

# 基于改进 YOLOv11 的钢管内壁缺陷检测方法研究

齐向晶<sup>a</sup>, 高宏伟<sup>b</sup>, 张熙函<sup>a</sup>

(沈阳理工大学 a. 自动化与电气工程学院, b. 国际工程学院, 沈阳 110159)

**摘要:** 钢管内壁缺陷检测对于保障产品质量与生产安全至关重要, 传统检测方法受光照、空间限制, 检测效率较低, 难以满足工业化需求。近年来, 深度学习技术在缺陷检测领域的应用取得了一定进展, 但仍面临光照变化、纹理干扰、缺陷与背景相似等挑战, 检测的准确性与实时性有待提升。为此, 提出一种基于改进 YOLOv11 的缺陷检测模型, 将 Slim-neck 模块与高效多尺度注意力 (efficient multi-scale attention, EMA) 机制相结合, 通过 Slim-neck 模块对标准卷积与深度可分离卷积提取的特征进行有效融合, 降低计算复杂度, 同时增强特征提取能力, 通过 EMA 机制的并行子结构设计减少顺序依赖, 提高检测效率。实验结果表明, 本文模型在自建数据集上检测的平均精度均值 (mAP@0.5) 较基础模型提升了 11.8%, 在计算开销和参数量方面亦展现出优异性能, 验证了其在钢管内壁缺陷检测中的有效性和应用价值。

**关键词:** 钢管内壁; 缺陷检测; YOLOv11; 特征提取; 特征融合

中图分类号: TP391.4 文献标志码: A DOI:10.3969/j.issn.1003-1251.2025.06.002

## Detection Method of Defects on Inner Wall of Steel Pipe Based on Improved YOLOv11

QI Xiangjing, GAO Hongwei, ZHANG Xihan

(Shenyang Ligong University, Shenyang 110159, China)

**Abstract:** To detect the defects on the inner wall of steel pipes is crucial for product quality and production safety. Traditional detection methods are limited by lighting conditions and spatial constraints, resulting in low detection efficiency and failure to meet industrial demands. In recent years, deep learning techniques have made certain progress in the field of defect detection. However, challenges such as lighting variations, texture interference, and the similarity between defects and the background still hinder detection accuracy and real-time performance. To address these issues, this paper proposes an improved YOLOv11-based defect detection model that integrates a Slim-neck module with an efficient multi-scale attention (EMA) mechanism. The Slim-neck module effectively fuses features extracted by standard convolution and depthwise separable convolution, reducing computational complexity while enhancing feature extraction capability. Additionally, the parallel substructure design of the EMA mechanism reduces sequential dependencies, thereby improving detection efficiency. Experimental results demonstrate that the proposed model improves the mean average precision (mAP@0.5) by 11.8% over the baseline model on a custom dataset. It also exhibits excellent performance in terms of computational overhead and parameter count, validating its effectiveness and application value in detection of defects on inner wall of steel pipes.

**Key words:** inner wall of steel pipe; detection of defects; YOLOv11; feature extraction; feature fusion

钢管是一种应用广泛的工业材料,其质量直接影响工业生产及工程的可靠性和安全性。钢管内壁缺陷种类较多、危害较大,对其进行有效检测是保证钢管质量的重要手段。

目前钢管内壁缺陷检测主要采用漏磁检测、超声导波检测和涡流检测等无损检测方法。漏磁检测技术通过磁化钢管并检测磁场泄漏来识别缺陷<sup>[1]</sup>,该方法能够快速、准确地检测出钢管表面及近表面缺陷,适用于大规模自动化检测,但对缺陷方向敏感、定位精度有限,难以满足高精度检测需求。唐琴等<sup>[2]</sup>提出了一种融合漏磁检测与电磁超声检测的复合传感结构,可同步检测钢试样的表面及内部缺陷。超声导波检测技术利用超声波传播特性检测内壁腐蚀和裂纹,具有检测范围广、效率高的优势<sup>[3]</sup>。吕国钢等<sup>[4]</sup>研究了一种电磁超声导波检测方法,并验证了其对小径管道裂纹检测的有效性。但该方法易受焊缝、支架影响,信号衰减较快,对小缺陷分辨率有限,导致检测准确度不高<sup>[5]</sup>。涡流检测是一种基于电磁感应原理的检测技术,在钢管内外壁缺陷检测方面取得了一定进展<sup>[6]</sup>,如焦靖淇等<sup>[7]</sup>提出了一种直流磁化下的涡流检测方法,提高了检测灵敏度,增强了对深层缺陷的识别能力。但该技术仅适用于表面和近表面缺陷,对厚壁钢管检测能力有限,且受材料电磁特性影响较大<sup>[8]</sup>。

