

文章编号:1671-1637(2024)06-0135-13

复杂背景下基于改进 YOLO v8s 的 混凝土桥梁裂缝检测方法

蒋仕新¹, 邹小雪¹, 杨建喜¹, 李昊², 黄雪梅¹, 李韧¹, 张廷萍¹,
刘新龙¹, 王笛¹

(1. 重庆交通大学 信息科学与工程学院, 重庆 400074; 2. 重庆交通大学 交通运输学院, 重庆 400074)

摘要:针对混凝土桥梁裂缝因背景复杂、细小模糊导致检测精度较差的问题,提出了一种基于改进 YOLO v8s 的混凝土桥梁裂缝检测算法,精确定位了裂缝位置信息,并测量了裂缝长、宽等关键参数;以 YOLO v8s 模型为基础,引入全维度动态卷积(ODConv)获取特征图中更丰富的上下文信息,以增强目标特征提取能力,提高对细小模糊裂缝的检测能力;采用改进通道注意力模块构建了级联双层特征改进注意力(C2f-MA)融合模块,以挖掘特征图中更多的纹理信息,进一步使网络更加关注裂缝特征,从而抑制无关背景信息的干扰,提高复杂背景下的裂缝检测效果;使用焦距交并比(WIoU)损失函数来解决低质量样本的识别问题,进一步优化了模型的收敛速度和检测准确率;在桥梁检测报告中筛选出存在裂缝细小模糊、阴影、人工画线、杂草等具有复杂背景的裂缝图像,通过人工标注的方式建立了复杂背景条件下桥梁裂缝图像数据集;以召回率、平均精度和模型存储容量作为量化评价指标,并依次通过对比试验及消融试验来对模型进行综合评估。研究表明:改进 YOLO v8s 算法的召回率、平均精度和模型存储容量分别为 0.829、0.893 和 11.14 MB,其综合评价指标优于基准方法 YOLO v8s 和其他目标检测模型,证明了提出的算法在复杂背景下具有良好的鲁棒性。

关键词:桥梁工程;改进 YOLO v8s 算法;深度学习;裂缝检测;目标检测

中图分类号:U446 **文献标志码:**A **DOI:**10.19818/j.cnki.1671-1637.2024.06.009

Concrete bridge crack detection method based on improved YOLO v8s in complex backgrounds

JIANG Shi-xin¹, ZOU Xiao-xue¹, YANG Jian-xi¹, LI Hao², HUANG Xue-mei¹,
LI Ren¹, ZHANG Ting-ping¹, LIU Xin-long¹, WANG Di¹

(1. School of Information Science and Engineering, Chongqing Jiaotong University, Chongqing 400074, China;

2. College of Traffic and Transportation, Chongqing Jiaotong University, Chongqing 400074, China)

Abstract: To address the issue of low detection accuracy for cracks in concrete bridges caused by complex backgrounds, as well as small and obscure features, the crack location information was accurately located, and key parameters such as crack length and width were measured based on

收稿日期:2024-05-11

基金项目:国家自然科学基金项目(62003063, 62103068);重庆市自然科学基金项目(CSTB2022NSCQ-MSX1599, cstc2020jcyj-msxmX0047);重庆市教育委员会科学技术研究项目(KJZD-K202400709, KJZD-M202300703, KJQN202100748)

作者简介:蒋仕新(1992-),男,重庆大足人,重庆交通大学副教授,工学博士,从事计算机视觉与桥梁健康监测研究。

引用格式:蒋仕新,邹小雪,杨建喜,等.复杂背景下基于改进 YOLO v8s 的混凝土桥梁裂缝检测方法[J].交通运输工程学报,2024,24(6):135-147.

Citation:JIANG Shi-xin, ZOU Xiao-xue, YANG Jian-xi, et al. Concrete bridge crack detection method based on improved YOLO v8s in complex backgrounds[J]. Journal of Traffic and Transportation Engineering, 2024, 24(6): 135-147.

the improved YOLO v8s algorithm for concrete bridge crack detection. Based on the YOLO v8s model, the omni-dimensional dynamic convolution (ODConv) was incorporated to capture richer contextual information in feature maps, enhancing the model's ability to extract the target features and detect small and obscure cracks. An improved channel attention module was used to develop the concatenated two-layer feature-modified attention (C2f-MA) fusion module, enabling the extraction of more texture information from feature maps. This modification further made the network focus on crack features, suppressing interferences from irrelevant background information and improving the crack detection performance in complex backgrounds. The weighted intersection over union (WIoU) loss function was introduced to address the challenge of low-quality sample recognition, optimizing the model's convergence speed and detection accuracy. Crack images with complex backgrounds such as small and obscure cracks, shadows, artificial lines, and weeds were screened in the bridge detection report. A bridge crack image dataset was established by manual annotation with complex background conditions. The model's performance was comprehensively evaluated through comparative and ablation experiments by taking recall, average precision and model storage capacity as quantitative evaluation indicators. Research results demonstrate that the improved YOLO v8s algorithm achieves recall, average precision of 0.829, 0.893 and 0.631, respectively, as well as the model storage capacity of 11.14 MB. Its comprehensive evaluation indicators outperforms the baseline YOLO v8s and other target detection models, validating that the proposed algorithm exhibits robust performance in complex backgrounds. 6 tabs, 12 figs, 30 refs.

Key words: bridge engineering; improved YOLO v8s algorithm; deep learning; crack detection; target detection

Author resume: JIANG Shi-xin(1992-), male, associate professor, PhD, shixinjiang@cqjtu.edu.cn.

