

基于YOLOv8-R50CBAM的枪管内壁缺陷检测研究

姜美竹, 李耀明*, 张 煌

(中北大学 机械工程学院, 山西 太原 030051)

摘 要: 枪管内壁缺陷检测中存在强纹理干扰、弱对比缺陷与复杂噪声的问题, 而现有方法在检测精确度等方面还存在不足, 因此本文提出了一种改进的YOLOv8-R50CBAM检测方法。通过对比ResNet50、MobileNetV1与VGG16, 确定以ResNet50作为主干网络。模型核心包括: 在主干网络嵌入CBAM注意力机制, 构建了前端特征提取模块CRB, 用来增强缺陷特征响应并抑制背景干扰; 在Neck部分引入基于PANet的双向路径聚合模块BPAF, 强化了高层语义与底层空间信息交互; 在推理阶段集成Soft-NMS、测试时增强TTA及置信度温度缩放校准, 形成了检测推理增强模块ODI, 用来提升密集缺陷召回率及置信度可靠性。实验结果表明, 所提模型在枪管内表面缺陷数据集上的精确率、F1-score和mAP分别达到89.3%、0.858和85.0%, mAP较基准模型提升了19.53个百分点。消融实验表明, CRB与BPAF模块可显著改善细小缺陷的检出与定位, ODI模块增强了模型在复杂样本上及输出置信度的可靠性。本文构建了一个融合注意力机制、双向多尺度融合与系统推理优化的检测模型框架, 为枪管内壁的自动化检测提供了参考, 其不足之处在于模型复杂度增加, 轻量化部署还需进一步研究。

关键词: 枪管缺陷检测; 改进YOLO; 注意力机制; Soft-NMS; 测试时增强

中图分类号: TJ303.4

文献标识码: A

doi: 10.62756/jnuc.issn.1673-3193.2025.12.0002

引用格式: 姜美竹, 李耀明, 张煌. 基于YOLOv8-R50CBAM的枪管内壁缺陷检测研究[J]. 中北大学学报(自然科学版), DOI:10.62756/jnuc.issn.1673-3193.2025.12.0002.

Jiang Meizhu, Li Yaoming, Zhang Huang. Research on Gun Barrel Inner Surface Defect Detection Based on YOLOv8-R50CBAM[J]. Journal of North University of China(Natural Science Edition), DOI:10.62756/jnuc.issn.1673-3193.2025.12.0002.

Research on Gun Barrel Inner Surface Defect Detection Based on YOLOv8-R50CBAM

Jiang Meizhu, Li Yaoming*, Zhang Huang

(School of Mechanical Engineering, North University of China, Taiyuan 030051, China)

Abstract: In the detection of defects on the inner surface of gun barrels, challenges such as strong textural interference, low-contrast defects, and complex noise exist, and existing methods still have shortcomings in terms of detection accuracy and other aspects. This paper proposed an improved YOLOv8-R50CBAM detection method. After comparing ResNet50, MobileNetV1, and VGG16, ResNet50 was selected as the backbone network. The core components of this model include: Embedding the CBAM attention mechanism into the backbone network to construct the front-end feature extraction module CRB, which enhanced defect feature

收稿日期: 2025-12-01

基金项目: 国家自然科学基金资助项目(52305516)

作者简介: 姜美竹(2001—), 女, 硕士生, 主要从事管孔类智能加工与检测的研究。

* 通信作者: 李耀明(1977—), 男, 教授, 博士, 主要从事先进制造、金属复合材料设计与制造等研究。E-mail: lym@nuc.edu.cn。

responses and suppressed background interference; Introducing a PANet-based bidirectional path aggregation module BPAF in the neck section to strengthen the interaction between high-level semantic and low-level spatial information; Integrating Soft-NMS during the inference stage, and enhancing TTA and confidence temperature scaling calibration during testing to form the detection inference enhancement module ODI, which was used to improve the recall rate of dense defects and the reliability of confidence scores. Experimental results show that the proposed model achieves a precision of 89.3%, an $F1$ -score of 0.858, and an mAP of 85.0% on the gun barrel inner surface defect dataset, with an mAP improvement of 19.53 percentage point over the baseline model. Ablation experiments demonstrate that the CRB and BPAF modules significantly improve the detection and localization of fine defects, while the ODI module enhances the model's detection capabilities on complex samples and the reliability of output confidence scores. This paper constructs a detection model framework that integrates attention mechanisms, bidirectional multi-scale fusion, and system inference optimization, providing a reference for the automated inspection of gun barrel inner surface. A limitation of this approach is the increased model complexity, and further research is needed to achieve lightweight deployment.

Key words: gun barrel defect detection; improved YOLO; convolutional block attention module; soft non-maximum suppression; test-time augmentation

0 引 言

枪支的性能与安全在很大程度上取决于其关键部件——枪管的质量。制造与使用过程中产生的内部缺陷(如热裂纹、锈蚀、点蚀)是影响枪支使用寿命的重大隐患^[1]。当前,依赖人工内窥镜的检测方法存在主观性强、效率低且一致性差等固有弊端,而传统机器视觉方法又难以适应复杂多变的缺陷形态。因此,发展针对枪管内壁缺陷的高性能、智能化的自动检测技术,对于提升军工检测的自主化水平、保障装备质量安全具有重要现实意义。近年来,随着机器视觉和深度学习的发展^[2-3],传统军工正在向智能化和现代化转型,许多学者都尝试将其应用于枪管内壁缺陷检测方面,以弥补传统检测的局限性。在技术演进方面,早期的研究主要集中于利用传统机器视觉方法。例如,邓兆鹏^[4]开发了一套基于机器视觉全景钻孔摄像系统,提出了一种新的两级树式图像分类方法,为深孔部件的自动化检测奠定了工程应用基础。然而,这类方法对于复杂背景下形态多变的缺陷泛化能力有限。近年来,以卷积神经网络为代表的深度学习方法展现出显著优势^[5]。对于两阶段检测算法, Ren 等^[6]提出了 Faster R-CNN 框架,通过引入区域提议网络实现了端到端的精准目标检测,其核心思想为后续众多缺陷检测研究提供了理论基础。对于 YOLO 单阶段检测

算法,因其高效率而在实时性要求较高的场合应用更广。Wang 等^[7]提出的 YOLOv7 模型,通过引入一系列可训练的“免费赠品”优化策略,在保持实时性的同时进一步提升了检测精度,代表了当前实时检测器的最新水平。学者们针对实际应用中的挑战进行了诸多优化。韩强^[8]针对小目标缺陷分辨率低和易重叠等一系列问题,对 YOLOv8 模型加以改进,通过增引入变形卷积模块和注意力机制有效提升了模型在复杂环境中缺陷识别的准确率与鲁棒性,该方法对枪管内壁检测具有很大的参考价值。

综上所述,尽管深度学习在通用目标检测领域已取得显著成果,但将其应用于枪管内壁这一特殊场景仍面临小样本、细微缺陷检测以及易受环境干扰导致误检漏检等挑战。本文的核心工作在于面向枪管内壁缺陷检测这一特定高难度场景,对 YOLOv8 模型进行一系列深度、有针对性的结构改进与策略集成,以构建一套专有的检测模型。具体而言,本文提出的 YOLOv8-R50CBAM 模型,在特征提取阶段,将 CBAM 嵌入残差瓶颈,构建了注意力增强的残差模块,以主动抑制背景纹理并增强缺陷特征响应;在特征融合阶段,设计了一种双向路径聚合结构,强化高层语义向低层的传递与低层定位信息向高层的回流,以优化多尺度特别是小缺陷的特征表达;在推理阶段,利用集成 Soft-NMS、测试时增强与置信度校准等多策略后处理机制,系统性地提升了复杂环境下的召回率与结果可靠性。

