

基于ATI-YOLO的铝合金焊缝表面缺陷检测

张正¹, 邹丽^{1,2}, 陈千荣¹

(1. 大连交通大学轨道智能工程学院; 2. 大连市区块链技术与应用重点实验室, 辽宁大连 116028)

摘要: 焊缝缺陷检测是材料加工和成型过程中的关键环节, 表面缺陷会直接影响焊接质量的评估。为提高小目标铝合金焊缝表面缺陷检测的精准度, 以深度学习算法YOLOv11n为基础提出ATI-YOLO模型。首先, 使用ADown-Conv结构代替原模型中的下采样方式, 通过引入自适应策略改善下采样过程中信息丢失和忽略关键特征的问题, 同时增强模型的灵活性。其次, 设计C2PSA-TA模块, 帮助模型从3种维度理解图像的缺陷信息, 提高焊缝缺陷检测效率。最后, 将LMPDIoU损失函数与Inner思想结合, 设计了InnerLMPDIoU损失函数, 以优化目标检测中的边界框回归问题, 提高模型的准确性。多组实验结果表明, ATI-YOLO模型相比原YOLOv11n模型在Precision、Recall、mAP50指标上分别提升了3.6%、2%、2.5%。此外, 在NEU-DET和GC10-DET数据集上的泛化实验结果也表明, ATI-YOLO模型在mAP50指标上分别提升了2.3%、2.8%, 充分证明了该改进模型的泛用性和有效性。

关键词: 铝合金焊缝; 表面缺陷检测; YOLOv11n; 自适应策略

DOI: 10.11907/rjtk.251049

中图分类号: TP317.4

文献标识码: A

文章编号: 1672-7800(2026)003-0212-09

扫描二维码阅读全文:



Aluminum Alloy Weld Surface Defect Detection Based on ATI-YOLO

ZHANG Zheng¹, ZOU Li^{1,2}, CHEN Qianrong¹

(1. School of Railway Intelligent Engineering, Dalian Jiaotong University;

2. Dalian Key Laboratory of Blockchain Technology and Application, Dalian 116028, China)

Abstract: Weld seam defect detection is a key step in material processing and forming, its results directly affect the evaluation of welding quality. To improve the precision of aluminum alloy weld seam defect detection, the ATI-YOLO model was proposed based on the deep learning algorithm YOLOv11n. First, the ADown-Conv structure was used to replace the downsampling method in the original model. An adaptive strategy was introduced to reduce information loss and prevent the neglect of key features during downsampling. This also enhanced the model's flexibility. Second, the C2PSA-TA module was designed to help the model understand defect information from three dimensions, improving weld seam defect detection efficiency. Finally, the LMPDIoU loss function was combined with the Inner concept to create the InnerLMPDIoU loss function. This optimized the bounding box regression problem in object detection, improving the model's accuracy. Experimental results showed that the ATI-YOLO model outperformed the original YOLOv11n model by 3.6, 2, and 2.5 percentage points in Precision, Recall, and mAP50, respectively. Moreover, generalization experiments on the NEU-DET and GC10-DET datasets showed that the ATI-YOLO model increased mAP50 by 2.3 and 2.8 percentage points, respectively. These results fully proved the versatility and effectiveness of the improved model.

Key Words: aluminum alloy weld seam; surface defect detection; YOLOv11n; adaptive strategy

0 引言

铝合金材质因具备良好的耐腐蚀性、较高的强度以及

便捷的维护使用特性, 在航空航天和轨道交通等关键领域得到了广泛应用。在实际工业应用中, 由于受到材料自身特性以及外部环境等因素影响, 焊接表面可能会产生各种程度的缺陷, 例如未熔合、裂纹等^[1]。这些缺陷可能威胁

收稿日期: 2025-02-19

基金项目: 辽宁省应用基础研究计划项目(2023JH2/101300236)

作者简介: 张正(2001-), 男, CCF学生会会员, 大连交通大学轨道智能工程学院硕士研究生, 研究方向为计算机视觉、图像处理; 邹丽(1980-), 女, 博士, CCF会员, 大连交通大学轨道智能工程学院教授、博士生导师, 研究方向为智能信息处理; 陈千荣(2001-), 女, 大连交通大学轨道智能工程学院硕士研究生, 研究方向为智能信息处理。本文通讯作者: 邹丽。

生产安全,削弱材料的疲劳强度和抗腐蚀能力。因此,为防止意外造成损失,焊后缺陷检测至关重要。传统目标检测算法通常依赖滑窗技术和手工特征提取方法,例如 Viola-Jones 人脸检测算法^[2]和 HOG+SVM 缺陷检测算法^[3]。然而,滑窗方法缺乏针对性,计算复杂且容易产生冗余,而手动设计的特征在应对多样性变化时表现出较弱的鲁棒性,难以胜任复杂背景下的焊缝缺陷检测任务。因此,随着人工智能技术的兴起,研究者们试图将深度学习技术应用于焊缝缺陷识别,旨在让检测效果更加精确、高效。

早期的深度学习模型主要为双阶段模型。2014年, Girshick 等^[4]提出基于卷积神经网络(Convolutional Neural Network, CNN)的 R-CNN 算法,在 VOC2007 数据集上的平均准确率(mAP)达到 58.5%,远超传统方法的 33.7%,实现了显著提升。相比传统的机器视觉算法, CNN 的优势在于能够自主完成特征提取,生成更具代表性的特征。由于 CNN 网络具备强大的特征提取能力,研究人员尝试将其应用于铝合金焊缝表面缺陷检测领域,并取得了良好效果,但双阶段模型的检测速度通常较慢,使其在工业领域的应用受到限制。近年来,深度学习驱动的目标识别技术日益成熟。YOLO(You Only Look Once)模型作为一种基于卷积神经网络(CNN)的单阶段目标检测技术,凭借高精度与快速响应的特点备受关注^[5]。其将目标检测转化为回归问题,利用网格划分与边界框定位来高效预测目标。目前, YOLO 已发展至 YOLOv11。程松等^[6]针对 YOLOv5s 模型在铝合金焊缝缺陷检测中漏检率较高的问题,对其结构进行了优化:用 DConv 替换原卷积模块,并采用 DIoU 和 CIoU 作为新的损失函数。改进后的模型在提升检测精度的同时,有效减少了参数量。黎源东等^[7]提出一种建立在 YOLOv8 基础上的 YOLOv8-LSD 模型,该方法加入 Shuffle-Attention 机制以提升特征提取能力,并在 Head 部分增设 160×160 小目标检测层,以增强小目标缺陷检测效果。通过在 NEU-DET 数据集上的实验验证,相较于原模型,其总体平均检测精度(mAP)提升了约 3.9%。

