

DOI:10.19651/j.cnki.emt.2518835

# 基于改进 RT-DETR 的转向节表面缺陷检测算法<sup>\*</sup>

张上<sup>1,2</sup> 朱帅<sup>1,2</sup> 张岳<sup>1,2</sup>

(1.三峡大学湖北省建筑质量检测装备工程技术研究中心 宜昌 443002;2.三峡大学计算机与信息学院 宜昌 443002)

**摘要:** 针对汽车转向节表面缺陷识别过程中存在的检测精度低、模型复杂度高及对缺陷边界信息关注不足等问题,本文提出一种改进 RT-DETR 的转向节表面缺陷检测算法 GSG-DETR。首先,设计多尺度边缘信息传递模块 GLOFT 改进主干网络,通过强化边缘信息的捕捉与传递,提高模型对缺陷边缘的敏感度。其次,在颈部网络中引入选择边信息聚集模块 SBA,构建低分辨率边界信息与深层语义特征的自适应融合机制,优化多尺度缺陷边界特征对齐策略。最后,采用 GroupNorm 结构化剪枝方法,剪除耦合层冗余网络,以降低模型参数量和计算量。实验结果表明,GSG-DETR 算法在转向节裂纹检测任务中的 mAP50 达到 88.2%,相比基准模型提高 2.0%,参数量和计算量分别下降 34.3% 和 32.1%,FPS 提升至 105.1 帧,整体优于其他改进算法。在 NEU-DET 数据集上进一步验证其泛化能力,改进算法 mAP50 较基准模型提升 4.3%。综上所述,GSG-DETR 不仅在检测精度表现出色,而且更符合实际应用。

**关键词:** RT-DETR;表面缺陷检测;转向节;边缘信息;通道剪枝

**中图分类号:** TP391.4;TN98 **文献标识码:** A **国家标准学科分类代码:** 510.4050

## Steering knuckle surface defect detection algorithm based on the improved RT-DETR

Zhang Shang<sup>1,2</sup> Zhu Shuai<sup>1,2</sup> Zhang Yue<sup>1,2</sup>

(1. Hubei Province Engineering Technology Research Center for Construction Quality Testing Equipment, China Three Gorges University, Yichang 443002, China; 2. College of Computer and Information Technology, China Three Gorges University, Yichang 443002, China)

**Abstract:** To address issues such as low detection accuracy, high model complexity, and insufficient attention to defect boundary information in the steering knuckle surface defect detection process, this paper proposes an improved RT-DETR-based steering knuckle surface defect detection algorithm GSG-DETR. First, a multi-scale edge information transfer module GLOFT is designed to improve the backbone network by enhancing the capture and transfer of edge information, thus increasing the model's sensitivity to defect boundaries. Next, a selective edge information aggregation module SBA is introduced into the neck network, constructing an adaptive fusion mechanism between low-resolution boundary information and deep semantic features, optimizing the alignment strategy of multi-scale defect boundary features. Finally, a GroupNorm-based structured pruning method is employed to eliminate redundant coupled layers, reducing the model's parameter count and computational complexity. Experimental results demonstrate that the GSG-DETR algorithm achieves an mAP50 of 88.2% in the steering knuckle crack detection task, a 2.0% improvement over the baseline model, with a 34.3% reduction in parameters and a 32.1% reduction in computational complexity, while the FPS increases to 105.1 frames. Further validation on the NEU-DET dataset shows that the improved algorithm yields a 4.3% increase in mAP50 compared to the baseline model. In summary, GSG-DETR not only excels in detection accuracy but also aligns better with practical applications.

**Keywords:** RT-DETR; surface defect detection; steering knuckle; edge information; pruning of the passage

## 0 引言

汽车转向节作为车辆转向系统中的重要承力部件,其

表面状况直接影响整车行驶安全。由于转向节采用高强度合金钢锻造而成,使得在热成型阶段易产生微裂纹<sup>[1]</sup>。若锻造过程中形成的裂纹未被及时检测并处理,势必对车辆

收稿日期:2025-05-16

\* 基金项目:湖北省数字经济试点示范建设专项(2312-420625-04-02-996363)、湖北省国家级大学生创新创业训练计划(S202311075047)、国家级大学生创新创业训练计划(202111075012,202011075013)项目资助

安全性产生不良影响。因此,提升转向节表面缺陷检测技术的精度,已成为汽车制造过程中不可忽视的重要环节<sup>[2]</sup>。

在传统的锻件表面缺陷检测中,人工检测依然是主要方法。然而,该方法存在劳动强度大、检测时间长等明显不足,且受限于操作人员经验差异及高强度工作环境,易导致漏检和误判,从而对产品的整体质量控制造成不利影响<sup>[3-4]</sup>。因此,急需引入更加精确且快速的检测技术,以确保在生产过程中能高效识别转向节表面裂纹。

随着深度学习技术的迅速发展,基于卷积神经网络(convolutional neural network, CNN)的检测方法逐渐成为主流<sup>[5-6]</sup>。目前的检测方法分为两类:一类是以 R-CNN<sup>[7]</sup>、Mask R-CNN<sup>[8]</sup>、Faster-RCNN<sup>[9]</sup>为代表的二阶段算法,另一类是以 SSD<sup>[10]</sup>、YOLO<sup>[11]</sup>系列为代表的单阶段算法,这两类算法具有较好的检测精度。但是,在实际工业检测中,当面对多层次结构、高关联特征以及复杂背景干扰的微裂纹缺陷时,应用 CNN 的目标检测算法在处理微裂纹缺陷时精度仍显不足。一方面, CNN 局部感受野的特点使算法难以捕捉微裂纹与复杂背景的全局语义关联,噪声干扰下易发生漏检误判;另一方面, CNN 依赖于非极大值抑制(non-maximum suppression, NMS)硬锚框机制,预定义的锚框尺寸难以适应微裂纹形态多样性。相比之下,基于 Transformer 的 RT-DETR(real-time detection transformer)算法通过引入全局注意力机制,使目标检测模型能有效捕获微裂纹与复杂背景之间的长距离依赖关系,从而克服传统 CNN 模型在处理复杂场景时的不足。同时, RT-DETR 采用端到端处理方式,直接在模型内部完成目标筛选与定位,省去传统 NMS 后处理操作,不仅降低计算时延,也显著提升密集裂纹检测场景下的检测精度与鲁棒性。

在实际目标检测应用中,张岳等<sup>[12]</sup>针对转向节表面缺陷,基于 YOLOv8s 提出轻量多尺度卷积模块 LMSCov,通过将单个卷积转化为不同卷积核的多个卷积,在小于单个卷积模块的计算量下实现跨多尺度提取目标特征,使得计算量降至 27.2 GFLOPs,解决了转向节表面缺陷检测速度慢,实时性差的问题,但检测精度和召回率表现不佳,分别为 83.4% 和 70.5%;张伟健等<sup>[13]</sup>基于 RT-DETR,设计 RepNCSPELAN4 模块,使骨干网络 RCANet 更轻量化且能更准确的提取数据特征,同时融合目标的上下文信息,增加对小目标的检测能力,使得改进后的 RT-DETR 模型参数量降低 9.5 M,但精度仅提升 1.5%,仍难以满足对钢材的高精度检测;针对车辆行驶过程中摄像头捕捉到的交通标志大多数是小目标,秦伦明等<sup>[14]</sup>为提高模型对小目标的识别精度,在 RT-DETR 的主干网络中引入正交通道注意力机制 Ortho attention,使交通标志检测精度提升 2.37%,但模型的参数量为 18.22 M,计算量为 53.4 GFLOPs,不利于部署在实际环境;张靖雯等<sup>[15]</sup>为解决玻璃绝缘子缺陷对比度低、多尺度问题,提出一种基于改进 RT-DETR 的轻量化玻璃绝缘子缺陷检测算法,通过引入轻量化骨干网络