1 YOLOv8n模型概述

YOLOv8n是Ultralytics公司开发的YOLOv8系列中最轻量级的模型,专用于边缘计算和实时应用场景优化。作为YOLOv8系列中最轻量级的模型,通过引入全新的骨干网络架构与高效的C2f模块,在保持较低计算复杂度的同时显著提升了特征提取能力。该模型采用无锚框(Anchor-Free)检测机制与解耦预测头(Decoupled Head)设计,有效简化了检测流程并优化了分类与定位任务的分工协作,面向枪管内壁缺陷检测这一特定工业场景,对模型部署的实时性、计算资源消耗及硬件兼容性有着严格要求^[9]。YOLOv8n作为轻量级架构,在保持足够检测精度的前提下,其约 3.2×10^6 的参数量与6MB的模型体积显著优于后续版本,更适合部署于算力受限

的工业检测设备;同时,该模型成熟的ONNX/TensorRT支持体系降低了实际部署难度,而较新的YOLO系列在模型复杂度与资源需求上的提升并未带来与本任务特征相匹配的精度增益,且其边缘计算兼容性尚未得到充分验证。因此,从工程应用的成本效益比出发,选取了YOLOv8n作为本研究的基线模型(baseline)^[10]。

2 枪管内壁缺陷检测方法

2.1 YOLOv8-R50CBAM模型

YOLOv8n网络结构主要由四部分构成:输入端(input)、骨干网络(backbone)、颈部网络(neck)和检测输出端(head)^[11]。该研究模型改进主要集中于骨干网络和颈部网络^[12]。YOLOv8-R50CBAM模型结构图如图1所示。

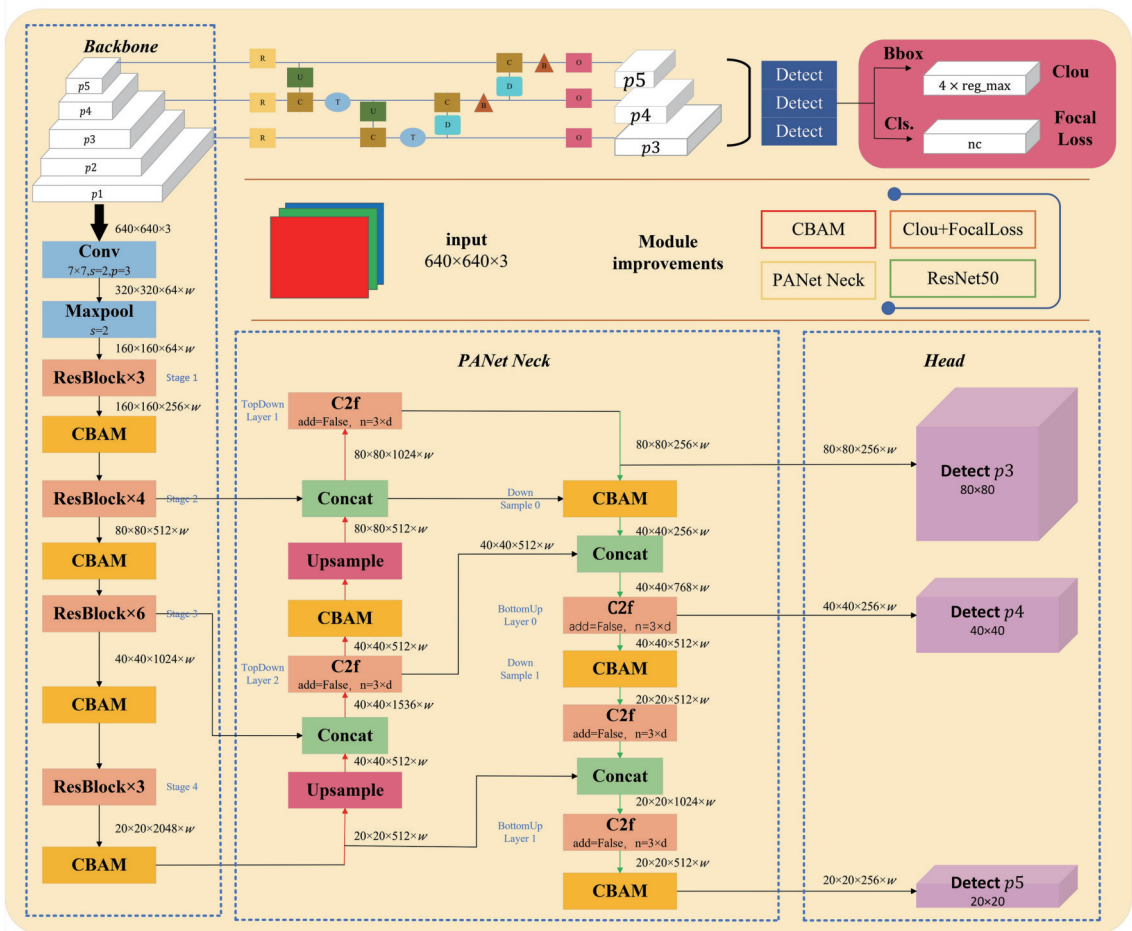


图1 YOLOv8-R50CBAM模型

Fig. 1 YOLOv8-R50CBAM Model

该模型构建了一个协同增强的检测框架。在特征提取层面,以ResNet50为骨架嵌入CBAM注意力模块,通过通道与空间重标定主动抑制背景

噪声并强化缺陷特征,从而输出多层次特征。在特征融合层面,基于PANet构建双向路径聚合模块(BPAF),实现了高层语义下传与低层定位信

息回流,为检测头提供兼具判别语境与几何细节的多尺度特征。在检测推理层面,优化多尺度检测与推理增强模块(ODI),采用Soft-NMS缓解密集缺陷误抑制,利用测试时增强提升输出稳定性,并通过置信度校准使预测分数更可靠。

2.2 主干网络替换

在单阶段目标检测器(如YOLO系列)中,主干网络是模型的核心组成部分,负责从输入图像中提取多层次、高语义信息的特征图。其设计直接决定了模型的特征表示能力、计算效率以及最终的检测性能。本研究选取了3种具有代表性的卷积神经网络VGG16、MobileNetV1和ResNet50作为YOLOv8n的主干网络方案。

VGG16^[13]由牛津大学Visual Geometry Group提出,是深度学习早期最具代表性的架构之一。其核心思想在于通过堆叠多个小型 3×3 卷积核来替代大型卷积核(如 5×5 、 7×7),在保持相同感受野的同时,显著减少了参数量并增加了非线性。VGG16结构规整、层级深邃(13个卷积层与3个全连接层),体现了通过增加网络深度来提升特征学习能力的经典思路。然而,其全连接层参数量巨大,且整体计算成本较高。为将其作为YOLOv8的主干,本研究移除了其原始的全连接层,选取了网络中部(如conv4_3)和后部(如conv5_3和最后的池化层之后)的输出,作为多尺度特征图的来源。由于VGG16的特征图通道数(如512)与YOLOv8颈部网络的预期输入可能存在差异,在特征图输出后添加了额外的 1×1 卷积层进行通道数调整,以匹配后续PAN-FPN结构所需的输入维度。

