基于多头注意力机制的改进YOLOv8的钢结构防腐涂层缺陷检测

俞 悦[1],蒋首超[1,2*],王亦君[1]
(1.同济大学 土木工程学院,上海 200092;2.同济大学 土木工程防灾减灾全国重点实验室,上海 200092)

摘 要: 钢结构耐腐蚀能力差,需要防腐涂层进行保护,涂层老化会导致防腐效果降低甚至失效,因此对防腐涂层的维护至关重要。本文构建了用于服役钢结构防腐涂层缺陷检测的数据集,并提出了一种基于多种注意力机制改进的YOLOv8算法,用于检测钢结构防腐涂层中的缺陷。对Backbone中的C2f模块加入了多尺度的可变形注意力机制,提高对小目标的检测能力。Head采用了融合自注意力机制的Dynamic Head,增强了尺度感知、空间感知和任务感知能力。采用优化分类损失函数Slide Loss来解决数据集中样本不均衡的问题,优化IoU损失函数使Shape IoU提升了多尺度目标的检测能力。所提出的改进算法解决了传统机器学习图像分类模型无法进行多缺陷检测、泛化性能不强的问题,与YOLOv8基线算法相比,改进算法提高了5.82%的精确率、5.01%的召回率和4.78%的m_{AP}指标。

关键词: 深度学习;改进的YOLOv8算法;多头注意力机制;计算机视觉;缺陷检测;防腐涂层;数据集
中图分类号: TP391,TU761.1^{+}3 **文献标志码:** A **文章编号:** 1671-9379(2026)08-0116-07
DOI: 10.13969/j.jzgjgjz.20241223001

Improved YOLOv8 Based on Multi-Head Attention Mechanism for Defect Detection of Anti-Corrosion Coatings on Steel Structures

YU Yue[1], *JIANG Shouchao*[1,2*], *WANG Yijun*[1]
(1. College of Civil Engineering, Tongji University, Shanghai 200092, China; 2. State Key Laboratory of Disaster Reduction in Civil Engineering, Tongji University, Shanghai 200092, China)
*Corresponding author: scjiang@tongji.edu.cn

Abstract: Steel structures have poor corrosion resistance and require anti-corrosion coatings for protection. Coating aging can reduce or even eliminate the anti-corrosion effects, making the maintenance of their anti-corrosion coatings crucial. This study independently constructs a dataset for defect detection of anti-corrosion coatings on steel structures in service, and proposes an improved YOLOv8 algorithm based on multiple attention mechanisms for detecting defects in anti-corrosion coatings on steel structures. A multi-scale deformable attention mechanism is added to the C2f module in Backbone to improve its ability to detect small targets. The Head adopts a Dynamic Head that integrates a self-attention mechanism to enhance scale, spatial, and task awareness. The classification loss function is optimized as Slide Loss to address sample imbalance in the dataset, and the IoU loss function is optimized as Shape IoU to enhance the detection capability of multi-scale targets. The proposed improved algorithm overcomes the limitations of traditional machine learning image classification models that cannot perform multi-defect detection and have weak generalization performance. Compared with the YOLOv8 baseline algorithm, the improved algorithm improves

收稿日期:2024-12-23

作者简介:
俞 悦(2000—),女,硕士研究生,主要从事钢结构涂层表面缺陷自动检测方法方面的研究。E-mail:2232332@tongji.edu.cn。
通信作者:
蒋首超(1971—),男,博士,副教授,主要从事智能建造、工程结构抗火理论与设计方法方面的研究。E-mail:scjiang@tongji.edu.cn。

precision by 5.82%, recall by 5.01%, and m_{AP} index by 4.78%, respectively.

Keywords: deep learning; improved YOLOv8 algorithm; multi-head attention mechanism; computer vision; defect detection; anti-corrosion coating; dataset

钢结构在现代各类建筑中应用广泛。钢材因自身金属化学特性耐腐蚀能力较差。腐蚀不仅会造成钢材强度、塑性变形等性能退化，同时会引发钢材表面不平整及蚀坑，从而导致应力分布改变与应力集中[1]，进而降低钢结构的安全性。防腐涂层涂装是钢结构防腐工程的关键环节，但受涂装工艺、结构长期暴露于外部环境等因素影响，钢结构涂层易出现各类物理、化学老化现象。剥落、裂缝、空鼓、锈蚀等缺陷不仅破坏了涂层完整性，还会显著降低甚至完全丧失防腐效果，威胁结构的耐腐蚀性能。为保障建筑结构安全，服役期钢结构防腐涂层的缺陷检测与维护尤为重要。传统人工检测方法工作量大，且易受检测人员主观因素影响；基于红外热成像技术的检测方法，易受环境温度、缺陷深度、热流密度等因素干扰，导致检测精度降低[2]。