RE-FasterNet,创新性的部分重复卷积和高效多尺度注意力机制,以提高特征提取效率和对小目标、低对比度缺陷的检测能力,所提算法使得检测精度提升 2.8%,模型大小下降 23.6%,计算量下降 13.1%,但参数量和计算量仍为 17.7 M 和 51.4 GFLOPs,轻量化效果有限;Yu 等<sup>[16]</sup>针对铁路钢材缺陷不规则分布、尺寸变化、环境光照变化和复杂背景问题,提出 FHB-DETR 算法,基于 Faster Block 设计 Faster CGLU 模块,采用部分卷积 PConv 处理 1/4 输入通道,使得计算量降至传统卷积的 1/16,引入卷积门控单元 CGLU 增强局部特征提取与噪声抑制提升小缺陷敏感度,使得计算复杂度降低 6%,参数量减少 25%,但精度提升依然有限,计算量仍为 53.5 GFLOPs。

上述研究大多集中于轻量化模块设计和注意力机制创新以降低模型规模、提升对小目标的检测能力。然而,在缺陷边界精细建模、跨层级特征有效传递与整合,以及模型精度与轻量化之间的平衡方面仍存明显不足。

针对上述问题,本文提出 GSG-DETR 转向节表面缺陷检测算法。该算法基于 RT-DETR-r18 模型改进,旨在通过增强模型对缺陷边缘信息的感知能力,并采取轻量化模块之外的方法,在显著提高复杂转向节表面缺陷检测精度的同时,维持模型的轻量化特性,从而更好地满足工业对高精度与高效率检测的实际部署需求。主要贡献如下:

1) 设计多尺度边缘信息传递模块(global edge information transfer, GLOFT),通过强化缺陷边缘区域特征的捕捉与跨层传递能力,使浅层特征中提取到的边缘信息传递到整个主干网络上,提升模型在复杂场景中对缺陷边界的敏感性。

2) 在颈部网络中引入选择边信息聚集模块(selective boundary aggregation, SBA),该模块自适应融合低分辨率边界信息与深层语义特征,优化多尺度缺陷的边界语义对齐策略,有效增强模型对微裂纹等小尺度缺陷的精准表征能力。

3) 采用结构化剪枝 GroupNorm 方法,将模型压缩简化。在保证检测精度的前提下,显著降低参数量和计算量,同时提高检测速度,使得算法在实际工业场景中利于部署。

## 1 RT-DETR 算法

RT-DETR<sup>[17]</sup>是一款基于 DETR 架构的高效实时目标检测器,旨在兼顾高精度与实时推理性能。该模型继承端到端检测的优势,并通过优化推理速度,提高目标识别效率。其网络结构如图 1 所示。RT-DETR 的整体结构由 ResNet-18 主干网络(backbone)、高效混合编码器(hybrid encoder)和解码器(decoder)3 个核心模块组成。

首先,主干网络负责从输入图像中提取基础特征。然后,高效混合编码器通过基于注意力的同尺度特征交互(attention-based intra-scale feature interaction, AIFI)在高层特征中施加自注意力计算,以降低计算复杂度。同时,基于 CNN 的跨尺度特征融合(CNN-based cross-scale feature

fusion,CCFF)逐步整合相邻层的特征,以提升目标表征能力。在进入解码器后,RT-DETR 结合 Transformer 结构,引入交并比(intersection over union, IoU)感知查询选择机制,动态筛选出最具代表性的特征作为解码器的初始查询。最后,解码器通过逐步优化增强目标表示,并借助辅助预测头

(auxiliary prediction head)生成最终的目标边界框和置信度。

相较于传统目标检测方法,RT-DETR 在计算资源占用相近的情况下,展现出更强的检测能力。它无需依赖 Mosaic 数据增强,即可在短时间内完成高效训练,同时实现高精度且快速的目标检测结果。

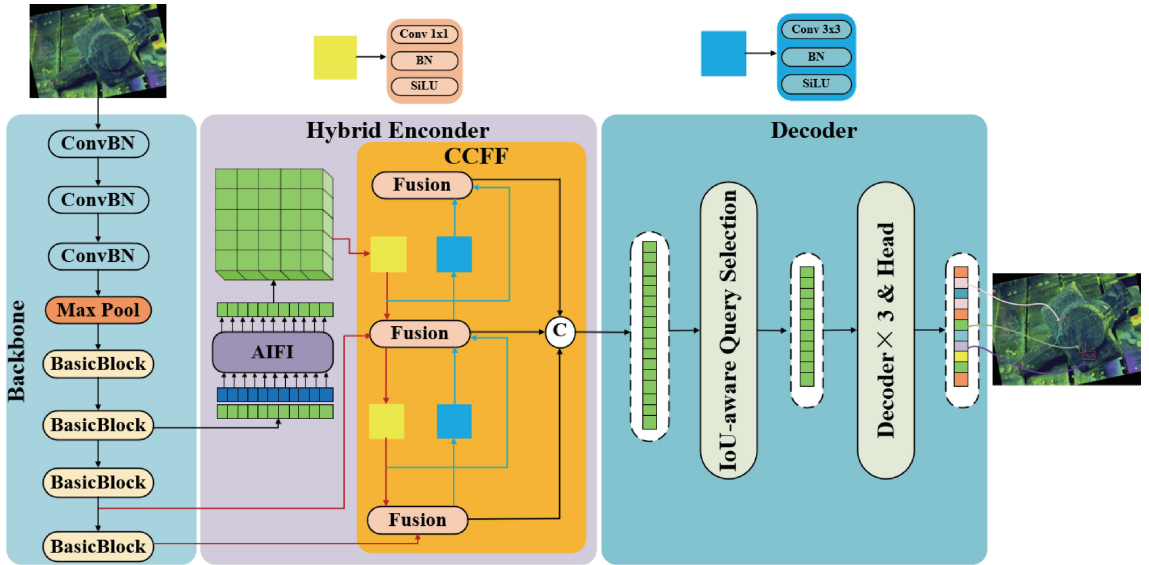


图 1 RT-DETR 网络结构

Fig. 1 RT-DETR network structure

## 2 GSG-DETR 算法

本文提出的算法在 RT-DETR-r18 基础上改进。具体网络结构如图 2 所示。首先,设计多尺度边缘信息模块 GLOFT 以改进主干网络,多尺度边缘信息生成器(multi-scale edge information generator, MSEIG)对边缘响应进行分层建模,并通过卷积边缘融合(convolution edge fusion, ConvEF)实现边缘信息和主干网络卷积特征的有效整合。接着,在颈部网络中引入多个 SBA 模块,融合低层细节信息和高层语义信息。最后,采用 GroupNorm 结构化剪枝方法以压缩算法规模。