随着深度学习技术的快速发展,其在计算机视觉领域的应用日益成熟,极大地推动了缺陷检测技术的进步。宋建辉等<sup>[9]</sup>提出了一种基于卷积块注意力模块(CBAM)的全卷积抓取姿态预测方法,该方法采用深度可分离卷积降低了参数量,增强了多尺度特征提取能力,提升了预测精度。吕艳辉等<sup>[10]</sup>提出了一种基于特征融合与双向简化门结构的卷积循环神经网络(CRNN)模型用于文本识别,增强了特征提取能力,提高了计算效率。蒲家鹏等<sup>[11]</sup>提出了一种改进的YOLOv4算法,通过减少网络参数量并优化特征提取方法,有效提升了航拍图像中小目标的检测精度。大量研究表明<sup>[12-13]</sup>,结合注意力机制的卷积神经网络(CNN)及其衍生架构在缺陷检测中具有良好的性能,可提高检测算法的鲁棒性。但CNN在捕捉全局模式方面存在不足,尤其在处理小尺寸或复杂背景的缺陷时表现不佳。为提升多尺度特征提取能力,付丽君等<sup>[14]</sup>将多源小波时频变换与CNN相结合,用于轴承故障诊断,提高了检测精度和鲁棒性。ResNet通过残差学习缓解了梯

度消失问题,但其深层结构可能导致梯度爆炸,影响网络稳定性,为此提出了FC-ResNet<sup>[15]</sup>(结合CBAM提升特征捕捉能力,减少计算成本)和FFC-ResNet<sup>[16]</sup>(结合傅里叶卷积增强全局特征提取能力)等改进算法,提升了模型的收敛性和适用性。YOLO因检测效率高受到青睐,但在密集或重叠缺陷检测上效果不佳,为此研究者引入注意力机制改进YOLO,以增强复杂场景下的检测能力<sup>[17]</sup>。总之,深度学习在缺陷检测领域取得了一定进展,但仍需进一步提升特征提取能力、网络稳定性和检测精度等,以应对工业应用中的挑战。

本文提出一种基于改进YOLOv11的钢管内壁缺陷检测方法。引入Slim-neck模块<sup>[18]</sup>优化YOLOv11的颈部网络,在主干网络引入高效多尺度注意力(efficient multi-scale attention,EMA)机制<sup>[19]</sup>,以增强模型对不同形态缺陷特征的敏感度,改善对多尺度、多形态缺陷的检测效果。使用自建数据集验证本文模型的性能。

## 1 模型介绍

相较于其他深度学习模型,YOLOv11在目标检测中具有精度高和鲁棒性好的特点,相较于其他版本的YOLO模型,YOLOv11在钢管内壁缺陷检测方面表现更为突出。为了进一步提升其检测性能,本文在YOLOv11的主干网络中引入EMA机制,以增强模型对关键特征的聚焦能力,提升检测精度,同时引入Slim-neck模块对YOLOv11的颈部网络进行优化,在保持模型实时性的前提下增强其对钢管内壁缺陷的识别能力。本文改进YOLOv11模型的网络结构如图1所示,图中SP-PF表示快速空间金字塔池化,C3k2模块是一种改进的卷积模块,GS卷积以及VoV-GSCSP模块是Slim-neck模块的组成部分,下文中将具体论述。

### 1.1 EMA 机制

传统CBAM与SE注意力机制存在通道降维造成信息损失、顺序依赖性强及计算复杂度高等问题,难以满足实际需求。EMA模块采用并行子结构设计,可减弱在小目标缺陷特征提取过程中的顺序依赖,提升推理效率,通过多尺度空间特征融合与跨空间信息聚合,有效增强模型对微小缺陷与复杂背景的判别能力。EMA的总体结构如图2所示。

