

# 基于增强可变形卷积的带钢表面缺陷检测网络

孙艳歌\*, 徐成龙, 郭华平, 张 莉

(信阳师范大学 计算机与信息技术学院, 河南 信阳 464000)

**摘要:**提出了一种面向带钢表面缺陷检测的新型网络架构,设计了基于协同注意力的增强可变形卷积模块,并将其作为插件集成至骨干网络;利用协同注意力机制,自适应调整卷积核形状,有效捕捉带钢表面不规则缺陷,显著提升骨干网络的特征提取能力。在NEU-DET数据集上的实验结果表明,该方法在平均精度(mAP)上达到81.6%。

**关键词:** 缺陷检测; 可变形卷积; 协同注意力

中图分类号: TP391.4

文献标志码: A

开放科学(资源服务)标识码(OSID):



## Strip steel surface defect detection network based on enhanced deformable convolution

SUN Yange\*, XU Chenglong, GUO Huaping, ZHANG Li

(College of Computer and Information Technology, Xinyang Normal University, Xinyang 464000, China)

**Abstract:** A novel network architecture for strip steel surface defect detection was proposed, which designed an enhanced deformable convolution module based on co-attention mechanism and integrated it into the backbone network as a plugin. By leveraging the co-attention mechanism, the shape of the convolution kernel was adaptively adjusted, which effectively captured irregular defects on the strip steel surface, and significantly improved the feature extraction capability of the backbone network. Experimental results on the NEU-DET dataset demonstrated that the proposed method achieved an average precision (mAP) of 81.6%.

**Key words:** defect detection; deformable convolution; collaborative attention

## 0 引言

带钢是航空航天、机械制造等领域不可或缺的原材料。生产过程中,因操作失误或工艺限制等因素,带钢表面可能产生裂纹、夹杂等缺陷,进而影响产品使用寿命和安全性<sup>[1]</sup>。因此,构建高效的检测网络对实现带钢表面缺陷精准检测具有重要意义<sup>[2]</sup>。

基于深度学习的方法已成功应用于带钢表面缺陷检测,其中基于卷积神经网络的方法最为常

用。例如, WANG等<sup>[3]</sup>提出了一种增强型Faster-RCNN<sup>[4]</sup>,采用基于可变形旋转网络的ResNet50<sup>[5]</sup>作为特征提取器,增强了骨干网络对带钢表面缺陷特征的提取能力。LIU等<sup>[6-7]</sup>提出了基于改进SSD模型的带钢表面缺陷检测方法,利用ResNet50进行特征提取,提高了检测精度。为应对带钢表面缺陷检测中尺度差异、形态分布不均带来的技术挑战, DU等<sup>[8]</sup>提出了AFF-Net,通过自适应采样与特征对齐机制,有效提升对复杂缺陷的关注能力。

收稿日期: 2024-11-28; 修回日期: 2025-01-07; \*通信联系人, E-mail: yangesun@xynu.edu.cn

基金项目: 国家自然科学基金项目(62062004); 河南省自然科学基金项目(222300420274); 信阳师范学院研究生科研创新基金项目(2021KYJJ56)

作者简介: 孙艳歌(1982—), 女, 河南平顶山人, 副教授, 博士, 主要从事机器学习与数据挖掘研究。

引用格式: 孙艳歌, 徐成龙, 郭华平, 等. 基于增强可变形卷积的带钢表面缺陷检测网络[J]. 信阳师范大学学报(自然科学版), 2026, 39(1): 46-51.

SUN Yange, XU Chenglong, GUO Huaping, et al. Strip steel surface defect detection network based on enhanced deformable convolution[J]. Journal of Xinyang Normal University (Natural Science Edition), 2026, 39(1): 46-51.