### 2.1 多尺度边缘信息传递模块(GLOFT)

在转向节表面缺陷检测任务中,RT-DETR 原始的 ResNet-18<sup>[18]</sup>主干网络虽能提取图像的低层次特征,并通过其深层次的卷积层捕捉更丰富的模式,但仍存在一些局限性。具体来说,ResNet-18 主干网络在处理转向节表面高频、细节丰富的图像特征时,其对于边缘细节和微小缺陷的敏感度较低,这是由于标准卷积层在特征提取过程中未能有效保留图像中的边缘信息,导致模型难以捕捉到转向节表面重要的缺陷特征,从而影响检测精度。

为解决这一问题,本文设计 GLOFT 模块以改进 ResNet-18 主干网络,如图 3 所示。该模块通过浅层特征生成多尺度边缘信息,并与不同尺度的特征进行融合。其核心目标在于增强浅层特征的边缘信息,同时避免从原始

图像中,直接提取边缘时可能引入的噪声。该模块包括 3 部分:1) SobelConv 卷积操作,通过微分算子提取梯度特征;2) MSEIG,采用空洞卷积金字塔进行跨尺度边缘增强;3) ConvEF,通过注意力机制实现边缘特征与语义特征的动态融合。SobelConv 通过 Sobel 边缘检测算子提取图像的边缘信息,捕捉水平方向  $G_x$  和垂直方向  $G_y$  的边缘特征。随后,将两者的结果相加,为多尺度边缘信息生成器提供更加精确的边缘特征。提取边缘特征过程如下:

$$SobelConv(\mathbf{X}) = \bigoplus_{c=1}^c (G_x^{(c)} + G_y^{(c)}) \quad (1)$$

式中:  $\mathbf{X} \in R^{B \times C \times H \times W}$  表示输入特征图,  $c$  为通道,  $\oplus$  表示元素相加。

基于 SobelConv 提取的边缘特征, MSEIG 进一步对边缘响应进行分层建模,通过多级最大池化层对边缘信息特征图进行下采样生成多个尺度的边缘特征图。然后,对每个池化后的特征图应用  $1 \times 1$  卷积进行通道数调整和特征增强。最终,生成多个尺度的边缘信息特征图,这些特征图包含在不同分辨率下的边缘信息。

第  $i$  个尺度的边缘特征表达式如下:

$$outputs_i = Conv_{1 \times 1}(MaxPool^i(SobelConv(\mathbf{X}))) \quad (2)$$

其中,  $i = 1, 2, \dots, n$ 。  $n$  表示多尺度层数,  $outputs_i$  为最终生成边缘特征。

ConvEF 将 MSEIG 生成的多尺度边缘信息特征图与其他卷积特征图进行融合,先将来自不同尺度的边缘特征图进行拼接,拼接后的特征图通过  $1 \times 1$  卷积进行通道融

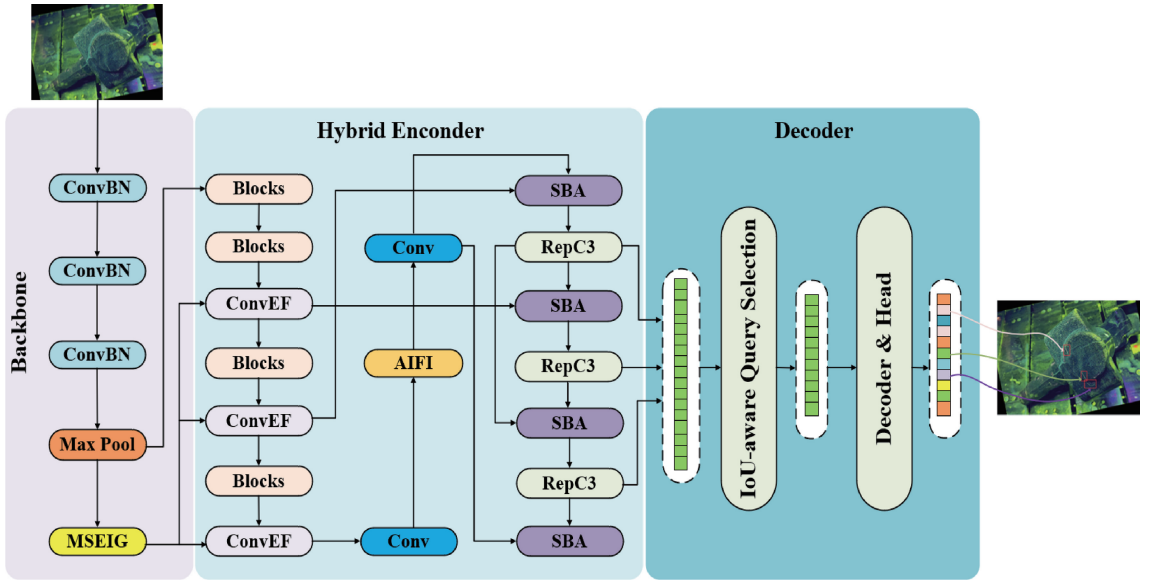


图 2 GSG-DETR 网络结构  
Fig. 2 GSG-DETR network structure

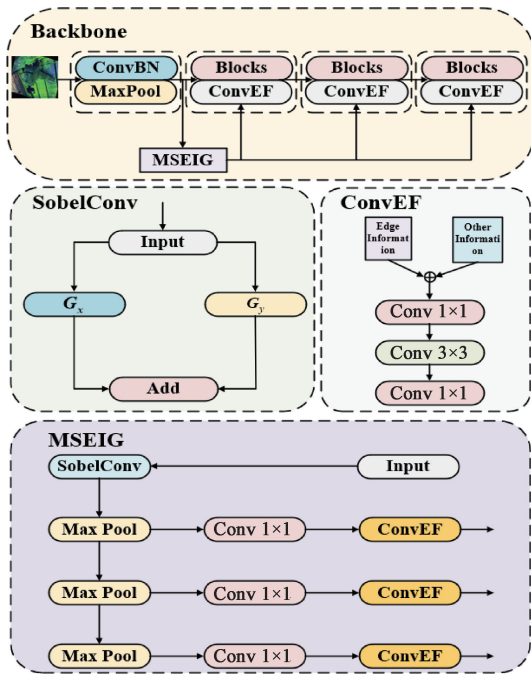


图 3 多尺度边缘信息传递模块图

Fig. 3 Global edge information transfer module

合,ConvEF 再通过  $3 \times 3$  卷积进一步提取融合后的特征,这一操作有助于从融合特征中提取有效的细节信息,增强网络对转向节裂纹边缘的感知能力。最后,ConvEF 通过其他  $1 \times 1$  卷积调整输出特征图的通道数,以确保输出的特征图维度符合后续网络的要求。融合过程如下:

$$X_{fusion} = Conv_{1 \times 1}(Conv_{3 \times 3}(Conv_{1 \times 1}(Concat(E_1, E_2, \dots, E_i)))) \quad (3)$$

