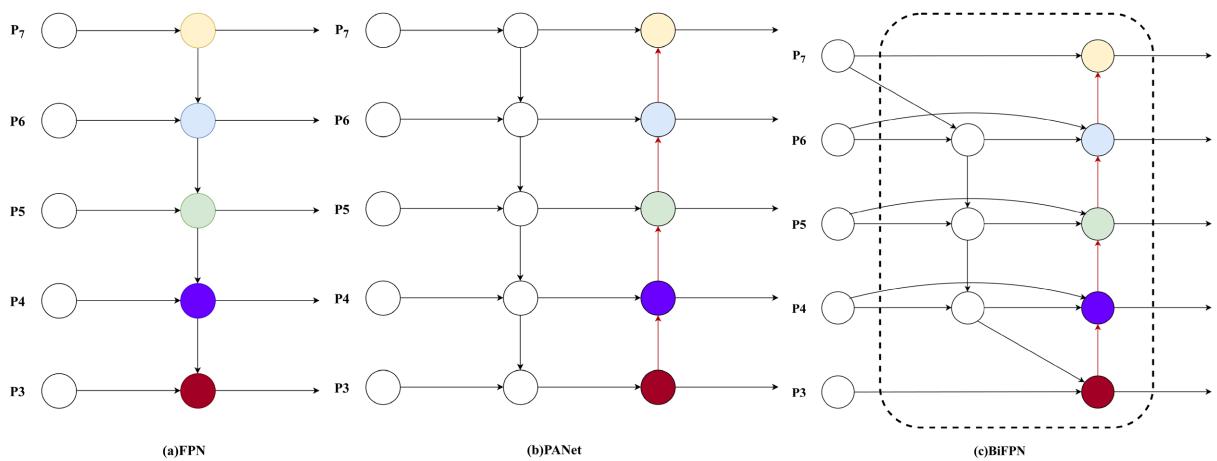


Figure 3. Architecture of the bidirectional feature pyramid network (BiFPN)

图 3. 双向特征金字塔网络图

BiFPN 通过双向路径拓扑与加权特征融合重构多尺度特征交互机制。在自顶向下的路径中, 高层语义特征通过上采样与低层特征融合, 增强了微小目标的细节响应[14]; 在自底向上的路径中, 低层高分辨率特征通过下采样与高层特征结合, 扩展了全局上下文感知范围。传统 FPN 对不同尺度特征采用直接相加或拼接方式, 忽略了不同层级特征的重要性差异[15]。BiFPN 引入可学习的通道权重参数, 通过归一化加权实现特征贡献度的动态校准。具体而言, 对于输入特征图  $P_i$  与  $P_j$ , 其融合权重  $\omega_i$  与  $\omega_j$  通过可训练参数生成, 并经过 Softmax 归一化处理, 最终融合公式(3)为:

$$P_{\text{fused}} = \frac{\omega_i \cdot P_i + \omega_j \cdot P_j}{\omega_i + \omega_j + \varepsilon} \quad (3)$$

该机制使模型能够自主调整不同层级特征的融合权重, 综上, BiFPN 通过双向跨尺度交互与动态权重机制, 为工业检测模型提供了高效、鲁棒的多尺度特征融合方案。

## 2.5. 损失函数

在工业缺陷检测任务中, 损失函数的设计直接影响了模型对微小目标、复杂背景及多尺度缺陷的感知能力[16][17]。本文综合交叉熵、FocalLoss 及 SIoU 损失函数的优势, 构建了多任务协同优化机制, 解决了工业场景中样本分布不均、定位精度低与分类置信度漂移等核心问题[18][19]。在定位优化方面, 模型引入 SIoU(Shape-AwareIoU)损失, 通过角度成本、距离成本与形状对齐机制提升边界框回归精度[20]。该损失函数将传统 IoU 指标扩展为四元组计算:

$$L_{\text{SIoU}} = 1 - \text{IoU} + \lambda_{\text{angel}} \cdot \cos(2\theta) + \lambda_{\text{dist}} \cdot \rho^2 + \lambda_{\text{shape}} \cdot \frac{\omega_h + \omega_w}{2} \quad (4)$$