基于卷积神经网络的检测器通常先借助主干网络从输入图像中提取有效的粗粒度特征表示,再通过颈部(neck)网络完成多尺度特征融合,为后续缺陷分类与定位任务奠定基础。因此,带钢表面缺陷检测模型的性能优劣,关键取决于主干网络的多尺度特征提取能力与缺陷表征精度。为突破这一核心技术要点,尝试在主干网络中嵌入新型结构模块,以强化其特征提取效能。

基于上述设计思路,提出一种面向带钢表面缺陷检测的新型网络架构,核心创新点如下:基于协同注意力机制设计增强型可变形卷积模块,并将其作为轻量化插件集成至主干网络,通过动态调整卷积核形态与采样策略,精准提取不同形状、

不规则分布的缺陷特征,进一步优化主干网络的特征表征能力。

## 1 基于增强可变形卷积的缺陷检测网络

### 1.1 网络结构

所提出的网络架构由主干网络、颈部网络及预测头三部分构成,如图1所示。借鉴YOLOv8的设计思路,考虑到Darknet53网络兼具强大的特征表征能力与较低的计算复杂度,选择其作为主干网络,负责带钢表面缺陷的粗粒度提取。该主干网络的核心组成单元包括CBS模块、C2f模块及SPPF模块,各模块协调作用完成初始特征的高效提取。

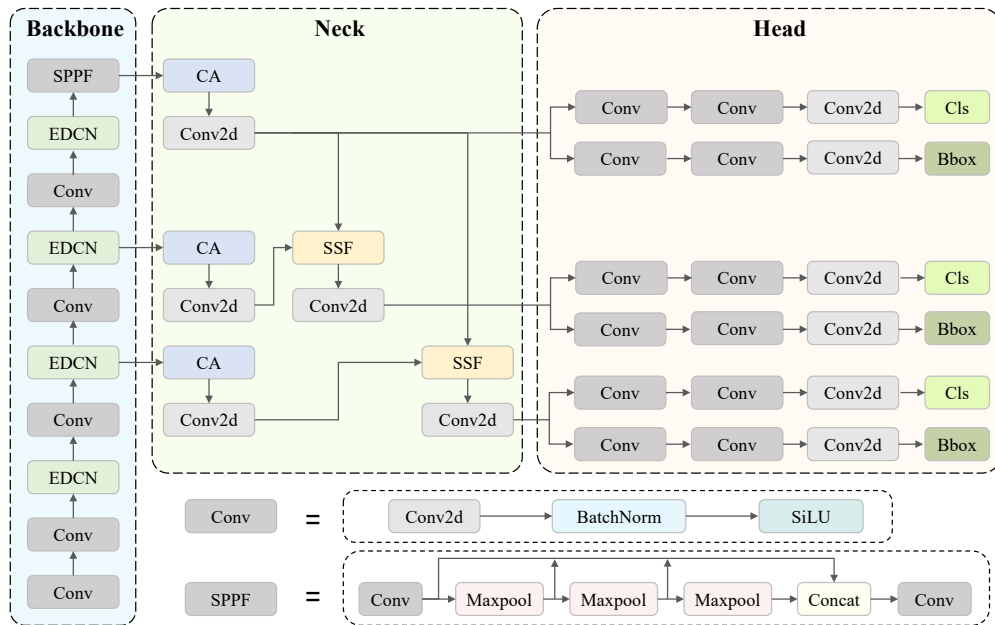


图1 网络结构

Fig. 1 Network structure

为进一步强化骨干网络(Backbone)的特征提取效能,在Darknet53中嵌入一种基于协同注意力的增强可变形卷积模块,如图1(左)所示。颈部(Neck)采用金字塔融合结构,对骨干网络提取的粗粒度特征进行多轮细化与跨尺度融合,以提升特征表征的丰富度与精准度,如图1(中)所示。预测头(Head)由CBS模块与Conv2d模块串联构成,通过层级化特征映射实现对带钢表面缺陷的准确定位与类别识别。

