

# 基于改进 YOLOv11n 的轻量级桥梁裂缝图像检测算法

孙伟<sup>1</sup>, 刘文江<sup>2\*</sup>

1. 山东交通学院轨道交通学院, 山东 济南 250357; 2. 山东交通学院航空学院, 山东 济南 250357

**摘要:**针对当前桥梁裂缝图像检测精度较低, 算法规模较大不便于部署在资源受限的边缘设备等问题, 提出一种改进 YOLOv11n (you only look once version 11 nano) 的轻量级桥梁裂缝图像检测算法, 通过融合 ShuffleNetV2 模块与跨尺度融合模块 (cross-scale fusion module, CCFM), 构建 Shuffle-CCFM 结构, 提高多尺度特征的融合能力并降低算法参数量; 在跨通道局部空间注意力 (cross-channel partial spatial attention, C2PSA) 模块中引入倒置残差块注意力机制 (inverted residual mobile block with attention, iRMB), 形成 C2PSA-iRMB 模块, 提高算法对复杂裂缝细节的识别能力, 并增强同一裂缝结构区域内空间长距离特征的关联建模能力; 在 C3k2 模块中集成小波卷积 (wavelet transform convolution, WTConv), 形成 C3k2-WTConv 模块, 提高模型在不同尺度下的特征提取能力; 采用动态上采样器 DySample 代替传统上采样模块, 根据特征图内容自适应调整采样位置, 提高上采样阶段的空间分辨率与细节还原能力。开展消融试验、对比试验和可视化检测效果试验验证改进 YOLOv11n 算法的检测性能, 试验结果表明: 相较于 YOLOv11n 算法, 引入 Shuffle-CCFM 结构、C2PSA-iRMB 模块、C3k2-WTConv 模块和 DySample 模块后的改进 YOLOv11n 算法的参数量  $N_p$ 、计算量  $N_f$ 、权重文件大小  $T$  分别减小 27.5%、23.8%、32.7%, 交并比阈值为 50 时平均精度均值  $E_{mAP50}$ 、交并比阈值从 50 增至 95 时平均精度均值  $E_{mAP50-95}$  和召回率  $R$  分别提高 1.6%、3.8%、0.4%, 算法轻量化和检测精度明显提高; 改进 YOLOv11n 算法对桥梁裂缝图像的检测精度和性能指标明显优于 YOLOv5n、YOLOv6n、YOLOv8n、YOLOv10n 等轻量级算法, 适合部署于计算资源受限的边缘设备; 改进 YOLOv11n 算法在桥梁裂缝可视化检测试验中对检测结果精确率有更高的置信度, 对尺寸微小、形态复杂的裂缝细节捕捉能力较强, 在复杂背景下具有较强的抗干扰能力。

**关键词:** 桥梁裂缝图像检测; YOLOv11n; ShuffleNetV2; CCFM; iRMB; WTConv; DySample

中图分类号: U446.2; TP183

文献标志码: A

文章编号: 1672-0032(2025)06-0095-11

引用格式: 孙伟, 刘文江. 基于改进 YOLOv11n 的轻量级桥梁裂缝图像检测算法[J]. 山东交通学院学报, 2025, 33(6): 95-105.

SUN Wei, LIU Wenjiang. Lightweight bridge crack image detection algorithm based on improved YOLOv11n[J]. Journal of Shandong Jiaotong University, 2025, 33(6): 95-105.

## 0 引言

随着桥梁服役年限的延长, 混凝土结构中易产生碳化、裂缝以及损伤剥落等多种病害, 其中以裂缝最常见<sup>[1]</sup>。裂缝可导致混凝土保护层脱落, 进而引发钢筋锈蚀, 削弱桥梁的耐久性, 降低结构承载力, 甚至引发桥梁垮塌等严重事故, 影响交通运行安全。目前, 识别桥梁裂缝主要依赖人工巡检, 效率较低, 成本

收稿日期: 2025-06-11

基金项目: 交通部交通运输行业重点科技项目(2022-ZD4-066)

第一作者简介: 孙伟(2001—), 男, 山东泰安人, 硕士研究生, 主要研究方向为目标识别和深度学习, E-mail: 1023235632@qq.com。

\*通信作者简介: 刘文江(1971—), 男, 山东潍坊人, 教授, 硕士研究生导师, 工学博士, 主要研究方向为检测技术与自动化装置、交通安全, E-mail: sdjt\_lwj@163.com。

较高,且存在较大的安全风险<sup>[2]</sup>。随着计算机视觉和深度学习理论的迅速发展,基于深度学习理论的数字图像目标自动检测技术广泛应用于裂缝检测领域<sup>[3]</sup>。相较传统人工检测,数字图像目标自动检测能简化繁琐的图像预处理步骤,量化分析裂缝特征,提高检测精度和效率。

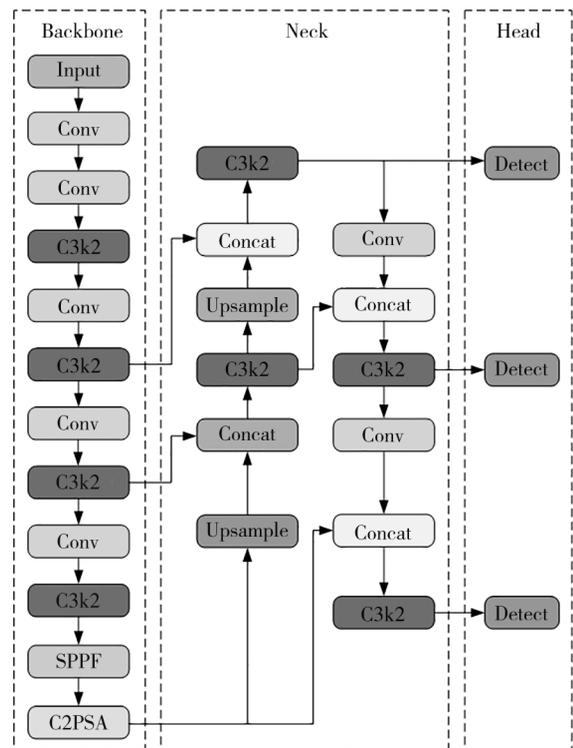
近年来,YOLO(you only look once)系列模型在图像处理领域展现较高的检测精度和实时性<sup>[4]</sup>,在桥梁裂缝检测领域受到广泛关注。刘永胜等<sup>[5]</sup>采用轻量级 MobileNetv2 网络作为单次多框检测器(single shot multibox detector, SSD)算法的骨干网络,引入通道注意力机制 SENet 结构,提高 SSD 算法在实际工程裂缝病害中的检测精度,该方法在复杂环境下检测精度较低。黄可原等<sup>[6]</sup>基于一阶段目标检测算法 YOLOv5(you only look once version 5),在 Convolution、Batch Normalization、SiLU 模块中融入卷积注意力机制模块以减少对图像特征提取的影响,在骨干网络末端添加坐标注意力模块提高定位精度,该方法在小规模图像数据集下检测精度较高,对极细裂缝的检测能力验证不足。王枫等<sup>[7]</sup>提出基于三维数字孪生的桥梁病害自动巡检方法,通过引入 Lion 优化器和优化损失函数改进 YOLOv8(you only look once version 8)算法,检测桥梁裂缝和露筋病害,该方法实现了桥梁病害的三维可视化和量化分析,但存在检测系统硬件成本高及复杂环境需人工干预的问题。现有桥梁病害检测方法存在检测精度与检测速度难以兼顾的问题,在裂缝特征提取与学习方面表现不足,难以应对实际复杂场景的桥梁裂缝检测任务<sup>[8]</sup>。