近年来，计算机视觉技术快速发展，凭借无接触、检测对象适应性广、覆盖范围大、作业环境安全便捷等优势，已在汽车、机械、工业制造等多个领域广泛应用[3]。XU等[4]提出了一种针对汽车涂料特征的蚁群优化边缘检测算法(APFACO)，通过结合局部与全局的更新策略，实现对车身漆膜六种缺陷的高效检测，识别率高达97.76%。吴旭东等[5]提出了一种基于光度学特征和SVM的识别方案，实现了铝合金压痕、划痕、污点、擦伤、坑蚀五种缺陷的分类，达到了96.5%的识别率。此外，研究者针对印制电路板(PCB)缺陷检测提出YOLOv8-PCB检测算法[6]，结合WIoUv3损失函数，提高了检测精度。轻量化网络模型SL-Unet[7]通过多尺度特征融合增强了对复杂背景下缺陷的检测能力。联合注意力引导特征融合网络JAFFNet[8]采用联合注意力引导特征融合，提升了对低对比度缺陷的检测能力。

计算机视觉在土木工程领域的研究也日益增多。GAO等[9]提出了一种基于CNN的钢材表面缺陷半监督学习检测方法，在训练过程中，模型对无标签的数据构造伪标签，并持续地调整伪标签权重优化模型，在钢铁生产车间中试点运行，识别率为86.72%。LIU等[10]采用U-Net网络结构对混凝土裂缝进行检测，使用特征融合来获得更高的精度。RAHMAN等[11]结合基于纹理特征的无监督图像分割方法和基于RGB特征的分类器，使用DeepLab网络实现了结构腐蚀的检测和分割，更准确地检测腐蚀边界。近年来，研究者进一步优化了缺陷检测模型，如基于YOLOv9的配电网绝缘子缺陷检测方法[12]，在提高检测精度的同时实现了81FPS的高效推理。融合注意力机制的轻量化航空发动机表面缺陷检测算法通过级联检测头策略提升了缺陷检测的精度[13]。此外，基于注意力机制的U-Net叶片缺陷检测方法在IoU指标上较标准U-Net检测方法提升了9.38%[14]，进一步提高了检测的可靠性。

尽管基于CNN和基于U-Net的检测方法已广泛用于缺陷检测，但近年来Transformer架构在该领域展现出巨大的潜力。ZHANG等[15]和ZHAO等[16]分别提出了DINO与RT-DETR两种基于Transformer架构的目标检测算法，通过改进编解码架构优化了缺陷检测的精度和推理效率。针对异常检测问题，VIEIRA等[17]设计AttentDifferNet，融合SENet和CBAM等注意力模块，实现了模型AUC指标的提升。目前土木工程领域的视觉检测对象大多聚焦于工厂环境下的建筑材料缺陷，或单一的裂缝和腐蚀问题，针对真实环境中钢结构防腐涂层缺陷的研究仍较为欠缺。首先，钢结构服役环境复杂，背景多样、光照变化等因素对缺陷检测影响较大；其次，防腐涂层缺陷常多种并存，且缺陷在尺寸、形态上差异显著。然而，传统的图像处理和机器视觉技术通常只能识别和分类单一缺陷，而基于Transformer的检测方法虽然在一般目标检测任务上表现优越，但在复杂工业场景下仍面临计算成本较高的问题。因此，如何结合CNN和Transformer的优势，设计兼具高效性和鲁棒性的缺陷检测模型，是一个值得深入研究的方向。

针对钢结构防腐涂层所处真实环境复杂以及缺陷多样的特点，本文提出了一种基于多头注意力机制的改进YOLOv8模型的深度学习目标检测算法，可实现对钢结构防腐涂层缺陷的识别与定位。通过图像检测实验表明，该改进算法具有出色的鲁棒性，即使在复杂背景下仍能保持较高的检测准确率；多尺度特征金字塔结构能够有效地检测不同尺度的缺陷，确保微小缺陷不会被遗漏；计算效率显著提升，检测速度更快；算法还解决了数据集中样本不均衡的问题，并增强了对小目标和多尺度目标的检测能力。