其中,  $\theta$  为预测框与真实框中心连线与水平轴的夹角,  $\rho$  为归一化距离偏差,  $\omega_h$  与  $\omega_w$  为高宽比差异系数。

## 3. 实验与分析

### 3.1. 数据集

为验证本文所提方法的有效性与泛化能力, 本研究选取了两个具有代表性的公开工业缺陷检测数据集进行测试评估, 分别为东北大学(NEU)发布的钢材表面缺陷基准数据集(NEU-DET)以及北京大学的 PCB 缺陷数据集。实验旨在通过不同类型、不同场景的缺陷数据上检验模型性能, 确保评估结果的全面性与可靠性。同时, 为系统评估本文所提出的 DBF-RTDETR 模型在工业缺陷检测任务中的综合性能, 本研究设计了全面的实验验证方案, 其核心包括消融实验与对比实验两个部分, 分别验证模型内部各改进模块的有效性及其在同类算法中的综合性能。

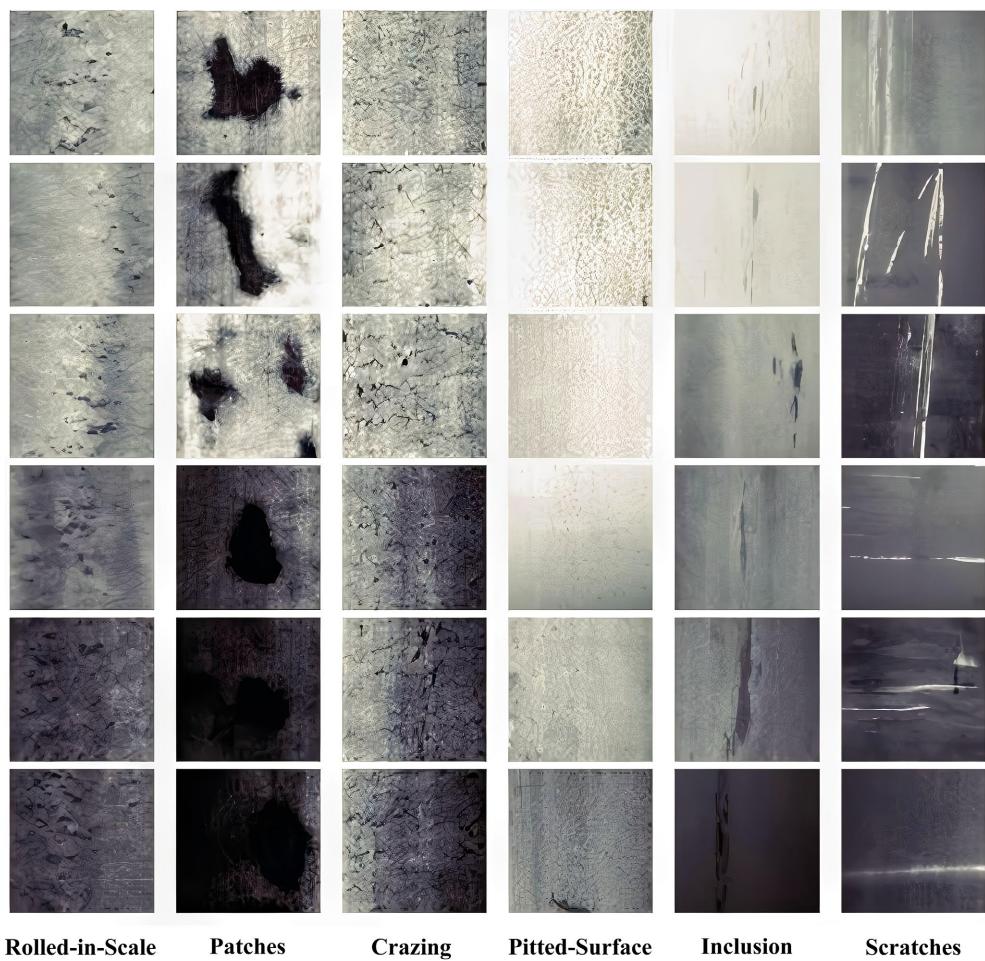
### 3.2. 数据集与实验环境

实验采用东北大学 NEU-steel-defect 钢材表面缺陷基准数据集和北京大学智能机器人开放实验室提供的 PCB 缺陷数据集。

东北大学钢材表面缺陷数据集涵盖了热轧钢带表面常见的六类典型缺陷: 轧制氧化皮、夹杂物、斑块、麻面、裂纹和划痕, 如图 4 所示。数据集中的所有原始图像均具有  $200 \times 200$  像素的固定分辨率。在数据构成上, 每类缺陷均包含 300 张样本图像, 使得数据集总规模达到 1800 张图像, 且各类别样本数量分布均匀, 这有助于避免模型训练过程中因类别不平衡而可能引发的偏差。在数据集划分方面, 为确保模型评估的可靠性, 本研究采用随机分层抽样方法, 将全部图像按照 8:1:1 的比例划分为三个独立子集。此种划分策略在保证模型有充足数据进行训练的同时, 也为模型性能的准确评估提供了可靠的数据基础。

北京大学智能机器人开放实验室提供的 PCB 缺陷数据集(PKU Market PCB), 为保障模型训练与评估的有效性, 本研究采用随机分层抽样方法, 将全部图像按 8:1:1 的比例划分为三个独立子集, 训练集包含 1109 张图像, 验证集包含 139 张图像, 测试集包含 138 张图像。