我国桥梁数量庞大,分布广泛,且检测任务繁重,亟需一种模型体积小、检测精度高和处理速度快的图像检测算法,实现高效可靠的桥梁病害检测。本文针对桥梁裂缝图像检测的实际需求,提出一种改进 YOLOv11n(you only look once version 11 nano)的轻量级图像检测算法,通过融合 ShuffleNetV2 模块与跨尺度融合模块(cross-scale fusion module, CCFM),以期提高多尺度特征的融合能力并降低模型参数量;在主干网络的跨通道局部空间注意力(cross-channel partial spatial attention, C2PSA)模块中引入倒置残差块注意力机制(inverted residual mobile block with attention, iRMB),以期提高模型对复杂裂缝细节的识别能力,并增强同一裂缝结构区域内空间长距离特征的关联建模能力;在颈部网络的 C3k2 模块中集成小波卷积(wavelet transform convolution, WTConv),以期提高模型在不同尺度下的特征提取能力;采用动态上采样器(dynamic sample, DySample)代替传统上采样模块,以期提高裂缝细节恢复能力,为桥梁裂缝图像检测提供参考。

## 1 YOLOv11n 算法

YOLOv11n、YOLOv11s、YOLOv11m、YOLOv11l 和 YOLOv11x<sup>[9]</sup> 5 个版本算法主要在模型参数规模、网络深度和计算复杂度上存在差异,适应不同应用场景和硬件平台的需求。通常模型越大,检测精度越高,但计算资源消耗越多。YOLOv11n 为轻量化版本,在目标检测领域显示较高的检测精度与计算效率,针对桥梁裂缝检测任务需求,本文选用 YOLOv11n 算法作为基础模型。YOLOv11n 算法由主干网络(Backbone Network)、颈部网络(Neck Network)和头部网络(Head Network)组成,结构如图 1 所示。

桥梁裂缝具有细小、狭长、对比度低等特征,在整幅图像中所占面积积极小,传统检测模型难以准确识别与定位这些微小目标,易出现漏检和误检现象。桥梁裂缝图像常伴随复杂背景干扰,混凝土纹理、光照变



注:Input 为特征图输入模块;Conv 为卷积模块;C3k2 为用于提取中间层的空间和通道特征模块;SPPF 为快速空间金字塔池化模块;Concat 为拼接模块;Upsample 为上采样模块;Detect 为根据提取的图像特征预测目标检测结果。

图 1 YOLOv11n 算法结构图

化、表面污染等<sup>[10]</sup>与裂缝特征在颜色、纹理等方面极其相似,影响裂缝特征提取效果,导致检测精度下降。针对桥梁裂缝这类小目标检测任务,需改进现有YOLOv11n算法,提高其特征提取能力。

作为轻量化版本的YOLOv11n算法虽已通过优化算法模型规模减小算法整体参数量,但部分网络结构仍存在参数较多、推理计算复杂等问题。在实际工程应用中,尤其是桥梁巡检等对实时性和部署环境要求较高的场景<sup>[11]</sup>,较大的模型参数量和计算量会降低检测速度,难以满足边缘设备或移动终端等资源受限环境下高效、快速的检测需求,需进一步改进与优化YOLOv11n算法结构,提高其计算效率。

## 2 改进YOLOv11n算法

为提高桥梁裂缝检测精度,减小算法模型参数量和计算量,对YOLOv11n算法进行如下改进。

1) 桥梁裂缝图像普遍呈现为尺寸微小、形态多变、与背景对比度低,检测任务对多尺度特征整合和小目标识别能力要求较高。通过融合ShuffleNetV2模块<sup>[12]</sup>与CCFM模块<sup>[13]</sup>,构建Shuffle-CCFM结构,提高多尺度信息整合和微小裂缝目标的检测能力,有效减小模型参数量,使算法更轻量化,便于在资源受限的边缘设备上部署。

2) 针对桥梁裂缝细小纹理及复杂形态对检测精确率的影响,在主干网络的C2PSA模块中引入iRMB模块<sup>[14]</sup>,形成C2PSA-iRMB模块,使算法能细致分辨裂缝的复杂细节,把图像中相隔较远但属于同一条裂缝的部分特征联系起来,更准确地识别裂缝的整体形状,提高对裂缝细节的捕捉能力和检测精度。

3) 桥梁裂缝在同一图像中存在跨度大和多尺度分布特征,YOLOv11n算法难以兼顾各类裂缝特征的有效提取和表达,在颈部网络的C3k2模块中集成WTConv<sup>[15]</sup>,形成C3k2-WTConv模块,通过多频率响应扩展卷积感受野,增强特征表达能力,有效提高模型在复杂场景下的适应性,控制模型参数规模。

4) 引入DySample<sup>[16]</sup>模块,弥补传统上采样模块在细节恢复能力方面的不足。DySample模块采用动态点重采样策略,根据特征图内容自适应调整采样位置,有效提高上采样阶段的空间分辨率与细节还原能力。改进YOLOv11n算法结构如图2所示。

### 2.1 Shuffle-CCFM结构

传统卷积神经网络算法采用加深网络结构或增加参数量的方法提高检测性能,但网络结构加深或参数量增加导致算法冗余度高,难以在资源受限的设备上部署<sup>[17]</sup>。ShuffleNetV2模块采用通道分割与通道混洗机制,可增加通道间信息交互频率,降低算法网络中的冗余特征信息,提高特征捕获效率,减小参数量;CCFM模块采用尺度空间理论和注意力机制,实现不同尺度特征间的深度融合,改善跨尺度信息的表达与融合能力<sup>[18]</sup>。主干网络ShuffleNetV2模块在提取图像特征时,输出各层级多尺度特征图并传送至颈部网络CCFM模块,由CCFM模块对图像特征进行跨尺度信息融合与深度优化,实现图像特征高效整合与多尺度增强,形成Shuffle-CCFM结构。Shuffle-CCFM结构能有效增强小尺度裂缝特征的表达能力和检测精确性,且使算法轻量化,提高检测效率。

ShuffleNetV2模块的网络架构由标准计算模块与下采样模块构成,网络架构如图3所示。在标准计算模块中,输入的特征图信息沿其通道维度被均等分割,分别进入左、右两个并行处理路径,右侧分支对分配的特征子集依次施加 $1 \times 1$ 卷积、 $3 \times 3$ 深度卷积和 $1 \times 1$ 卷积操作,左侧分支则对输入特征维持其原始形态,不执行任何变换操作,两分支的输出特征张量在通道维度上进行拼接操作和通道混洗操作