上述改进 YOLO 算法已广泛应用于铝合金焊缝表面缺陷检测,但在检测小目标铝合金焊缝表面缺陷时效果欠佳,且精度较低。尽管 YOLOv11 具备较高的检测效率,但小目标检测始终是 YOLO 系列算法的挑战。而传统的目标检测算法又具有精度低、速度慢的缺点。为此,本文针对小目标检测问题,提出一种基于 YOLOv11 的改进模型 ATI-YOLO,在保证检测准确率的同时,尽可能提高检测效率。本文的主要工作如下:①使用 ADown-Conv 模块代替原本的下采样方式,实现模型轻量化,提高小目标缺陷检测精度,并降低参数量与计算量;②改进 YOLOv11 的 C2PSA 模块,引入三重注意力机制(Triple Attention)^[8],构建 C2PSA-TA 模块,以增强网络对缺陷信息的理解;③使用一种基于 MPDIoU^[9]的边界框回归损失函数 LMPDIoU,并结合 Inner 思想设计 InnerLMPDIoU 损失函数,优化边界

框回归,以提高 YOLOv11 模型的准确性。

1 方法介绍

1.1 YOLOv11 模型

YOLOv11 是最新的目标检测模型,在 YOLO Vision 2024(YV24)大会上展现了实时检测技术的重大突破。该版本优化了结构与训练策略,在精度、速度和效能上均有显著提升。其创新设计集成了先进的特征提取技术,在保持参数精简的同时,增强了细节捕捉能力,提高了检测的准确性^[10]。YOLOv11 通过宽度和深度系数调整模型规模,提供 n、s、m、l、x 5 种尺寸。其中, YOLOv11n 体积最小,检测速度最快。因此,本文选择使用 YOLOv11n 模型。YOLOv11 网络结构如图 1 所示。

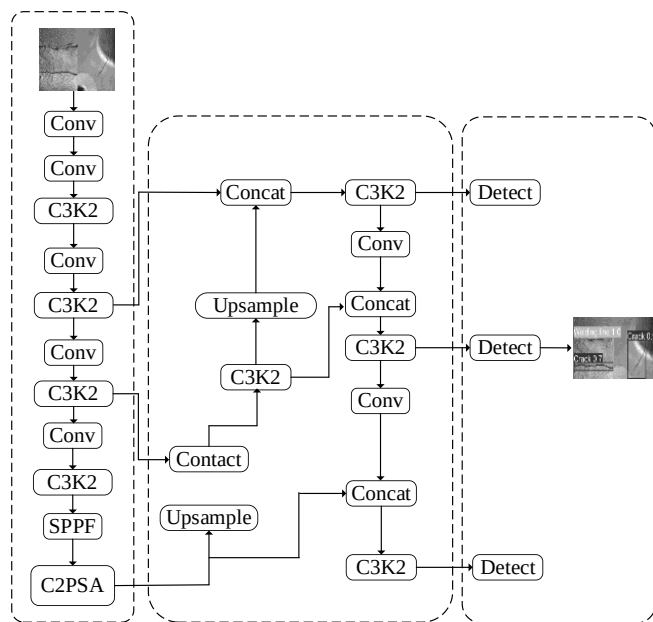


Fig. 1 YOLOv11 network structure

图1 YOLOv11 网络结构

YOLOv11 由主干、颈部和头部组成,整体结构类似 YOLOv8,但进行了多项优化,包括: C2f 模块被 C3k2 模块替换,结合可变卷积核和通道分离策略,增强了特征提取能力;引入 C2PSA 机制,将 C2(C2f 前身)与多头注意力结合,提高了多尺度特征提取性能;在解耦头的分类检测头中增加两层 DWConv,大幅降低了参数量和计算量。此外,由于模型缩小,因此调整深度和宽度放缩倍数以保持性能平衡。

1.2 ATI-YOLO 模型

针对焊缝缺陷图像目标分布分散、尺寸多样及特征不明显等问题, YOLOv11n 检测效果有限,而直接增大模型会导致计算量激增。为此,本文提出 ATI-YOLO,在兼顾性能与资源消耗的前提下优化模型结构。ATI-YOLO 模型结构如图 2 所示。

首先,采用 ADown-Conv 模块替代固定的下采样方式,

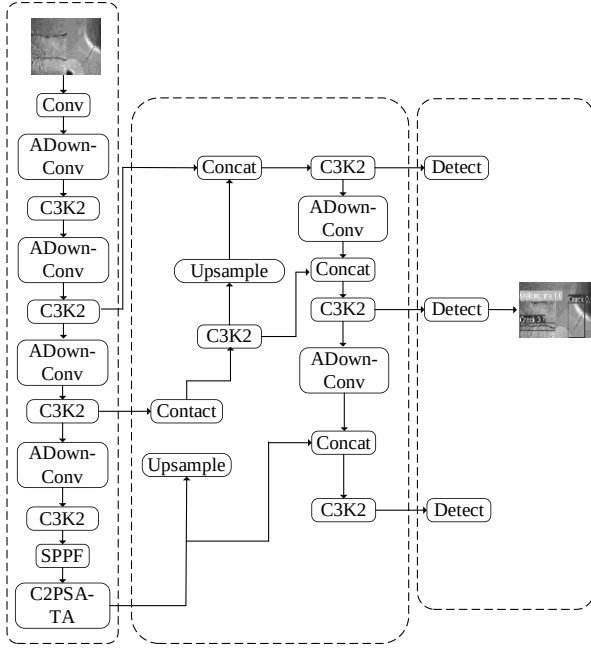


Fig. 2 ATI-YOLO network structure

图2 ATI-YOLO模型结构

使其能自适应调整下采样策略,并结合注意力机制优化空

间信息提取。其次,引入三重注意力机制(Triplet Attention),将C2PSA替换为C2PSA-TA模块,增强对不同特征类别的感知能力,提高局部与全局特征捕捉能力,使模型鲁棒性更强,且可扩展,可适用于多模态任务。在铝合金焊缝表面缺陷检测中,该改进方式展现了优势。最后,结合Inner思想改进LMPDIoU损失函数,提出InnerLMPDIoU损失函数,以优化边界框回归任务,加速收敛并提升检测精度。