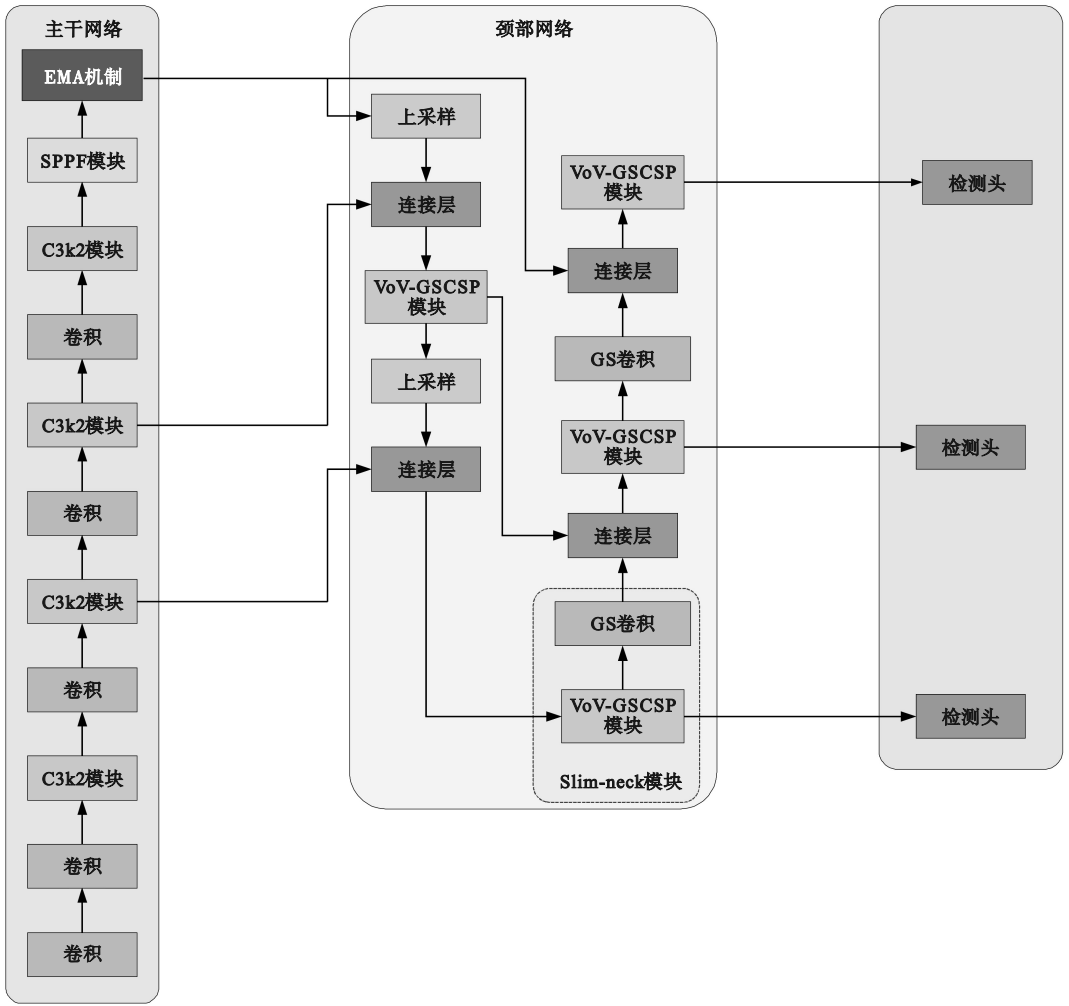


图 1 改进 YOLOv11 模型的网络结构

Fig. 1 The network structure of the improved YOLOv11 model

## 1) 图征群组

对于任意给定的输入特征映射  $X$  ( $X \in \mathbf{R}^{C \times H \times W}$ , 其中  $C$  为通道数,  $H$  为特征图高度,  $W$  为特征图宽度), EMA 将其在通道维度方向上划分为  $G$  个子特征, 以学习不同的语义, 其中组样式可表示为  $X = [X_0, X_1, \dots, X_i, \dots, X_{G-1}]$ ,  $X_i \in \mathbf{R}^{C // G \times H \times W}$ . 在不失一般性的情况下, 令  $G$  远小于  $C$ , 并假设学习到的注意力权重描述符用于加强每个子特征中感兴趣区域的特征表示。

## 2) 并行子网

EMA 通过三条并行路径提取分组特征图的注意力权重描述符, 以高效捕获多尺度的空间信息。在  $1 \times 1$  卷积分支中, 两个一维全局平均池化操作分别沿两个空间方向编码通道信息, 并通过连接编码特征共享同一卷积核, 从而避免降维操作。通过将  $1 \times 1$  卷积输出分解为两个向量, 并使用 Sigmoid 函数拟合二维 Binomial 分布, 实现跨

通道交互。在  $3 \times 3$  卷积分支中, 单个  $3 \times 3$  卷积核捕获局部跨通道交互特征, 进一步丰富特征表示。在恒等映射分支中, 直接将输入特征图绕过卷积与池化操作, 以恒等方式传递至融合阶段, 保留原始空间信息并增强特征稳定性。通过上述方式, EMA 不仅有效编码通道间的信息, 还可保留精确的空间结构信息, 从而提升特征图的表达能力。

