

doi: 10.7690/bgzdh.2024.12.008

基于改进 YOLOv5 的钢材表面缺陷检测

刘祉燊¹, 张晓玲¹, 刘珂宇², 刘晓军¹, 刘晓静³

(1. 山东石油化工学院智能制造与控制工程学院, 山东 东营 257061; 2. 中国石化集团共享服务公司东营分公司, 山东 东营 257029; 3. 中石化经纬有限公司胜利测井公司, 山东 东营 257096)

摘要: 针对钢材表面缺陷检测中小目标缺陷检测效果不理想、特征提取不充分的问题, 以 YOLOv5 算法为基础, 提出一种 YOLOv5s-ADW 算法。将自注意力与卷积混合模块 (a mixed model of self-attention and convolution, ACmix) 融入主干网络层, 增强模型的特征敏感度; 在特征融合层中加入可变形大内核注意力机制 (deformable large kernel attention, D-LKA), 增强模型对图像中不规则缺陷的捕捉能力; 将原损失函数替换为 Wise-IoU 损失函数, 降低数据集中低质量示例对模型检测效果的影响并提升小目标缺陷检测能力, 在 NEU-DET 上进行实验验证。实验验证结果表明: YOLOv5s-ADW 算法的平均精度均值 (mean average precision, mAP) 达到 88.3%, 相较原始模型提升了 14.4%; 小目标缺陷和漏检率高的缺陷平均精度 (average precision, AP) 也有较大提升, 相比其他主流算法, 能够更好解决上述问题。

关键词: YOLOv5; 钢材表面缺陷检测; 小目标缺陷检测; 特征提取
中图分类号: TP391 **文献标志码:** A

Steel Surface Defect Detection Based on Improved YOLOv5

Liu Zhishen¹, Zhang Xiaoling¹, Liu Keyu², Liu Xiaojun¹, Liu Xiaojing³

(1. School of Intelligent Manufacturing and Control Engineering, Shandong Institute of Petroleum and Chemical Technology, Dongying 257061, China; 2. Sinopec Group Shared Services Co., Ltd. Dongying Branch, Dongying 257029, China; 3. Sinopec Jingwei Co., Ltd. Shengli Logging Company, Dongying 257096, China)

Abstract: Aiming at the problems of unsatisfactory defect detection effect and insufficient feature extraction of small and medium targets in steel surface defect detection, a YOLOv5s-ADW algorithm based on YOLOv5 algorithm was proposed. a mixed model of self-attention and convolution (ACmix) module is integrated into the backbone network layer to enhance the feature sensitivity of the model. The deformable large kernel attention (D-LKA) mechanism is added to the feature fusion layer to enhance the ability of the model to capture irregular defects in images. The original loss function was replaced by Wise-IoU loss function to reduce the influence of low-quality examples on the model detection effect and improve the detection ability of small target defects, and the experimental verification was carried out on NEU-DET. Experimental results show that the mean average precision (mAP) of the YOLOv5s-ADW algorithm reaches 88.3%, which is 14.4% higher than that of the original model. The average precision (AP) of small target defects and defects with high missed detection rate is also greatly improved, which can better solve the above problems compared with other mainstream algorithms.

Keywords: YOLOv5; Steel surface defect detection; Small target defect detection; feature extraction

0 引言

钢材表面缺陷检测一直是工程技术领域中的重要研究课题^[1]。随着计算机设备算力的提升, 深度学习算法被应用于钢材表面缺陷检测中, 其中, YOLOv5 检测算法由于检测速度快和模型体积小等特点, 在钢材表面缺陷检测任务中表现较为优秀。