MobileNetV1^[14]由Google研究团队提出,是专门为移动端和嵌入式设备设计的高效架构,核心创新是“深度可分离卷积”,其将标准卷积分解为深度卷积和逐点卷积两个步骤,从而大幅减少了计算量和参数量。MobileNetV1的设计思想是在尽可能小的性能损失下,追求最佳速度与能效比,为实时目标检测在边缘计算场景中的应用奠定了基础。在将其集成至模型中时,主要是利用其深度卷积层和逐点卷积层构建特征金字塔的基础。考虑到YOLO检测头需要3个不同尺度的特征图,本研究从MobileNetV1的多个深度可分离卷积模块之后抽取了3层不同下采样率的特征输出。为了弥补深度可分离卷积在特征表示能力上可能的不足,并在不同

特征层间融合信息,在抽取的特征层之后引入了轻量化的特征融合模块,以增强语义信息。

ResNet50由微软研究院提出,其革命性贡献在于引入了“残差学习”机制。通过捷径连接实现恒等映射,ResNet有效缓解了深度神经网络中普遍存在的梯度消失和网络退化问题,使得构建数百甚至上千层的超深网络成为可能。ResNet50^[15]作为该系列的经典代表,通过残差块堆叠,极大地增强了特征的复用与传播能力,在精度和深度之间取得了卓越的平衡。在集成过程中,将ResNet50的多个残差阶段(stage)的输出直接对应为YOLOv8颈部网络所需的输入。ResNet50的res5c输出下采样率为32,但通道数高达2048,远超基线模型。为此,本研究设计了一个过渡模块,该模块包含一个 1×1 卷积,将通道数降至512,后接一个步长为1的 3×3 卷积以保持空间信息。同时,由于res4f的输出下采样率为16,与目标尺度略有偏差,在其适配路径中加入了最近邻上采样操作,以确保特征图空间尺寸的精确对齐。这种设计充分利用了ResNet的深度特征,并通过精确的通道与尺度适配,将其无缝集成到检测框架中。

本研究主干网络设计方案覆盖了经典深度模型、轻量化架构和残差学习这3种不同设计范式,可全面评估主干网络对缺陷检测性能的影响。

2.3 颈部网络的适应性优化与增强

为有效融合改进后主干网络所提取的多尺度、高维特征,并充分服务于后续的缺陷检测头,本研究对原YOLOv8n的颈部网络(Neck)进行了针对性的适应性优化与结构增强^[16]。

在特征融合层面,引入双向路径聚合的多尺度特征融合模块BPAF,以PANet的双向信息流为基础,将高层语义向低层传递与低层定位信息向高层回流同时纳入网络结构。高层语义下传使浅层细节获得更清晰的判别语境,能够更有效服务于小裂纹、小点蚀等微弱目标的发现。低层定位信息回流则补足深层特征中容易被稀释的几何线索,提高边界框回归的稳定性,减轻密集纹理与反光条件下定位漂移的风险。多尺度融合输出与检测尺度形成明确的对应关系,使检测头能够在不同感受野上分工协作,从结构层面缓解同一尺度同时承担语义与细节冲突的问题。

在检测与推理层面,提出多尺度检测与推理增强模块ODI,保持YOLO多尺度检测头的基本范式不变,重点优化解码、后处理与推理增强策略,使输

出更符合缺陷检测的工程需求。推理流程从解码到阈值过滤再到抑制合并,形成了清晰链路,Soft-NMS以分数衰减替代硬删除,缓解了密集候选框环境下的小缺陷误抑制问题,降低了漏检的结构性概率。测试时增强通过多尺度推理与可选的水平翻转,引入受控的视角扰动与尺度扰动,使同一缺陷在不同观察条件下获得更稳定的一致性输出。置信度温度缩放校准进一步修正模型分数的过度自信倾向,减少阈值对结果的敏感性,使输出分数更接近可比较的概率尺度,为实验中的阈值选择、指标对比与工程参数调优提供更稳定的基础。

2.4 损失函数优化

2.4.1 改进损失总体形式

本研究在训练脚本中以Improved-YOLOLoss替换原YOLOLoss,核心思路保持了YOLOLoss原有的总体组织方式,仍以多尺度输出逐层累计的方式计算总损失。改动集中在两处,一处是边界框回归度量的替换,另一处是分类分支的重加权。实现层面采用继承机制,ImprovedYOLOLoss直接继承YOLOLoss,并在初始化阶段额外引入focal_loss、focal_alpha、focal_gamma等控制参数,使分类损失能够在保持接口一致的前提下实现可控切换。训练循环中对每个检测尺度 l 分别调用yolo_loss(l , preds, targets),再将各尺度损失累加,这一调用方式与forward签名严格对齐,保证损失模块在多尺度结构下具有确定的可复现性。

改进后的总体形式仍可写为多项加权求和的结构,只是权重含义更偏向工程可调。回归项采用基于CIoU的几何一致性度量,分类项采用FocalLoss进行难例聚焦,置信度项沿用YOLOLoss的原有框架与计算路径,从而避免在同一版本中同时改动过多项而导致误差来源不可追踪。将这种设计用于枪管缺陷任务的原因是,缺陷类别之间的形态相近,且类别频次差异较大,分类梯度容易被易样本主导,回归梯度又容易被对齐误差牵引。把改动集中在bbox_iou和cls_loss两处,能够在不破坏主训练框架稳定性的前提下,针对性压缩两类最常见的误差链条。

2.4.2 边界框回归损失

在本研究的实现中,bbox_iou函数以中心点与宽高形式表示框,先计算交并比IoU,再计算最小外接框对角线尺度上的中心距离惩罚,最后引入宽高比一致性项 v 与自适应权重 α 。中心距离以欧氏距离

平方表达,外接框对角线用于归一化,使距离惩罚具备尺度不变性,宽高比一致性采用反正切差的平方形式。CIoU损失函数的定义为

$$\mathcal{L}_{\text{CIoU}} = 1 - \text{IoU} + \frac{\rho^2(b, b^{gt})}{c^2} + \alpha v, \quad (1)$$

式中: IoU 为预测框与真实框的交并比; $\rho(b, b^{gt})$ 为预测框中心点 b 与真实框中心点 b^{gt} 之间的欧氏距离; c 为覆盖预测框和真实框的最小外接矩形的对角线长度; v 为宽高比一致性度量。

$$v = \frac{4}{\pi^2} \left(\arctan \frac{\omega^{gt}}{h^{gt}} - \arctan \frac{\omega}{h} \right)^2, \quad (2)$$

$$\alpha = \frac{v}{(1 - \text{IoU}) + v}, \quad (3)$$

式中: ω 和 h 分别为预测框的宽度和高度; ω^{gt} 和 h^{gt} 分别为真实框的宽度和高度。

在训练框架的组织上,bbox_iou提供的是融合中心距离与宽高比约束的回归度量,回归损失在YOLOLoss原有路径中以一减形式参与优化,使得几何误差能够被显式惩罚。这样的替换对枪管缺陷任务更有针对性,原因很直接,小缺陷框的IoU对像素级偏移极其敏感,单纯依赖IoU的面积反馈容易出现梯度稀疏,加入中心距离项后,即使两框重叠较小,优化仍能获得明确的方向性信号。条状裂纹对长宽比的约束更强,宽高比惩罚项能够在回归框被背景纹理牵引时,提供额外的几何纠偏力,从而减少框形态漂移带来的定位不稳定。