1.3 ADown-Conv

焊接铝合金易产生裂纹和未焊透等缺陷,因形态相似,难以区分。而传统Conv模块特征提取能力有限,难以准确识别和定位缺陷,且因依赖单一尺度卷积核,难以适应多尺度检测。为此,本文提出以下采样ADown模块替换YOLOv11中的下采样卷积,并设计ADown-Conv模块,以增强小目标特征提取和下采样能力。ADown-Conv模块由多个卷积层和池化操作组成,结合了平均池化和最大池化策略,可有效缩小目标尺寸并增加特征通道,提升目标检测效率。其创新的双分支结构通过特征图进行分割和差分处理,增强了对不同尺寸缺陷的识别能力。ADown-Conv网络结构如图3所示。

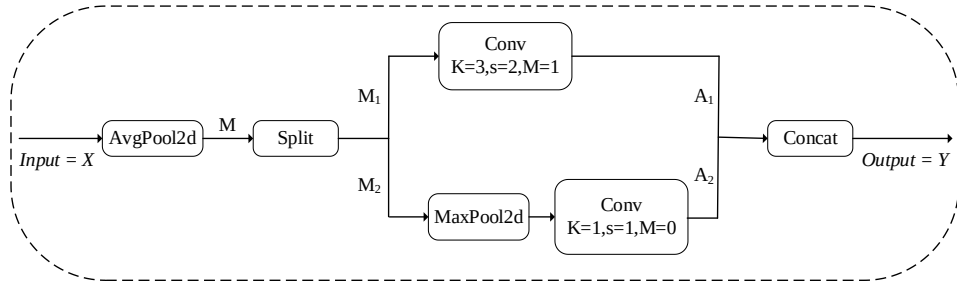


Fig. 3 ADown-Conv network structure

图3 ADown-Conv网络结构

下采样ADown-Conv模块先对输入特征图 X 进行平均池化(AvgPool),得到低分辨率特征图 M ,再将其均匀分割(Split)成 M_1 和 M_2 两部分,分别对应模块的左、右分支,过程如式(1)所示。左分支通过 3×3 卷积核($Conv_{3 \times 3}$)对 M_1 进行下采样,增强对局部细节的捕捉,得到 A_1 ,过程如式(2)所示。右分支先进行最大池化(MaxPool)以突出显著特征,再通过 1×1 卷积($Conv_{1 \times 1}$)以实现特征压缩和融合,得到 A_2 ,过程如式(3)所示。最终在通道维度拼接(Concat)两分支特征图,输出特征图 Y ,从而有效融合多种特征,提高铝合金焊缝缺陷检测能力,输出表达式如式(4)所示。

$$M_1, M_2 = \text{Split}(\text{AvgPool}(X)) \quad (1)$$

$$A_1 = \text{Conv}_{3 \times 3}(M_1) \quad (2)$$

$$A_2 = \text{Conv}_{1 \times 1}(\text{MaxPool}(M_2)) \quad (3)$$

$$Y = \text{Concat}(A_1, A_2) \quad (4)$$

尽管ADown-Conv模块在多尺度缺陷检测中具有优势,但也存在特征丢失和通用性不足等局限性。过度下采样可能导致细节信息丧失,并且对大尺寸或复杂形状的缺

陷适应性较差,限制了其在其他场景中的泛化能力。

1.4 C2PSA-TA

低分辨率可能导致中低层特征语义不足,影响目标定位精度,尤其在复杂场景和小目标检测中表现明显。同时,在特征分割和信息传递过程中,邻域组合操作可能削弱低层特征的表达能力,影响检测速度。为解决这些问题,本文提出C2PSA-TA模块,通过三重注意力机制(Triplet Attention)和三分支结构捕捉跨维度交互关系,利用旋转操作与残差变换构建维度间的依赖性,从多个视角分析输入数据,从而提高检测效果。三重注意力机制原理如图4所示。

三重注意力机制由3个平行的分支组成,前两个分支用于捕捉通道 C 与空间维度 H, W 之间的相互作用,第三个分支通过Z-Pool和卷积(Conv)操作生成空间注意力权重。3个分支的输出通过平均聚合产生最终效果。其中,Z-Pool层通过平均池化和最大池化特征降低张量深度,在保留丰富表示的同时减轻计算负担,如式(5)所示:

$$Z - \text{Pool}(\chi) = [\text{MaxPool0d}(\chi), \text{AvgPool0d}(\chi)] \quad (5)$$

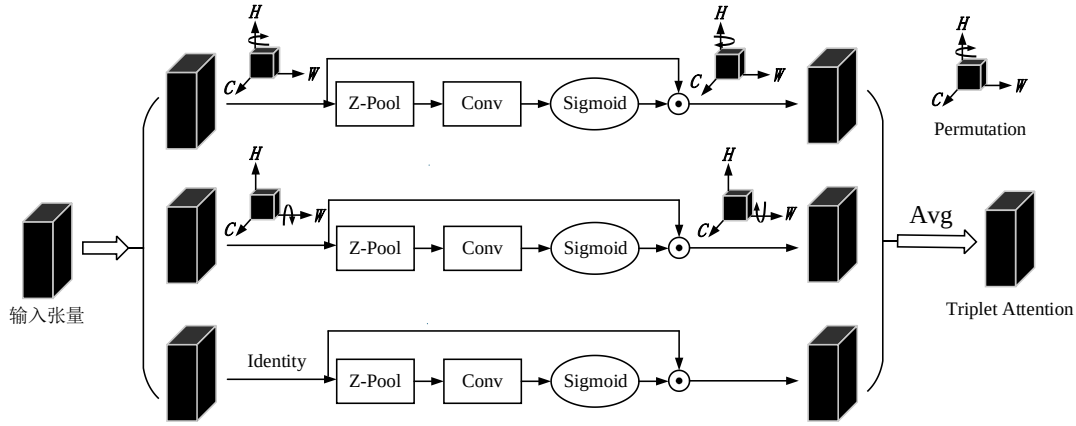


Fig. 4 The principle of triplet attention

图4 三重注意力机制原理

其中, $\text{MaxPool0d}(\chi)$ 、 $\text{AvgPool0d}(\chi)$ 分别代表最大池化和平均池化, $0d$ 为第零维交叉点。