## 3) 跨空间学习

本文提出一种跨空间信息聚合方法, 旨在不同空间维度方向上实现更丰富的特征融合。首先, 采用二维全局平均池化对  $1 \times 1$  卷积分支的输出进行全局空间信息编码, 并将其输出直接转换为与  $3 \times 3$  卷积分支的特征维度相匹配的形式。该过程利用全局池化操作有效建立空间维度上的长程依赖关系, 全局池化公式为

$$z_c = \frac{1}{H \times W} \sum_{j=1}^H \sum_{i=1}^W x_c(i, j) \quad (1)$$

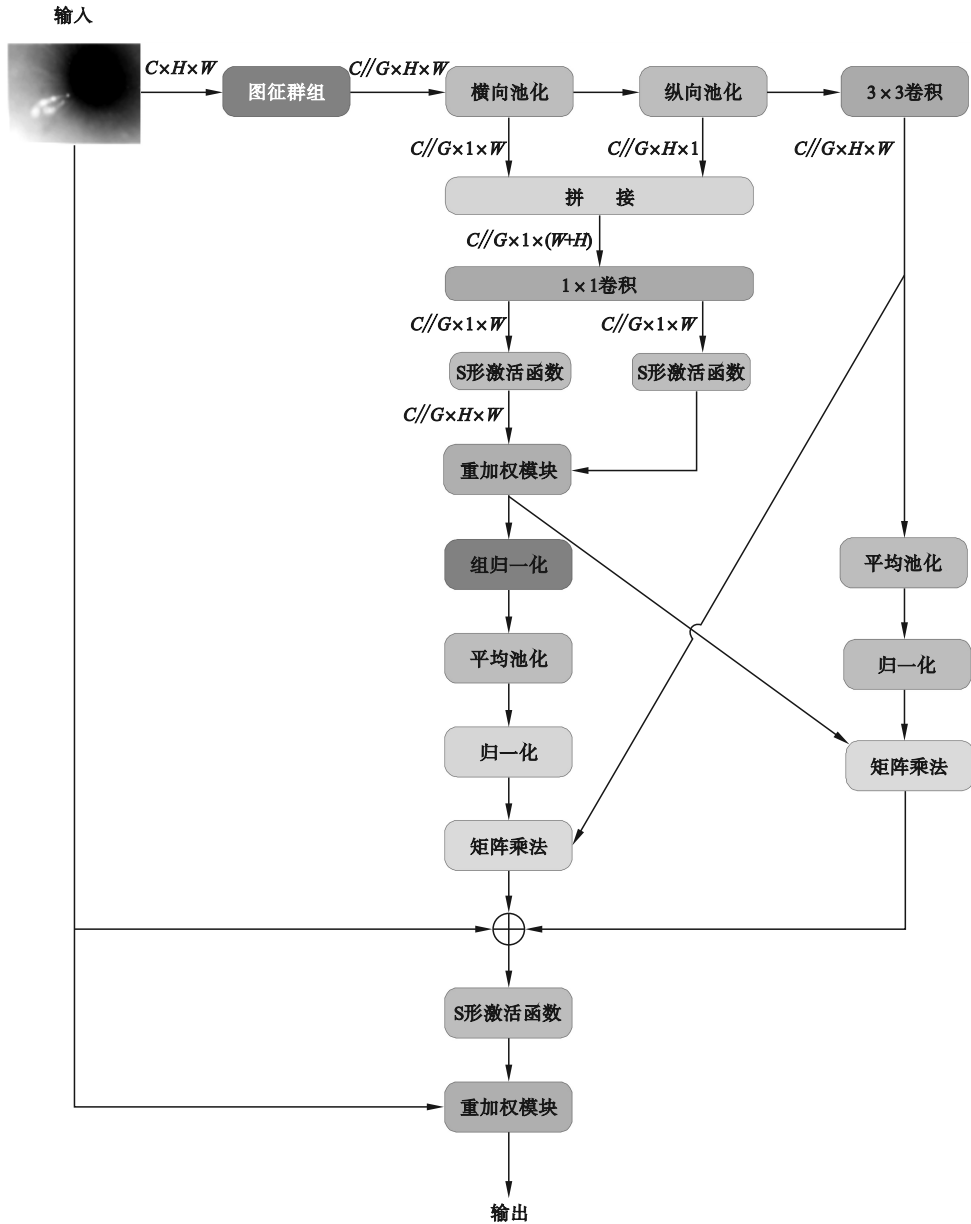


图2 EMA结构

Fig.2 EMA structure

式中: $z_c$ 表示通道 $c$ 处的全局池化输出; $x_c(i,j)$ 表示输入特征图在通道 $c$ 的像素点 $(i,j)$ 的像素值。然后,使用Softmax非线性函数拟合全局池化后的输出,以提高计算效率。最后,通过矩阵点积运算,得到第一个空间注意力图,用于收集不同尺度的空间信息。同样, $3 \times 3$ 卷积分支通过二维全局平均池化编码全局空间信息得到第二个空间注意力图, $1 \times 1$ 卷积分支的输出在通道特征联合激活机制之前变换为对应维度。生成的两个空间注意力图经融合后,采用S形激活函数(Sigmoid)进行归一化处理,从而捕捉像素级的成对关系,并突出全局上下文信息。

## 1.2 Slim-neck 模块

在目标检测任务中,通常在主干网络逐步压缩图像空间信息并扩展通道信息,易导致部分语义信息丢失。传统的颈部网络结构,如特征金字塔网络和路径聚合网络,虽可通过多尺度特征融合提升检测能力,但依赖额外计算模块,增加了计算复杂度与推理时间。为此,本文引用Slim-neck模块,其中包括GS卷积(GSConv)模块、GSbottle-neck模块以及VoV-GSCSP模块。

GSConv采用深度可分离卷积(DWConv)替代标准卷积(Conv),旨在保留更多通道连接的同时降低计算负担。以每秒浮点运算次数(FLOPs)

表示计算量,则标准卷积计算量  $F_1$  为