在深度学习模型训练中, 超参数的配置对模型的收敛速度、计算资源消耗及最终性能至关重要。本研究选用 AdamW 作为优化器, 该优化器通过引入解耦权重衰减, 有助于提升模型的泛化能力。训练过程中, 初始学习率设定为  $1 \times 10^{-4}$ , 并采用余弦退火策略对其进行动态调整, 该策略能够有效平滑地降低学习率, 辅助模型在训练后期更稳定地逼近局部最优点。同时, 权重衰减系数设置为  $1 \times 10^{-4}$ , 以约束模型复杂度, 防止过拟合。整个模型的训练轮数为 200 个周期, 批量大小设置为 4。其余参数均遵循深度学习框架的默认设置。



**Figure 4.** Defect category distribution in steel surface defect dataset  
**图 4.** 钢材表面缺陷数据集中缺陷的类别

### 3.3. 评价指标

为客观评估本研究提出的改进 RT-DETR 模型在缺陷检测任务中的综合性能, 需将其与原始 RT-DETR 模型以及其他主流检测模型在相同实验设置下进行系统对比。为客观评估目标检测模型的综合性能, 本研究采用精确率(Precision)、召回率(Recall)、平均精度(Average Precision, AP)与平均精度均值(mean Average Precision, mAP)作为核心评估指标, 其计算公式如下所示: 精确率( $P$ ), 用于衡量模型预测为正例的结果中真正为正例的比例, 反映检测结果的可靠性, 计算如式(5)所示, 其中, 真正例(True Positive, TP)是指模型正确识别出的正类样本数量, 即模型预测的边界框与真实标注框的交并比(IoU)不低于设定阈值(如 0.5), 且类别预测正确。假正例(False Positive, FP)则包括两类情形: 一是定位错误, 即边界框与真实标注的 IoU 低于阈值; 二是分类错误, 即边界框定位准确但类别预测错误。召回率( $R$ ), 用于衡量所有实际为正例的样本中被模型正确预测出来的比例, 反映模型对正类样本的覆盖能力, 计算如式(6)所示, 其中, 假负例 FN (False Negative)是指那些实际属于正类别的样本(例如图像中存在特定目标), 但被模型错误地判定为负类别的情况, 这类错误表明模型未能识别出本应识别出的目标, 通常对应于漏检现象。平均精度(AP), 是精确率 - 召回率曲线下的面积, 能够综合反映模型在不同召回率水平下的精确率表现。平均精度均值(mAP)则是所有类别 AP 的平均值, 常用于多类别目标检测任务中衡量模型的整体检测精度, 特别是在交并比(IoU)阈值为 0.5 时的 mAP(即 mAP@0.5)是常见的评估指标。除了检测精度指标, 推断速度