### 1.2 增强可变形卷积

标准卷积存在一个显著的限制性:其仅在固

定的位置对输入特征图进行采样,限制了卷积神经网络对全局信息的感知与捕捉能力。可变形卷积(Deformable Convolution, DCN)<sup>[9]</sup>作为增强卷积神经网络变换建模能力的有效方法,通过引入可学习的偏移量动态扩展空间采样位置。DCN的数学表达式为:

$$y(p_0) = \sum_{p_n \in R} w \cdot x(p_0 + p_n + \Delta p_n), \quad (1)$$

式中: $p_0$ 为感受野 $R$ 的中心点, $p_n$ 表示相对于中心点的原始偏移量, $\Delta p_n$ 表示输入特征映射经过卷积产生的偏移, $y$ 为输出, $w$ 为权重系数。

与式(1)中通过卷积直接计算偏移量的方式不同,引入可变形卷积在增强骨干网络特征表达能力的基础上,进一步提出一种协同注意力(Co-Attention, CoA)模块。该模块通过联合考量特征图的空间维度信息与通道维度信息,动态优化并调整可变形卷积的采样偏移量,使卷积核能够更精准地适配带钢表面不规则缺陷的形态与分布。CoA 模块具体包含空间注意力分支(Spatial Attention Branch, SAB)与通道注意力分支(Channel Attention Branch, CAB),其结构与增强可变形卷积(Enhanced Deformable Convolution, EDCN)的整体架构如图 2 所示。

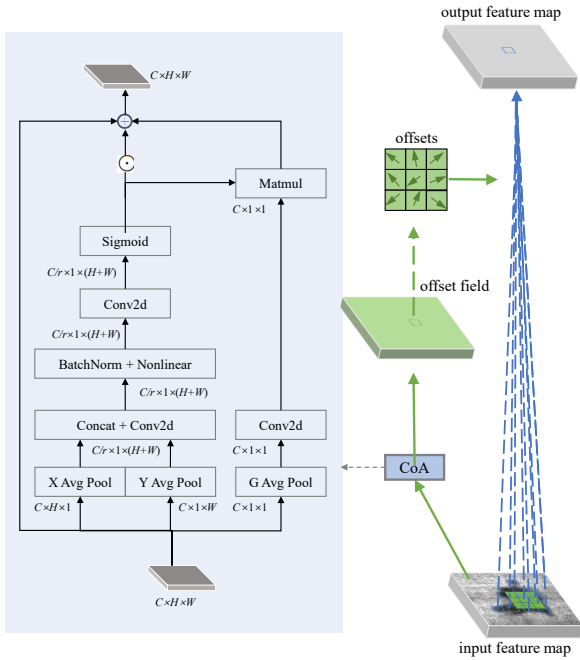


图 2 增强可变形卷积(EDCN)的结构

Fig. 2 Structure of enhanced deformable convolution

通过 SAB 与 CAB 的协同作用计算偏移量:

$$\Delta p_n = \text{Conv2d}(\text{SAB}(x) \oplus \text{CAB}(x)). \quad (2)$$

### 1.3 损失函数

使用 Wise-IoU<sup>[10]</sup> 作为损失函数 ( $L_{\text{WIoUv1}}$ ), 定义为:

$$L_{\text{WIoUv1}} = R_{\text{WIoU}} L_{\text{IoU}}, \quad (3)$$

$$R_{\text{WIoU}} = \exp\left(\frac{(x - x_{gt})^2 + (y - y_{gt})^2}{(W_g^2 + H_g^2)^*}\right), \quad (4)$$

式中:  $R_{\text{WIoU}}$  的核心作用是通过动态调整权重, 迫使模型更加关注与目标中心点对齐的预测边界;  $L_{\text{IoU}}$  是标准 IoU 损失, 用于量化预测边界框与真实边界框的重叠程度, 其值为两框交集面积与并集面积

之比;  $(x, y)$  和  $(x_{gt}, y_{gt})$  分别表示预测框与真实边界框的中心坐标;  $W_g$  和  $H_g$  对应同时包含真实边界框与预测框的最小封闭区域的宽度和高度; 上标星号(\*)表示在计算过程中, 为避免  $W_g$  和  $H_g$  产生阻碍模型收敛的梯度异常, 将二者从计算图中分离处理, 确保训练过程的稳定性。