$$F_1 = C_{in} \times C_{out} \times K \times K \times H \times W \quad (2)$$

式中:  $C_{in}$  和  $C_{out}$  分别为输入卷积和输出卷积的通道数;  $K$  为卷积核大小。可分离卷积计算量  $F_2$  为

$$F_2 = C_{in} \times K \times K \times H \times W + C_{in} \times C_{out} \times H \times W \quad (3)$$

由于  $C_{out}$  远小于  $C_{in}$ , 故采用深度可分离卷积可大幅降低计算开销。

GSConv 结构如图 3 所示。GSConv 通过通道重排(Shuffle)操作将 Conv 生成的特征渗透到 DWConv 生成的特征中,使两者特征均匀融合,增强通道间信息流动,且无额外计算负担。在此基础上,本文引用 GSbottleneck 结构,如图 4(a) 所示。相较于传统 Bottleneck 结构,GSbottleneck 通

过 DWConv 增强非线性特征表达能力,同时保留通道连接,有效提升特征提取能力,其核心操作表示为

$$Y = \text{Concat}(\text{Shuffle}(\text{DWConv}(\text{Conv}_{1 \times 1}(X))), X) \quad (4)$$

式中:  $Y$  表示输出特征;  $\text{Conv}_{1 \times 1}(\cdot)$  表示  $1 \times 1$  卷积降维操作;  $\text{DWConv}(\cdot)$  表示深度可分卷积操作;  $\text{Shuffle}(\cdot)$  表示通道洗牌操作;  $\text{Concat}(\cdot)$  表示特征拼接操作。

此外,GSbottleneck 引入一次性聚合策略,将前后阶段特征直接连接,有效缓解特征冗余,通过特征分割与融合进一步降低推理延时、减少冗余计算,形成轻量化的 VoV-GSCSP 结构,如图 4(b) 所示。