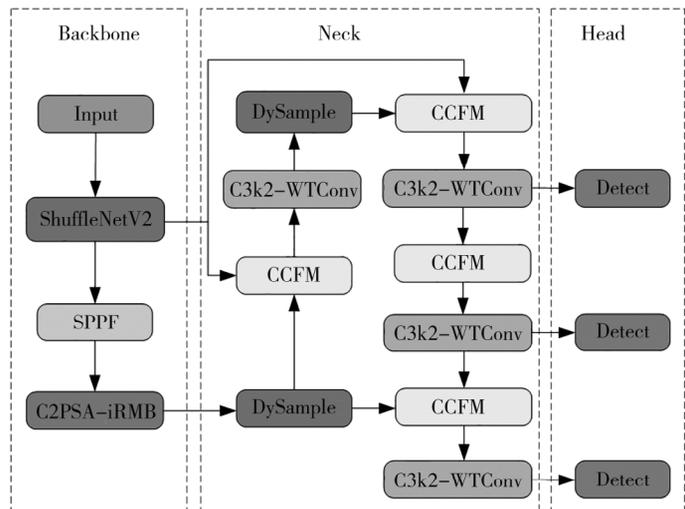
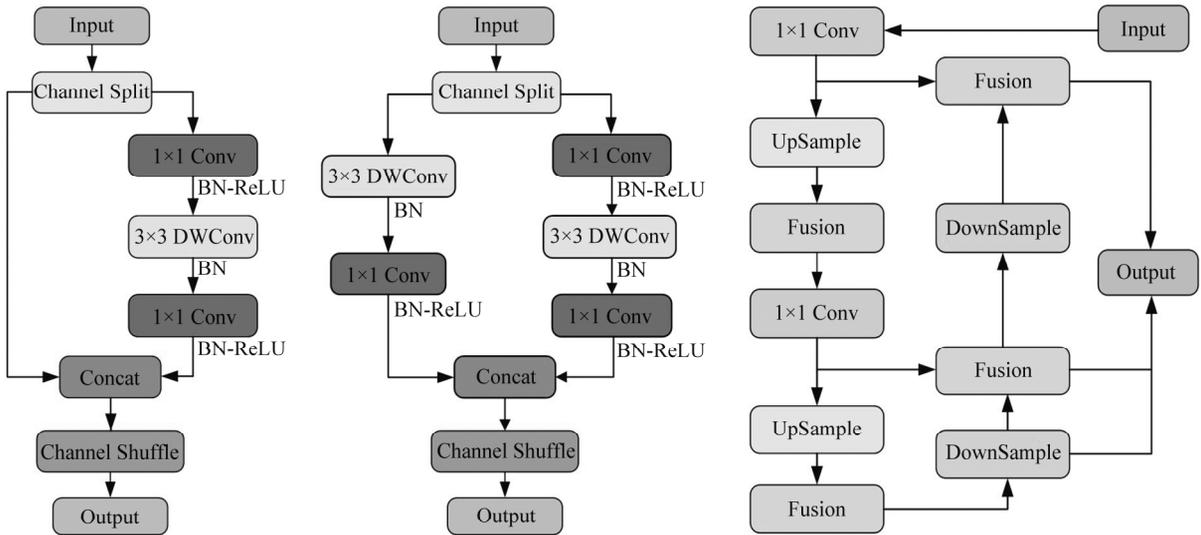


图2 改进YOLOv11n算法结构图

(Channel Shuffle)。在下采样模块中,输入的特征张量沿其通道维度被均等分割,右侧分支对分配的特征子集依次施加  $1 \times 1$  卷积、 $3 \times 3$  深度卷积和  $1 \times 1$  卷积操作,为实现空间维度的下采样,左侧分支对分配的特征子集依次施加  $3 \times 3$  深度卷积和  $1 \times 1$  卷积操作,降低特征图的空间分辨率,两分支的输出特征张量在通道维度上进行拼接操作和通道混洗操作,整合多尺度特征信息并促进通道间的特征流动。

CCFM 模块结合通道与空间注意力机制优化信息表达,网络架构如 4 所示。在颈部网络的融合路径中,CCFM 网络结构引入具有轻量化结构的  $1 \times 1$  卷积操作,能有效提取并融合初始阶段的特征图信息,降低模型的计算复杂度。在特征融合过程中,再引入具有轻量化结构的  $1 \times 1$  卷积操作,实现不同尺度特征图间的横向融合,增强模型的非线性特征表达能力。



a) 标准计算模块

b) 下采样模块

注: Channel Split 为通道拆分操作; DWConv 为深度卷积操作; BN (batch normalization) 为批量归一化; ReLU (rectified linear unit) 为修正线性单元; Output 为处理结果输出。

图 3 ShuffleNetV2 模块网络架构图

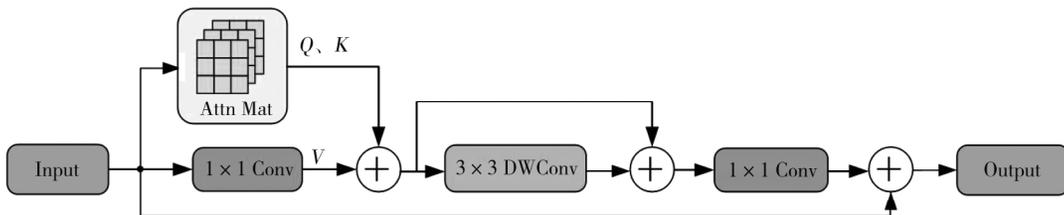
注: Fusion 为横向融合操作; DownSample 为下采样操作。

图 4 CCFM 模块网络架构图

### 2.2 C2PSA-iRMB 模块

桥梁裂缝特征通常表现为细小、狭长、纹理复杂、背景干扰明显等特点,要求图像检测算法能捕捉长距离特征依赖关系。传统卷积结构局限于局部空间感受野,难以有效捕获长距离的特征关系<sup>[19]</sup>。在主干网络末端的 C2PSA 模块中引入 iRMB 模块,形成 C2PSA-iRMB 模块,增强算法对微弱特征(如裂缝细节)的敏感性和捕获能力,避免网络复杂度过高导致过拟合风险。iRMB 模块继承 Transformer 架构自注意力机制与倒置残差结构的优点,自注意力机制实现任意空间位置间的特征交互,增强算法的全局特征捕获能力;倒置残差结构降低模型冗余性,保障信息在网络中高效流通并有效缓解梯度消失问题。

iRMB 模块结构如图 5 所示。iRMB 模块为融合  $1 \times 1$  卷积、 $3 \times 3$  深度卷积与自注意力机制的混合架构。在该架构中, $1 \times 1$  卷积应用于通道维度的压缩与扩展,能优化模型的计算效能。 $3 \times 3$  深度卷积应用于捕获局部空间特征,自注意力机制应用于建立特征间的长距离全局依赖关系。