三重注意力机制结构如图5所示。

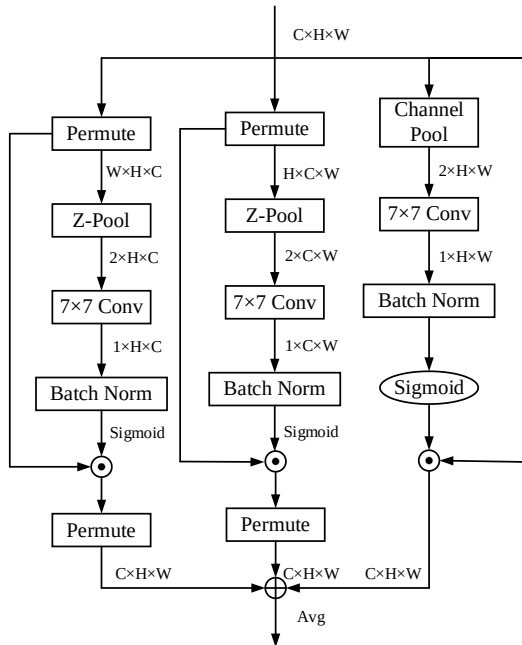


Fig. 5 The structure of triplet attention

图5 三重注意力机制结构

结合图4与图5, 当给定输入为 $C \times H \times W$ 时, 3个分支的具体实现步骤如下:

给定一个输入张量 $X \in R^{C \times H \times W}$, 将其传递给3个分支的每一个分支。

(1) 第一个分支将输入张量 X 沿 H 轴逆时针旋转 90° 进行 Z-Pool 操作后, 经过 7×7 卷积、Batch Norm 和 Sigmoid 激活函数生成空间注意力权重, 再通过残差变换和旋转操作将注意力权重与原始特征图相乘, 得到维度为 $(C \times H \times W)$ 的特征图。

(2) 第二个分支将输入张量 X 沿 W 轴逆时针旋转 90° 进行 Z-Pool 操作后, 经过 7×7 卷积、Batch Norm 和 Sigmoid

激活函数生成空间注意力权重, 再通过残差变换和旋转操作将注意力权重与原始特征图相乘, 得到维度为 $(C \times H \times W)$ 的特征图。

(3) 第三个分支通过 Channel Pool 处理输入张量 X , 生成形状为 $(2 \times H \times W)$ 的特征图, 经过 7×7 卷积、Batch Norm 和 Sigmoid 激活函数生成空间注意力权重, 再通过残差变换和旋转操作将注意力权重与原始特征图相乘, 得到维度为 $(C \times H \times W)$ 的特征图。

(4) 最后, 将3个分支的输出结果进行平均汇总, 生成形状为 $(C \times H \times W)$ 的精细张量。

虽然 C2PSA-TA 模块在增强特征表达和跨维度交互方面表现突出, 但也存在计算开销大、针对特定任务适应性差等局限性。其需要进行多个操作, 会增加计算量和内存消耗。尽管在多尺度和小目标检测中表现优异, 但在大目标或复杂背景下的任务中可能不如传统 CNN 结构灵活, 存在过拟合或泛化能力不足的风险。

1.5 InnerLMPDIoU

损失函数在 YOLO 训练中可起到优化模型性能、衡量预测与实际数据差异的作用, 目的是尽量降低损失值以提高预测准确性。常见的 IoU 损失函数包括: ①GIoU: 通过最小外接矩形解决无交集问题^[11]; ②DIoU: 综合考虑框的重叠、距离和尺度^[12]; ③CIoU: 在目标形状和尺寸不规则时表现更优; ④EIoU 则进一步考虑了宽高比^[13]。然而, 当预测框与真实框的长宽比相同但尺寸不同时, 这些损失函数无法有效进行优化。为此, MPDIoU 基于水平矩形的最小点距离进行计算, 综合考虑重叠、距离和宽高偏差, 提供更精准的损失评估。MPDIoU 的影响因素如图6所示。

其中, 实线框表示预测框, 虚线框表示真实框, w 、 h 分别表示边界框的宽和高, x_1^{prd} 、 y_1^{prd} 分别表示预测框的左上点坐标和右下点坐标, x_1^{gt} 、 y_1^{gt} 分别表示真实框的左上点坐标和左下点坐标, d_1 、 d_2 分别表示两框之间的最小点距离。计算公式如式(6)–(9)所示:

$$d_1^2 = (x_1^{prd} - x_1^{gt})^2 + (y_1^{prd} - y_1^{gt})^2 \quad (6)$$

$$d_2^2 = (x_2^{prd} - x_2^{gt})^2 + (y_2^{prd} - y_2^{gt})^2 \quad (7)$$

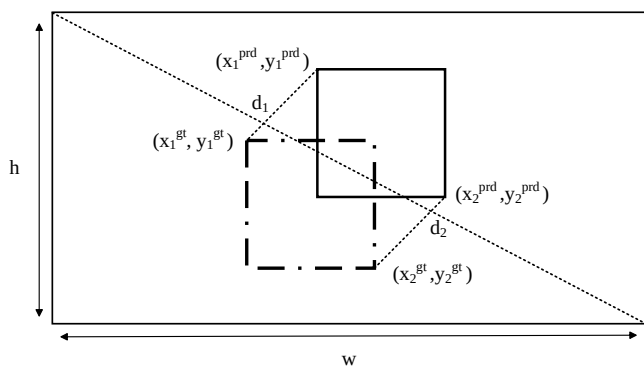


Fig. 6 Influencing factors of MPDIoU

图6 MPDIoU影响因素

$$MPDIoU = IoU \frac{d_1^2}{w^2 + h^2} \frac{d_2^2}{w^2 + h^2} \quad (8)$$

$$L_{MPDIoU} = 1 - MPDIoU \quad (9)$$

其中, $MPDIoU$ 表示最小距离交并比, IoU 表示交并比, L_{MPDIoU} 表示 $MPDIoU$ 损失值。

本文将 InnerIoU 思想引入 $MPDIoU$, 设计了新的损失函数 Inner $MPDIoU$ ^[14]。InnerIoU 通过引入比例因子调节辅助边界框大小来计算损失, 优化边界框回归。通常, 这一比例因子的取值范围为 $[0.5, 1.5]$ 。当比例因子大于 1 时, 辅助边界框扩展至大于真实边界, 有助于加速低 IoU 样本的回归过程; 当比例因子小于 1 时, 辅助框缩小, 回归梯度更大, 有助于加速高 IoU 样本的回归过程。将 InnerIoU 融入 $MPDIoU$ 可提升目标检测精度及泛化能力, 但由于在不同的 IoU 阶段, 辅助边界框的缩放可能导致梯度更新不稳定, 从而导致计算复杂度高, 并引发了稳定性问题, 需要通过实验进行验证。