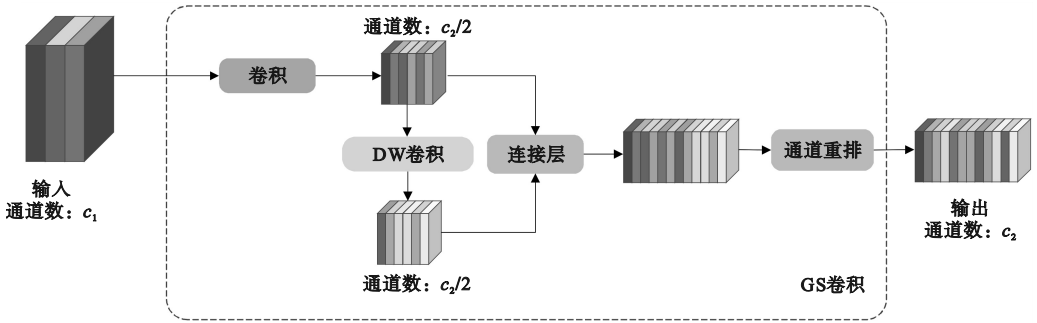


图 3 GSConv 结构

Fig.3 GSConv structure

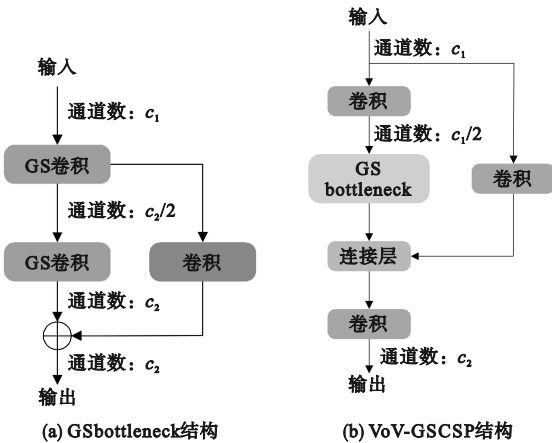


图 4 GSbottleneck 结构和 VoV-GSCSP 结构

Fig.4 GSbottleneck and VoV-GSCSP structure

综上, Slim-neck 模块在计算复杂度、信息保留与特征融合方面较传统颈部结构展现出明显优势,使特征表达更紧凑,降低因光照多样性带来的噪声冗余,尤其适用于实时性与资源约束要求较高的钢管内壁缺陷检测任务。

## 2 实验部分

### 2.1 数据采集与数据集制作

本文使用内窥镜在钢管内部进行缺陷图像的采集。考虑到不同光照条件对缺陷识别的影响,分别在亮光和暗光两种条件下进行数据采集,示意图如图 5 所示。在数据采集之前,首先对钢管内壁进行彻底清洁,对内窥镜头进行检查,确保内壁表面无灰尘或杂物、内窥镜头干净无遮挡,以保证图像采集的质量和精度。然后将内窥镜稳固地放置在钢管的圆心位置,确保内窥镜位于正中间,激光发出的射线位于靶纸中央时开始采集数据。采集过程中,首先开启光源,内窥镜沿钢管的轴线从近端到远端匀速缓慢运动,采集亮光条件下的图像数据,再关闭光源,内窥镜沿相反方向再次匀速缓慢运动,采集暗光条件下的图像数据。通过上述方式保证数据集涵盖不同光照条件下的钢管内壁缺陷图像,为后续缺陷检测模型训练提供全面且具有代表性的数据。

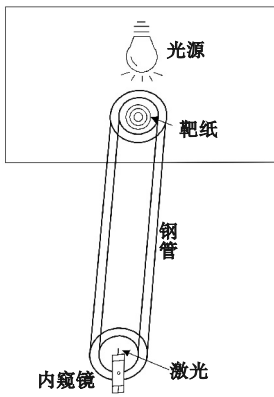


图5 数据采集示意图

Fig. 5 Schematic diagram of data collection

采集完缺陷图像后,首先进行数据增强处理,主要对缺陷部分进行增强,以便后续进行缺陷检测。然后对缺陷分类标注,将处理好的缺陷图像储存为 Text 数据格式,完成钢管内壁缺陷图像数据集的构建。数据集含有轧折(RS)、离层(Layer)、结疤(Scab)、氧化皮(Sc)、裂纹(Cr)共5类缺陷,其中轧折、离层类缺陷共1664张图像,结疤、氧化皮、裂纹类缺陷共652张图像,按照7:2:1的比例将数据集划分为训练集、验证集和测试集。

## 2.2 实验结果及分析

模型验证基于 Windows 11 操作系统,使用 CUDA 11.0 和 CUDNN 11.3 下的 Pytorch 1.9.0 框架实现,硬件设备采用 NVIDIA RTX 4090 GPU。所有实验均在自建数据集上进行。为全面评估模型性能,采用多个评价指标:使用平均精度均值 mAP@0.5 衡量检测精度;使用参数量衡量模型的容量与表达能力;使用帧率(FPS)反映模型的推理速度,展示模型在实时检测中的效率。

### 2.2.1 消融实验

逐步将 EMA 和 Slim-neck 模块应用于基础模型 YOLOv11s 中,验证各模块的性能,实验中仅对模块进行添加或者移除,其余所有参数均保持一致。在自建数据集上的消融实验结果如表1所示。由表1可知:单独使用 EMA 时,模型的 mAP@0.5 由 0.451 升至 0.463,但帧率由 7.0 降至 5.8;Slim-neck 模块对模型性能提升比较显著,单独引入 Slim-neck 模块时,mAP@0.5 提升至 0.493,帧率提升至 10.0;同时引入 EMA 和 Slim-neck 模块后,mAP@0.5 进一步提升至 0.504,但帧率降至 9.4。相比基础模型,本文改