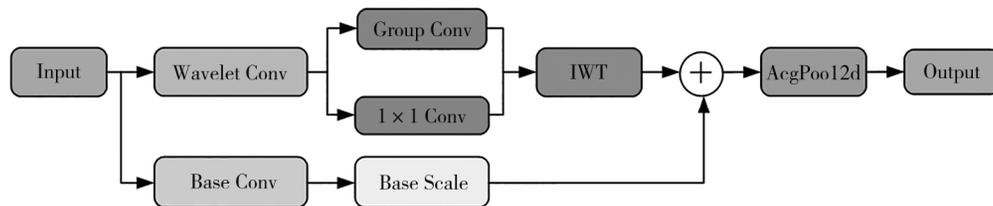
注: Attn Mat 为注意力矩阵; Q, K, V 分别为注意力机制中的查询 (Query)、键 (Key) 和值 (Value)。

图 5 iRMB 模块结构图

### 2.3 C3k2-WTConv 模块

裂缝图像通常在多尺度空间内表现出不同的特征频率和空间细节,传统卷积多基于空间域进行特征提取,对含不同频率特征的裂缝图像辨别能力有限,导致尺度适应性不足<sup>[20]</sup>。小波卷积(WTConv)模块基于小波变换理论,通过多频率分解方式有效捕获图像在不同尺度下的频率特征。小波变换能在频率域内将信号分解到不同频带,增强算法模型对高频细节(裂缝边缘特征)与低频结构特征(裂缝本体区域)的区分能力。此外,小波域内的卷积操作能更高效地利用多尺度频带信息,提高裂缝特征的泛化性和鲁棒性<sup>[21]</sup>。在颈部网络的C3k2模块中集成WTConv,形成C3k2-WTConv模块,具体操作为在C3k2模块瓶颈层单元中以WTConv替代传统的空间卷积操作,当瓶颈层的输入与输出通道数一致时,WTConv用作主卷积层,实现对输入特征的多尺度小波分解与空间重建。该设计无需调整C3k2模块的整体架构,通过替换模块融合WTConv,增强算法对多尺度和细粒度特征的表达能力。

WTConv模块结构如图6所示,WTConv模块采用小波变换卷积(Wavelet Conv)操作,实现多尺度、多频域的信息分解与融合,提高算法模型的特征表达能力。输入特征图经小波变换操作,将不同尺度的频带(高频和低频成分)分别进行卷积处理后提取深层特征,利用分组卷积(Group Conv)或 $1 \times 1$ 卷积操作减轻参数处理负担,融合(如拼接或相加)不同尺度频带特征,实现多尺度信息的整合,通过逆小波变换(Inverse Wavelet Transform, IWT)将融合的多频域特征进行空间重建,保持特征信息的空间一致性。为保证小波变换卷积与基础卷积(Base Conv)输出的特征图空间尺寸一致,采用二维平均池化(AvgPool2d)下采样操作处理特征图信息,使输出特征图空间维度一致。



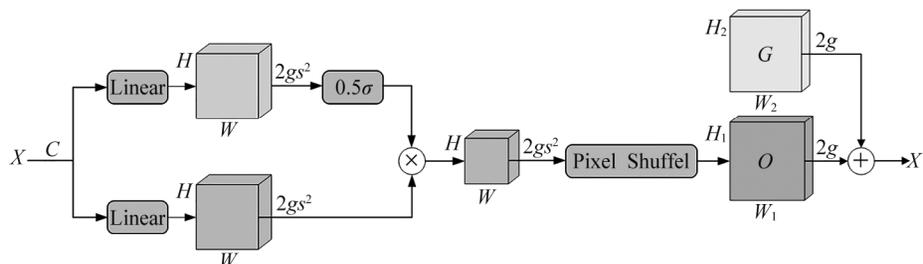
注: Base Scale 为特征图基准缩放比例,是多尺度处理框架中的核心参数。

图6 WTConv 模块结构图

### 2.4 DySample 模块

在裂缝检测任务中,上采样阶段直接决定小尺度特征的细节恢复质量。传统上采样模块多为固定规则的插值方法,空间重建的适应性较差,易丢失细节特征,制约检测性能<sup>[22]</sup>。DySample模块基于动态采样理论,通过特征内容自适应调整采样位置,根据输入特征的局部变化自适应提高空间采样密度,提高特征图空间重构的精确度,增强细节特征的恢复效果。

Dysample模块的采样点生成器结构如图7所示。采样点生成器对输入特征图采样集 $X$ 通过线性变换(Linear)和像素洗牌技术(Pixel Shuffle),结合范围因子(静态范围因子或动态范围因子)调节偏移特征图偏移量 $O$ ,与作用域特征图原始网格位置 $G$ 相加得到输出特征图采样集 $X'$ 。



注:  $C, H, W$  分别为输入特征图的通道数、高度和宽度;  $g$  为采样对应点数量;  $s$  为上采样比例因子;  $\sigma$  为 Sigmoid 函数, 用于生成范围因子;  $O$  为偏移特征图偏移量;  $H_1, W_1$  分别为偏移特征图的高度和宽度;  $G$  为作用域特征图原始网格位置;  $H_2, W_2$  分别为作用域特征图的高度和宽度。

图7 DySample 模块采样点生成器结构图

### 3 试验与结果分析

#### 3.1 试验准备

试验运行环境配置:高性能计算机操作系统为 Ubuntu22.04, CPU 为虚拟机 10vCPU, GPU 为 RTX3090, 内存为 60 GB, 显存为 24 GB; 并行计算平台为 CUDA12.4; 深度学习框架为 PyTorch2.1.0; 编程软件为 python3.10.14。设置试验训练次数为 300, 迭代为 64 批次, 学习率为 0.01, 优化器为 SGD (Stochastic Gradient Descent), 输入图片分辨率为 640 像素×640 像素。

试验数据集来源于 Roboflow 官方网站 (<https://roboflow.com>) 公开发布的桥梁裂缝图像数据集, 通过软件 LabelImg 标注 2 934 张桥梁裂缝图像中的裂缝, 采用随机划分的方法将数据集按 6:2:2 划分为训练集、测试集和验证集。

#### 3.2 评估指标

在桥梁裂缝图像检测算法中, 选取精确率  $P$ 、召回率  $R$ 、平均精度  $E_{mAP}$ 、权重文件大小  $T$ 、参数量  $N_p$ 、计算量  $N_f$  作为评估指标<sup>[23-26]</sup>。