是实际应用中的关键性能指标, 通常以每秒帧数(Frames Per Second, FPS)来衡量。FPS 值越高, 表明模型的处理速度越快, 实时检测能力越强。通过综合考量 mAP、精确率、召回率以及 FPS 等指标, 可以对模型的检测准确性和实时性进行全面评估。

$$P = \frac{TP}{TP + FP} \quad (5)$$

$$R = \frac{TP}{TP + FN} \quad (6)$$

$$AP = \int_0^1 P(R) dR \quad (7)$$

$$mAP = \frac{\sum_{i=1}^N AP_i}{N} \quad (8)$$

### 3.4. 消融实验

本研究以 RT-DETR (HGNetV2)作为基线模型, 设计并执行了一组严谨的消融实验, 系统评估本文所提出的算法改进策略对模型性能的贡献。该实验方案旨在通过控制变量的方式, 细致量化分析每个改进模块对最终检测性能的独立影响与协同效应。实验通过依次引入各个改进模块, 观察其对检测性能的影响。最终, 将所有四个改进模块(A + B + C + D)同时集成到基线模型中, 构建出完整的改进模型, 并评估其综合性能。**表 1** 定量记录了在相同实验环境下, 基准模型及逐步引入 DC-CSPDarknet 骨干网络(模块 A)、动态瓶颈模块(模块 B)、加权双向特征金字塔网络(模块 C)以及 SIOU 损失函数(模块 D)后, 模型在多项关键性能指标上的具体表现。这些指标涵盖了检测精度与效率两个核心维度, 具体包括精确率(Precision,  $P$ )、召回率(Recall,  $R$ )、平均精度均值(mean Average Precision, mAP)以及帧率(Frames Per Second, FPS)。

**Table 1.** Ablation study results with performance metrics

**表 1.** 消融实验各项指标对比

模型	$P/(\%)$	$R/(\%)$	$mAP50/(\%)$	$FPS/(f/s)$
RT-DETR	71.5	57.6	63.7	72
RT-DETR + A	76.7	66.4	68.4	79
RT-DETR + A + B	80.3	77.9	70.1	80
RT-DETR + A + B + C	82.6	80.3	73.7	85
RT-DETR + A + B + C + D(ours/本文模型)	<b>84.1</b>	<b>83.7</b>	<b>74.2</b>	<b>86</b>

由**表 1** 中数据可知, 原始 RT-DETR 模型的精确率( $P$ )为 71.5%, 召回率( $R$ )为 57.6%, 平均精度均值(mAP)为 63.7%, 推理速度为 72 FPS。以此为基础, 逐步引入各项改进后, 实验结果显示每一项改进均对模型性能产生了积极的促进作用。例如, 引入 DC-CSPDarknet 骨干网络(模块 A)有望在保持较高推理速度的同时增强特征提取能力, 从而可能提升召回率与 mAP; DBB 模块(模块 B)的加入旨在通过增强卷积结构的多样性来提升模型的表征能力; BiFPN 模块(模块 C)致力于更高效地融合多尺度特征, 特别是增强对小目标的检测能力; 而 SIOU 损失函数(模块 D)则关注于优化边界框回归的准确性和收敛速度。为系统评估本文提出的 DBF-RTDETR 模型的综合性能, 在完成全部模块集成与参数优化后, 于标准测试集上进行了最终验证。实验结果表明, 该改进模型在多项核心评估指标上均达到了当前最优水平, 具体表现为:

精确率达到 84.1%，召回率达到 83.7%，平均精度均值为 74.2%，同时推理速度保持在 86 FPS。与原始 RT-DETR 基线模型相比，本文所提模型展现出全面的性能提升。在检测精度方面，精确率显著提升了 13.1 个百分点，召回率大幅提高了 26.1 个百分点，平均精度均值提升了 10.5 个百分点。与此同时，模型的计算效率也得到改善，推理速度提升了 14 FPS。当所有改进模块协同工作时，模型在检测精度与效率方面达到了最佳平衡，各项指标均获得显著提升，充分验证了本文所提改进策略的有效性与协同效应。

### 3.5. 对比实验

为系统评估本文提出的 DBF-RTDETR 模型在目标检测任务中的综合性能，本研究设计了一套横向对比实验方案。为系统评估本文所提出模型的综合性能，本研究选取了六种具有代表性的较为先进目标检测架构作为对比基线。这些模型涵盖了单阶段检测(YOLO 系列)与基于 Transformer 的端到端检测(RT-DETR)两种主流技术路线，以确保比较的全面性。所有对比实验均在相同的硬件与软件训练环境下进行，采用统一的数据集，并为每个参与对比的模型应用其官方推荐或广泛使用的最优超参数设置，以确保比较的公平性与结果的可信度。