对于 YOLOv5 检测算法精度低以及钢材表面缺陷检测目标小的特点, 谷长江等^[2]融入 ULSAM 注意力机制和 DO-Conv 模块以提升对小目标的特征

提取能力和检测精度, mAP 为 76.6%; 李鑫等^[3]采用 Ghost Bottleneck 结构替换 C3 和 CBS 模块, 引入了 SE 注意力机制和深度卷积, 并增加了小目标检测层, 使得该方案的 mAP 达到了 76.8%; 吴敌等^[4]加入小目标检测层和 CBAM 注意力机制, 并将部分 CSP 层替换为 transformer encoder block, mAP 达到 74.8%。根据以上改进结果可以看出: 钢材表面缺陷检测中存在的小目标缺陷检测效果不理想、特征提取不充分的问题没有得到良好的解决。另外, 在制作钢材表面缺陷数据集时, 存在因缺陷锚框不

收稿日期: 2024-06-23; 修回日期: 2024-07-20

基金项目: 山东省高等学校科技计划项目 (J18KA359)

第一作者: 刘祉燊 (2001—), 男, 山西人。

通信作者: 张晓玲 (1982—), 女, 山东人, 硕士。

准确导致数据集出现低质量示例的情况，会造成模型检测性能降低，影响缺陷检测效果。

笔者基于 YOLOv5s 提出一种 YOLOv5s-ADW 算法，降低数据集中低质量示例对模型检测效果的影响并提升对小目标缺陷的检测能力。在数据集 NEU-DET 上的实验结果表明，该算法能够实现提升钢材表面缺陷检测效果。

1 YOLOv5s 模型介绍

笔者将 YOLOv5s 作为原始模型。其数据输入端采用 Mosaic 数据增强、自适应锚框等技术对原始数据进行图像预处理，使其转化为 640×640 像素的图像。主干网络采用 CSPNet 网络结构，包括 Conv、C3、SPPF 等模块，对检测目标进行特征提取。特征融合层将主干网络提取的特征进行融合，采用特征金字塔网络 (FPN) 和金字塔注意力网络 (PAN) 相

结合的结构加强网络特征融合能力。预测层主要采用 CIoU 作为边界框的损失函数^[5]，通过预测类别、坐标、置信度等信息对目标进行检测。

2 YOLOv5s 算法改进

对 YOLOv5s 有以下改进：1) 将主干网络层 SPPF 模块前的一个 C3 模块替换为自注意力与卷积混合 (ACmix) 模块，增强模型对目标的特征敏感度，以便更好地进行特征提取；2) 用可变形大核注意力 (D-LKA) 模块替换特征融合层中的 C3 模块，使模型增加对复杂图像的适应性，降低缺陷漏检率；3) 将原始模型中的 CIoU 损失函数替换为 Wise-IoU 损失函数，降低数据集中低质量示例对钢材缺陷检测的影响，并提升对小目标缺陷的检测效果。将改进后的 YOLOv5s 算法称为 YOLOv5s-ADW 算法，其算法结构如图 1 所示。