精确率

$$P = N_{TP} / (N_{TP} + N_{FP}) \times 100\%$$

式中:  $N_{TP}$  为算法检测正确的目标数,  $N_{FP}$  为算法检测错误的目标数。

召回率

$$R = N_{TP} / (N_{TP} + N_{FN}) \times 100\%$$

式中  $N_{FN}$  为算法漏检的目标数。

平均精度均值  $E_{mAP}$  为所有样本检测精确率的平均值,  $E_{mAP}$  越大, 算法的检测性能越好, 通常采用交并比 (intersection over union, IoU) 阈值为 50 时的平均精度均值  $E_{mAP50}$  及 IoU 阈值从 50 增至 95 时的平均精度均值  $E_{mAP50-95}$  作为算法检测结果预测精度的评价指标,  $E_{mAP50}$  和  $E_{mAP50-95}$  越大, 算法检测结果预测精度越高。参数量  $N_p$  反映算法模型的复杂度, 参数量越大, 算法的表达力越强, 理论上可捕捉更复杂的特征信息, 导致算法的计算复杂度越高, 检测时消耗的计算时间越长。计算量  $N_f$  是衡量 CPU、GPU 等硬件设备在处理图形和深度学习算法的性能指标, 反映算法模型的计算负担, 计算量越大, 要求硬件设备的浮点运算能力越强。权重文件大小  $T$  为深度学习算法参数文件的存储空间, 反映算法模型的存储开销,  $T$  主要取决于算法参数的数量和每个参数的存储位宽。

#### 3.3 消融试验

为验证改进 YOLOv11n 算法对桥梁裂缝的检测性能, 探究 Shuffle-CCFM 结构、C2PSA-iRMB 模块、C3k2-WTConv 模块和 DySample 模块对 YOLOv11n 算法的改进效果, 以 YOLOv11n 算法为基准算法, 开展 5 组消融试验, 试验结果如表 1 所示。

表 1 消融试验结果

基准算法	Shuffle-CCFM	C2PSA-iRMB	C3k2-WTConv	DySample	$E_{mAP50}/\%$	$E_{mAP50-95}/\%$	$R/\%$	$N_p/10^6$	$N_f/10^9$	$T/MB$
	×	×	×	×	94.3	68.3	90.1	2.18	6.3	5.2
	√	×	×	×	94.0	66.7	88.6	1.73	4.8	3.7
YOLOv11n	√	√	×	×	95.0	68.1	90.5	1.75	4.9	3.7
	√	√	√	×	94.8	70.0	90.7	1.58	4.7	3.5
	√	√	√	√	95.8	70.9	90.5	1.58	4.8	3.5

注: √为在基础算法中采用该结构或模块, ×为在基础算法中未采用该结构或模块。

由表 1 可知: 相较于 YOLOv11n 算法, 引入 Shuffle-CCFM 结构后,  $N_p$ 、 $N_f$ 、 $T$  分别减小 20.6%、23.8%、

28.8%,算法模型轻量化效果明显,但 $E_{mAP50}$ 、 $E_{mAP50-95}$ 和 $R$ 出现小幅度减小,表明引入 Shuffle-CCFM 结构后显著减小算法复杂度、计算负担和存储空间,但略限制特征表达能力,尤其在跨尺度融合初期可能损失部分细节信息;继续引入 C2PSA-iRMB 模块后, $N_p$ 、 $N_f$ 比引入 Shuffle-CCFM 结构仅增大 1.2%、2.1%, $E_{mAP50}$ 、 $E_{mAP50-95}$ 和 $R$ 明显提高,分别提高 1.1%、2.1%、2.1%,表明倒置残差注意力机制捕获长距离依赖特征信息效果明显,能高效提取并整合裂缝细节与长距离空间关联信息;继续引入 C3k2-WTConv 模块后, $N_p$ 、 $N_f$ 、 $T$ 比引入 Shuffle-CCFM 结构、C2PSA-iRMB 模块分别减小 9.7%、4.1%、5.4%, $E_{mAP50-95}$ 、 $R$ 分别提高 2.8%、0.2%, $E_{mAP50}$ 仅降低 0.2%,表明小波卷积模块通过频率域的多尺度分析有效分离不同尺度的裂缝特征,增强算法模型对复杂尺度变化的稳定性和泛化能力,评估指标 $E_{mAP50-95}$ 明显提高;继续引入 DySample 模块后, $N_p$ 、 $T$ 保持不变, $N_f$ 比引入 Shuffle-CCFM 结构、C2PSA-iRMB 模块、C3k2-WTConv 模块增大 2.1%, $E_{mAP50}$ 、 $E_{mAP50-95}$ 分别提高 1.1%、1.3%, $R$ 仅降低 0.2%,表明动态上采样器通过自适应采样增强主要特征的细节恢复,整体检测精度明显提高,由于算法在高置信度区域的特征表达突出,部分边缘或低对比度的微小裂缝目标被抑制,导致召回率略降。与基准算法 YOLOv11n 相比,引入 Shuffle-CCFM 结构、C2PSA-iRMB 模块、C3k2-WTConv 模块和 DySample 模块后的改进 YOLOv11n 算法的 $N_p$ 、 $N_f$ 、 $T$ 分别减小 27.5%、23.8%、32.7%,算法轻量化效果明显; $E_{mAP50}$ 、 $E_{mAP50-95}$ 和 $R$ 分别提高 1.6%、3.8%、0.4%,算法检测精度明显提高。

### 3.4 对比试验

为进一步验证改进后的 YOLOv11n 算法是否提高桥梁裂缝检测精度及算法轻量化效果,采用相同试验运行环境配置和数据集对目前广泛使用的 YOLOv5n、YOLOv6n、YOLOv8n、YOLOv10n 等轻量化算法分别进行桥梁裂缝图像检测试验,对比试验结果如表 2 所示。

表 2 不同算法下桥梁裂缝图像检测对比试验结果

算法	$E_{mAP50}/\%$	$E_{mAP50-95}/\%$	$R/\%$	$N_p/10^6$	$N_f/10^9$	$T/MB$
YOLOv5n	90.3	60.8	88.5	2.18	5.8	4.7
YOLOv6n	93.6	68.7	89.5	4.16	11.5	8.6
YOLOv8n	93.9	66.4	91.0	2.68	6.8	5.7
YOLOv10n	88.8	62.3	84.2	2.69	8.2	5.8
改进 YOLOv11n	95.8	70.9	90.5	1.58	4.8	3.5

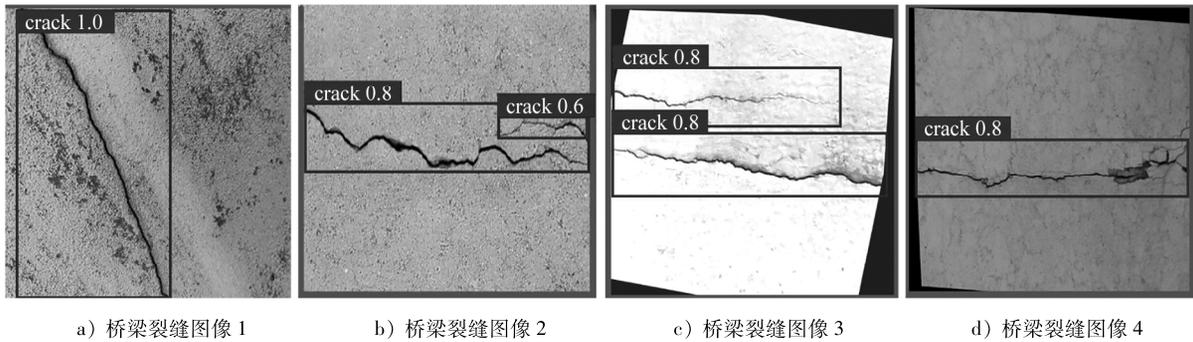
由表 2 可知:改进 YOLOv11n 算法的 $E_{mAP50}$ 比 YOLOv5n、YOLOv6n、YOLOv8n、YOLOv10n 算法分别提高 6.1%、2.4%、2.0%、7.9%, $E_{mAP50-95}$ 分别提高 16.6%、3.2%、6.8%、13.8%, $N_p$ 分别减小 27.5%、62.0%、41.0%和 41.3%; $N_f$ 分别减小 17.2%、58.3%、29.4%、41.5%; $T$ 分别减小 25.5%、59.3%、38.6%、39.7%。改进 YOLOv11n 算法的 $R$ 比 YOLOv5n、YOLOv6n、YOLOv10n 算法分别提高 2.3%、1.1%、7.5%,比 YOLOv8n 算法仅降低 0.5%。改进 YOLOv11n 算法对桥梁裂缝图像的检测精度和性能指标明显优于 YOLOv5n、YOLOv6n、YOLOv8n、YOLOv10n 算法,算法轻量化效果明显,便于部署到硬件资源有限的环境中。