表 2 所示为在钢材表面缺陷检测数据集对比试验结果，可见本文所研究的方法在数据集中表现优秀。在钢材检测数据集中， $P$ 、 $R$ 、mAP50 上实现了最佳性能，分别达到了 84.1%、83.7% 和 74.2%，平均帧数 FPS 为 86f/s 仅次于 YOLOv8n，但同样优于其他模型。表 3 所示为在 PCB 缺陷检测数据集的对比试验结果。本文的 DBF-RTDETR 模型在准确率与 mAP 上均取得最优秀的性能，分别为 95.7% 与 90.1%，在召回率方面也仅次于 RT-DETR-X 的 86.8%。

**Table 2.** Comparative study of steel surface defect detection performance

**表 2. 对比试验结果(钢材表面缺陷检测)**

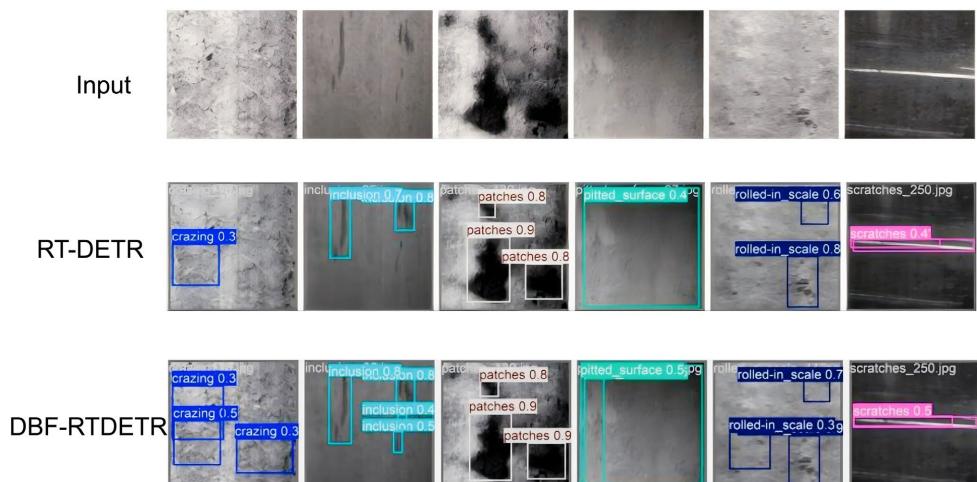
模型	$P/(\%)$	$R/(\%)$	mAP50/(\%)	FPS/(f/s)
RT-DETR-r18	71.5	57.6	63.7	72
RT-DETR-X	71.1	64.6	67.2	48
YOLOv5n	67.0	75.4	67.2	51
YOLOv8n	72.8	70.9	72.8	<b>112</b>
YOLOv10n	83.0	67.0	72.5	81
Ours (DBF-RTDETR)	<b>84.1</b>	<b>83.7</b>	<b>74.2</b>	86

**Table 3.** Comparative study of PCB surface defect detection performance

**表 3. 对比试验结果(PCB 表面缺陷检测)**

模型	$P/(\%)$	$R/(\%)$	mAP50/(\%)
RT-DETR-r18	89.5	82.9	88.1
RT-DETR-X	90.6	<b>86.8</b>	88.3
YOLOv5n	88.0	76.0	85.0
YOLOv8n	87.4	82.1	88.3
YOLOv10n	81.0	78.7	84.0
Ours (DBF-RTDETR)	<b>95.7</b>	78.5	<b>90.1</b>