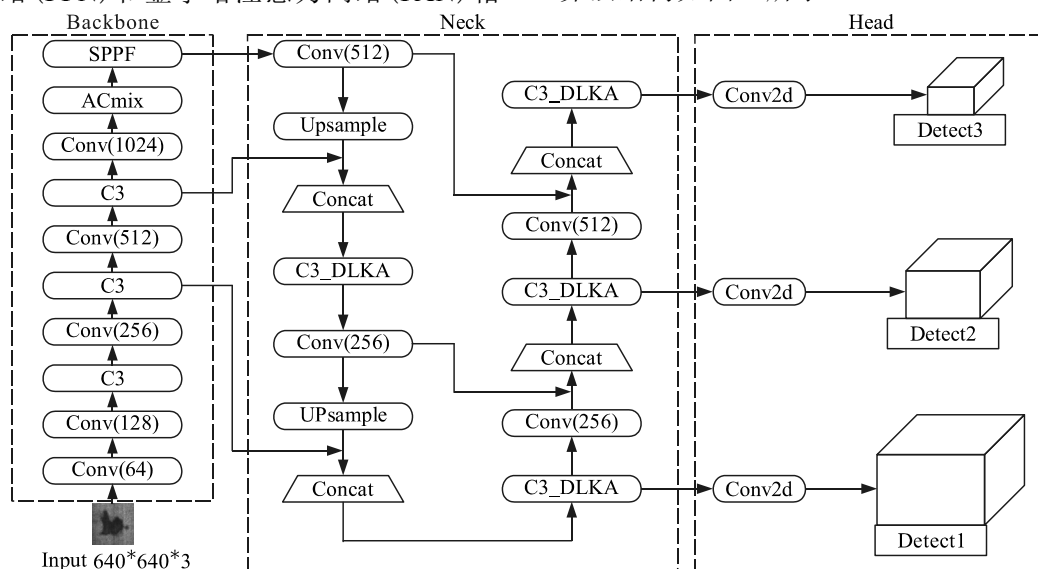


图 1 改进后 YOLOv5s 网络结构

2.1 ACmix 模块

为了提高网络对小目标缺陷的特征敏感度，笔者在主干网络层加入自注意力与卷积混合模块 (ACmix)。该模块将自注意力机制和卷积进行了混合^[6]，增加了网络的特征敏感度，可以使网络通过特征图获得更加全面的特征信息，并且由于其可以从局部感受野上收集信息，使得其能够更准确的区分钢材表面的小目标缺陷，提升检测精度和查全率。由于 ACmix 模块的参数量、计算复杂度与特征通道数有关，故为了避免出现计算量大幅增加而检测精度提升不明显的资源浪费情况，笔者只将主干网络层 SPPF 模块前的 C3 模块替换为 ACmix 模块。ACmix 模块的结构如图 2 所示。

ACmix 模块主要包含 2 部分：1) 将输入特征经由 3 个 1×1 的卷积操作被重塑为 N 块，获得 3 个 $3 \times N$ 的特征图，作为第 2 部分的输入特征集。2) 按照自注意力机制和卷积 2 种路径分别进行信息收集。卷积路径首先通过通道层的全连接对通道进行扩张，生成 N^2 个 1×1 的特征图，然后使用类似于卷积的可学习的分组卷积结构来实现空间偏移，使其能够从局部感受野上收集信息，最后经过聚合生成新的特征 F_{conv} ；自注意力机制路径则通过 1×1 的卷积变换计算出 query、key、value 后，对其进行注意力权重的计算和拼接不同头的操作，得到特征 F_{att} 。收集到的信息附加可学习的标量 α 和 β ，相加得到最终输出结果 F_{out} ，公式为：

$F_{out} = \alpha F_{att} + \beta F_{conv}$ (1)
最终得到的 F_{out} 使得大目标缺陷和小目标缺陷

的特征较平衡进入特征融合网络，以此增加小目标的检测精度。

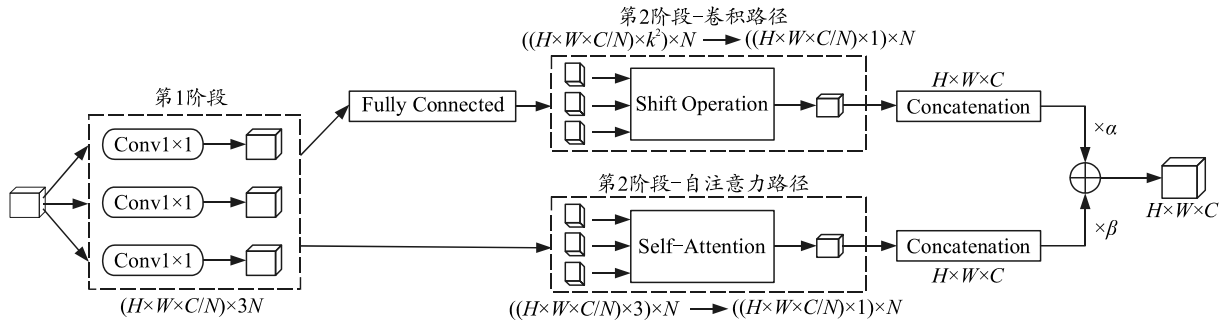


图 2 ACmix 模块结构

2.2 D-LKA 模块

为了增强对缺陷的捕捉能力，笔者在特征融合层中融入可变形大内核注意力模块^[7]。该模块通过大卷积核模拟类似自注意力机制的感受野，并引入了一种基于数据感受野的动态适应，通过可变形卷积来灵活调整采样网格，增强模型对图像中大小随机且形状不规则缺陷的捕捉能力，从而提升检测精度。2D 版本结构如图 3 所示。