1　基于多头注意力机制的改进模型

1.1　YOLOv8网络结构

YOLOv8的主干网络(Backbone)延用CSPDarknet架构，核心的CBS卷积块由卷积层、批量归一化层和SiLU激活函数层构成。SPPF金字塔池化层由CBS卷积块、最大池化层和连接层组成，并将最大池化层的连接方式从SPP模块的“并联”改为“串联”，从而实现全局特征与局部特征在特征图维度上的融合，减小了计算量的

同时增大了感受野。

此外，YOLOv8引入了新的C2f模块来替代YOLOv5中的C3模块，C2f模块包含两个CBS卷积块：第一个卷积块扩展了输入通道，并通过Split操作将特征图分割后输入Bottleneck模块，可将多个残差连接并将所有输出进行汇总；第二个卷积块则将扩展后的特征图合并回原始通道数。这些改进使C2f模块获得了更多的梯度流信息。

YOLOv8作为YOLO系列中首个无锚模型，可以直接预测目标中心的位置，而不再依赖于锚框的偏移量。这种无锚检测方法减小了计算量，提高了NMS后处理的速度。同时，分类头与检测头在Head网络部分实现了解耦，进一步优化了检测性能。

1.2 可变形注意力机制

为了增强YOLOv8模型在处理尺寸不规则的涂层缺陷时的特征提取能力，本文对C2f模块进行了基于可变形注意力(Deformable Attention)的改进，并将其命名为DA_C2f。可变形注意力[18]是一种新型的自注意力机制，通过学习偏移量进而调整了自注意力机制中Key和Value的位置。鉴于在视觉注意力机制中，不同的Query通常具有相似的注意力图，即每个Query可以共享经过变形的Key和Value，这使得模型能够更灵活地聚焦于图像中的目标区域，从而捕捉到信息价值更高的特征。

本文将可变形注意力机制整合到Bottleneck模块中的第二个CBS卷积块中，并设置Shortcut=True，最终的输出为原始输入加上经过可变形注意力模块的输出。

1.3 基于自注意力的动态检测头

目标检测头面临的挑战主要有三类：一是尺度感知，即目标通常具有不同的尺度特征；二是空间感知，即不同目标的空间形态各不相同；三是任务感知，即边界框、中心点和角点等描述子的约束信息完全不同。多数网络的Head改进仅能解决其中一两类问题，而动态检测头(Dynamic Head)[19]能同时兼顾这三类问题，因此本文引入了动态检测头。在YOLOv8中部署Dynamic Head，分支设计变得更加灵活，只要将预测类型添加到其末端即可。

1.4 损失函数优化

YOLOv8在目标检测中采用无锚框结构，直接输出分类结果和边界框预测结果，其损失函数由分类损失和边界框回归损失构成。钢结构防腐涂层缺陷严重程度差异大、检测任务复杂，且存在缺陷样本数量不均衡问题。

1.4.1 分类损失

在多分类任务中，YOLOv8的分类损失采用了基于N个目标的二元交叉熵。交叉熵损失计算了分类预测概率分布与真实标签概率分布的交叉熵，以此来衡量模型的性能。尽管交叉熵损失作为经典的损失函数之一被广泛应用，但在本文建立的数据集上表现不佳，难以使损失函数值收敛到较低水平。

为了让模型更关注分类中的困难样本，并解决样本难度不均衡的问题，本文引入Slide Loss作为分类损失函数。Slide Loss在小目标检测和困难样本检测中具有一定的改进效果[20]。Slide Loss首先通过目标边界框与预测框的IoU值来区分“简单样本”和“困难样本”。然后，基于所有边界框IoU值的平均值设定阈值μ。最后，通过引入Slide函数对IoU值在区间$(\mu-0.1, \mu)$内的难识别样本增加权重，从而让模型更加关注困难样本。

1.4.2 边界框回归损失

原始YOLOv8的边界框回归损失函数采用了CIoU损失函数。虽然CIoU损失函数通过添加中心点归一化距离损失项和形状损失项来减少Ground-truth与Anchor Box之间的宽高比差异，但仍未充分考虑边界框的形状和尺度等固有属性对其回归效果的影响[21]。为了解决这一不足，本文采用Shape IoU，它引入了与目标大小相关的缩放因子，以及边界框宽度和高度，使得Shape IoU能够更好地考虑这些固有属性，增强回归效果。