2.4.3 分类损失FocalLoss

枪管内表面图像的背景纹理常呈现规则加工痕迹,网络在训练早期会快速学会排除明显背景,于是大量负样本成为低损失样本,其梯度却仍可能在数量上压制少量难例。FocalLoss的意义在于把交叉熵中的易样本权重主动压低,让训练更聚焦难例,尤其聚焦那些置信度不高却更接近真实缺陷的样本。该损失在密集检测领域被提出,并通过引入调制因子实现对难例的聚焦控制。其数学表达式为

$$\mathcal{L}_{\text{Focal}} = -\alpha_i (1 - p_i)^\gamma \log(p_i). \quad (4)$$

式中: p_i 为模型对真实类别的预测概率; α_i 为类别平衡权重; γ 为聚焦参数,用于控制简单样本权重的降低程度。

$$p_i = \begin{cases} p, & y=1 \\ 1-p, & \text{其他} \end{cases}. \quad (5)$$

$$\alpha_i = \begin{cases} \alpha, & y=1 \\ 1-\alpha, & \text{其他} \end{cases}. \quad (6)$$

在本研究中,经过网格搜索确定最优超参数为 $\alpha=0.25$, $\gamma=2.0$,此时Focal Loss可有效缓解类别不平衡问题。改进后的总体损失函数由三部分组成:

$$\mathcal{L}_{\text{total}} = \lambda_{\text{box}} \mathcal{L}_{\text{CIoU}} + \lambda_{\text{cls}} \mathcal{L}_{\text{Focal}} + \lambda_{\text{obj}} \mathcal{L}_{\text{obj}}, \quad (7)$$

式中: $\mathcal{L}_{\text{obj}}$ 为置信度损失,保持YOLOv8n的二元交叉熵损失不变。

2.5 卷积块注意力模块的集成

除了替换整个主干网络架构,在现有网络中嵌入注意力机制是另一种增强模型特征表示能力的有效策略。卷积块注意力模块(Convolutional Block Attention Module, CBAM)是一种轻量级、即插即用的前馈注意力机制,它通过自适应地细化特征图在通道和空间两个维度上的权重,使模型能够更加关注于信息丰富的特征,同时抑制不重要的特征^[17]。CBAM依次应用通道注意力模块和空间注意力模块,实现了一种高效的注意力机制。

卷积块注意力模块(CBAM)的整体前向传播过程表示为

$$F' = M_c(F) \otimes F, \quad (8)$$

$$F'' = M_s(F') \otimes F', \quad (9)$$

式中: $F \in R^{C \times H \times W}$ 为输入特征图; $M_c \in R^{C \times 1 \times 1}$ 为通道注意力权重向量; $M_s \in R^{1 \times H \times W}$ 为空间注意力权重矩阵; $\otimes$ 表示逐元素相乘。

通道注意力权重 M_c 的计算公式为

$$\begin{aligned} M_c(F) &= \sigma(\text{MLP}(\text{AvgPool}(F))) + \\ &\text{MLP}(\text{MaxPool}(F)) = \sigma(W_1(W_0(F_{\text{avg}}^c))) + \\ &W_1(W_0(F_{\text{max}}^c))), \end{aligned} \quad (10)$$

式中: σ 为Sigmoid激活函数; $W_0 \in R^{r \times c}$ 和 $W_1 \in R^{r \times c}$ 为多层感知机(MLP)的权重矩阵, r 为压缩比率。

空间注意力权重 M_s 的计算公式为

$$\begin{aligned} M_s(F') &= \\ &\sigma(f^{7 \times 7}([\text{AvgPool}(F'); \text{MaxPool}(F')])) = \\ &\sigma(f^{7 \times 7}([F_{\text{avg}}^s; F_{\text{max}}^s])), \end{aligned} \quad (11)$$

式中: $f^{7 \times 7}$ 为滤波器尺寸为 7×7 的卷积操作; $[\cdot; \cdot]$ 为沿通道维度的拼接操作。

CBAM采用串行注意力结构,先进行通道注意力,再进行空间注意力。实现层面并不追求复杂的全局建模,它更强调轻量与可插拔。通道注意力用全局平均池化与全局最大池化提取通道统

计,再经共享的两层 1×1 卷积形成压缩与恢复,最后经Sigmoid得到通道权重。代码中reduction取16,这是一种兼顾表达与开销的折中。对枪管内表面而言,这一步更像在选择对缺陷敏感的纹理基元通道,同时压制反光、阴影等易扰动通道。

得到通道增强后的特征 F' 后,空间注意力在通道维上进行平均与最大汇聚,形成两张单通道响应图,拼接后再用 7×7 卷积生成空间权重。 7×7 的选择意味着更大的局部聚合范围,它对细长裂纹的连续响应与锈蚀片状区域的连通响应更友好。实现时,先cat再conv再Sigmoid,最后与输入逐元素相乘完成空间加权。

卷积块注意力模块的网络结构如图2所示,其中,MaxPool为最大池化层;Shared MLP为共享多层感知机;s为Sigmoid激活函数;Conv为卷积层。

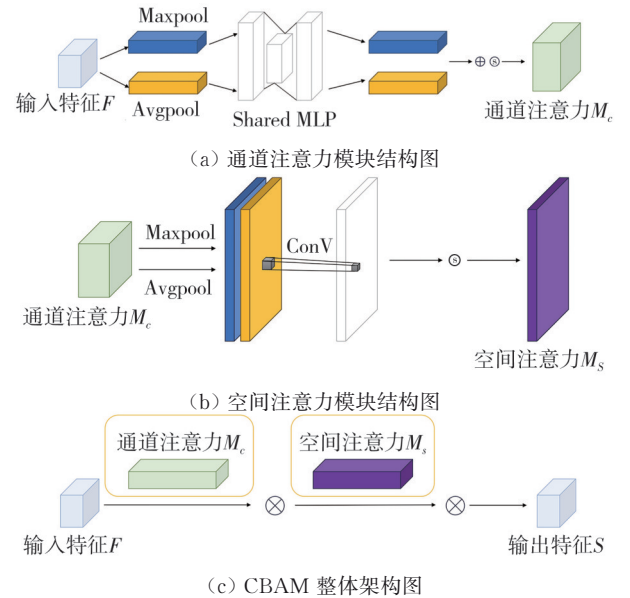


图2 卷积块注意力模块CBAM架构的示意图

Fig. 2 Architectural diagram of the CBAM

通道与空间的串行并不是形式上的堆叠,更像是一个先筛选语义通道,再定位空间区域的两步校准。对缺陷检测这种既要分清类别,又要找准位置的任务,这种先语义后定位的顺序常常更稳定。尤其在枪管内壁场景中,背景纹理与缺陷纹理高度同质,先在通道层面上抑噪,再在空间层面聚焦,会比单独进行更可靠。

3 实验结果及分析

3.1 实验环境

为了确保本次实验数据的可靠性与可重复性,所有的实验均在相同的计算环境中完成。本

次实验使用的计算机操作系统为 Windows 11, 配备 NVIDIA GeForce GTX 1650 Ti GPU 及 Intel (R) Core (TM) i5-9300H CPU 处理器, 主频为 2.40 GHz, 内存为 16G。实验环境方面, 采用 PyTorch 2.5.1 作为深度学习框架, 搭配 CUDA 12.4 和 cuDNN 9.0 进行 GPU 加速计算, 模型训练基于 Python 3.11.9 环境, 并利用 Intel MKL 2023.1.0 数学核心库优化张量运算。计算机视觉任务使用 TorchVision 0.20.1 库进行图像预处理与数据增强。