输入图像经由卷积初始层后，维度减小为 $H/4 \times W/4 \times C$ ；随后，由 4 个 MaxViT 块分 4 个阶段对图像进行特征提取，每个阶段结束后，图片进行

降采样 1/2；图片进入解码器后，由 4 个 D-LKA 层进行语义分割，每层包含 2 个 D-LKA 块，D-LKA 块由 LayerNorm、可变形 LKA 和 MLP 组成。通过积分残差连接确保了有效的特征传播，并且可以推广到更深层，其公式为：

$$x_1 = D-LKA-Attn(LN(x_{in})) + x_{in}; \quad (2)$$

$$x_{out} = MLP(LN(x_1)) + x_1; \quad (3)$$

$$MLP = Conv_1(GeLU(Conv_d(Conv_1(x)))) \quad (4)$$

式中： x_{in} 为输入特征；LN 为层归一化；D-LKA-Attn 为可变形 LKA 注意力； $Conv_d$ 为深度卷积； $Conv_1$ 为线性层；GeLU 为激活函数。

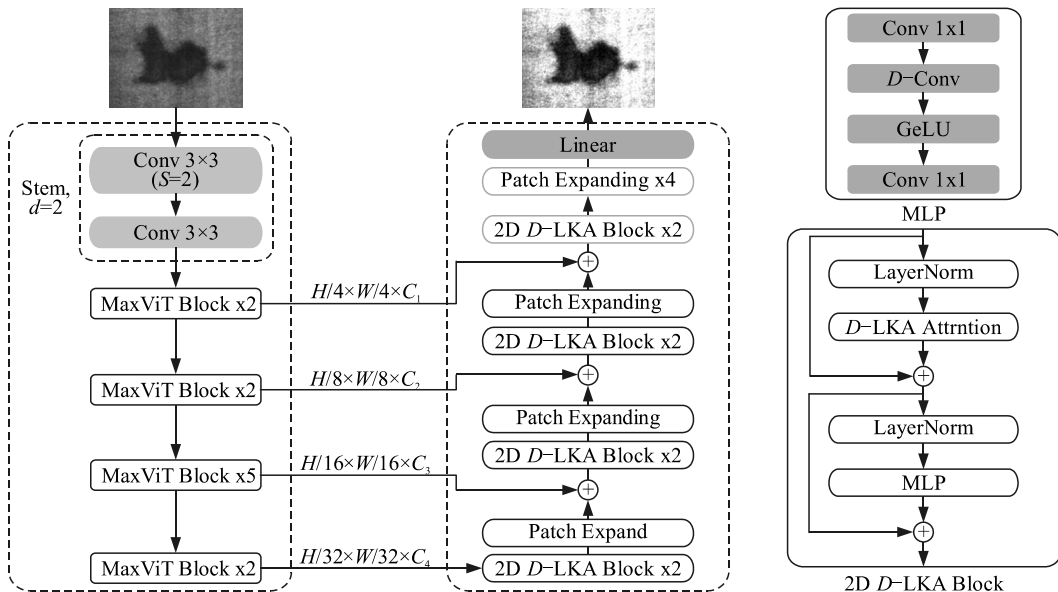


图 3 D-LKA_2D 版本结构

经过 DLKA 层后，由 Patch 扩展层进行分辨率上采样，把图像维度进行还原，同时减小通道维度；线性层负责生成最终的输出。

2.3 Wise-IoU 损失函数

YOLOv5s 网络模型的损失函数由分类损失、置信度损失和边界框损失组成，目标检测任务中的边

界框表征目标物大小和位置，所以边界框损失计算对检测效果影响较大^[8]。Zhang 等^[9]的研究表明 CIoU 在长宽比的引入上存在问题，并指出回归样本的回归质量过高或过低会对损失函数造成不良影响。Tong 等^[10]提出的 Wise-IoU 解决了上述问题，并针对非单调静态聚焦机制的潜力没有被充分利用