为了简化表达，我们在后文中将改进后的模型简称为Attention_YOLOv8。

2 实验与分析

2.1 数据集准备

目前，国内外针对防腐涂层缺陷检测的研究较少，尚未有公开的涂层缺陷数据集支持深度学习模型的训练。为了解决这一问题，本文自主构建了一个用于服役钢结构防腐涂层缺陷检测的数据集。通过拍摄桥梁、车站、住宅等多种钢结构防腐涂层表面的高分辨率图像，共收集了961张样本图像。数据集涵盖了剥落、空鼓、裂纹、锈蚀四类表面缺陷，图像比例范围从1∶1至3∶4不等。图1展示了数据集的部分样本。

数据集标注的规范性和准确性直接影响算法的性能。为了确保标注质量，本文对上述961张样本采用LabelImg软件进行标注，标注过程中遵循如下原则：

(1) 在标注缺陷时，应避免边界框过小，以防止池化操作导致特征消失；

(2) 对于密集缺陷，采用合并处理，避免边界框过于紧凑；

(3) 边界框应靠近缺陷边缘，但避免与边缘完全重合，以防止数据增强过程中出现误差；



图1　防腐涂层缺陷数据集样例

Fig. 1　Illustrations of anti-corrosion coating defect dataset

(4) 对于重叠度较高的缺陷，应出于实际工程应用的考虑，对其中最严重的缺陷进行标注，如对于涂层剥落或裂纹后发生腐蚀的情况，将其标注为腐蚀缺陷。

由于拍摄的缺陷图像具有较高的分辨率，而YOLOv8模型常用的训练样本尺寸为640像素×640像素。为避免因图像尺寸过大占用大量内存导致训练时间延长，同时防止过度缩小图像导致小目标特征丢失，降低模型召回率和精确度，本文在模型训练前对数据集进行了滑动窗口分割处理。最终，将缺陷样本统一切割成1 200×1 200大小的子图，共生成8 193张样本图像。采用8∶2比例划分训练集与验证集，即训练样本6 554张、验证样本1 639张。样本标签为YOLO格式的txt文件。通过统计，各类缺陷的数量和占比如表1所示。

表1　数据集中各类缺陷的数量和占比

Table 1　The number and proportion of various defects in the dataset

缺陷类型	数量	在缺陷总数中占比/%	与样本数之比
剥落	7 445	23.99	0.91
空鼓	2 629	8.47	0.32
裂纹	3 275	10.55	0.40
锈蚀	17 684	56.98	2.16

2.2　实验环境及配置

实验中硬件配置的CPU为Intel i7 12700F，显卡为技嘉RTX 4090魔鹰24G；软件配置的系统为Windows 10，集成开发环境为Pycharm Community，数据科学工具为Anaconda；实验环境为Python 3.8，CUDA 11.8，Pytorch 2.2.1。

2.3　评价指标

本文用于评价模型性能的指标主要有精确率(Precision，P)、召回率(Recall，R)，均基于混淆矩阵完成计算。对于分类模型，混淆矩阵如表2所示。

表2　混淆矩阵

Table 2　Confusion matrix

混淆矩阵		预测类别	
		正类别	负类别
真实类别	正类别	True Positive (T_P)	False Negative (F_N)
	负类别	False Positive (F_P)	True Negative (T_N)

根据混淆矩阵可计算模型评价指标，如式(1)和式(2)所示：

$$P=\frac{T_P}{T_P+F_P} \tag{1}$$

$$R=\frac{T_P}{T_P+F_N} \tag{2}$$

由于引入了锚框、交并比IoU的概念，因此增加了平均精确率A_P(Average Precision)来衡量每个类别的检测精度。A_P用于衡量不同的检测置信度阈值下的模型检测能力，可以通过对准召曲线求积分得到，计算公式如式(3)所示。所有类别的平均精确率的平均值m_{AP}(Mean Average Precision)是目标检测中最常用的性能评价指标，用于衡量N个目标检测模型的平均精度，对不同的类别计算A_P并取平均值，即得到m_{AP}，计算公式如式(4)所示。

$$A_P=\int_0^1 p(r)\mathrm{d}r \tag{3}$$

$$m_{AP}=\frac{1}{N}\sum_{i=0}^{N}A_{Pi} \tag{4}$$

2.4　实验结果

YOLOv8模型有n、s、m、l和x五个版本。YOLOv8m在检测精度上显著优于较小版本模型的检测精度，且与较大版本模型的检测精度相比差异不大，同时其计算资源消耗更低，因此本文选择了YOLOv8m作为基线网络模型，并在此基础上进一步优化网络结构和训练方法。