2 仿真实验与结果分析

2.1 实验环境

采用 PyTorch 作为深度学习框架, 实验硬件配置为: 12 vCPU Intel(R) Xeon(R) Silver 4214R CPU @ 2.40GHz 处理器、NVIDIA GeForce RTX 3080 Ti GPU 显卡, 编译环境为: Python3.8+torch1.11.0+CUDA11.3。模型具体训练参数见表 1。

2.2 实验数据集

本文采用由现场工作人员拍摄的小目标铝合金表面

Table 1 Model training parameters

表1 模型训练参数

参数名称	数值
图像尺寸	640×640
训练周期	300
批量大小	16
初始学习率	0.01
权重衰减系数	5×10^{-4}
学习率动量	0.937
训练优化器	SGD

焊缝缺陷图像, 这些图像记录了焊接机器人在完成焊接作业后的焊缝状态。所使用的焊机型号为 fronius 的 TPS 600i 型机器人焊接设备, 焊接方式采用 MIG 氩弧焊技术。在焊接过程中, 设定电压为 22.3 V, 电流为 150 A, 将焊枪的移动速度控制在 45 cm/min。拍摄完毕后, 将拍摄的图片按编号分类存储, 确保命名清晰, 并对拍摄图像进行初步标注, 记录缺陷位置、类型和相关描述。该数据集包括 4 种缺陷类别, 分别为裂纹 (Crack)、气孔 (Porosity)、飞溅 (Spatters) 与未熔合 (Welding line)。裂纹呈线状或曲线状, 具有明显的断裂特征, 且宽度较小, 深度较深, 沿着焊缝或焊接接头方向分布; 气孔呈圆形或椭圆形, 通常分布均匀; 飞溅形态不规则, 通常为小的金属颗粒或颗粒团; 未融合形态不规则, 通常局部区域没有完全融合, 出现焊缝与母材间的间隙。图像分辨率为 2 560×1 920。为了应对数据集存在的类别不平衡问题, 实验通过随机翻转、亮度变换和旋转缩放等数据增强方法, 生成了 5 820 张各类别样本数量均衡的缺陷图像。数据集部分示例图如图 7 所示。为保证模型训练效果, 将数据集按照 8:1:1 的比例划分为训练集 4 656 张、验证集 582 张和测试集 582 张图片。由于该数据集中的所有图像均在一线人员工作处拍摄, 属于原始一手图像, 因此更真实地反映了焊缝缺陷的实际情况。

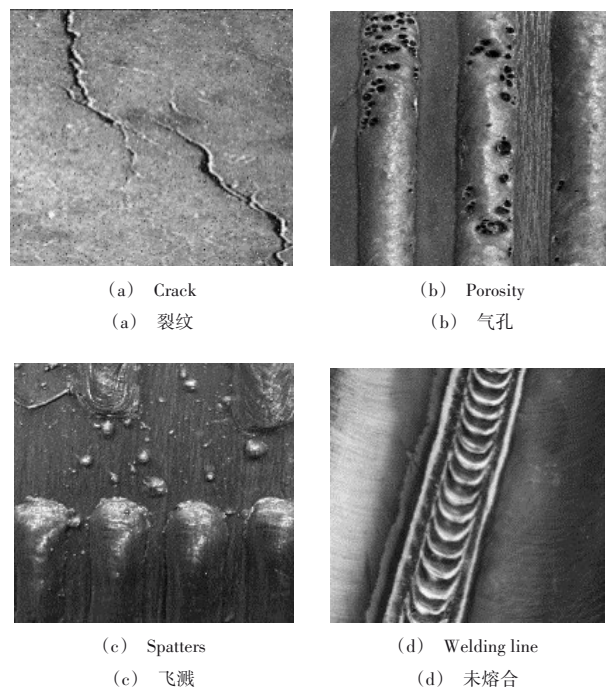


Fig. 7 Defect schematic diagram

图7 缺陷示意图

2.3 评价指标

在目标检测领域, 模型的性能评估至关重要, 而检测的准确性是评估的核心指标之一。本文选择平均精确率 (Mean Average Precision, mAP)、准确率 (Precision)、召回率 (Recall)、模型计算量 (GFLOPs) 以及模型参数量 (Parameters) 作为评估指标。相关计算公式如式 (10) 一式

(12)所示:

$$P = \frac{T_p}{T_p + F_p} \tag{10}$$

$$R = \frac{T_p}{T_p + F_N} \tag{11}$$

$$mAP = \frac{1}{n} \sum_{i=1}^n \int_0^1 P(R) d(R) \tag{12}$$

其中, T_p 为模型正确预测的正样本数量, F_p 为模型错误预测的负样本数量, F_N 为模型错误预测为负样本,但实

际为正样本的数量; n 为缺陷类型的总数; 准确率 P 评估模型预测正类别的准确性; 召回率 R 衡量模型对正类别样本的覆盖能力; 平均精确率 mAP 为各类别精确率的平均值; 模型参数量为可训练参数的总数, 影响模型的复杂度; 计算量为模型每秒的浮点运算次数, 反映运行速度和性能。

2.4 消融实验

为了进一步验证本文提出的改进模块对模型性能的影响, 共设计了 8 组消融实验, 并对每组实验数据进行统计处理与分析。实验结果如表 2 所示。

Table 2 Ablation experiment results

表 2 消融实验结果

模型	ADown-Conv	C2PSA-TA	InnerLMPDIoU	P/%	R/%	mAP50/%	Params/M	GFLOPs
YOLOv11n	-	-	-	83.7	77.2	82.9	2.58	6.3
A	√	-	-	86.1	77.9	84.4	2.1	5.1
B	-	√	-	86.9	77.6	84.3	2.53	6.3
C	-	-	√	84.7	78.4	84.1	2.58	6.3
D	√	√	-	87.8	77.8	84.8	2.05	5.1
E	√	-	√	85.7	78.3	84.5	2.1	5.1
F	-	√	√	86.2	76.9	83.4	2.53	6.3
G	√	√	√	87.3	79.2	85.4	2.05	5.1