进模型的 mAP@0.5 提升了 11.8%,帧率提升了 34.3%,验证了各模块在提升精度的同时有效控制了模型复杂度。

表1 消融实验结果

Table 1 Ablation experimental results

EMA	Slim-neck	mAP@0.5	参数量	帧率(FPS)
×	×	0.451	9 414 735	7.0
✓	×	0.463	9 219 215	5.8
×	✓	0.493	9 981 871	10.0
✓	✓	0.504	9 786 315	9.4

### 2.2.2 对比实验

采用本文提出的模型与其他主流的一阶段检测模型在自建数据集上进行对比实验,结果如表2所示。可见,本文模型在钢管缺陷检测中的各项性能指标均优于其他轻量化检测模型。相比 YOLOv10s 模型,本文模型在参数量仅增加 21.7% 的情况下,mAP@0.5 提升了 15.3%,帧率提升了 38.2%;相比 YOLOv8s 模型,本文模型的 mAP@0.5 提升了 12.5%,参数量减少了 12.0%,帧率提升了 51.6%;相比 YOLOv5s 模型,本文模型的 mAP@0.5 提升了 22.3%,参数量减少了 22.2%,帧率提升了 20.5%。上述结果表明,本文提出的改进模型在提高检测精度的同时有效控制了模型的复杂度,并显著增强了推理效率。

表2 对比实验结果

Table 2 Comparative experimental results

模型	输入图像尺寸	mAP@0.5	参数量	帧率(FPS)
YOLOv5s		0.412	12 578 641	7.8
YOLOv8s		0.448	11 127 519	6.2
YOLOv10s	640 × 640	0.437	8 038 830	6.8
本文改进模型		0.504	9 786 315	9.4

### 2.2.3 可视化结果分析

为进一步证明本文提出模型在钢管内壁缺陷检测任务中的有效性,将检测结果可视化,并与 YOLOv11s 模型进行比较,可视化结果如图6所示。实验结果表明,相较于 YOLOv11s 模型,本文模型在降低缺陷漏检、误检以及细微缺陷识别等方面表现更优,显著提升了检测的准确性与鲁棒性,更适用于钢管内壁缺陷检测的实际应用场景。

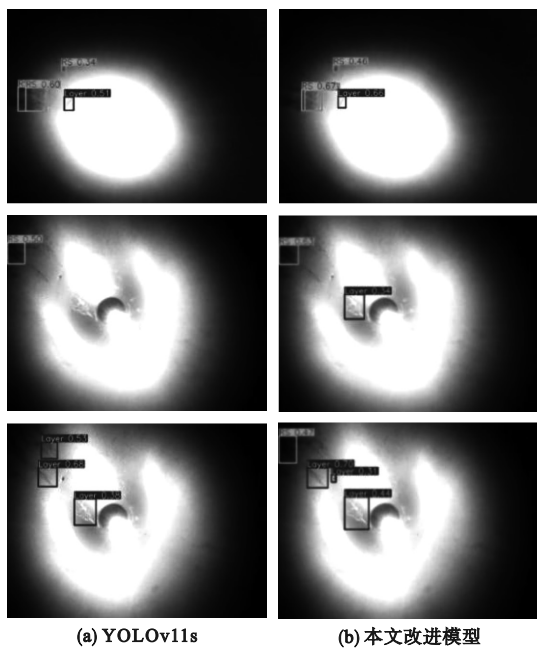


图 6 可视化结果比较

Fig. 6 Comparison of visualization results

### 3 结论

提出了一种基于改进 YOLOv11 的钢管内壁缺陷检测模型。引入 EMA 机制,通过其并行子结构设计减少了顺序依赖,提升了计算效率;引入 Slim-neck 模块,融合了标准卷积和深度可分卷积提取的特征,避免了额外计算负担,确保了模型在各种设备上的适配性。在自建数据集上的实验结果表明,本文改进模型的 mAP@0.5 较 YOLOv11s 模型提升了 11.8%,在参数量和计算量上亦表现优异,在准确性和计算效率之间达到了有效平衡,提供了一种具有良好应用前景的钢管内壁缺陷检测解决方案。