3.2 数据集的制作与预处理

3.2.1 数据集的制作

本研究所用图像数据来源于公开数据集 <https://www.kaggle.com/datasets/anumfatima/hollow-cylindrical-defect-detection-dataset>, 数据集^[18]共 1 071 张三通道图像, 其中枪管图像共 275 张, 分辨率为 $1\,600 \times 900$, 包含五类缺陷, 侵蚀(Pitting)、点蚀(Chipping)、污点(Dirt)、热裂纹(Thermal cracking)^[19]、锈蚀(Rusting)^[20], 如图 3 所示。该数据集最初为炮管和枪管内部目标检测任务而构建, 由于数据性质特殊, 是为数不多的较为可信的公开数据集。然而, 由于其标注标准与本研究任务存在显著差异, 如原数据集采用全图标注, 缺乏对微细

缺陷的标注, 存在漏标错标的情况, 图片分辨率过高, 增加了模型计算负担, 样本数量有限, 难以满足深度学习模型对大规模数据的需求, 直接使用难以满足本研究对精度和专业性的要求。

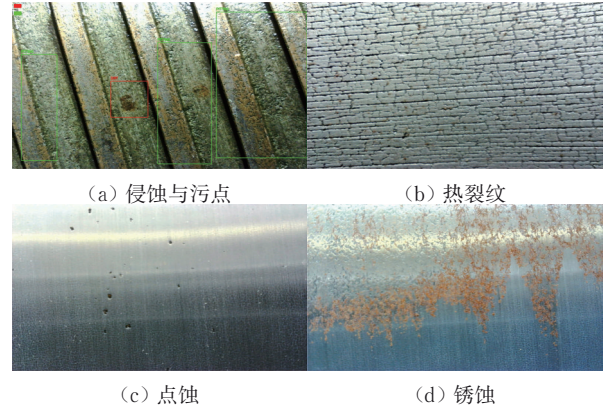


图 3 枪管内壁缺陷类型示意图

Fig. 3 Schematic diagram of gun barrel inner surface defect types

为此, 本文在此基础上自主构建了一个更加细致的枪管内壁缺陷检测数据集, 对其中枪管的 275 张图像使用 LabelImg 重新标注, 数据集采用 PASCAL VOC 格式进行标注, 即为每张图像生成一个对应的 XML 文件, 文件中包含了目标的类别信息以及基于像素的边界框坐标, XML 文件与对应的图像文件名完全相同, 每张图像至少含有一个缺陷标注。将枪管内壁缺陷数据集按 8:1:1 的比例随机划分为训练集、验证集和测试集。

表 1 枪管内壁缺陷数据集分布表

Tab. 1 Distribution of the gun barrel inner surface defect dataset

数据集及比例	锈蚀(Rusting)	热裂纹(Thermal cracking)	点蚀(Chipping)	侵蚀(Pitting)	污点(Dirt)
训练集 (80%)	1 496	1 370	1 490	1 010	4 322
验证集 (10%)	187	171	186	126	540
测试集 (10%)	187	171	186	127	541

3.2.2 预处理

由于数据集类型特殊, 有些类别的缺陷数量较少, 如污点、点蚀等, 所以采用了针对枪管内壁缺陷检测的定向数据增强方法, 图 4 展示了具体的方法路线。通过滤波和数据增强^[21]来增大数据量, 其中数据增强方法主要包括几何变换^[22]、色彩空间调整、噪声注入和环境模拟等多元化技术。具体操作中, 在几何变换方面应用了水平翻转、垂直翻转、随机缩放、位移缩放旋转等操作; 在色彩处理方面实施了随机亮度对比度调整、伽马校正、色调饱和度和值变换以及通道混淆等操作; 同时引入了运动模糊、高斯模糊、ISO 噪声、图像压缩噪声等多种噪声模拟。

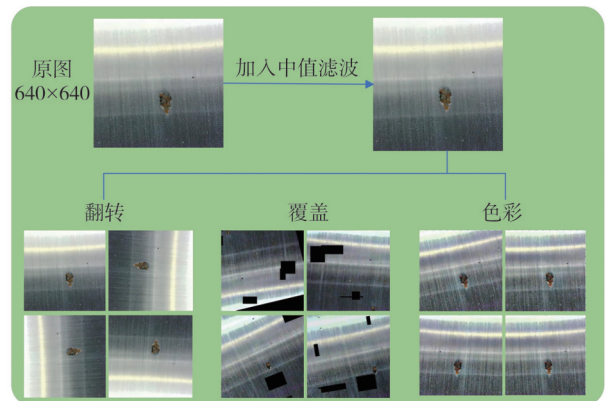


图 4 数据增强方法技术路线图

Fig. 4 Technical roadmap for data augmentation methods

该方案能够将关键缺陷类别的样本量显著提升至目标规模,如污点类别从451例扩充至1 263例,同时严格保持边界框标注的有效性,为高精度缺陷检测模型提供了高质量、强鲁棒性的训练数据基础。

3.3 参数配置

输入图片尺寸统一为 640×640 ,训练周期为500轮。在优化器设置方面,采用随机梯度下降(SGD)算法,动量因子设置为0.9,权重衰减系数为 5×10^{-4} 。学习率调度采用余弦退火策略,初始值设为0.05,最小值设为 5×10^{-4} 。针对不同的训练阶段(冻结主干网络与全网络微调),实验根据批量大小对学习率进行动态缩放,其计算公式为

$$l_r = l_{r_{\text{base}}} \times \left(\frac{\text{batch_size}}{\text{nbs}} \right), \quad (12)$$

式中: l_r 为实际使用的学习率; $l_{r_{\text{base}}}$ 为基础学习率(本文设置为 5×10^{-2}); batch_size 为当前训练的批量大小(本文设置为16); nbs 为基准批量大小(本文设为64)。

3.4 评估指标

本文选取了精确率(Precision)、召回率(Recall)、精确率和召回率的调和平均数(F1)、平均类别精度(mean Average Precision, mAP)作为核心评价指标,具体的计算公式为

定义二分类混淆矩阵的基本元素

$$\text{真正例: } TP = |\{i | \hat{y}_i = 1, y_i = 1\}|. \quad (13)$$

$$\text{假正例: } FP = |\{i | \hat{y}_i = 1, y_i = 0\}|. \quad (14)$$

$$\text{真负例: } TN = |\{i | \hat{y}_i = 0, y_i = 0\}|. \quad (15)$$

$$\text{假负例: } FN = |\{i | \hat{y}_i = 0, y_i = 1\}|. \quad (16)$$

基于上述定义,精确率和召回率的计算公式为

$$P = \frac{TP}{TP + FP} = \frac{TP}{\text{所有检测结果}}, \quad (17)$$

$$R = \frac{TP}{TP + FN} = \frac{TP}{\text{所有真实结果}}. \quad (18)$$

对于每个类别 c ,给定一组检测结果按置信度降序排列,计算精确率-召回率曲线。

$$AP(c) = \sum_{k=1}^n P_{\text{interp}}(k) \cdot \Delta r(k), \quad (19)$$

式中: $P_{\text{interp}}(k)$ 为第 k 个召回率点处的插值精确率。

$$P_{\text{interp}}(r) = \max_{\tilde{r} \geq r} P(\tilde{r}). \quad (20)$$

对于多类别目标检测, mAP的计算公式为

$$mAP = \frac{1}{C} \sum_{c=1}^C AP(c), \quad (21)$$

式中: C 为类别总数。

3.5 实验结果分析

3.5.1 主干网络替换实验

为探究不同特征提取架构对枪管内壁缺陷检测性能的影响,本研究系统性地对比了以MobileNetV1、VGG16和ResNet50为主干网络的YOLOv8n模型。表2~表4分别详细列出了3种模型在五类典型缺陷(热裂纹、点蚀、污点、侵蚀、锈蚀)上的平均精度(AP)、F1分数、召回率与精确率。ResNet50凭借其深层残差结构取得了最优异的综合性能, mAP达到65.47%,显著优于MobileNetV1(44.07%)和VGG16(62.99%)。