## 2 结果与分析

实验采用 PyTorch 深度学习框架搭建网络模型, 依托 2 块 RTX 2080 Ti 显卡进行并行训练, 并通过 CUDA 加速计算, 以提升模型训练效率与收敛速度。

### 2.1 数据集

为验证所提方法的有效性, 实验采用东北大学公开的带钢表面缺陷标准数据集<sup>[12]</sup> (NEU-DET)。该数据集包含裂纹(Cr)、内含物(In)、斑块(Pa)、点蚀表面(Ps)、轧制氧化皮(Rs)及划痕(Sc)共 6 类典型钢表面缺陷, 部分示例如图 3 所示。每类缺陷含 300 幅图像, 共计 1800 幅样本, 分辨率均为  $200 \times 200$  像素。数据集按 8:2 的比例随机划分为训练集与验证集。

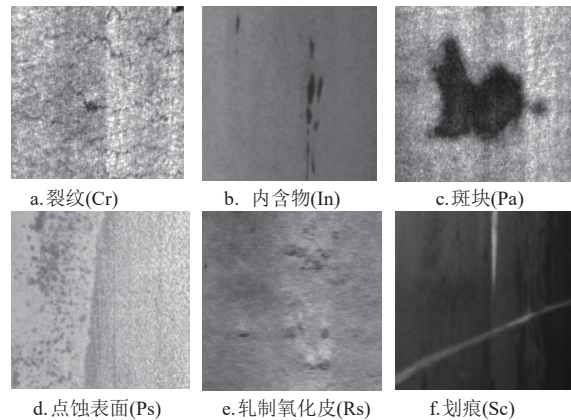


图 3 缺陷示例图像

Fig. 3 Sample image of defect

### 2.2 模型评价指标

采用平均精度均值 (Mean Average Precision, mAP) 和 F1-score 作为缺陷检测结果的主要评价指标。

(1) mAP 是目标检测领域的核心评估指标, 需为每个类别计算精度均值。其计算公式如下:

$$\text{AP} = \int_0^1 p(r) dr, \quad (5)$$

$$mAP = \frac{1}{n} \sum_{i=1}^n AP_i. \quad (6)$$

(2) 精度(又称查准率),用于描述预测为正例的样本中,实际也为真正例的比例,其计算公式为:

$$P = \frac{TP}{TP + FP} \times 100\%, \quad (7)$$

式中:TP表示模型预测为正的样本,FP表示模型预测为负的正样本。

(3) 召回率:目标召回率针对特定目标类别定义,即预测正确的目标框数与所有真实目标框数的比值,其计算公式为:

$$R = \frac{TP}{TP + FN} \times 100\%, \quad (8)$$

式中:FN为模型预测为正的负样本。

(4)  $F1$ -score:准确率与召回率的调和平均数,常用于分类模型的性能评估。

$$F1\text{-score} = \frac{2 \times P \times R}{P + R}. \quad (9)$$

### 2.3 不同检测模型对比实验

为验证所提方法的有效性,选择9种典型目标检测方法作为对比基准,具体包括SSD-300<sup>[7]</sup>、Faster-RCNN<sup>[4]</sup>、RetinaNet<sup>[13]</sup>、YOLOv5<sup>[14]</sup>、YOLOv8<sup>[8]</sup>、PP-YOLOE<sup>[15]</sup>、RTMDET<sup>[16]</sup>、DETR<sup>[17]</sup>及RT-DETR<sup>[18]</sup>。表1展示了各对比方法在NEU-DET数据集上的检测结果。