本研究选取了精确率(Precision,  $P$ )、召回率(Recall,  $R$ )以及交并比阈值为 0.5 时的平均精度均值(Map50)作为核心评价指标。 $P$  与  $R$  指标的高水平同步提升, 说明模型在有效控制误检(低假阳性)的同时, 也显著降低了漏检(低假阴性)的发生概率。而 Map50 指标的优异表现, 则进一步证实了模型对于数据集中存在的多类别缺陷均能保持稳定的高精度识别能力, 体现出其优异的泛化性能。图 5 通过直观对比不同算法在典型样本上的检测效果, 清晰展示了本文方法在复杂背景干扰下对小目标缺陷的精准定位与识别能力, 进一步佐证了其在实际工业应用中的有效性与鲁棒性。因此, 本文提出的 DBF-RTDETR 算法在工业缺陷检测任务中展现出卓越的综合性能。其在精确率、召回率以及平均精度均值(Map50)等核心评价指标上均取得了优异的表现, 同时保持了较高的推理速度(FPS)。该模型在检测精度与计算效率之间取得了良好平衡, 不仅能够精准识别各类复杂缺陷, 尤其在小目标和复杂背景干扰场景下表现出较强的鲁棒性, 而且其轻量化设计和高效的计算特性使其能够满足实际工业产线对实时检测与快速部署的严格要求。



**Figure 5.** Qualitative comparison of visual detection results

**图 5. 检测结果对比图**

本研究选取了多种具有代表性的先进目标检测模型进行对比实验, 全面评估本文所提出的 DBF-RTDETR 模型的检测性能。参与对比的模型包括单阶段检测器 YOLO 系列中的 YOLOv5n、YOLOv8n、YOLOv10n, 以及基于 Transformer 架构的 RT-DETR 系列中的 RT-DETR-r18 和 RT-DETR-X。所有实验均在公开的钢材表面缺陷数据集与 PCB 缺陷检测数据集上进行, 以确保评估的客观性与普适性。实验结果表明, 本文提出的 DBF-RTDETR 模型在多项关键评估指标上均表现出显著优势。相较于其他参比模型, DBF-RTDETR 在复杂背景干扰下对细微缺陷的检出率与识别置信度均有明显提升, 尤其是在应对钢材表面裂纹以及 PCB 微小瑕疵等挑战性场景时, 其检测性能更为优异。这验证了模型结构改进在增强特征提取能力与多尺度特征融合方面的有效性。

## 4. 结论

本文旨在应对现有基于 RT-DETR 的工业产品表面缺陷检测模型在复杂背景干扰、较高内存占用以及小目标检测精度不足等方面的主要挑战。为此, 本研究提出一种改进的 DBF-RTDETR 检测方法。该方法通过三个关键层面的创新实现性能提升: 首先采用 DC-CSPDarknet 结构作为骨干网络, 通过引入交叉阶段部分网络(CSP)机制增强特征提取能力并降低计算复杂度; 其次设计多样分支块(DBB), 结合可变形卷积与注意力机制提升模型对不规则缺陷的空间自适应能力; 最后引入加权双向特征金字塔网络(BiFPN)加强多尺度特征融合, 显著提升小目标的检测精度。这些改进使模型在保持较高推理速度的同时, 有效

提升了复杂工业场景下的检测性能。实验结果表明，相较于原始 RT-DETR 算法，本文所提模型在钢材表面缺陷与 PCB 缺陷两类数据集上均表现出显著的性能提升，尤其在复杂工业背景下的鲁棒性与对小缺陷的敏感度方面改善明显。综上所述，本研究为工业产品表面缺陷检测提供了一种可行且高效的解决方案，在检测精度、模型可解释性及计算效率之间取得了良好平衡。未来的研究工作将聚焦于模型的进一步轻量化设计及其在多样化工业场景中的实际应用推广。