的问题做了改进。

Wise-IoU 损失函数能够降低数据集中低质量示例对钢材缺陷检测的影响，而其计算过程中使用的相对面积加权也对小目标缺陷检测有一定帮助，故本文中使用 Wise-IoU 作为边界框损失函数。Wise-IoU 参数如图 4 所示。

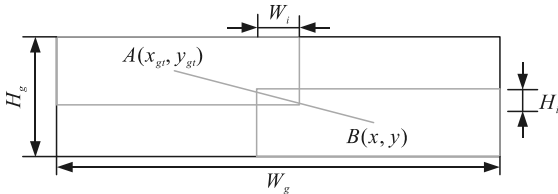


图 4 Wise-IoU 参数

上图中： A 为目标框，宽高分别为 w_{gt} 和 h_{gt} ； B 为锚框，宽高分别为 w 和 h ； W_i ， H_i 为目标框和锚框并集的宽高。则 Wise-IoU 损失函数可表示为：

$$L_{IoU} = 1 - IoU; \quad (5)$$

$$R_{WIoU} = \exp(((x - x_{gt})^2 + (y - y_{gt})^2) / (W_g^2 - H_g^2)^*); \quad (6)$$

$$L_{WIoU} = r R_{WIoU} L_{IoU}. \quad (7)$$

式中： IoU 为锚框和目标框交并比； R_{WIoU} 为距离损失；上标*为将 W_g 和 H_g 从计算图中分离，防止 R_{WIoU} 产生阻碍收敛的梯度，当 $R_{WIoU} \in [1, e)$ 时，一般质量锚框的 L_{IoU} 将会明显增大，当 $L_{IoU} \in [0, 1]$ 时，高质量锚框的 R_{WIoU} 会明显降低，且在锚框与目标框能够很好的重合时，降低对中心点距离的关注度； r 为梯度增益，锚框质量过低过高，都会使得动态 FM 将小的梯度增益 r 分配给锚框，使得边界框回归集中到一般质量的锚框，而锚框的质量评价标准是不断更新的，故 Wise-IoU 可以在任何时刻做出满足当时状况的梯度增益分配方案。

3 实验设置及结果分析

3.1 数据集介绍

笔者采用东北大学的公开数据集 NEU-DET 进行钢材缺陷检测实验。该数据集包含裂纹 (Cracking)、

夹杂 (Inclusion)、斑块 (Patches)、点蚀面 (Pitted Surface)、氧化轧皮 (Rolled-in Scale) 和划痕 (Scratches) 共 6 种缺陷，每种缺陷有 300 张大小为 200×200 像素图片，将数据集按照 8:2 的比例划分为训练集与验证集。目标区域尺寸占比小于原图的 10% 的样本为小目标^[11]，因此，钢材表面的每种缺陷都包含小目标，夹杂和氧化轧皮中小目标比例较高。

3.2 实验环境配置

本实验的深度学习框架为 PyTorch 2.2.2，编译环境为 Python3.8，操作系统为 Windows10，CPU 为 Intel Xeon Processor E5-2680，GPU 为 NVIDIA GeForce RTX 3060(12 G)。实验训练策略：批处理大小为 12，迭代次数为 300，初始学习率为 0.01，学习率动量为 0.937，衰减系数为 0.000 5。