以 YOLOv11n 作为基准实验, 模型 A 加入 ADown-Conv 模块, 模型 B 加入 C2PSA-TA 模块, 模型 C 加入 InnerLMPDIoU 损失函数。实验结果表明, 模型 A 既提高了对小目标缺陷的检测精度, 又降低了参数量和计算量, P 、 mAP 分别提高了 2.4%、1.5%, 参数量与计算量分别降低了 0.48 M、1.2 GFLOPs; 模型 B 使网络可以更全面地理解缺陷信息, mAP 提高了 1.4%, 尤其是 P 提高了 3.2%; 模型 C 优化了边界框回归, 在不改变参数和计算量的情况下提升了性能; 模型 D 融合了 ADown-Conv 模块和 C2PSA-TA 模块, P 、 mAP 分别提高了 4.1%、1.9%, 参数量与计算量分别降低了 0.53 M 和 1.2 GFLOPs; 模型 E 和 F 分别融合 ADown-Conv 与 InnerLMPDIoU 损失函数、C2PSA-TA 模块与 InnerLMPDIoU 损失函数, 实验结果均验证了改进模块的有效性。G 模型即为本文提出的 ATI-YOLO 模型, P 、 R 、 mAP 分别提高了 3.6%、2%、2.5%, 参数量与计算量分别降低了 0.53 M 和 1.2 GFLOPs, 综合对比效果最好, 更适用于铝合金焊缝表面缺陷检测。消融实验的直观展示如图 8 所示。

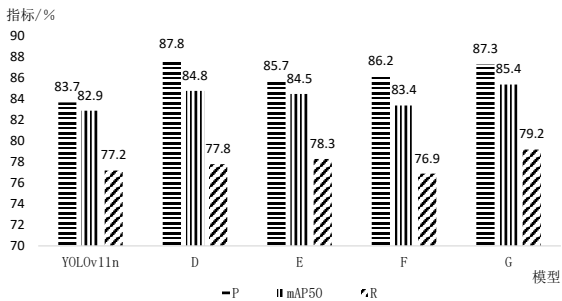


Fig. 8 Bar chart of evaluation metrics for ablation experiments

图 8 消融实验的评价指标柱状图

2.5 对比实验

本节通过对比不同的目标检测模型, 验证本文改进方法的有效性。对比范围涵盖了双阶段目标检测算法 Faster R-CNN^[15]、SSD^[16], 单阶段目标检测算法 YOLOv11n、YOLOv10n、YOLOv9-t、YOLOv8n、YOLOv7-tiny、YOLOv5n, 以及同领域研究者改进的 YOLOv11 模型 YOLOv11_ECA^[17]。此外, 还包括一些知名企业和科研团队设计的新模型, 例如百度提出的 RT-DETR-l^[18] 和华为研发的 Gold YOLO^[19]。其中, RT-DETR-l 是一种基于 DETR 架构的实时端到端检测器, 而 Gold YOLO 通过 GD 机制增强了信息融合能力, 提高了多目标和小目标的检测性能^[20]。本文设计了 12 组对比实验, 具体结果如表 3 所示。

实验结果显示, Faster R-CNN、SSD、YOLOv9-t、RT-DETR-l 和 Gold YOLO-n 等模型因计算量与参数量较大,

Table 3 Comparative experiment results

表 3 对比实验结果

模型	P/%	R/%	mAP50/%	Params/M	GFLOPs
Faster R-CNN ^[15]	81.9	79.4	78.3	7.21	22.5
SSD ^[16]	76.7	73.3	75.4	5.63	16.4
YOLOv11n	83.7	77.2	82.9	2.58	6.3
YOLOv10n	82.2	74.6	81.7	2.69	8.2
YOLOv9-t	85.4	75	82.1	10.73	45.8
YOLOv8n	84.9	76.7	82.9	2.68	6.8
YOLOv7-tiny	84.1	74.8	81.8	2.25	6.1
YOLOv5n	83.7	74.6	81.1	2.18	5.8
YOLOv11_ECA ^[17]	84.8	78.6	84.2	2.59	6.3
RT-DETR-l ^[18]	82.5	75.3	80.3	32.81	108
Gold YOLO-n ^[19]	86.1	78.4	83.2	19.76	59.4
ATI-YOLO	87.3	79.2	85.4	2.05	5.1

不适合对性能要求高的应用场景。尽管 YOLOv11_ECA 相比 YOLOv11n 在 mAP50、精确率(P)和召回率(R)上有所提升,但性能仍不如 ATI-YOLO 模型,且结构更为复杂。ATI-YOLO 相比 YOLOv11n、YOLOv10n、YOLOv8n、YOLOv7-tiny、YOLOv5n 模型, mAP50 获得了提升,并显著降低了计算量和参数量,证明了其在轻量化与性能方面的双重优势,因此具有广泛的应用前景。

2.6 不同类型缺陷精度对比

上文整体介绍了 ATI-YOLO 模型的优越性,本节则重点分析不同类型缺陷的检测效果,统计并对比了 YOLOv11n 和 ATI-YOLO 模型的检测精度,如表 4 所示。从表中可以看出,每种类型缺陷的 mAP50 与 mAP50-95 都有不同程度的提升,证明了改进后的模型能够检测到更多小目标缺陷,表明 ATI-YOLO 模型对于小目标缺陷具有更高的敏感性。

Table 4 Accuracy comparison for different types of defects

表 4 不同类型缺陷精度对比 (%)

缺陷	YOLOv11n		ATI-YOLO	
	mAP50	mAP50-95	mAP50	mAP50-95
Crack	91.7	69.3	93.1	73.3
Porosity	71.3	40.7	75.1	44.6
Spatters	71	41	75.2	45.7
Welding line	97.8	89.8	98.3	90.8
平均值	82.9	60.2	85.4	63.6

2.7 检测结果及热力图可视化分析

本节通过直观展示模型的预测能力,对比了 ATI-YOLO 模型与几种表现较好的模型。对比模型包括基准模型 YOLOv11n、同领域研究人员基于 YOLOv11 改进的 YOLOv11_ECA 模型、华为公司设计的新模型 Gold YOLO-n,以及当前广泛使用的 YOLOv8n。为了确保预测结果的全面性,选取了 4 组图片,每组图片分别对应不同类型的缺陷,如图 9 所示。