式中:  $E_i$  表示第  $i$  层边缘特征,  $X_{fusion}$  为融合特征图。

最终,整个 GLOFT 模块的特征生成过程可通过以下统一表达来表示:

$$y = ConvEF(MSEIG(SobelConv(X))) \quad (4)$$

通过这一设计,GLOFT 有效克服传统卷积网络在边缘建模方面的不足,增强目标的边缘结构特征表达,从而提高目标检测任务中对缺陷边界定位的准确性。

### 2.2 选择边信息聚集(SBA)

在转向节表面缺陷检测任务中,表面裂纹在位置和形状上具有高度多变性,细粒度轮廓描绘和位置校准对小目标准确检测至关重要。传统的特征金字塔网络 FPN (feature pyramid networks,FPN)通过简单的上下采样和拼接方式融合深层与浅层特征,虽能捕捉不同尺度的特征,但该融合方式导致在细节上对裂纹信息造成损失。此外,深层特征有丰富的语义信息,但空间分辨率低,难以捕捉细节;浅层特征空间分辨率高,但缺乏语义信息,难以区分复杂背景和微小缺陷。若直接将低层特征与高层特征融合,易导致信息不对称。

为解决这一问题,本文在颈部网络中引入 SBA<sup>[19]</sup> 模块,详细结构如图 4 所示。该模块通过自适应加权和精细的低层细节信息和高层语义信息融合,以避免信息冗余,并更细粒度地描绘裂纹轮廓,重新校准目标的位置。这种融合方式不仅保留低层特征的细节,增强对微小缺陷的敏感性,还通过精细化加权提升高层特征的语义表达能力,从而提高转向节表面缺陷的检测精度。

不同于直接拼接或加权相加的传统特征融合方法<sup>[20]</sup>,本文设计一种新的重校准注意力单元块重校准注意力单元块(re-calibration attention unit,RCAU)。该单元在融合之前,自适应地提取浅层和深层信息。浅层和深层信息通过不同的方式输入到两个 RCAU 模块中,从而弥补深层语



### 3 实验数据与分析

#### 3.1 实验环境

本研究使用的详细实验配置和环境如表 1 所示。

实验过程中,图像输入大小为  $640 \times 640$ ,训练次数设定为 200 次, batch size 设置为 8。使用 AdamW 优化器对模型参数进行自适应调整,初始学习率设定为  $1 \times 10^{-4}$ 。

表 1 实验配置和环境

Table 1 Experimental configuration and environment

名称	版本
CPU	AMD Ryzen 5 5600 3.50 GHz
GPU	NVIDIA RTX 4060 Ti, 8 GB
编程语言	Python 3.10
操作系统	Windows
深度学习框架	Pytorch 2.0.0, CUDA 11.8

#### 3.2 实验数据集

本文使用的数据集 FDMPI 采集自湖北三环锻造有限公司转向节生产线的探伤车间,涵盖多视角、多物距以及不同光照条件等复杂工况场景,具备较强的工业代表性。原始数据集共包含 362 张分辨率为  $1024 \times 768$  的裂纹图像。

在数据划分方面,为在样本数量有限的情况下兼顾训练充分性与评估可靠性,本文将原始数据按 4:1 比例划分为训练集(290 张)和验证集(72 张),并额外设置与验证集规模相同的独立测试集(72 张)。该划分策略能最大程度保留训练样本,提高模型学习效率,同时确保评估过程的客观性与泛化性。为进一步提升模型在复杂场景下的鲁棒性,训练集中引入了由 Meta 开发的 AugLy<sup>[22]</sup> 图像增强框架,对样本进行多维度增强操作,涵盖几何变换、光照模拟与噪声注入等策略。增强后训练集扩展至 3 147 张图像,包含 4 063 个标注为缺陷(defect)的目标;验证集与测试集则保持原始图像分布,以保证评估结果的可比性与真实性。

最终构建的数据集包括:增强后的训练集(3 147 张/4 063 个目标)、原始验证集(72 张/87 个目标)和测试集(72 张/85 个目标)。实验结果表明,该数据增强方案使样本多样性提升达 985%,显著提升了深度学习模型在复杂工业环境下的泛化能力与鲁棒性。

#### 3.3 评估指标

为评估算法在缺陷检测任务中的性能,选取精确率(Precision)、召回率(Recall)、平均类别精度(mAP)、参数量(Parameters)、计算量(GFLOPs)以及帧数(FPS)作为性能评估指标。其中, Precision、Recall 和 mAP 的计算公式如下:

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

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

$$mAP = \frac{1}{N} \sum_{i=0}^N AP_i \quad (12)$$

其中,  $TP$  表示正确预测为正例的目标数量,即真实为正样本的目标被正确识别的数量;  $FP$  表示错误预测为正例的目标数量,即实际为负样本的目标被误判为正样本的数量;  $FN$  表示错误预测为负例的目标数量,即真实为正样本的目标未被正确识别的数量。  $mAP$  作为衡量模型综合检测性能的指标,表示所有类别精度的均值,能有效反映模型在多类别检测任务中的整体表现。

#### 3.4 剪枝结果分析

为有效降低模型的参数量和计算量,并提升目标检测的识别速度。本文采用 GroupNorm 剪枝方法对优化后的网络进行剪枝。以改进的 GSG-DETR 为基准模型,通过设置不同的 Speed-up ( $Speed-up = MACs(base)/MACs(pruned)$ ),持续调整 Speed-up 参数,实现对网络不同程度的剪枝,从而在精度和效率之间找到最佳平衡。由表 2 可知,采用 GroupNorm 剪枝方法使参数量和计算量大幅减小的同时,仍保证了检测精度。其中 Speed-up 为 1.25 时,剪枝后模型大小 16.2 M,计算量为 49.4 GFLOPs,检测精度保持不变; Speed-up 设置为 1.5,在 mAP 略微下降 0.1% 时,参数量和计算量分别大幅降至 13.2 M, 38.7 GFLOPs,且 FPS 提高至 105.1 fps。 Speed-up 分别为 1.75、2.0、2.25 和 2.5 时,模型参数量和计算量虽有显著下降,但 mAP 也随之大幅度下降,不利于对转向节缺陷更精准的检测。

因此,当 Speed-up 值为 1.5 时,在检测精度和检测速度之间取得最佳平衡。

#### 3.5 消融实验

为深入分析不同改进模块对模型整体性能的影响,本文在相同的实验环境和超参数设置下,分别用 GLOFT 模块和 SBA 模块对主干网络和颈部网络进行改进,并采用通道剪枝 GroupNorm( $Speed-up=1.5$ )进行消融实验。实验共设计 7 组,实验结果如表 3 所示。

组 A 为未经任何改进的 RT-DETR-r18 模型,作为基准模型与其他优化组进行对比。

组 B 加入 GLOFT 模块后,相较于基准模型,算法的精确率、召回率和 mAP 分别提升 0.3%、1.5% 和 0.6%。同时,该模块通过微分算子实际降低基础参数量 0.2 M,计算量同步下降 0.1 GFLOPs。这表明 GLOFT 模块不仅能有效增强模型对缺陷边缘区域多尺度特征的捕捉能力,其微分算子还可优化基础网络参数量。

组 C 采用 SBA 模块后,精确率、召回率和 mAP 分别提升了 0.6%、1.8% 和 1.0%。该组实验说明,SBA 通过重校准注意力复用参数,仅增加 0.4 M 参数量即实现特征融合,有效融合低层次的细节信息与高层次的语义

信息。

组 D 在同时使用 GLOFT 和 SBA 模块后,精确率、召回率和 mAP 显著提升,分别提升 1.1%、1.9% 和 2.1%,mAP 达到最高,为 88.3%。说明两个模块既能强化缺陷边缘区