3.3 评价指标

采用准确率 P (precision)、召回率 R (Recall)、平均精度 (AP) 和平均精度均值 (mAP) 作为模型的性能评价指标。P、R 和 AP 的计算公式如下：

$$P = TP / (TP + FP) 100\% \quad (8)$$

$$R = TP / (TP + FN) 100\%; \quad (9)$$

$$AP = \int_0^1 P(R) dR; \quad (10)$$

$$mAP@0.5 = \frac{1}{6} \sum_{i=0}^6 AP(i). \quad (11)$$

另外，评价指标还包括模型体积 (weight)，其表示训练完成后生成的模型权重文件的大小，模型体积过大会导致存储、传输和推理速度等方面的瓶颈，限制其在移动端设备上的部署和应用。

3.4 实验结果分析

3.4.1 消融实验

为了对比不同方案模型的性能，笔者进行了消融实验。实验结果如表 1 所示。

表 1 消融实验结果对比

实验方案	ACmix	D-LKA	Wise-IoU	AP						mAP
				Cracking	Inclusion	Patches	Pitted Surface	Rolled-in Scale	Scratches	
YOLOv5s	-	-	-	0.279	0.829	0.970	0.712	0.726	0.914	0.739
改进 1	√	-	-	0.255	0.850	0.958	0.958	0.765	0.929	0.786
改进 2	-	√	-	0.342	0.827	0.969	0.872	0.801	0.903	0.786
改进 3	-	-	√	0.306	0.796	0.964	0.848	0.702	0.904	0.753
改进 4	√	√	-	0.385	0.897	0.980	0.967	0.829	0.977	0.839
改进 5	√	-	√	0.384	0.912	0.975	0.959	0.854	0.961	0.841
改进 6	-	√	√	0.478	0.924	0.984	0.958	0.847	0.976	0.866
YOLOv5s-ADW	√	√	√	0.515	0.958	0.987	0.958	0.912	0.967	0.883

实验 YOLOv5s 为本文中改进前的原始模型，

改进 1 和 2 分别为在原始模型的基础上加入了

ACmix 模块、D-LKA 模块后的训练结果，可以看出加入 ACmix 模块与 D-LKA 模块后 mAP 值有较明显的提升，改进 4 为在原始模型的基础上加入 ACmix 模块和 D-LKA 模块，mAP 值提升了 10.0%，其中，裂纹的 AP 值提升了 10.6%，点蚀面的 AP 值提升了 25.5%，其他缺陷的 AP 值相较原始模型也均有较大提升，表明 ACmix 模块与 D-LKA 模块配合能够使得模型特征提取能力更强。

改进 3 为原始模型使用 Wise-IoU 损失函数后的数据，相较原始模型 mAP 值提升效果不明显，故笔者在替换 Wise-IoU 损失函数后，分别加入 ACmix 模块和 D-LKA 模块，即为改进 5 和 6。从结果来看每个种类的 AP 值相较于改进 1 和 2 都有提升，能够证明 Wise-IoU 损失函数在降低数据集中低质量示例的消极影响方面有一定作用。

实验 YOLOv5s-ADW 将这 3 个模块同时融入原始模型中，对比改进 4，除点蚀面和划痕的 AP 值略有下降外，其他缺陷 AP 值均有提升；对比原始模型，所有缺陷的 AP 值有较大提升，其中漏检率较高的裂纹和点蚀面的 AP 值分别提升了 23.6%和 24.6%，夹杂和氧化轧皮 2 种小目标比例较高的缺陷，分别提升了 12.9%和 18.6%，另外，mAP 值提升了 14.4%，达到了 88.3%。数据表明，本文中的改进算法实现了对 YOLOv5s 模型性能的提升，更好地解决了小目标缺陷检测效果不理想、特征提取不充分的问题。