表2 基于MobileNetV1主干网络的YOLOv8模型检测性能
Tab. 2 Detection performance of the YOLOv8 model based on the MobileNetV1 backbone network

缺陷类别	AP	F1-score	召回率/%	精确率/%
热裂纹	36.07	0.29	17.85	76.87
点蚀	54.17	0.65	68.96	61.53
污点	46.04	0.46	32.65	75.29
侵蚀	38.20	0.40	28.61	67.02
锈蚀	45.86	0.46	34.15	71.90
mAP	44.07			

表3 基于VGG16主干网络的YOLOv8模型检测性能
Tab. 3 Detection performance of the YOLOv8 Model based on the VGG16 backbone network

缺陷类别	AP	F1-score	召回率/%	精确率/%
热裂纹	62.53	0.63	54.82	75.43
点蚀	69.85	0.76	75.36	76.00
污点	63.60	0.67	56.12	82.09
侵蚀	57.38	0.59	48.49	75.02
锈蚀	61.57	0.65	61.55	67.99
mAP	62.99			

表4 基于ResNet50主干网络的YOLOv8模型检测性能
Tab. 4 Detection performance of the YOLOv8 model based on the ResNet50 backbone network

缺陷类别	AP	F1-score	召回率/%	精确率/%
热裂纹	66.82	0.68	63.19	72.6
点蚀	70.47	0.77	78.50	75.58
污点	60.07	0.64	50.00	90.74
侵蚀	62.94	0.64	53.40	80.26
锈蚀	67.04	0.68	65.42	71.81
mAP	65.47			

由表2~表5的对比可以清晰地看出, ResNet50在全部五类缺陷检测上均保持了最高或接近最高的AP值,尤其在热裂纹、锈蚀和点蚀这三类关键缺陷上取得了超过65%的AP,展现了其强大的特征表征能力。VGG16模型性能均衡,但在处理侵蚀这类低对比度缺陷时表现不佳。轻

量级的MobileNetV1由于网络深度与表征能力的限制,在所有缺陷类别上的检测性能均显著落后于另外两种模型。

实验结果表明,ResNet50在枪管内表面缺陷检测任务中性能最优,主要原因包括:1)残差连接有效缓解了梯度消失,保障了模型收敛的稳定性,有利于复杂背景下微弱缺陷特征的提取;2)瓶颈结构(Bottleneck)平衡了模型容量与计算效率,兼顾了实时性与检测精度;3)深层网络具备较强的层次化特征表达能力,可有效提取高层语义信息,提升了对密集、不规则及低对比度缺陷(如点蚀、裂纹)的识别能力。

VGG16采用连续卷积层堆叠结构,在深度增加时易出现性能饱和与退化现象,难以从复杂背景中有效提取细微缺陷的深层非线性特征。此外,其全连接层参数量大,计算开销高,但对检测精度的增益有限。

MobileNetV1采用深度可分离卷积来实现轻量化,但牺牲了特征表示的丰富性。在处理强背景噪声与微弱缺陷信号时,其有限的特征表达能力成了性能提升的瓶颈,导致检测精度明显低于其他对比模型。

ResNet50凭借残差连接有效克服了深层网络的性能退化问题,其特征提取能力显著优于VGG16的单纯堆叠结构和MobileNetV1的轻量化设计。针对枪管内表面复杂的纹理干扰与微弱缺陷特征,ResNet50在推理效率与特征表征能力之间取得了最佳平衡,因此被选为本研究的主干网络。

3.5.2 主干网络改进实验

为系统评估本文所提出的改进策略,包括CBAM注意力机制与CloU+Focal Loss损失函数的有效性,在相同的实验环境下,对3种基线主干网络分别进行了优化,并进行了对比实验。优化后的基于3种主干网络的YOLOv8n模型检测性能对比图5所示。

从各类缺陷的检测性能来看,ResNet50在缺陷(热裂纹、点蚀、污点、侵蚀、锈蚀)上均保持稳定且较高的AP值,显示出其对多尺度、多形态缺陷的强适应能力。ResNet50凭借其深层网络结构,能够更好地捕捉细微的纹理变化和形态特征,从而在所有缺陷类别上实现均衡的高精度检测。表5详细列出了优化前后模型的综合性能指标。

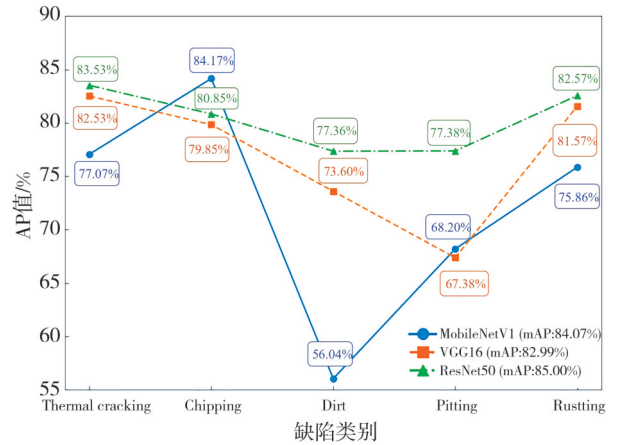


图5 优化后的基于3种主干网络的YOLOv8n模型检测性能对比图

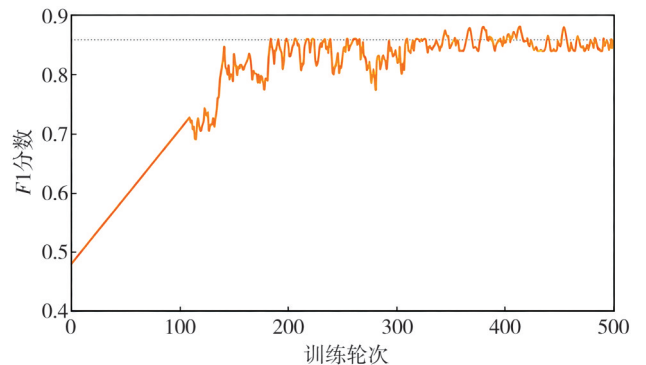
Fig. 5 Optimized comparative performance chart of YOLOv8n models with three different backbone networks

表5 3种主干网络优化前后性能对比

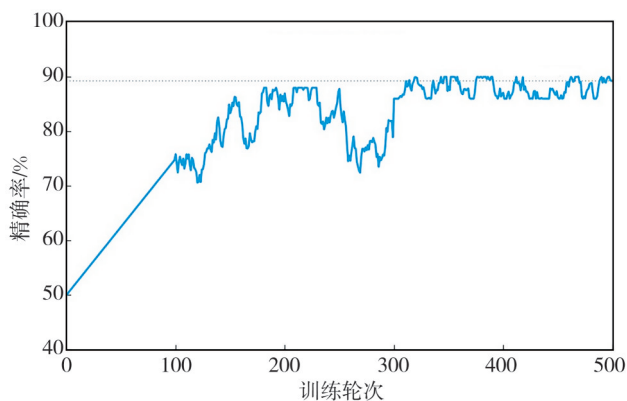
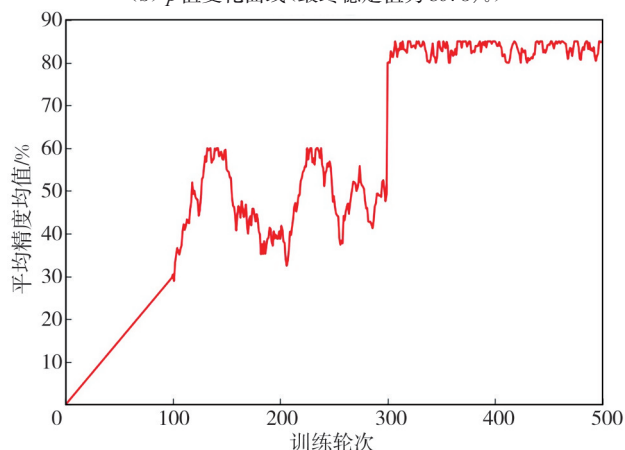
Tab. 5 Performance comparison of three backbone networks before and after optimization