域特征的捕捉与跨层传递能力又能自适应融合低分辨率边界信息与深层语义特征。证明相融后,能极大提高对转向节微裂纹的检测精度。然而,两者也增加了网络结构的复杂度,导致计算量和参数量分别上升 1.6 M 和 5 GFLOPs。

表 2 不同 Speed-up 剪枝实验结果

Table 2 Pruning experiment results corresponding to different acceleration ratios

Speed-up	Precision/%	Recall/%	mAP0.5/%	Parameters/M	FLOPs/G	FPS(bs=8)
1.0	87.9	84.2	88.3	21.7	62.0	81.3
1.25	87.2	84.5	88.3	16.2	49.4	97.5
<b>1.5</b>	<b>87.3</b>	<b>84.7</b>	<b>88.2</b>	<b>13.2</b>	<b>38.7</b>	<b>105.1</b>
1.75	86.7	81.2	86.9	10.9	32.8	110.3
2.0	85.2	80.8	85.2	9.0	27.6	113.6
2.25	84.8	79.1	84.4	7.9	23.3	115.5
2.5	83.2	78.5	82.7	6.7	20.1	121.6

表 3 消融实验结果

Table 3 Ablation experiment results

组别	GLOFT	SBA	GroupNorm	Precision/%	Recall/%	mAP0.5/%	Params/M	FLOPs/G	FPS(bs=8)
A				86.8	82.3	86.2	20.1	57.0	97.8
B	✓			87.1	83.8	86.8	19.9	56.9	99.2
C		✓		87.4	84.1	87.2	20.5	59.5	86.4
D	✓	✓		<b>87.9</b>	84.2	<b>88.3</b>	21.7	62.0	81.3
E	✓		✓	86.8	83.5	86.6	13.9	<b>36.8</b>	101.9
F		✓	✓	87.0	83.9	87.1	15.6	40.6	99.6
G	✓	✓	✓	87.3	<b>84.7</b>	88.2	<b>13.2</b>	38.7	<b>105.1</b>

组 E 和组 F 通过引入 GroupNorm 剪枝方法,参数量分别降至 13.9 M 和 15.6 M,计算量分别降至 36.8 GFLOPs 和 40.6 GFLOPs,FPS 提升至 101.9 fps 和 99.6 fps。这表明,剪枝方法能大幅减小模型的计算开销,并提升了检测速度,然而,准确率依然未达到最佳。

组 G 在加入 GLOFT 模块和 SBA 模块并结合 GroupNorm 剪枝(Speed-up=1.5)后,mAP 相比组 D 虽略下降 0.1%,但仍高达 88.2%,相比组 A 的 86.2% 提高 2%。同时,精确率和召回率分别达到 87.3% 和 84.7%,均优于组 A。归因于 GLOFT 强化的边缘特征可识别性与 SBA 保障的跨层语义对齐有效补偿剪枝误差,参数量降为 13.2 M,相比组 A 的 20.1 M 减少 34.3%;计算量从 57.0GFLOPs 降低至 38.7GFLOPs,减少 32.1%。同时,FPS 提升至 105.1 fps,超过组 A 的 97.8 fps。使得检测精度和检测速度保持最佳平衡。

综上所述,GSG-DETR 模型不仅在检测精度上取得显著突破,而且计算效率实现大幅提升,充分展示优化策略在提升模型性能方面的成功应用,使得该模型可在工业实际中发挥重要作用。

### 3.6 对比实验

为综合评估所提出改进模型的检测性能,本文设计了与多种主流及最新目标检测算法的对比实验。选取的对比算法包括:基于 CNN 的单阶段检测器(如 YOLOm 系列及其轻量化改进 YOLOv8s)、基于 CNN 的两阶段检测器 Faster R-CNN、基于 Transformer 的主流钢材缺陷检测模型 FHB-DETR、最新目标检测算法 DEIM<sup>[23]</sup> 和 RT-DETR 系列模型,以及本文提出的改进模型 GSG-DETR,实验环境和参数设置保持一致的前提下,在 FDMPI 数据集进行对比实验,结果如表 4 所示。在精度方面,GSG-DETR 实现 84.7% 的召回率与 88.2% 的 mAP,为所有模型中最高,相较轻量化 YOLOv8s 提升 4.8%,优于 YOLOv9m 与 FHB-DETR 分别 3.7% 和 0.4%,显示出优异的检测性能。在模型复杂度上,GSG-DETR 仅需 13.2 M 参数,远低于 YOLOv8m(25.8 M)、YOLOv9m(20.0 M) 和 FHB-DETR (18.2 M),且计算量为 38.7 GFLOPs,显著优于 YOLOv8m(78.9 GFLOPs)、YOLOv9m(76.5 GFLOPs) 和 FHB-DETR (56.3 GFLOPs),接近轻量化改进 YOLOv8s(27.2 GFLOPs),展现出色的模型紧凑性与计算

效率。此外,模型在  $\text{batchsize}=8$  时实现 105.1 FPS,能够满足工业实时检测的要求。

为进一步体现本文模型在高精度和轻量化间的综合优势,本文引入单位参数精度比( $\text{mAP}0.5/\text{Params}$ )作为评估指标,并绘制图 6 进行直观比较。图 6 中通过 3 条不同形状曲线揭示模型性能:点线显示本文模型以 88.2% 的 mAP 居顶;虚线表明模型仅需 13.2 M 参数量,优于大多

数模型;实线量化其精度-参数量平衡效率,本文模型单位参数精度比为 6.68(%/M),显著优于 Faster-RCNN 等其他模型,与改进 YOLOv8s 仅低 0.39(%/M)。同时结合表 4 数据验证,相比于轻量化改进 YOLOv8s,本文模型仅以 1.4 M 的参数增量代价,实现 4.8% 的 mAP 显著突破,有效实现检测精度与轻量化的平衡。

表 4 对比实验结果

Table 4 Comparative experimental results

模型	Precision/%	Recall/%	mAP0.5/%	Params/M	FLOPs/G	FPS(bs=8)
Faster-RCNN	78.2	64.8	73.7	56.3	218.2	78.3
YOLOv8m	85.9	77.6	84.1	25.8	78.9	92.4
YOLOv9m	86.3	78.2	84.5	20.0	76.5	98.7
YOLOv10m	85.5	78.7	84.3	16.5	63.4	101.5
改进 YOLOv8s	86.7	70.5	83.4	<b>11.8</b>	<b>27.2</b>	107.1
DEIM	85.4	84.1	86.6	32.3	93.7	<b>111.2</b>
RT-DETR-r18	86.8	82.3	86.2	20.1	57.0	97.8
RT-DETR-r34	85.6	84.3	86.7	31.7	88.1	87.6
FHB-DETR	<b>87.3</b>	83.6	87.8	18.2	56.3	96.4
Ours	87.1	<b>84.7</b>	<b>88.2</b>	13.2	38.7	105.1