如表1所示,所提方法的综合性能最优,mAP为81.5%, $F1$ -score为77.1%。相较于SSD-300、Faster-RCNN、RetinaNet、YOLOv5、YOLOv8、PP-YOLOE、RTMDET、DETR及RT-DETR,mAP分别提升7.8、2.1、14.1、4.5、4.1、6.7、8.4、10.3及8.8个百分点, $F1$ -score对应提升14.0、19.2、15.9、11.9、2.6、4.8、5.7、12.3及6.1个百分点。

结果表明,EDCN通过协同注意力机制与可变形卷积的深度融合,自适应地调整感受野,进而能够在不同尺度和形状的缺陷区域中提取更精准的特征。协同注意力机制能够动态地捕捉和增强关键区域的信息,尤其在缺陷形态复杂、背景干扰大的情况下,能有效抑制无关背景信息,突出缺陷本身的特征。可变形卷积则通过自适应地调整卷积核的采样位置,使网络灵活适应不同形状缺陷,尤其在缺陷边缘和形状变异较大的情况下,提升了特征提取的准确性。

表1 本文模型与主流目标检测模型对比结果

Tab. 1 Comparisons between the proposed model and the mainstream models

模型	mAP/%	$F1$ -score/%
SSD-300	73.7	63.1
Faster-RCNN	79.4	57.9
RetinaNet	67.4	61.2
YOLOv5	77.0	75.2
YOLOv8	77.4	74.5
DETR	71.2	64.8
RT-DETR	72.7	71.0
RTMDET	73.1	71.4
PP-YOLOE	74.8	72.3
本文模型	81.5	77.1

### 2.4 消融实验

为进一步验证所提方法的有效性,针对网络模型开展消融实验。通过在骨干网络引入EDCN模块并优化损失函数,以评估各改进方法对模型性能的影响,实验结果如表2所示。

表2 消融实验结果

Tab. 2 Results of ablation experiments

模型	EDCN	$W_{IoU}$	mAP/%
基线模型	×	×	74.5
优化模型1	✓	×	78.1
优化模型2	×	✓	76.8
本文模型	✓	✓	81.6

由表2数据可知,以未引入EDCN模块且未优化损失函数的模型作为基准(mAP=74.5%),仅在骨干网络中加入EDCN模块时,模型平均精度(mAP)提升至78.1%,较基准提升3.6个百分点,验证了EDCN模块通过动态调整感受野、精准捕捉缺陷特征的核心价值,使模型对缺陷信息的提取能力显著增强;仅优化损失函数时,模型mAP较基准提升2.3个百分点,表明优化后的损失函数能更高效地引导模型学习,提升定位与分类的精准度;而同时引入EDCN模块与优化损失函数时,较基准模型提高了7.1个百分点。

### 2.5 泛化实验

为进一步验证所提方法的有效性,在公开可用的德国异常检测数据集DAGM<sup>[19]</sup>上进行了泛化实验。该数据集包含10类纹理背景缺陷,图像尺寸为512×512像素。表3列出了对比方法的实验

结果。

由表 3 数据可知,所提模型在 mAP 与 F1-score 指标上显著优于其他对比模型,mAP 指标达到 98.0%,F1-score 指标高达 97.3%。

## 2.6 可视化

为直观验证所提方法的性能,图 4 展示了模型的检测结果。

如图 4 所示,所提方法能够有效地检测出形状不规则的缺陷(如 In(第 2 列)、Pa(第 3 列)、Sc(末列))以及边缘不清晰的缺陷(如 Cr(第 1 列)、Ps(第 4 列)、Rs(倒数第 2 列))。从可视化结果可见,模型对不同形态、不同边缘特征的缺陷均展现出较强的识别与定位能力,进一步验证了所提方法在复杂缺陷检测场景中的有效性与可靠性。

表 3 泛化性实验对比结果

Tab. 3 Generalization experiment comparison results

模型	mAP/%	F1-score/%
SSD-300	95.0	90.1
Faster-RCNN	95.4	83.8
Retinanet	87.8	84.3
YOLOv5	94.6	94.7
YOLOv8	95.8	93.9
DETR	80.1	75.7
RT-DETR	91.8	87.2
RTMDet	93.1	90.5
PP-YOLOE	92.9	89.7
本文模型	98.0	97.3