2.4.1　对比实验

本文选择了目标检测领域的经典模型进行对比实验，实验结果如表3所示。从实验结果可以看出，Attention_YOLOv8m在精确率、召回率和平均精确率均

值上均具有显著优势。同时，Attention_YOLOv8m的参数量处于中等偏上水平且帧率较高，反映了其具备较强的表示能力、较快的推理速度，适合在需要高精度和高吞吐量的应用中使用。

表3 不同目标检测模型的对比实验结果

Table 3 Comparative experimental results of different object detection models

模型	P/%	R/%	m_{AP}/%	参数量	帧率
Faster R-CNN	79.63	66.15	57.80	41.2 M	~20 FPS
YOLOv5m	80.57	69.21	70.16	21.4 M	156 FPS
YOLOv8m	83.12	70.37	73.24	25.9 M	408 FPS
Attention_YOLOv8m	88.94	75.38	78.02	48.8 M	251 FPS

2.4.2 消融实验

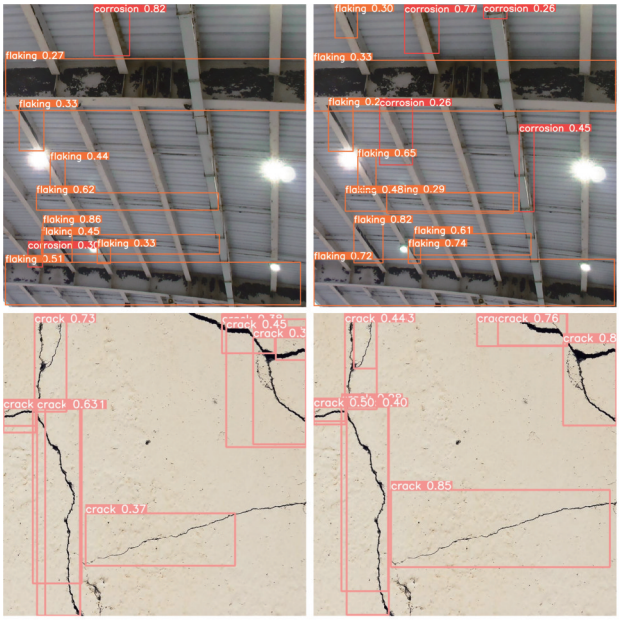
消融实验采用“控制变量法”对改进模型进行评估。由于本文提出的Attention_YOLOv8m模型引入了可变形注意力、动态检测头和损失函数优化三个改进模块，因此需要通过消融实验来验证每个改进点的有效性。表4展示了消融实验的结果。结果表明，在本文的数据集中，可变形注意力模块主要提高了准确率，而动态检测头则显著提高了召回率。图2至图4为钢结构防腐涂层检测模型在本文自建数据集的消融实验结果示例。图中“corrosion”代表腐蚀缺陷，“flaking”代表剥落缺陷，“crack”代表裂纹缺陷，“blistering”代表空鼓缺陷。

表4 钢结构防腐涂层检测模型的消融实验结果

Table 4 The ablation experiment results of the anti-corrosion coating detection model for steel structures

可变形注意力	动态检测头	损失函数优化	P/%	R/%	m_{AP}/%
×	×	×	83.12	70.37	73.24
√	×	×	84.52	69.17	73.15
×	√	×	83.90	71.43	74.28
×	×	√	84.29	70.47	73.56
√	√	×	86.72	73.76	76.63
×	√	√	85.65	74.72	75.12
√	×	√	86.38	72.89	76.07
√	√	√	88.94 (+5.82)	75.38 (+5.01)	78.02 (+4.78)

与原始YOLO v8相比，增加可变形注意力模块后准确率从83.12%提升至84.52%，提升了1.4%。图2展示了几个结果实例。可变形注意力模块能够更好地捕捉和聚焦于小目标及其周围的重要特征，因此在检测结果中，小目标的识别率通常会有所提高，漏检的情况减少。如图2第一组结果，多个小目标在添加了可变性注意力模块之后被检测出。由于可变形注意力模块更能适应目标的变形特征，检测结果中的边界框往往更紧密、更准确。如图2第二组结果，边界框与目标的真实边界贴合得更好，同时置信度也有所提高。