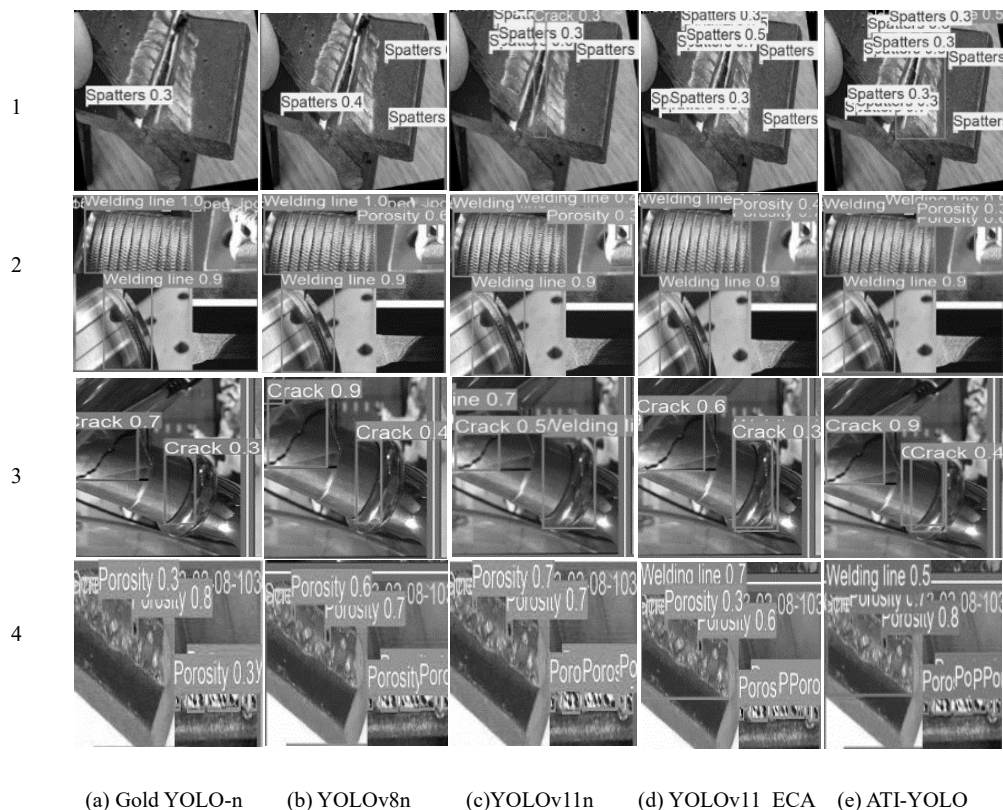


Fig. 9 The prediction performance of different models on various types of defects

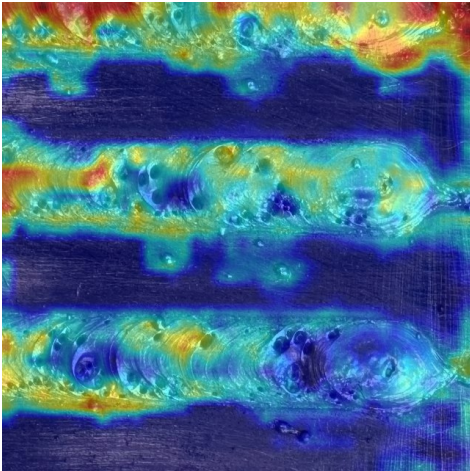
图 9 各模型对不同类型缺陷的预测情况

从图中可以看出,在第 1 组图片的预测中, YOLOv11_ECA 模型和 ATI-YOLO 模型相比其他模型可检测到更多的 Spatters 缺陷,而仅有 ATI-YOLO 模型额外准确识别了 Welding line 缺陷;在第 2 组图片的预测中,只有 ATI-YOLO 模型成功在 Porosity 缺陷周围识别出 Welding line 缺陷;在第 3 组图片的预测中,ATI-YOLO 模型没有将 Crack 缺陷误判为 Welding line 缺陷,且能准确检测出图片中的所有 Crack 缺陷;在第 4 组图片的预测中, YOLOv11_ECA 模型和 ATI-YOLO 模型均能准确识别出图片

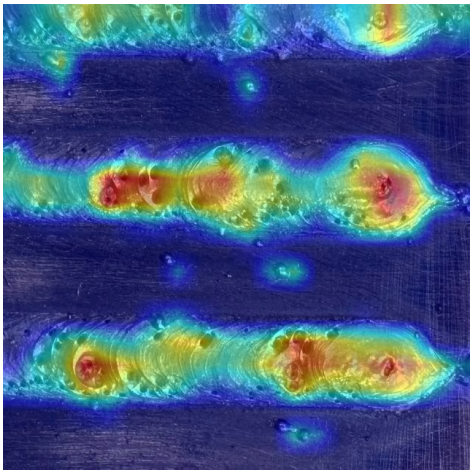
中的所有缺陷,但 ATI-YOLO 模型的置信度更高。从预测结果可以看出,本文提出的 ATI-YOLO 模型相比其他同类模型具有更高的准确性,并在小目标缺陷检测中表现出色。

同时,本节展示了在特定场景下表现优异的 C2PSA-TA 注意力机制模块与 YOLOv11n 的热力图可视化对比结果,如图 10 所示。梯度加权类激活映射(Grad-CAM)是一

种可视化方法,通过生成热力图帮助理解深度卷积神经网络(CNN)在分类决策中关注的图像区域,突出显示对决策最重要的部分。



(a) YOLOv11n



(b) C2PSA-TA

Fig. 10 Heatmap visualization results

图 10 热力图可视化结果

在目标检测结果分析中,通过热力图可直观地显示图像各部分对预测结果的贡献,红色越深表示模型对该区域的关注度越高。图 10(a)表明,YOLOv11n 模型不仅关注气孔缺陷所在区域,而且对飞溅缺陷有所响应,说明该模型可能对无关区域产生过度关注,增加了漏检或误检风险。图 10(b)显示,C2PSA-TA 模型能更精准地聚焦于气孔缺陷区域,说明该注意力机制增强了模型对缺陷区域的关注度,减少了注意力分散现象,从而提升了检测的准确率。

2.8 泛化实验

通过以上实验分析,充分验证了 ATI-YOLO 模型在铝合金焊缝缺陷检测数据集上的出色性能。为进一步研究该模型的泛化能力,本文选取了两个相近领域的开源数据集进行实验。数据集名称及下载地址如下:①NEU-DET 钢材表面缺陷数据集(<http://faculty.neu.edu.cn/songkechen/>