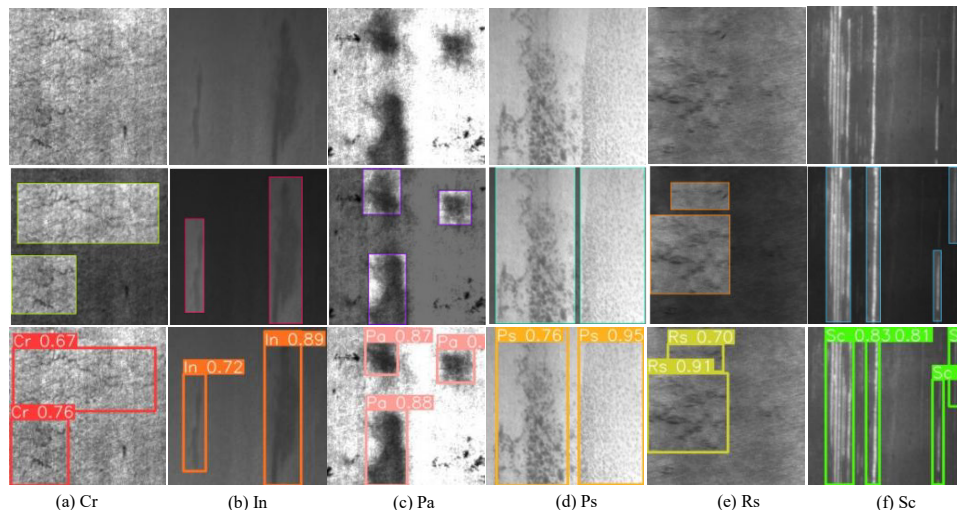


图 4 改进后模型的可视化

Fig. 4 Visualization of the improved model

## 3 结束语

带钢表面缺陷检测是保证带钢质量的关键。提出一种新型的带钢表面缺陷检测方法,以增强特征提取能力。具体而言,设计一种改进的变形卷积,利用协同注意力机制捕获带钢表面的不规则缺陷,从而增强了骨干网络的特征提取能力。

## 参考文献:

- [1] 徐科,杨朝霖,周鹏. 热轧带钢表面缺陷在线检测的方法与工业应用[J]. 机械工程学报, 2009, 45(4): 111-114, 124.  
XU Ke, YANG Zhaolin, ZHOU Peng. Technology of online surface inspection for hot rolled steel strips and its industrial application[J]. Journal of Mechanical Engineering, 2009, 45(4): 111-114, 124.

在 NEU-DET 和 DAGM 数据集上进行了大量实验,结果表明:与其他主流方法相比,所提方法的 mAP 值最高,分别为 81.6% 和 98.0%,表明其在带钢表面缺陷检测中的有效性。