### 3.5 可视化检测效果试验

为更直观地验证改进算法的有效性,对比分析改进 YOLOv11n 算法与 YOLOv11n、YOLOv10n、YOLOv8n、YOLOv6n、YOLOv5n 在桥梁裂缝图像检测任务中的可视化检测效果,如图 8~13 所示。

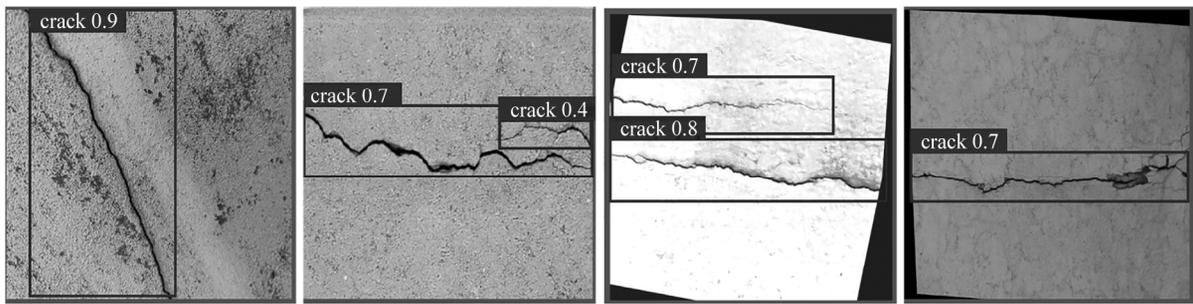
由图 8~13 可知:对桥梁裂缝图像 1,改进 YOLOv11n 算法检测的裂缝(crack)置信度为 1.0,高于其他检测算法,表明改进后的算法对检测结果精确率有更高的置信度;对桥梁裂缝图像 2 和 3,存在一条明显的主裂缝和一条尺寸较小、形态复杂的细微裂缝,改进 YOLOv11n 算法均能将两条裂缝检测出来且 crack 置信度均较高,其他检测算法对尺寸较小的细微裂缝的 crack 置信度较低,表明改进后的 YOLOv11n 算法对尺寸微小、形态复杂的裂缝细节捕捉能力较强;对桥梁裂缝图像 4,改进 YOLOv11n 算法准确地识

别裂缝与背景纹理,框选出真实裂缝,检测框与裂缝轮廓贴合紧密,且评估为较高的 crack 置信度, YOLOv11n、YOLOv10n、YOLOv8n、YOLOv6n 未正确框选出真实裂缝, YOLOv5n 虽正确框选出真实裂缝,但评估的 crack 置信度较低,表明改进后的算法在复杂背景下具有较强的抗干扰能力, YOLOv11n 等其他算法在裂缝相似背景纹理干扰下产生误判或降低识别可信度。综上所述,改进 YOLOv11n 算法在桥梁裂缝图像检测任务中对检测结果精确率有更高的置信度,对尺寸微小、形态复杂的裂缝细节捕捉能力较强,在复杂背景下具有较强的抗干扰能力。



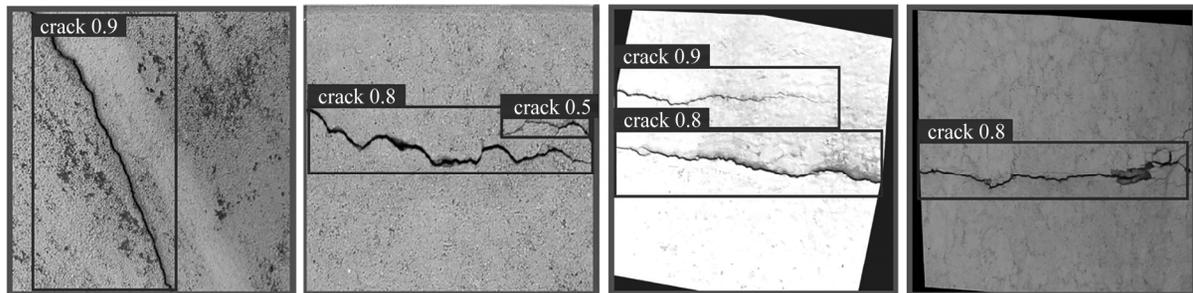
a) 桥梁裂缝图像 1      b) 桥梁裂缝图像 2      c) 桥梁裂缝图像 3      d) 桥梁裂缝图像 4

图 8 改进 YOLOv11n 算法对桥梁裂缝图像的可视化检测效果



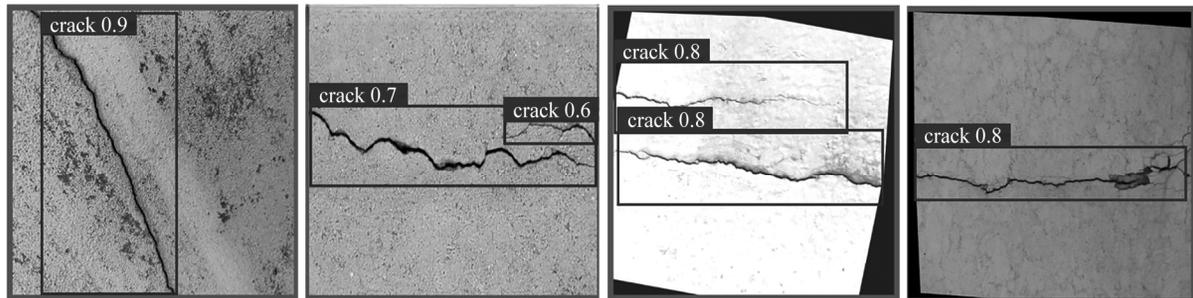
a) 桥梁裂缝图像 1      b) 桥梁裂缝图像 2      c) 桥梁裂缝图像 3      d) 桥梁裂缝图像 4