a) YOLOv8 b) YOLOv8+Deformable Attention

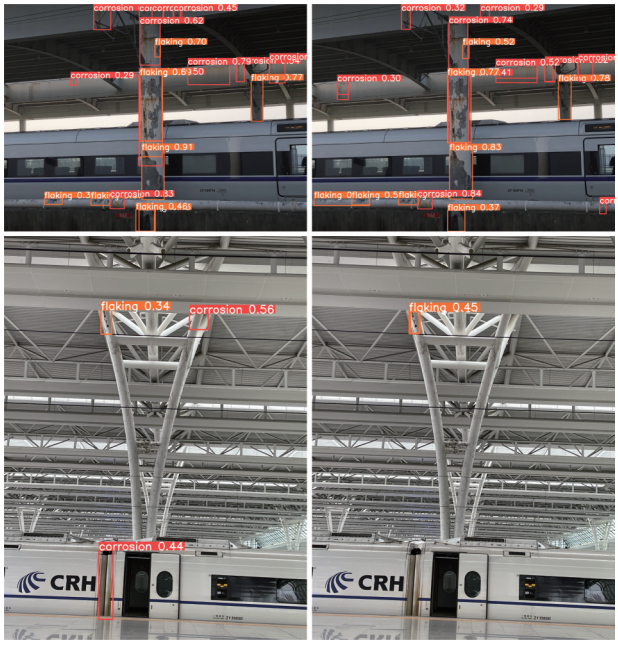
图2 可变形注意力模块消融实验结果

Fig. 2 Ablation experiment results of deformable attention

增加动态检测头模块能够提高模型的召回率，召回率从70.37%提升至71.43%，提升了1.06%。图3展示了几个结果实例。动态检测头模块通过动态调整特征提取和融合的方式，能够更好地处理不同尺度的目标，因此大中小目标的检测精度都会提高，尤其是中小尺度目标的召回率显著提升。同时，由于动态检测头模块提高了模型的空间感知能力，检测结果中的目标位置更为精确。即使在拥挤或复杂背景中，模型也能更准确地定位和分割目标，减少误检或漏检。如图3第一组结果，在高铁站台这样的复杂场景中，大量缺陷同时出现在远、近场景中，增加动态检测头模块之后多个未检测到的中小尺度目标被召回。

由于添加了动态检测头模块，模型对各种场景输入的适应性更强，检测结果更为一致和稳定。在复杂场景中，动态检测头模块能更好地聚焦于关键特征，同时忽略不相关的背景信息，减少由于噪声导致的误检。如图3第二组结果，原始YOLOv8出现了误检的情况，在增加动态检测头模块之后，误检不存在了。

在YOLOv8中将分类损失函数替换为Slide Loss，能

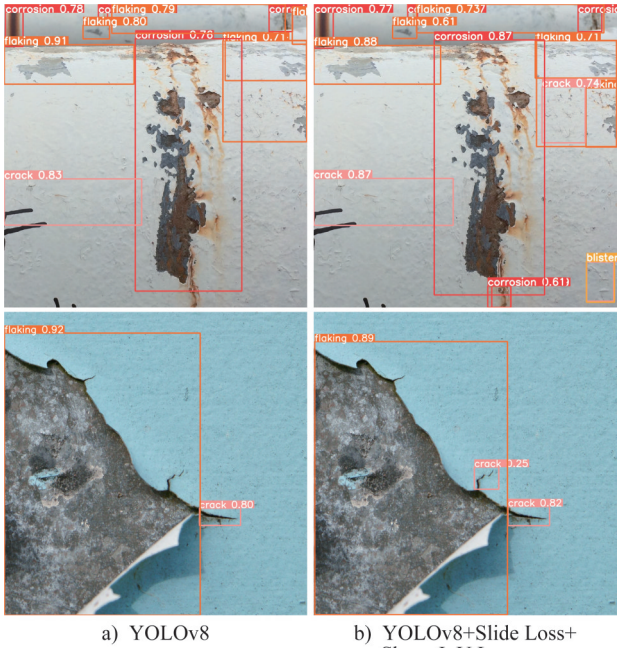


a) YOLOv8　　b) YOLOv8+Dynamic Head

图3　动态检测头模块消融实验结果

Fig. 3　Ablation experiment results of dynamic head

够避免模型识别偏向于常见类别，而忽视低频类别，增加检测结果中对低频类别的关注，提升识别率。边界框回归损失替换为Shape IoU Loss可以提高多尺度目标检测的效果。在Slide Loss和Shape IoU Loss的共同作用下，检测结果的改善作用更加明显。如图4所示，裂缝等小目标、剥落与空鼓等特征不明显的难分辨缺陷召回率被提升，受背景与噪声干扰导致的误检情况减少。