模型架构	mAP@0.5 (优化前)/%	mAP@0.5 (优化后)/%	提升幅度/百分点
MobileNetV1	44.07	58.50	14.43
VGG16	62.99	76.20	13.21
ResNet50	65.47	85.00	19.53

对比结果显示,3种网络的mAP@0.5指标分别实现了显著增长,其中,轻量级的MobileNetV1提升14.43百分点,平衡型的VGG16提升13.21百分点,而更深层的ResNet50提升幅度最为突出,达到19.53百分点。这一梯度式的提升效果清晰地表明,改进策略的有效性与主干网络固有的特征提取能力呈正相关。基础性能越强大的网络,越能够更好地利用注意力机制与改进的损失函数所带来的优化信息,从而获得更大的性能增益。实验结果表明,本研究所提出的改进策略(集成CBAM注意力机制与优化损失函数)对各类主干网络性能均具有显著提升效果,充分验证了改进方案的有效性与普适性。本研究提出的模型YOLOv8-R50CBAM的性能如图6所示。



(a) F1-Score收敛曲线(最终稳定值为0.858)

(b) p 值变化曲线(最终稳定值为 89.3%)

(c) mAP 收敛曲线(最终稳定值为 85.0%)

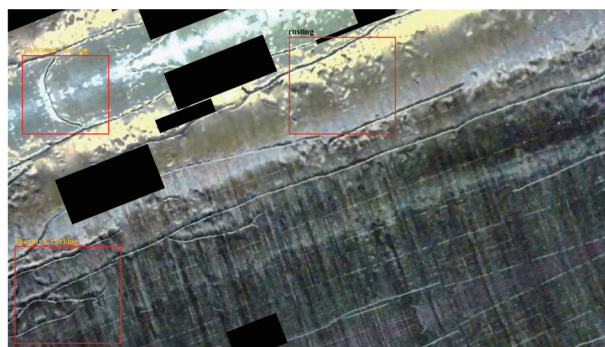
图 6 YOLOv8-R50CBAM 模型性能曲线图

Fig. 6 Performance curves of the YOLOv8-R50CBAM model

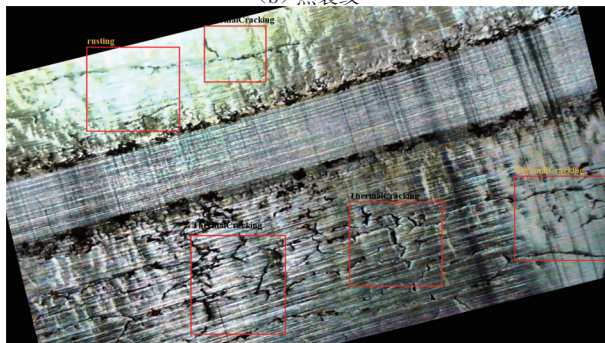
训练过程分析表明, YOLOv8-R50CBAM 模型具有良好的收敛特性和稳定的性能提升作用。F1-Score 总体上随训练轮次上升, 虽然在 300 轮次附近出现短期波动, 但后续快速恢复稳定; 精确率持续增长且与召回率保持平衡, 表明模型在保持高检出能力的同时并未牺牲检测精度。平均精度(mAP)曲线从 0 开始快速提升, 最终收敛于 85.0% 附近, 验证了模型架构的特征提取能力和训练策略的有效性。实验结果系统证明了该模型在枪管内壁缺陷检测任务中具有优秀的学习能力和泛化潜力。图 7 展示了 YOLOv8-R50CBAM 模型在枪管内壁缺陷检测中的可视化结果。



(a) 侵蚀与污点



(b) 热裂纹



(c) 锈蚀

图 7 YOLOv8-R50CBAM 检测模型的可视化结果

Fig. 7 Visualization results of the YOLOv8-R50CBAM detection model

3.5.3 消融实验

本研究的 YOLOv8-R50CBAM 模型采用消融试验来验证 CBAM 模块、改进的 PANet、CIoU 损失函数、Focal Loss 分类损失在枪管内壁缺陷检测识别上的优势, 试验结果如表 6 所示。

表 6 消融实验结果对比

Tab. 6 Ablation study comparative results

R50	MV1	VGG16	CBAM	PANet+	CIoU	FL	mAP/%
✓	×	×	✓	×	✓	✓	84.02
✓	×	×	×	✓	✓	✓	83.94
✓	×	×	✓	✓	×	✓	84.41
✓	×	×	✓	✓	✓	×	84.23
×	×	✓	✓	✓	✓	✓	82.99
×	✓	×	✓	✓	✓	✓	84.07
✓	×	×	✓	✓	✓	✓	85.00

注: 表格展示了各组件的启用(✓)与禁用(×)状态及对应的 mAP 值; R50=ResNet50, MV1=MobileNetV1, FL=Focal Loss, PANet+=改进 PaNet。

实验结果表明, 采用 ResNet50 主干网络配合改进 PaNet 特征金字塔、CBAM 注意力机制、CIoU 损失函数和 Focal Loss 分类损失的完整配置方案取得了最优性能, mAP 达到 85.0%, 验证了各组件协同优化的有效性。相较于基准配置(ResNet50+改进 PaNet 结构+CBAM+CIoU+Focal Loss), 移除改进 PaNet 结构使 mAP 降低了

0.98 百分点,表明双向特征融合机制能有效增强多尺度缺陷特征的表征能力;移除CBAM注意力模块导致mAP下降了1.06百分点,证实了注意力机制在复杂背景干扰下的特征选择优势。不同主干网络的对比实验显示,相较于MobileNetV1和VGG16, ResNet50的mAP分别提升了0.93百分点和2.01百分点,体现了深层残差网络在特征提取方面的优越性。损失函数消融实验进一步表明,CIoU损失相较于传统IoU损失提升了0.59百分点, Focal Loss相较于标准交叉熵损失提升了0.77百分点,验证了改进损失函数在边界框回归和类别不平衡问题处理上的有效性。因此,各组件的协同优化显著提升了模型在复杂背景下的检测能力。

4 结 论

针对强背景干扰下的枪管内壁缺陷检测难题,本文提出了一种YOLOv8-R50CBAM检测模型。核心工作围绕网络结构与推理流程展开系统性改造,显著提升了检测性能。模型在3个层面进行了创新性设计。在特征提取层面,提出注意力增强的残差模块CRB,通过在ResNet50的残差瓶颈中嵌入CBAM,实现对背景噪声的抑制与缺陷纹理的稳定增强。在特征融合层面,设计双向路径聚合模块BPAF,基于PANet构建双向信息流,有效提升小缺陷的可见性与边界框回归的一致性。在检测与推理层面,构建推理增强模块ODI,引入Soft-NMS缓解密集缺陷漏检,采用测试时增强策略提升输出稳定性,并通过置信度校准使预测分数更可靠。训练中采用结合CIoU与Focal Loss的改进损失函数,优化了定位精度与类别不平衡问题。