[zh_CN/zhym/263269/list/index.htm](http://faculty.neu.edu.cn/songkechen/));②GC10-DET 金属表面缺陷检测数据集(<https://github.com/lvxiaoming2019/GC10-DET-Metallic-Surface-Defect-Datasets>)。

通过对比 YOLOv11n 和 ATI-YOLO 模型在这两个相似领域数据集上的表现,结果显示 ATI-YOLO 模型在 P(精确率)、R(召回率)和 mAP50 等评价指标上均优于 YOLOv11n,参数量和计算量也有所减少,充分证明了本文模型具有优异的泛化能力。具体实验结果见表 5。

Table 5 Generalization experiment results

表 5 泛化实验结果

数据集	模型	P/%	R/%	mAP50/%	Params/M	GfLOPs
数据集 1	YOLOv11n	73.5	73.5	79.6	2.58	6.3
	ATI-YOLO	76.4	74.7	81.9	2.05	5.1
数据集 2	YOLOv11n	63.2	60.7	68.5	2.59	6.3
	ATI-YOLO	66.3	65.4	71.3	2.07	5.1

3 结语

本文使用一种基于 YOLOv11 的改进模型 ATI-YOLO,旨在提高对小目标铝合金焊缝表面缺陷的检测效果。该模型首先将 YOLOv11 的下采样卷积替换为 ADown-Conv 模块,在降低特征维度的同时保留关键信息,可实现模型轻量化并提升检测精度。其次,设计 C2PSA-TA 模块,采用三支结构捕捉跨维度关系,并结合旋转操作与残差变换,增强模型的多维度感知能力,提高复杂场景下的小目标检测精度。最后,将 LMPDIoU 损失函数与 Inner 思想相结合设计 InnerLMPDIoU 损失函数,优化边界框回归,提升 YOLOv11 模型的性能。实验结果表明,本文改进模型相比原模型,整体平均检测精度(mAP)提高了 2.5%。下一步研究将重点关注如何在多种真实场景下部署模型,并持续优化网络结构,以进一步提高测试效率。

参考文献:

[1] FAN D, HU A D, HUANG J K, et al. X-ray image defect recognition method for pipe weld based on improved convolutional neural network[J]. Transactions of the China Welding Institution, 2020, 41(1): 7-11.
樊丁,胡按得,黄健康,等.基于改进卷积神经网络的管焊缝X射线图像缺陷识别方法[J].焊接学报,2020,41(1):7-11.

[2] VIOLA P, JONES M J. Robust real-time face detection[J]. International Journal of Computer Vision, 2004, 57: 137-154.

[3] HU D, LYU B, WANG J J, et al. Study on HOG-SVM detection method of weld surface defects using laser visual sensing[J]. Transactions of the China Welding Institution, 2023, 44(1): 57-62, 70.
胡丹,吕波,王静静,等.焊缝表面缺陷激光视觉传感 HOG-SVM 的检测方法[J].焊接学报,2023,44(1):57-62,70.

[4] GIRSHICK R, DONAHUE J, DARRELL T, et al. Rich feature hierarchies for accurate object detection and semantic segmentation [C]//2014 IEEE Conference on Computer Vision and Pattern Recognition, 2014: 580-587.

[5] REDMON J, DIVVALA S, GIRSHICK R, et al. You look only once: uni-

- fied real-time object detection [C]//2016 IEEE Conference on Computer Vision and Pattern Recognition, 2016: 779–788.
- [6] CHENG S, YANG H G, XU X Q, et al. Improved lightweight X-ray aluminum alloy weld defects detection algorithm based on YOLOv5[J]. Chinese Journal of Lasers, 2022, 49(21): 136–144.
- 程松, 杨洪刚, 徐学谦, 等. 基于 YOLOv5 的改进轻量型 X 射线铝合金焊缝缺陷检测算法[J]. 中国激光, 2022, 49(21): 136–144.
- [7] LI Y D, HE Z M. A steel surface defect detection algorithm based on improved YOLOv8[J]. Software Guide, 2024, 23(11): 200–205.
- 黎源东, 贺智明. 一种改进 YOLOv8 的钢材表面缺陷检测算法[J]. 软件导刊, 2024, 23(11): 200–205.
- [8] CUI Y, YU Z K, HAN J C, et al. Dual-triple attention network for hyperspectral image classification using limited training samples[J]. IEEE Geoscience and Remote Sensing Letters, 2021, 19: 1–5.
- [9] MA S, XU Y. Mpdou: a loss for efficient and accurate bounding box regression[DB/OL]. <https://arxiv.org/pdf/2307.07662>.
- [10] KHANAM R, HUSSAIN M. YOLOv11: an overview of the key architectural enhancements[DB/OL]. <https://www.arxiv.org/pdf/2410.17725>.
- [11] REZATOFIGHI H, TSOI N, GWAK J Y, et al. Generalized intersection over union: a metric and a loss for bounding box regression[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2019: 658–666.
- [12] ZHENG Z H, WANG P, LIU W, et al. Distance-IoU loss: faster and better learning for bounding box regression[C]//Proceedings of the AAAI Conference on Artificial Intelligence, 2020: 12993–13000.
- [13] ZHANG Y F, REN W Q, ZHANG Z, et al. Focal and efficient IOU loss for accurate bounding box regression[J]. Neurocomputing, 2022, 506: 146–157.
- [14] ZHANG H, XU C, ZHANG S J. Inner-IoU: more effective intersection over union loss with auxiliary bounding box[DB/OL]. <https://arxiv.org/pdf/2311.02877>.
- [15] REN S Q, HE K M, 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, 2016, 39(6): 1137–1149.
- [16] LIU W, ANGUELOV D, ERHAN D, et al. SSD: single shot multibox detector[C]//European Conference on Computer Vision, 2016: 21–37.
- [17] LI Y D, YAN H, LI D, et al. Robust miner detection in challenging underground environments: an improved YOLOv11 approach[J]. Applied Sciences, 2024, 14(24): 11700.
- [18] ZHAO Y, LYU W Y, XU S L, et al. Detsr beat yolos on real-time object detection[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2024: 16965–16974.
- [19] WANG C C, HE W, NIE Y, et al. Gold-YOLO: efficient object detector via gather-and-distribute mechanism[C]//Proceedings of the 37th International Conference on Neural Information Processing Systems, 2023: 51094–51112.
- [20] CARION N, MASSA F, SYNNAEVE G, et al. End-to-end object detection with transformers [C]//European Conference on Computer Vision, 2020: 213–229.

(责任编辑:黄健)