a) YOLOv8　　b) YOLOv8+Slide Loss+ Shape IoU Loss

图4　损失函数模块消融实验结果

Fig. 4　Ablation experiment results of loss functions

3　结论

本文提出了基于计算机视觉的钢结构防腐涂层缺陷检测方案，推动了钢结构防腐涂层缺陷自动化检测落地，改善了人工检测存在的主观干扰、环境限制等问题。主要结论如下：

（1）针对剥落、裂纹、空鼓和腐蚀四类在钢结构使用寿命期间可能出现且对防腐蚀性能有直接影响的缺陷，构建了一个相应的数据集。由于目前尚无成熟的开源涂层缺陷检测数据集，本文通过拍摄获取了真实环境中钢结构涂层表面高分辨率缺陷图像，并以此为基础建立了钢结构防腐涂层缺陷检测数据集。每个样本包含多种缺陷，并通过边界框进行缺陷定位。

（2）本文提出了一种基于多种注意力机制的改进的Attention_YOLOv8检测算法。通过引入可变形注意力模块改进骨干网络、融合自注意力机制的动态检测头，优化了分类损失函数和边界框回归损失函数，解决了传统机器学习图像分类模型无法进行多缺陷检测且泛化性能不足的问题。与YOLOv8基线算法相比，改进后的算法在精确率上提高了5.82%，召回率提高了5.01%，m_{AP}指标提高了4.78%。

但本文数据集仍存在一定的局限性。由于数据主要来源于实际工程项目的现场拍摄，受限于现场条件和收集方式，未能详细记录每张图像对应的环境条件（如气候区域、使用年限）以及涂层材料和厚度等信息。这在一定程度上可能影响模型的代表性和泛化能力。未来的研究将考虑扩充数据集，以涵盖更多不同环境条件和材料类型的钢结构样本，从而进一步提高模型的适应性和鲁棒性。

参考文献：

[1] SOUFEIANI L, FOLIENTE G, NGUYEN K T Q, et al. Corrosion protection of steel elements in façade systems-A review [J]. Journal of Building Engineering, 2020, 32: 101759. DOI:10.1016/j.jobe.2020.101759.

[2] 吕大财，陈向东，朱飞，等. 钢结构涂装红外热成像技术检测的影响因素分析[J]. 无损探伤，2020，44(5)：16-20. DOI:10.13689/j.cnki.cn21-1230/th.2020.05.004.
LYU Dacai, CHEN Xiangdong, ZHU Fei, et al. Analysis of infrared thermal imaging technology on steel structure coating detection [J]. Nondestructive Testing Technology, 2020, 44(5): 16-20. DOI:10.13689/j.cnki.cn21-1230/th.2020.05.004.(in Chinese)

[3] 罗东亮，蔡雨萱，杨子豪，等. 工业缺陷检测深度学习方法综述[J]. 中国科学：信息科学，2022，52(6)：1002-1039.
LUO Dongliang, CAI Yuxuan, YANG Zihao, et al. Survey on industrial defect detection with deep learning[J]. Scientia Sinica (Informationis), 2022, 52(6): 1002-1039. (in Chinese)

[4] XU J B, ZHANG J D, ZHANG K P, et al. An APF-ACO algorithm for automatic defect detection on vehicle paint[J]. Multimedia Tools and Applications, 2020, 79(35): 25315-25333. DOI:10.1007/s11042-020-09245-2.

[5] 吴旭东，闫志鸿，胡飞涛，等. 基于光度学特征与SVM的热轧铝板表面缺陷检测[J]. 热加工工艺，2023，52(6)：157-162. DOI:10.14158/j.cnki.1001-3814.20212961.
WU Xudong, YAN Zhihong, HU Feitao, et al. Surface defect detection of hot-rolled aluminum sheet based on photometric features and SVM [J]. Hot Working Technology, 2023, 52(6): 157-162. DOI:10.14158/j.cnki.1001-3814.20212961.(in Chinese)

[6] 王艳，罗健，陶金，等. 基于YOLOv8-PCB的印制电路板缺陷检测[J]. 激光与光电子学进展，2025，62(4)：0412002.
WANG Yan, LUO Jian, TAO Jin, et al. Defect detection of printed circuit boards based on YOLOv8-PCB[J]. Laser & Optoelectronics Progress, 2025, 62(4): 0412002. (in Chinese)

[7] 周永钦，王勇，王瑛. 基于多尺度特征及注意力机制的轻量化PCB缺陷检测方法[J]. 计算机与现代化，2024(2)：88-92，107.
ZHOU Yongqin, WANG Yong, WANG Ying. Lightweight PCB defect detection method based on multi-scale features and attention mechanism[J]. Computer and Modernization, 2024(2): 88-92, 107. (in Chinese)

[8] JIANG X H, YAN F, LU Y, et al. Joint attention-guided feature fusion network for saliency detection of surface defects [J]. IEEE Transactions on Instrumentation and Measurement, 2022, 71: 2520912. DOI: 10.1109/TIM.2022.3218547.