## 参考文献

- [1] 邵延华, 张铎, 楚红雨, 等. 基于深度学习的 YOLO 目标检测综述[J]. 电子与信息学报, 2022, 44(10): 3697-3708.
- [2] 宋智伟, 黄新波, 纪超, 等. 基于 Flexible YOLOv7 的输电线路绝缘子缺陷检测和故障预警方法[J]. 高电压技术, 2023, 49(12): 5084-5094.
- [3] 李松, 史涛, 井方科. 改进 YOLOv8 的道路损伤检测算法[J]. 计算机工程与应用, 2023, 59(23): 165-174.
- [4] Wang, H., Zhang, Y. and Wu, W. (2025) OHEH-RTDETR: An Improved RT-DETR Detection Model Based on Frequency Layered Processing and Advanced Feature Selection. *Measurement Science and Technology*, **36**, Article 046211. <https://doi.org/10.1088/1361-6501/adc1f3>
- [5] 吴斌斌, 张礼华, 刘军伟, 等. 基于改进的 EP-RTDETR 小目标 PCB 表面缺陷检测[J]. 制造技术与机床, 2025(3): 139-148.
- [6] Wang, S., Jiang, H., Li, Z., Yang, J., Ma, X., Chen, J., et al. (2024) PHSI-RTDETR: A Lightweight Infrared Small Target Detection Algorithm Based on UAV Aerial Photography. *Drones*, **8**, Article 240. <https://doi.org/10.3390/drones8060240>
- [7] Bai, J., Jiang, T., Zou, Z., Wang, Y. and Xue, T. (2025) Detecting Object in Remote Sensing Image Based on Improved RT-DETR. *International Journal of Pattern Recognition and Artificial Intelligence*, **39**, Article 2555011. <https://doi.org/10.1142/s0218001425550110>
- [8] 程哲宇. 基于改进 RTDETR 的 FPC 缺陷检测[D]: [硕士学位论文]. 南昌: 南昌工程学院, 2024.
- [9] 周文骏, 黄硕, 张宁, 等. 基于 MA-DETR 的 SAR 图像飞机目标检测[J]. 光学精密工程, 2024, 32(18): 2814-2822.
- [10] Zhao, Y., Lv, W., Xu, S., Wei, J., Wang, G., Dang, Q., et al. (2024) DETRS Beat Yolos on Real-Time Object Detection. 2024 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), Seattle, 16-22 June 2024, 16965-16974. <https://doi.org/10.1109/cvpr52733.2024.01605>
- [11] 韩慧健, 邢怀宇, 张云峰, 等. 基于 Transformer 多元注意力的钢材表面缺陷视觉检测[J]. 郑州大学学报(工学版), 2025, 46(5): 69-76.
- [12] 秦伦明, 梅温泉, 崔昊杨, 等. 改进 YOLOv8s 的遥感图像目标检测算法[J]. 激光与光电子学进展, 2025, 62(10): 317-325.
- [13] 雷志鹏, 彭川, 许子涵, 等. 基于改进 Deformable DETR 模型的多源局部放电识别方法及其应用[J]. 中国电机工程学报, 2024, 44(15): 6248-6261.
- [14] Wang, S., Liu, M., Dong, S. and Chen, S. (2025) ERL-RTDETR: A Lightweight Transformer-Based Framework for High-Accuracy Apple Disease Detection in Precision Agriculture. *Concurrency and Computation: Practice and Experience*, **37**, e70276. <https://doi.org/10.1002/cpe.70276>
- [15] Tian, F., Song, C. and Liu, X. (2025) Small Target Detection in Coal Mine Underground Based on Improved RTDETR Algorithm. *Scientific Reports*, **15**, Article No. 12006. <https://doi.org/10.1038/s41598-025-96638-8>
- [16] Xie, Z. and Zou, X. (2024) MFAD-RTDETR: A Multi-Frequency Aggregate Diffusion Feature Flow Composite Model for Printed Circuit Board Defect Detection. *Electronics*, **13**, 3557. <https://doi.org/10.3390/electronics13173557>
- [17] Wang, Y., Wu, B., Zhang, L., Wang, Z., Liu, J., Dong, J., et al. (2025) Enhanced PCB Defect Detection via HSA-RTDETR on RT-DETR. *Scientific Reports*, **15**, Article No. 31783. <https://doi.org/10.1038/s41598-025-11394-z>
- [18] Li, B., Fang, J. and Zhao, Y. (2025) RTDETR-REFA: A Real-Time Detection Method for Multi-Breed Classification of Cattle. *Journal of Real-Time Image Processing*, **22**, Article No. 38. <https://doi.org/10.1007/s11554-024-01613-7>
- [19] Maulana, A.W., Wibowo, S.A., Sunarya, U., Rahmania, R. and Insani, A. (2024) Performance Improvement of YOLOV8-RtdeTR Method Based Retail Product Detection. *Journal of Image and Graphics*, **12**, 292-301. <https://doi.org/10.18178/joig.12.3.292-301>
- [20] 李宗刚, 宋秋凡, 杜亚江, 等. 基于改进 DETR 的机器人铆接缺陷检测方法研究[J]. 铁道科学与工程学报, 2024, 21(4): 1690-1700.