### 参考文献 (References):

- [1] 耿浩,杨理践,项行鸿. 钢管高速漏磁信号检测与分析[J]. 仪表技术与传感器,2021(11):88-93,98.  
GENG H, YANG L J, XIANG X H. High speed magnetic leakage signal detection and analysis of steel pipe[J]. Instrument Technique and Sensor,2021(11):88-93,98. (in Chinese)
- [2] 唐琴,高斌,薛沁文,等. 基于漏磁检测-电磁超声多物理融合传感器技术研究[J/OL]. 机械工程学报,2025:1-11 [2025-03-15]. <https://kns.cnki.net/kcms/detail/11.2187.TH.20250303.1726.010.html>.  
TANG Q, GAO B, XUE S W, et al. Research on multiphysics fusion sensor technology based on magnetic flux leakage-electromagnetic acoustic transducer[J/OL]. Journal of Mechanical Engineering,2025:1-11 [2025-03-15]. <https://kns.cnki.net/kcms/detail/11.2187.TH.20250303.1726.010.html>. (in Chinese)

- cnki.net/kcms/detail/11.2187.TH.20250303.1726.010.html. (in Chinese)
- [3] 范效礼,柴楠,时亚南,等. 电磁检测技术在油气管道焊缝缺陷检测中的应用与进展[J]. 无损检测,2025,47(1):79-85.  
FAN X L, CHAI N, SHI Y N, et al. Application and progress of electromagnetic testing technologies in the detection of weld defects in oil and gas pipelines[J]. Nondestructive Testing,2025,47(1):79-85. (in Chinese)
- [4] 吕国钢,孙世超,杨广源,等. 超声导波检测技术在煤气管道管壁腐蚀检测中的研究和应用[J]. 鞍钢技术,2025(1):47-53.  
LYU G G, SUN S C, YANG G Y, et al. Study on ultrasonic guided wave detection technology detecting wall corrosion of gas pipelines and its application[J]. Angang Technology,2025(1):47-53. (in Chinese)
- [5] LI X P, SUN Y J, LIU X Y, et al. Adaptive multi-scale Bayesian framework for MFL inspection of steel wire ropes[J]. Machines,2024,12(11):801.
- [6] HUANG S P, WANG Z Q, YANG J H, et al. Adaptive fast Walsh-Hadamard transform for magnetic flux leakage signal of broken wire damage extraction under noise background[J]. Nondestructive Testing and Evaluation,2025,40(2):564-584.
- [7] 焦靖淇,王平. 油气管道内表面涡流无损检测系统设计[J]. 机械制造与自动化,2025,54(1):16-19.  
JIAO J Q, WANG P. Design of eddy current non-destructive testing system for inner surface of oil and gas pipelines[J]. Machine Building & Automation,2025,54(1):16-19. (in Chinese)
- [8] CHEN T, ZHANG C Z, LYU C, et al. Study on eddy current testing method of stress in load-bearing steel structure based on magnetoelastic effect[J]. Journal of Instrumentation,2025,20(2):P02009.
- [9] 宋建辉,顾天宇,刘砚菊,等. 基于注意力机制的全卷积抓取姿态预测方法[J]. 沈阳理工大学学报,2023,42(2):1-7.  
SONG J H, GU T Y, LIU Y J, et al. A fully convolutional grasp posture prediction method based on attention mechanism[J]. Journal of Shenyang Ligong University,2023,42(2):1-7. (in Chinese)
- [10] 吕艳辉,刘明鑫. 面向文本识别的 CRNN 模型的改进[J]. 沈阳理工大学学报,2024,43(4):27-31.  
LÜ Y H, LIU M X. Improvement of CRNN model for text recognition[J]. Journal of Shenyang Ligong University,2024,43(4):27-31. (in Chinese)
- [11] 潘家鹏,王雪梅,高宏伟. 基于改进 YOLOv4 的航拍图像目标检测研究方法研究[J]. 沈阳理工大学学报,2023,42(3):46-53.  
PU J P, WANG X M, GAO H W. Research of aerial image object detection method based on improved YOLOv4[J]. Journal of Shenyang Ligong University,2023,42(3):46-53. (in Chinese)
- [12] 文峰,陈禹铭. 基于改进 YOLOv4 的电子器件表面缺陷检测技术研究[J]. 沈阳理工大学学报,2021,40(2):1-7,21.  
WEN F, CHEN Y M. Research on surface defect detection technology of electronic components based on improved YOLOv4[J]. Journal of Shenyang Ligong University,2021,40(2):1-7,21. (in Chinese)