[9] GAO Y P, GAO L, LI X Y, et al. A semi-supervised convolutional neural network-based method for steel surface defect recognition [J]. Robotics and Computer-Integrated Manufacturing, 2020, 61: 101825. DOI: 10.1016/j.rcim.2019.101825.

[10] LI S Y, ZHAO X F, ZHOU G Y. Automatic pixel-level multiple damage detection of concrete structure using fully convolutional network [J]. Computer-Aided Civil and Infrastructure Engineering, 2019, 34(7): 616-634. DOI: 10.1111/mice.12433.

[11] RAHMAN A, WU Z Y, KALFARISI R. Semantic deep learning integrated with RGB feature-based rule optimization for facility surface corrosion detection and evaluation [J]. Journal of Computing in Civil Engineering, 2021, 35(6): 04021018. DOI:10.1061/(ASCE)CP.1943-5487.0000982.

[12] 杨绵芳，陈伯建，李扬笛，等. 基于改进YOLOv9算法的配电网绝缘子缺陷检测[J]. 上海电力大学学报，2024，40(6)：571-578.
YANG Mianfang, CHEN Bojian, LI Yangdi, et al. Insulator defect detection in distribution network based on improved YOLOv9 method [J]. Journal of Shanghai University of Electric Power, 2024, 40(6): 571-578. (in Chinese)

[13] 李耀华，刘传召. 融合注意力机制的轻量化航空发动机表面缺陷检测算法[J]. 航空动力学报，2025，40(6)：20240096. DOI:10.13224/j.cnki.jasp.20240096.
LI Yaohua, LIU Chuanzhao. Lightweight aero-engine surface defect detection algorithm fused with attention mechanism[J]. Journal of Aerospace Power, 2025, 40(6): 20240096. DOI:10.13224/j.cnki.jasp.20240096.(in Chinese)

[14] 祁雷，李宁，梁伟，等. 基于注意力机制的U-Net叶片缺陷图像分割[J]. 中国安全科学学报，2024，34(5)：139-146. DOI:10.16265/j.cnki.issn1003-3033.2024.05.0767.
QI Lei, LI Ning, LIANG Wei, et al. Blade defect detection of U-Net network based on attention mechanism[J]. China Safety Science Journal, 2024, 34(5): 139-146. DOI: 10.16265/j.cnki.issn1003-3033.2024.05.0767.(in Chinese)

[15] ZHANG H, LI F, LIU S L, et al. DINO: DETR with improved denoising anchor boxes for end-to-end object detection [EB/OL]. (2022-03-07) [2025-11-18]. https://arxiv.org/abs/2203. 03605.

[16] ZHAO Y A, LV W Y, XU S L, et al. DETRs beat YOLOs on real-time object detection [C]//2024 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Seattle: IEEE, 2024.

[17] VIEIRA E SILVA A L, SIMÕES F, KOWERKO D, et al. Attention modules improve image-level anomaly detection for industrial inspection: A DifferNet case study[C]//2024 IEEE/CVF Winter Conference on Applications of Computer Vision (WACV). Waikoloa: IEEE, 2024.

[18] XIA Z F, PAN X R, SONG S J, et al. Vision transformer with deformable attention[C]//2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). New Orleans: IEEE, 2022.

[19] DAI X Y, CHEN Y P, XIAO B, et al. Dynamic head: Unifying object detection heads with attentions[C]//2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Nashville: IEEE, 2021.

[20] YU Z P, HUANG H B, CHEN W J, et al. YOLO-FaceV2: A scale and occlusion aware face detector [J]. Pattern Recognition, 2024, 115 (11): 110714. DOI: 10.1016/j.patcog.2024.110714.

[21] ZHANG H, ZHANG S J. Shape-IoU: More accurate metric considering bounding box shape and scale[EB/OL]. (2023-12-29)[2025-11-18]. https://arxiv.org/abs/2312.17663.