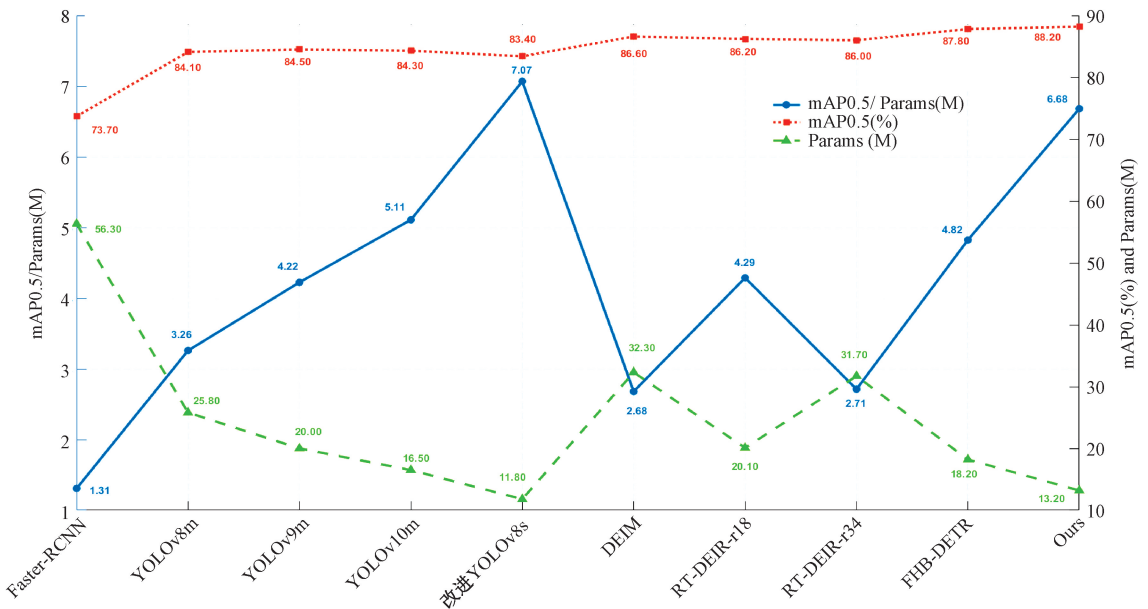


图 6 各模型检测性能与模型复杂度对比

Fig. 6 Comparison chart of the detection performance and model complexity of each model

### 3.7 可视化分析

为直观验证 GSG-DETR 算法与 YOLOv10 m 及 RT-DETR-r18 在转向节微裂纹检测任务中的准确性差异,本文从测试集中选取部分检测结果进行可视化对比,在稀疏缺陷场景、密集缺陷场景及复杂背景场景均优于 YOLOv10 m 及 RT-DETR-r18,检测效果如图 7 所示。在

图 7 第 1 行所示的稀疏缺陷场景中,本文算法相较基准模型及 YOLOv10 m 展现出更高的检测精度,并能生成更加精准的检测框。在图 7 第 2 行的密集缺陷场景中,GSG-DETR 能有效识别和定位多个微裂纹,通过精准的裂纹边缘捕捉与高置信度框选,显著减少漏检与误检的情况。在图 7 第 3 行复杂背景场景中,虽然背景中包含铭文及其他

干扰元素,GSG-DETR 仍能有效抑制背景干扰,准确识别 缺陷。

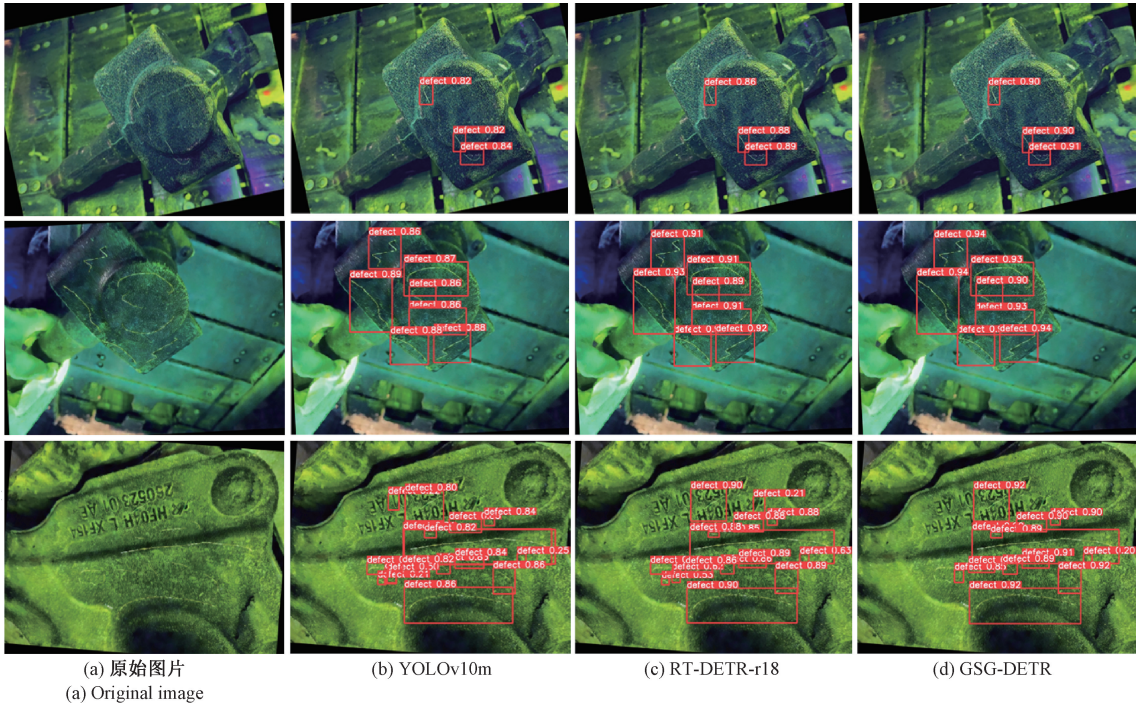


图 7 原始图片、YOLOv10m、RT-DETR-r18 和 GSG-DETR 模型对比

Fig. 7 Comparison of the original image, YOLOv10m, RT-DETR-r18 and GSG-DETR models

综上所述,在稀疏缺陷、密集缺陷及复杂背景场景下,本文算法均展现出优异的检测性能,进一步验证其在转向节表面缺陷检测任务中的准确性。

为对比不同模型对转向节裂纹区域的关注度,并为热成型过程中的缺陷控制提供指导,本文还采用 Grad-CAM<sup>[24]</sup>方法生成热力图进行可视化分析,如图 8 所示。