未来工作中,将对网络进行更深入的优化,提升网络的特征辨识能力。同时,将优化网络中的共享超参数,以获得最优超参数组合。

- [ 2 ] 马燕婷,赵红东,阎超,等.改进YOLOv5网络的带钢表面缺陷检测方法[J].电子测量与仪器学报,2022,36(8):150-157.  
MA Yanting, ZHAO Hongdong, YAN Chao, et al. Strip steel surface defect detection method by improved YOLOv5 network[J]. Journal of Electronic Measurement and Instrumentation, 2022, 36(8): 150-157.
- [ 3 ] WANG Shuai, XIA Xiaojun, YE Lanqing, et al. Automatic detection and classification of steel surface defect using deep convolutional neural networks[J]. Metals, 2021, 11(3): 388.
- [ 4 ] REN Shaoqing, HE Kaiming, GIRSHICK R, et al. Faster R-CNN: Towards real-time object detection with region proposal networks[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2017, 39(6): 1137-1149.
- [ 5 ] HE Kaiming, ZHANG Xiangyu, REN Shaoqing, et al. Deep residual learning for image recognition[C]//2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Las Vegas, 2016: 770-778.
- [ 6 ] LIU Xiaoyue, GAO Jie. Surface defect detection method of hot rolling strip based on improved SSD model[C]//Database Systems for Advanced Applications. DASFAA 2021 International Workshops, Taipei, 2021: 209-222.
- [ 7 ] LIU Wei, ANGUELOV D, ERHAN D, et al. SSD: Single shot MultiBox detector[C]//Computer Vision-ECCV 2016, Amsterdam, 2016: 21-7.
- [ 8 ] DU Yongzhao, CHEN Haixin, FU Yuqing, et al. AFF-Net: A strip steel surface defect detection network via adaptive focusing features[J]. IEEE Transactions on Instrumentation and Measurement, 2024, 73: 1-14.
- [ 9 ] DAI Jifeng, QI Haozhi, XIONG Yuwen, et al. Deformable convolutional networks[C]//2017 IEEE International Conference on Computer Vision (ICCV), Venice, 2017: 764-773.
- [10] 刘凤春,张靖,薛涛,等.YOLO-VDCW:一种新的轻量化带钢表面缺陷检测算法[J].中国冶金,2024,34(6):125-135.  
LIU Fengchun, ZHANG Jing, XUE Tao, et al. YOLO-VDCW: A new light weight surface defect detection algorithm for stripsteel[J]. China Metallurgy, 2024,34(6):125-135.
- [11] LEE S H, LIN S R, CHEN S F. Identification of tea foliar diseases and pest damage under practical field conditions using a convolutional neural network[J]. Plant Pathology, 2020, 69(9): 1731-1739.
- [12] SONG Kechen, YAN Yunhui. A noise robust method based on completed local binary patterns for hot rolled steel strip surface defects[J]. Applied Surface Science, 2013, 285(Part B): 858-864.
- [13] 于敏,屈丹,司念文.改进的RetinaNet目标检测算法[J].计算机工程,2022,48(8):249-257.  
YU Min, QU Dan, SI Nianwen. Improved RetinaNet algorithm for object detection[J]. Computer Engineering, 2022, 48(8): 249-257.
- [14] 邱天衡,王玲,王鹏,等.基于改进YOLOv5的目标检测算法研究[J].计算机工程与应用,2022,58(13):63-73.  
QIU Tianheng, WANG Ling, WANG Peng, et al. Research on object detection algorithm based on improved YOLOv5[J]. Computer Engineering and Applications, 2022,58(13):63-73.
- [15] 姜杰,张立民,刘凯,等.基于改进PP-YOLOE和ByteTrack算法的红外船舶目标检测跟踪方法[J].兵器装备工程学报,2024,45(11):291-297.  
JIANG Jie, ZHANG Limin, LIU Kai, et al. Research on infrared ship target detection and tracking method based on improved PP-YOLOE and ByteTrack algorithms[J]. Journal of Ordnance Equipment Engineering, 2024, 45(11): 291-297.
- [16] 张体浪,周梦颖,陆源,等.基于改进RTMDet-ins的城市排水管网缺陷检测算法研究[J].给水排水,2025,61(3):147-154.  
ZHANG Tilang, ZHOU Mengying, LU Yuan, et al. Research on defect detection algorithm of urban drainage network based on improved RTMDet-ins[J]. Water & Wastewater Engineering, 2025, 51(3): 147-154.
- [17] CARION N, MASSA F, SYNNAEVE G, et al. End to end object detection with transformers[C]//Computer Vision-ECCV 2020: 16th European Conference, Glasgow, 2020: 213-229.
- [18] ZHAO Yian, LYU Wenyu, XU Shangliang, et al. Detsr beat YOLOs on real-time object detection[C]//2024 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), Seattle, 2024: 16965-16974.
- [19] RIGOLL G. Pattern recognition: 30th DAGM symposium munich[M]. Berlin: Springer, 2008.

责任编辑:郭红建 陈松楠