图 9 YOLOv11n 算法对桥梁裂缝图像的可视化检测效果



a) 桥梁裂缝图像 1      b) 桥梁裂缝图像 2      c) 桥梁裂缝图像 3      d) 桥梁裂缝图像 4

图 10 YOLOv10n 算法对桥梁裂缝图像的可视化检测效果



a) 桥梁裂缝图像 1      b) 桥梁裂缝图像 2      c) 桥梁裂缝图像 3      d) 桥梁裂缝图像 4

图 11 YOLOv8n 算法对桥梁裂缝图像的可视化检测效果

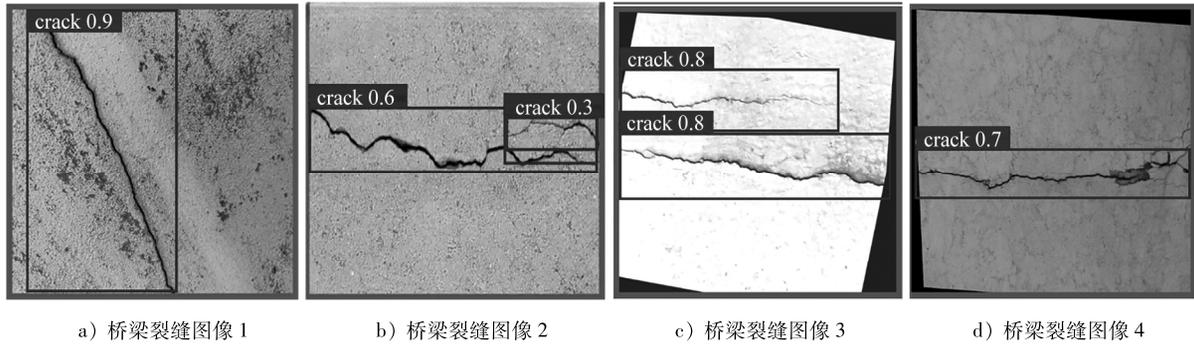


图 12 YOLOv6n 算法对桥梁裂缝图像的可视化检测效果

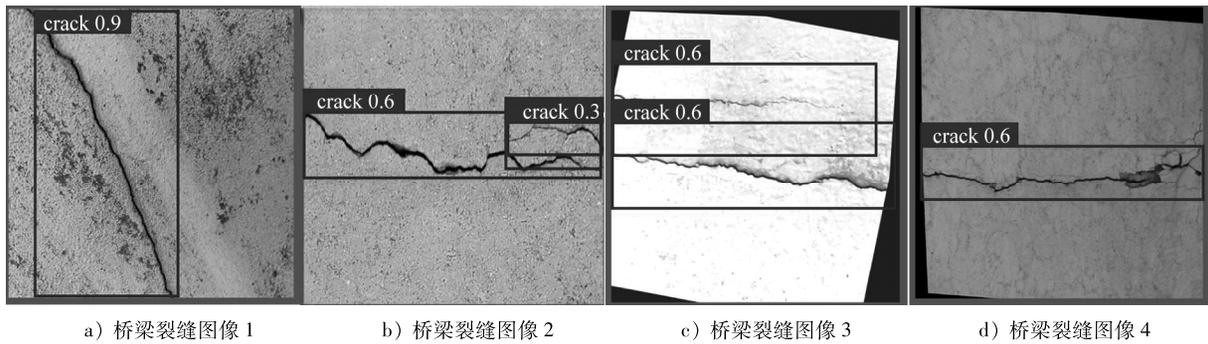


图 13 YOLOv5n 算法对桥梁裂缝图像的可视化检测效果

## 4 结论

为解决桥梁结构安全与健康监测维护中裂缝病害检测问题,提高桥梁裂缝检测精度,提出一种改进 YOLOv11n 的轻量级桥梁裂缝图像检测算法。

1) 构建 Shuffle-CCFM 结构,提高多尺度信息整合和微小裂缝目标的检测能力,减小模型参数量,使算法适合部署于资源受限的边缘设备;在 C2PSA 模块中引入 iRMB 模块,形成 C2PSA-iRMB 模块,使算法能细致分辨裂缝的复杂细节,准确识别裂缝的整体形状,进一步提高对裂缝细节的捕捉能力和检测精度;在 C3k2 模块中集成 WTConv,形成 C3k2-WTConv 模块,通过多频率响应扩展卷积感受野,增强特征表达能力,提高模型在复杂场景下的适应性,并进一步控制模型参数规模。在颈部网络引入 DySample 模块,根据特征图内容自适应调整采样位置,提高上采样阶段的空间分辨率与细节还原能力。

2) 进行消融试验,相较于 YOLOv11n 算法,引入 Shuffle-CCFM 结构、C2PSA-iRMB 模块、C3k2-WTConv 模块和 DySample 模块后,改进 YOLOv11n 算法的  $N_p$ 、 $N_f$ 、 $T$  分别减小 27.5%、23.8%、32.7%,算法轻量化效果明显; $E_{\text{mAP}50}$ 、 $E_{\text{mAP}50-95}$  和  $R$  分别提高 1.6%、3.8%、0.4%,算法检测精度明显提高。

3) 采用相同试验运行环境配置和数据集进行桥梁裂缝检测对比试验,改进 YOLOv11n 算法对桥梁裂缝图像的检测精度和性能指标明显优于 YOLOv5n、YOLOv6n、YOLOv8n、YOLOv10n 等轻量级算法,适合部署于计算资源受限的边缘设备。

4) 进行可视化检测效果试验,改进 YOLOv11n 算法在桥梁裂缝图像检测任务中对检测结果精确率有更高的置信度,对尺寸微小、形态复杂的桥梁裂缝细节捕捉能力较强,在复杂背景下具有较强的抗干扰能力。

未来可进一步研究改进 YOLOv11n 算法在复杂场景下的泛化能力与边缘设备部署适应性。

### 参考文献:

- [1] 林土淦,韦俊,刘学军,等.基于 CLAHE 与 U-Net 模型的无人机桥梁裂缝检测方法[J].西部交通科技,2024(9):