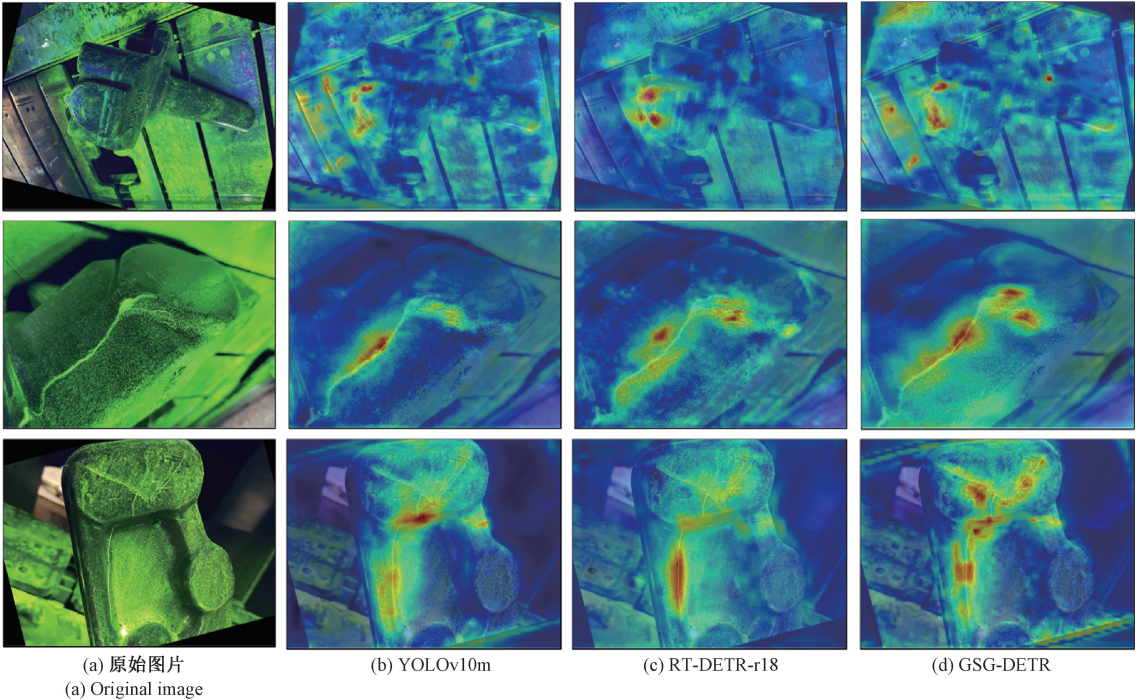


图 8 原始图片、YOLOv10m、RT-DETR-r18 和 GSG-DETR 模型热力图对比

Fig. 8 Comparison of thermal maps of the original, YOLOv10m, RT-DETR-r18 and GSG-DETR models

在热力图中,红色区域代表模型对特征点高度关注。GSG-DETR 在转向节不同部位对小缺陷的响应敏感,可精准捕捉裂纹边缘,并有效抑制背景干扰。热力图揭示,锻件缺陷主要集中于 3 个关键区域:边缘区域、表面中心区域及厚薄过渡区域。这些部位由于应力集中或材料流动不均,成为缺陷发生的高风险区域。通过热力图对比分析,GSG-DETR 更能显示裂纹聚集区域,在准确性方面优于 RT-DETR-r18 和 YOLOv10 m。此外,热力图还为转向节热成型过程中的缺陷控制提供有效的工艺优化依据。

### 3.8 NEU-DET 数据集实验

为进一步验证 GSG-DETR 在钢材表面缺陷检测任务中的泛化能力,本文在 NEU-DET<sup>[25]</sup> 数据集上与 RT-DETR-r18 及改进的 RT-DETR 方法实验,NEU-DET 部分图片案例如图 9 所示,泛化实验结果如表 5 所示。从整体性能指标来看,GSG-DETR 在精确率、召回率和 mAP 3 个核心评价指标上均优于对比模型。其中,GSG-DETR 相较原始 RT-DETR-r18,mAP 提升 4.3%,相较改进 RT-DETR 提升 3.3%,显著增强了检测精度和鲁棒性。在模型复杂度方面,GSG-DETR 的参数量为 13.0 M,虽略高于改进 RT-DETR 的 10.4 M,但远低于原始 RT-DETR-r18 的 19.9 M;计算量为 38.6 GFLOPs,处于三者之间,综合体现出较好的精度与效率平衡。

从各类缺陷的检测精度来看,GSG-DETR 在开裂类别的精度达到 37.2%,高于改进 RT-DETR 的 35.4%和原始 RT-DETR-r18 的 29.2%,表明其对边界模糊、结构不规则的小目标裂纹具有更强的感知能力;在点蚀表面和轧制氧化皮两个小目标类别中,检测精度分别达到 98.2%和 79.7%,在所有对比模型中表现最优,体现出 GSG-DETR

在微小区域缺陷检测方面的突出性能。

相比之下,改进的 RT-DETR 虽通过 SlideVarifocal-Loss 加强对难样本的关注,并取得了在内含物类缺陷上最高的识别精度,但其在小目标缺陷如 Ps、Rs 的检测能力仍不及 GSG-DETR。这一结果进一步验证了本文模型中边缘信息传递与多尺度边界语义对齐机制的有效性。

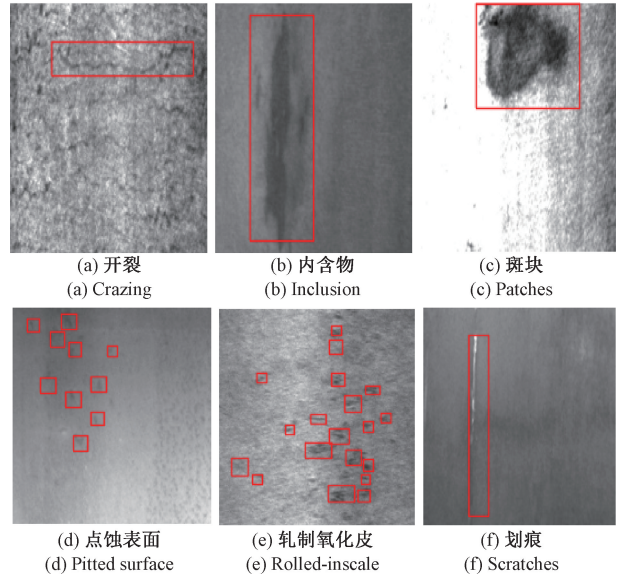


图 9 NEU-DET 部分图片案例

Fig. 9 Examples of some images in the NEU-DET

综上所述,本文提出的算法通过捕捉缺陷边缘区域的特征、并自适应地融合低分辨率边界信息与深层语义特征,有效提升检测精度。同时,结合 GroupNorm 剪枝算法,显著降低模型的参数量和计算量,使得模型更契合工业部署。

表 5 泛化实验结果

Table 5 Generalization experiment results

模型	Precision/%	Recall/%	mAP0.5/%	Params/M	FLOPs/G	Cr/%	In/%	Pa/%	Ps/%	Rs/%	Sc/%
RT-DETR-r18	66.5	72.0	71.9	19.9	56.9	29.2	57.9	96.0	94.6	58.9	94.7
改进的 RT-DETR	69.1	73.8	72.9	<b>10.4</b>	<b>32.4</b>	35.4	<b>80.2</b>	92.5	80.2	57.7	91.5
Ours	<b>74.7</b>	<b>76.5</b>	<b>76.2</b>	13.0	38.6	<b>37.2</b>	67.1	88.0	<b>98.2</b>	<b>79.7</b>	87.0

## 4 结 论

针对目前汽车转向节检测算法对小目标缺陷检测精度低、缺陷边界信息关注不足,以及模型自身计算量、参数量较大造成检测速度慢的问题,本文提出一种新的转向节表面缺陷检测 GSG-DETR 算法。通过引入多尺度边缘信息传递模块和选择边信息聚集模块,有效提升了模型在复杂背景和微小缺陷场景中的检测精度。特别是 GLOFT 模块能够强化缺陷边缘区域特征的捕捉与跨层传递,从而提高模型对缺陷边界的敏感度;SBA 模块通过自适应融合低分辨率边界信息与深层语义特征,优化多尺度缺陷的边界

语义对齐策略,有效增强模型对微裂纹等小尺度缺陷的精准表征能力。

其次,GroupNorm 在保证检测精度的同时,显著降低模型的参数量和计算量,提高了检测速度,满足工业环境中的实时检测需求。通过对模型的剪枝实验与消融实验,本文验证在不同 Speed-up 下,模型在检测精度、参数量、计算量以及 FPS 等方面的平衡。特别是在 Speed-up 为 1.5 时,检测精度与模型压缩之间达到最佳平衡。