3.4.2 对比实验

为了进一步评估本文中模型的性能，在相同数据集的前提下，将本文中模型与目前主流模型及其改进版本进行对比，对比结果如表 2 所示。

表 2 模型性能对比

组别	评价指标			
	P	R	Weight/MB	mAP@0.5
Faster R-CNN	0.621	0.723	108.0	0.657
YOLOv3	0.566	0.624	117.0	0.599
YOLOv5s	0.782	0.692	14.5	0.739
文献[2]	0.762	0.703	-	0.766
文献[3]	0.689	0.756	12.2	0.768
文献[4]	0.754	0.675	-	0.748
YOLOv5s-ADW	0.897	0.802	16.3	0.883

由上表可知，YOLOv5s-ADW 算法的精度值和召回率分别达到了 89.7%和 80.2%，较原始模型分别提升了 11.5%和 11.0%，并且超过了所有对比实验；在模型体积上，改进 YOLOv5s 算法的大小为 16.3 MB，远低于 Faster R-CNN、YOLOv3，只比原始模型大 1.8 MB，由于文献[3]使用 Ghost Bottleneck

结构减少了参数量，使得模型参数量降低了 2.3 MB。总的来说，本文中算法以增大模型体积 1.8 MB 的代价，使得 mAP 值提升了 14.4%，以及精度值和召回率有了显著提高。

4 结束语

该算法将 ACmix 模块和 D-LKA 模块融入 YOLOv5s 模型，将原损失函数替换为 Wise-IoU 损失函数。实验结果表明：本文中算法相较于 YOLOv5s 原始模型 mAP 值提升了 14.4%，达到了 88.3%，其中小目标缺陷和漏检率较高缺陷的 AP 值有较大提升，实现了对 YOLOv5s 模型性能的提升。

参考文献：

- [1] 任中伟, 方维维, 许文元, 等. 面向轨道缺陷检测的联邦学习轻量化模型训练技术研究[J]. 铁道学报, 2023, 45(4): 77-83.
- [2] 谷长江, 高法钦. 改进 YOLOv5s 的钢材表面缺陷检测[J]. 软件工程, 2023, 26(8): 31-34.
- [3] 李鑫, 汪诚, 李彬, 等. 改进 YOLOv5 的钢材表面缺陷检测算法[J]. 空军工程大学学报(自然科学版), 2022, 23(2): 26-33.
- [4] 吴敌, 李明辉, 马文凯, 等. 基于改进 YOLOv5 的钢材表面缺陷检测[J]. 陕西科技大学学报, 2023, 41(2): 162-169.
- [5] ZHENG Z, WANG P, LIU W, et al. Distance-IoU loss: Faster and better learning for bounding box regression[J]. New York Proceedings of the AAAI conference on artificial intelligence, 2020, 34(7): 12993-13000.
- [6] PAN X, GE C, LU R, et al. On the integration of self-attention and convolution[C]. Conference on Computer Vision and Pattern Recognition, New Orleans, LA, USA, 2022: 805-815.
- [7] AZAD R, NIGGEMEIER L, H Ü TTEMANN M, et al. Beyond self-attention: Deformable large kernel attention for medical image segmentation[C]. 2024 IEEE/CVF Winter Conference on Applications of Computer Vision (WACV), Waikoloa, HI, USA, 2024: 1287-1297.
- [8] 崔晶楠, 黄春艳, 李艳玲. 改进 YOLOv5s 算法的钢材表面缺陷检测[J/OL]. 电子科技. <https://doi.org/10.16180/j.cnki.issn1007-7820.2024.12.008>.
- [9] ZHANG Y F, REN W, ZHANG Z, et al. Focal and efficient IOU loss for accurate bounding box regression[J]. Neurocomputing, 2022, 506: 146-157.
- [10] TONG Z, CHEN Y, XU Z, et al. Wise-IoU: bounding box regression loss with dynamic focusing mechanism[EB/OL]. [2023-1-24]. <https://arxiv.org/pdf/2301.10051>.
- [11] 梁鸿, 王庆玮, 张千, 等. 小目标检测技术研究综述[J]. 计算机工程与应用, 2021, 57(1): 17-28.