- 1-3.
- [2] 姬文鹏,贺骥,常勇,等.基于轻量化卷积神经网络的桥梁混凝土裂缝检测方法及其验证[J].公路交通技术,2023,39(1):125-132.
- [3] AI D H, JIANG G Y, LAM S K, et al. Computer vision framework for crack detection of civil infrastructure: a review[J]. Engineering Applications of Artificial Intelligence, 2023,117:105478.
- [4] REDMON J, DIVVALA S, GIRSHICK R, et al. You only look once: unified, real-time object detection[C]//Proceedings of IEEE Conference on Computer Vision and Pattern Recognition. Las Vegas, NV, USA:IEEE, 2016:779-788.
- [5] 刘永胜,熊吉光,游志杰,等.基于改进 SSD 算法的结构表面裂缝病害检测[J].华东交通大学学报,2023,40(6):1-7.
- [6] 黄可原,赵毅,胡楠,等.基于注意力机制的 YOLOv5 网络对混凝土桥梁裂缝识别的研究[J].科技通报,2024,40(9):71-76.
- [7] 王枫,吴华勇,赵荣欣,等.基于无人机桥梁表面病害巡检与三维数字孪生[J].城市道桥与防洪,2023(10):229-232.
- [8] HU S C, LIU X, WANG W J, et al. A universal structure of YOLO series small object detection models[C]//Computer Vision-ACCV 2024. Singapore: Springer, 2025:468-484.
- [9] KHANAM R, HUSSAIN M. YOLOv11: an overview of the key architectural enhancements[EB/OL]. (2024-10-23) [2025-5-10]. <https://doi.org/10.48550/arXiv.2410.17725>.
- [10] 方虹苏,石鑫雨,贺门涛,等.基于深度学习的桥梁裂缝病害检测算法研究[J].公路交通技术,2025,41(2):111-118.
- [11] 党良慧,张玉金,姜月武,等.基于多特征优选的图像拼接算法及系统设计[J].智能计算机与应用,2021,11(7):60-65.
- [12] MA N N, ZHANG X Y, ZHENG H T, et al. ShuffleNet V2: practical guidelines for efficient CNN architecture design [C]//Computer Vision-ECCV 2018. Cham, Germany: Springer International Publishing, 2018:122-138.
- [13] ZHAO Y A, LVW Y, XU S L, et al. DETRs beat YOLOs on real-time object detection[C]//Proceedings of IEEE/CVF Conference on Computer Vision and Pattern Recognition. Seattle, WA, USA: IEEE, 2024:16965-16974.
- [14] ZHANG J N, LI X T, LI J, et al. Rethinking mobile block for efficient attention-based models [C]//Proceedings of IEEE/CVF International Conference on Computer Vision. Paris, France: IEEE, 2023:1389-1400.
- [15] FINDER S E, AMOYAL R, TREISTER E, et al. Wavelet convolutions for large receptive fields[EB/OL]. (2024-07-15) [2025-05-10]. <https://doi.org/10.48550/arXiv.2407.05848>.
- [16] LIU W Z, LU H, FU H T, et al. Learning to upsample by learning to sample [C]//Proceedings of IEEE/CVF International Conference on Computer Vision. Paris, France: IEEE, 2023:6004-6014.
- [17] 朱俊彬.基于卷积神经网络的路面裂缝识别研究[D].贵阳:贵州大学,2024.
- [18] DING K Y, DING Z H, ZHANG Z B, et al. Sed-yolo: a novel object detection method for efficient road crack detection [J]. Multimedia Systems, 2024,30(6):1-16.
- [19] 程鑫森,张雪松,曹冰洁,等.改进 RT-DETR 的小目标检测方法研究[J].计算机工程与应用,2025,61(15):144-155.
- [20] 张伟旗,蔡佳楠,艾兵,等.基于重校准与上下文引导的无人机视角下小目标检测算法[J/OL].重庆工商大学学报(自然科学版). (2025-05-13) [2025-05-25]. <https://kns.cnki.net/kcms/detail/50.1155.N.20250513.1108.002.html>.
- [21] GUO H R, WANG B, FANG T, et al. A novel ViT model with wavelet convolution and SLAttention modules for underwater acoustic target recognition[J]. Journal of Marine Science and Engineering, 2025,13(4):13040634.
- [22] 钱藏龙,汤文兵.基于 SD-YOLOv11 的航拍道路病害检测模型[J].湖北民族大学学报(自然科学版),2025,43(2):224-230.
- [23] 刘印,龚长友,徐国栋.基于改进 YOLOv10 的轻量化目标检测算法[J].自动化与信息工程,2025,46(1):29-35.
- [24] 张兴盛,阮久宏,沈本兰,等.基于改进 YOLOv5s 算法的轨道扣件缺陷检测[J].山东交通学院学报,2025,33(2):10-18.
- [25] LUO Y C, CI Y S, JIANG S X, et al. A novel lightweight real-time traffic sign detection method based on an embedded device and YOLOv8[J]. Journal of Real-Time Image Processing, 2024,21(2):1-10.

[26] 许云涛,焦培刚,刘家齐.基于YOLOv5m的电机换向器缺陷检测[J].山东交通学院学报,2024,32(2):10-18.

## Lightweight bridge crack image detection algorithm based on improved YOLOv11n

SUN Wei<sup>1</sup>, LIU Wenjiang<sup>2\*</sup>

1. School of Rail Transportation, Shandong Jiaotong University, Jinan 250357, China;

2. School of Aeronautics, Shandong Jiaotong University, Jinan 250357, China

**Abstract:** To address the problems of low accuracy in current bridge crack image detection and large algorithm scale that is inconvenient for deployment on resource-constrained edge devices, a lightweight bridge crack image detection algorithm based on improved YOLOv11n (you only look once version 11 nano) is proposed. By integrating the ShuffleNetV2 module with CCFM (cross-scale fusion module), a Shuffle-CCFM structure is constructed to enhance multi-scale feature fusion capability while reducing algorithm parameters. The iRMB (inverted residual mobile block with attention) is introduced into the C2PSA (cross-channel partial spatial attention) module to form the C2PSA-iRMB module, which improves the algorithm's recognition capability for complex crack details and enhances the correlation modeling capability of spatially distant features within the same crack structure region. WTConv (wavelet transform convolution) is integrated into the C3k2 module to form the C3k2-WTConv module, improving the model's feature extraction capability at different scales. DySample is adopted to replace the traditional upsampling module, adaptively adjusting sampling positions according to feature map content to enhance spatial resolution and detail restoration capability during the upsampling stage. Ablation experiments, comparative experiments, and visualization detection effect experiments are conducted to evaluate the detection performance of the improved YOLOv11n algorithm. The experimental results show that: compared with the YOLOv11n algorithm, after introducing the Shuffle-CCFM structure, C2PSA-iRMB module, C3k2-WTConv module, and DySample module, the improved YOLOv11n algorithm's params  $N_p$ , computation cost  $N_f$ , and weight file size  $T$  are reduced by 27.5%, 23.8%, and 32.7%, respectively, while mean average precision at intersection over union threshold of 50  $E_{\text{mAP}50}$ , mean average precision at intersection over union threshold from 50 to 95  $E_{\text{mAP}50-95}$ , and recall  $R$  increase by 1.6%, 3.8%, 0.4% respectively, demonstrating significant improvements in algorithm lightweighting and detection accuracy. The improved YOLOv11n algorithm's detection accuracy and performance indicators for bridge crack images are significantly superior to lightweight algorithms such as YOLOv5n, YOLOv6n, YOLOv8n, and YOLOv10n, making it suitable for deployment on edge devices with limited computational resources. The improved YOLOv11n algorithm demonstrates higher confidence in detection result precision in bridge crack visualization detection experiments, exhibits stronger capability in capturing details of minute-sized and morphologically complex cracks, and possesses stronger anti-interference capability in complex backgrounds.

**Keywords:** bridge crack image detection; YOLOv11n; ShuffleNetV2; CCFM; iRMB; WTConv; DySample

(责任编辑:边文超)