通过对比实验,本文提出的 GSG-DETR 算法在转向节表面缺陷检测任务中,mAP 达到 88.2%,较现有主流算法 RT-DETR 和 YOLO 系列在多项性能指标上表现更为

优秀。此外,泛化实验还验证了该模型在钢材表面缺陷检测中具备泛化能力,表明其在其他缺陷检测任务中的有效性。

综上,GSG-DETR 算法不仅在转向节表面裂纹检测中取得显著突破,还在工业化部署和实时应用中展示了强大的性能和可行性,具有重要的工程应用价值。目前,本研究仅停留在算法研究阶段,未来将把该算法部署到实际应用场景中。

## 参考文献

- [1] 汪西. 某汽车转向节锻造工艺优化及淬火裂纹控制策略研究[D]. 重庆:重庆大学,2023.  
WANG X. Optimization of forging process of an automobile steering knuckle and research on control strategy of quenching cracks [D]. Chongqing: Chongqing University, 2023.
- [2] 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 raildefect non-destructivetesting and evaluation[J]. Chinese Journal of Scientific Instrument, 2019, 40 (2):11-25.
- [3] 邓志鹏,何施茗,杨根,等. 基于深度学习的纹理表面缺陷检测方法综述[J]. 计算机集成制造系统,2025,31(3):721-745.  
DENG ZH P, HE SH M, YANG G, et al. Survey of texture surface defect detection method based on deep learning [J]. Computer Integrated Manufacturing Systems, 2025, 31(3): 721-745.
- [4] 沈冰星,黄洪琼. 基于 ESE-YOLO 的钢带表面缺陷检测研究[J]. 电子测量与仪器学报,2025,39(8):126-135.  
SHEN B X, HUANG H Q. Research on surface defect detection of steel strip based on ESE-YOLO[J]. Journal of Electronic Measurement and Instrumentation, 2025,39 (8):126-135.
- [5] 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.
- [6] 朱栋,胡伟笑,赵腾. 太阳能电池片图像校正与表面缺陷检测[J]. 电子测量技术,2024,47(8):126-133.  
ZHU D, HU W X, ZHAO T. Image calibration and surface defect detection of solar photovoltaic cells[J]. Electronic Measurement Technology, 2024, 47 (8): 126-133.
- [7] GIRSHICK R, DONAHUE J, DARRELL T, et al. Rich feature hierarchies for accurate object detection and semantic segmentation[C]. IEEE Conference on Computer Vision and Pattern Recognition, 2014: 580-587.
- [8] TERVEN J, CORDOVA-ESPARZA D M, ROMERO-GONZÁLEZ J A. A comprehensive review of YOLO architectures in computer vision: From YOLOv1 to YOLOv8 and yolo-nas [J]. Machine learning and knowledge extraction, 2023, 5(4): 1680-1716.
- [9] CAO C, WANG B, ZHANG W, et al. An improved faster R-CNN for small object detection[J]. IEEE Access, 2019, 7: 106838-106846.
- [10] KANG S H, PARK J S. Aligned matching: Improving small object detection in SSD[J]. Sensors, 2023, 23(5): 2589.
- [11] JIANG P, ERGU D, LIU F, et al. A review of YOLO algorithm developments [J]. Procedia Computer Science, 2022, 199: 1066-1073.
- [12] 张岳,张上,王恒涛,等. 基于跨尺度特征提取的锻件表面裂纹检测算法[J]. 计算机集成制造系统,2023,39(3):102-115.  
ZHANG Y, ZHANG SH, WANG H T, et al. Crack detection algorithm for forging surface based on cross-scale feature extraction [J]. Computer Integrated Manufacturing Systems, 2023, 39(3): 102-115.
- [13] 张伟健,徐向英,章永龙. 改进 RT-DETR 的钢材表面缺陷检测算法[J]. 机械科学与技术,2025,42(2):45-53.  
ZHANG W J, XU X Y, ZHANG Y L. Improved RT-DETR algorithm for detecting surface defects in steel[J]. Mechanical Science and Technology for Aerospace Engineering, 2025, 42(2): 45-53.
- [14] 秦伦明,张云起,崔昊杨,等. 基于改进 RT-DETR 的极端天气下交通标志检测方法[J]. 电子测量技术,2025,48(9):56-64.  
QIN L M, ZHANG Y Q, CUI H Y, et al. Improved RT-DETR based method for traffic sign recognition in extreme weather [J]. Electronic Measurement Technology, 2025, 48(9): 56-64.
- [15] 张靖雯,孙坚,徐红伟,等. 基于改进 RT-DETR 的玻璃绝缘子缺陷检测算法[J]. 电子测量技术,2025,48(4):78-87.  
ZHANG J W, SUN J, XU H W, et al. Defect detection algorithm for glass insulators based on improved RT-DETR [J]. Electronic Measurement Technology, 2025, 48(4): 78-87.
- [16] YU CH H, CHEN X W. Railway rutting defects

- detection based on improved RT-DETR[J]. Journal of Real-Time Image Processing, 2024, 21(4): 146.
- [17] ZHAO Y, CHEN B, LIN J, et al. DETRs beat YOLOs on real-time object detection [J]. ArXiv preprint arXiv:2304.08069, 2023.
- [18] HE K M, ZHANG X Y, REN SH Q, et al. Deep residual learning for image recognition [C]. IEEE Conference on Computer Vision and Pattern Recognition, 2016: 770-778.
- [19] TANG F L, XU ZH X, HUANG Q M, et al. DuAT: Dual-aggregation transformer network for medical image segmentation [C]. Chinese Conference on Pattern Recognition and Computer Vision, 2023: 343-356.
- [20] AHMAD K, KHAN K, AL-FUQAHA A. Intelligent fusion of deep features for improved waste classification [J]. IEEE Access, 2020, 8: 96495-96504.
- [21] FANG G, MA X, SONG M, et al. Depgraph: Towards any structural pruning [C]. IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2023: 16091-16101.
- [22] PAPAKIPOS Z, BITTON J. AugLy: Data augmentations for adversarial robustness[C]. IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2022: 156-163.
- [23] HUANG SH H, LU ZH CH, CUN X D, et al. DEIM: DETR with Improved Matching for Fast Convergence[J]. ArXiv preprint arXiv:2412.04234, 2024.
- [24] SELVARAJU R R, COGSWELL M, DAS A, et al. Grad-CAM: Visual explanations from deep networks via gradient-based localization[C]. IEEE International Conference on Computer Vision, 2017: 618-626.
- [25] HE Y, SONG K CH, MENG Q G, et al. An end-to-end steel surface defect detection approach via fusing multiple hierarchical features[J]. IEEE Transactions on Instrumentation and Measurement, 2019, 69(4): 1493-1504.

### 作者简介

张上, 博士, 副教授, 主要研究方向为物联网、目标检测、计算机应用和图像处理。

E-mail: zhangshang@ctgu.edu.cn

朱帅, 硕士研究生, 主要研究方向为目标检测和嵌入式技术。

E-mail: zhu\_shuai0927@163.com

张岳(通信作者), 硕士研究生, 主要研究方向为图像识别和智能制造。

E-mail: zhangyue980202@ctgu.edu.cn