在拥有五类缺陷的数据集上, YOLOv8-R50CBAM实现了85.00%的mAP, F1分数稳定在0.858,训练过程收敛良好。各类别中, Thermal cracking、Rusting与Chipping检测性能优异,而Dirt与Pitting等弱对比散点型缺陷更具挑战,符合实际规律。实验表明,所提结构具有必要性,将主干网络替换为MobileNetV1和VGG时, mAP分别降至84.07%和82.99%,证明轻量化模型在保持整体性能时,对细碎纹理缺陷的敏感性减小,而深层残差结构在复杂背景下更具表征优势。

本文的核心贡献并非单一指标的突破,而是为枪管内壁缺陷检测的工业场景构建了一条结构清晰、可复现、可解释的缺陷检测技术路径,通过模块化的设计系统性地解决了从特征提取、融合到后处理的关键问题,可以作为工程部署与后续扩展研究的基础。本文模型仍存在一定局限,对于极小与边界模糊的缺陷(如侵蚀缺陷)的检测仍存在上限。同时, Soft-NMS与TTA等策略虽然提升了精度,但增加了推理耗时,不利于实时部署。此外,模型的跨设备、跨批次泛化能力有待进一步验证。后续可围绕以上局限进行研究。

利益冲突声明/Conflict of Interests

所有作者声明不存在利益冲突。

All authors declare no relevant conflict of interests.

参考文献:

- [1] Mizanuzzaman M, Nyeem H, Kader M F. Automatic abnormality detection and estimation in small arms barrel: a new software-based system for bangladesh army[EB/OL]. (2022-02-10) [2025-12-01]. <https://www.techrxiv.org/doi/full/10.36227/techrxiv.19123271.v1>.
- [2] 张稳. 枪械内膛疵病检测系统的设计与实现[D]. 北京: 北京邮电大学, 2013.
- [3] 张顺, 龚怡宏, 王进军. 深度卷积神经网络的发展及其在计算机视觉领域的应用[J]. 计算机学报, 2019, 42(3): 453-482.
Zhang Shun, Gong Yihong, Wang Jinjun. The development of deep convolution neural network and its applications on computer vision[J]. Chinese Journal of Computers, 2019, 42(3): 453-482. (in Chinese)
- [4] 邓兆鹏. 前视全景钻孔摄像系统的图像分类及视频拼接技术研究[D]. 青岛: 山东科技大学, 2019.
- [5] 陈超, 齐峰. 卷积神经网络的发展及其在计算机视觉领域中的应用综述[J]. 计算机科学, 2019, 46(3): 63-73.
Chen Chao, Qi Feng. Review on development of convolutional neural network and its application in computer vision[J]. Computer Science, 2019, 46(3): 63-73. (in Chinese)
- [6] Ren S, He K, 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.
- [7] Wang C Y, Bochkovskiy A, Liao H Y M. YOLOv7: Trainable bag-of-freebies sets new state-of-the-art for real-time object detectors [C]//2023 IEEE/CVF Con-

- ference on Computer Vision and Pattern Recognition (CVPR), 2023: 7464-7475.
- [8] 韩强. 面向小目标检测的改进 YOLOv8 算法研究[D]. 长春: 吉林大学, 2023.
- [9] 郭丽峰, 黄俊杰, 吴禹竺, 等. 基于 YOLOv8n 改进的水稻病害轻量化检测[J]. 农业工程学报, 2025, 41(8): 156-164.
- Guo Lifeng, Huang Junjie, Wu Yuzhu, et al. Detecting rice diseases using improved lightweight YOLOv8n[J]. Transactions of the Chinese Society of Agricultural Engineering, 2025, 41(8): 156-164. (in Chinese)
- [10] 熊伟, 黄玉谦, 彭鑫旭. 基于改进 YOLOv8 模型的 PCB 缺陷检测算法[J]. 电子测量技术, 2025, 48(18): 159-167.
- Xiong Wei, Huang Yuqian, Peng Xinxu. PCB defect detection model based on improved YOLOv8 algorithm[J]. Electronic Measurement Technology, 2025, 48(18): 159-167. (in Chinese)
- [11] 梁胤杰, 南新元, 蔡鑫, 等. 基于数据增广与改进 YOLOv8 的桥梁缺陷检测[J]. 广西师范大学学报(自然科学版), 2025, 43(3): 84-97.
- Liang Yinjie, Nan Xinyuan, Cai Xin, et al. Bridge defect detection based on data augmentation and improved YOLOv8[J]. Journal of Guangxi Normal University (Natural Science Edition), 2025, 43(3): 84-97 (in Chinese)
- [12] 孙华华. 基于机器视觉的身管内壁质量智能检测[D]. 西安: 西安工业大学, 2023.
- [13] 周思康, 张明柱, 王宁宁, 等. 基于 ECA-VGG16 和迁移学习的砂轮磨损状态识别研究[J]. 机床与液压, 2025, 53(19): 103-109.
- Zhou Sikang, Zhang Mingzhu, Wang Ningning, et al. Research on wear state recognition of grinding wheel based on ECA-VGG16 and transfer learning[J]. Machine Tool & Hydraulics, 2025, 53(19): 103-109. (in Chinese)
- [14] 周建新, 何洋. 基于改进 MobileNetV2 的钢板表面缺陷检测[J]. 组合机床与自动化加工技术, 2024(9): 183-187.
- Zhou Jianxin, He Yang. Steel plate surface defect detection based on improved MobileNetV2[J]. Modular Machine Tool & Automatic Manufacturing Technique, 2024(9): 183-187. (in Chinese)
- [15] 盛强. 基于内窥图像的精密小深孔表面缺陷检测及其形貌重构研究[D]. 西安: 西安理工大学, 2024.
- [16] 姚若禹, 郑世玲, 史怡璇, 等. 基于改进 YOLOv8n 的钢材表面缺陷检测算法[J]. 聊城大学学报(自然科学版), 2025, 38(2): 177-189.
- Yao Ruoyu, Zheng Shiling, Shi Yixuan, et al. Steel surface defect detection algorithm based on improved YOLOv8n[J]. Journal of Liaocheng University (Natural Science Edition), 2025, 38(2): 177-189. (in Chinese)
- [17] 梁喜凤, 魏志卫. 基于改进 CycleGAN 与 YOLOv8 的夜间番茄茎、枝分割方法[J]. 农业工程学报, 2025, 41(8): 147-155.
- LIANG Xifeng, WEI Zhiwei. Segmenting tomato stems and branches at night time using improved CycleGAN and YOLOv8[J]. Transactions of the Chinese Society of Agricultural Engineering, 2025, 41(8): 147-155. (in Chinese)
- [18] Shafi I, Mazahir A, Fatima A, et al. Internal defects detection and classification in hollow cylindrical surfaces using single shot detection and MobileNet[J]. Measurement, 2022, 202: 111836.
- [19] Lawton B. Thermo-chemical erosion in gun barrels[J]. Wear, 2001, 251(1): 827-838.
- [20] Shanmugamani R, Sadique M, Ramamoorthy B. Detection and classification of surface defects of gun barrels using computer vision and machine learning[J]. Measurement, 2015, 60: 222-230.
- [21] Shao X J, Zhu S J, Cao L J. Research on intelligent bore peek and measurement system based on machine vision technology for gun barrel[J]. Vibroengineering Procedia, 2016, 8: 414-421.
- [22] 袁泽辉. 基于机器视觉的管道内表面缺陷检测技术[D]. 上海: 华东理工大学, 2020.