Foundation items: National Natural Science Foundation of China (62003063, 62103068); Natural Science Foundation of Chongqing (CSTB2022NSCQ-MSX1599, cstc2020jcyj-msxmX0047); Science and Technology Research Project of Chongqing Municipal Education Commission (KJZD-K202400709, KJZD-M202300703, KJQN202100748)

0 引 言

桥梁是现代交通基础设施中不可或缺的重要组成部分,承载着人们的交通需求和经济发展重任。桥梁在投入使用后,受车辆行驶过程中轮胎带来的磨损冲击、大自然极端恶劣天气的侵蚀及修建初期遗留的结构缺陷等影响,需要交通养护部门定期检查桥面损伤。同时,部分老旧桥梁建设标准与当前不断增加的桥梁运力要求的负载不匹配,进一步恶化了桥面损伤程度^[1]。材料老化、车辆超载、运营环境恶劣等因素的影响,会造成混凝土桥梁出现裂缝、腐蚀、剥落等不同类别的结构损伤^[2],其中裂缝是混凝土桥梁表面损坏的典型标志。裂缝的检测与预防是桥梁养护的关键^[3],及时检测并修复裂缝可以避免桥梁裂缝继续恶化,在降低维修成本的同时,能保证正常交通秩序与驾驶人安全。

长期以来,桥梁裂缝检测主要依靠人工巡检,劳动强度大,耗时长,准确性低。基于机器学习^[4]和传统图像处理技术^[5]的裂缝提取与分析方法在实际应用中面临裂缝背景复杂、形态细小且不清晰的挑战,检测结果容易受到严重影响^[6]。随着人工智能和计算机硬件的快速发展,深度学习在计算机视觉领域广泛应用,基于卷积神经网络(Convolutional Neural Network, CNN)^[7]的目标检测技术成为目前裂缝检测领域中的研究热点。CNN能实现有效的特征提取^[8],但在识别桥梁表面裂缝时,存在如污渍、涂鸦、人工画线以及杂草等多种背景干扰因素,CNN在特征提取中难以区分裂缝和背景噪声,会影响到捕捉裂缝关键特征的能力,因此,部分研究者利用不同结构的卷积神经网络得到了许多不同的裂缝检测模型。Deng等^[9]将可变形卷积嵌入检测网络中,提高了检测器对平面外裂缝的检测精度;晏班夫

等^[10]为增强基于图像处理的路面表观病害检测效率和准确性,提出了基于更快速区域卷积神经网络(Faster Region-based Convolutional Neural Network, Faster R-CNN)与形态法的路面病害识别方法;余加勇等^[11]提出了一种基于无人机和用于生成目标分割掩码卷积神经网络(Mask Region-based Convolutional Neural Network, Mask R-CNN)的桥梁结构裂缝智能识别方法,将无人机和 Mask R-CNN相结合,有效实现了桥梁结构表面裂缝的高效识别和检测。与以 Faster R-CNN^[12]为典型代表的两阶段目标检测算法相比,单次检测多框检测器(Single Shot MultiBox Detector^[13], SSD)和 YOLO^[14]属于单阶段目标检测算法,这类算法提升了目标检测的效率和速度。彭雨诺等^[15]提出了基于改进 YOLO 的双网络桥梁表观病害快速检测算法,实现了裂缝病害的检测。目前,基于深度学习的目标检测算法在识别桥梁裂缝方面仍面临一些挑战,包括裂缝特征分布混乱、背景复杂、图像不清晰等问题,因此,如何提高裂缝检测的准确性仍需要进一步研究。

基于深度学习的目标检测算法主要集中于混凝土桥梁表面病害(如露筋、剥落、渗水)的检测,针对桥梁裂缝细小模糊,所处阴影、杂草、人工画线等干扰物影响的相关研究依然存在不足。鉴于此,本文结合改进 YOLO v8s 目标检测算法,提出一种复杂背景下混凝土桥梁裂缝检测方法,以解决桥梁裂缝存在细小模糊、多干扰物影响等问题,提高桥梁裂缝检测准确性,同时使用图像分割算法分割已检测到的裂缝;然后进行裂缝长宽测量,检测裂缝的量化信息;此外,本文构建了针对复杂背景下混凝土桥梁裂缝检测任务数据集,用于模型训练;最后,通过与多个目标检测算法进行对比试验,验证了提出的算法的有效性和优越性。

1 YOLO v8 目标检测算法

目标检测领域中,YOLO 目标检测算法发展迅速,Hussain^[16]在 2023 年提出了基于 YOLO v5 改进的目标检测网络 YOLO v8,该网络由主干网络、特征融合层和检测层组成,相较于 YOLO v5 网络结构,主要是在主干网络、预测头上进行改进。主干网络依旧使用跨阶段局部(Cross Stage Partial, CSP)提取主要特征,但将 YOLO v5 结构中的卷积特征融合模块 C3 替换为轻量化特征融合模块 C2f 模块,进一步实现网络轻量化,同时获得更加丰富

的梯度流信息;瓶颈层主要实现一系列特征融合与组合,采用特征金字塔网络(Feature Pyramid Network, FPN)-金字塔注意力网络(Pyramid Attention Network, PAN),即利用上采样操作对深层特征进行尺度放大,通过下采样操作对浅层特征进行尺度缩小,并将二者进行跨层融合连接,得到特征明显的加强特征,提升了网络的特征提取能力;检测头部分采用了目前主流的解耦头结构,使得分类和目标识别任务能够独立进行,有效发挥各自的优势,同时采用无锚框思想减少耗时与算力需求^[17]。YOLO v8 网络采用马赛克^[18]数据增强方法来提升网络性能并增强模型的鲁棒性,马赛克数据增强方法通过将随机选取的 4 个图像进行裁剪并拼接为一张新图像的方式来丰富训练集数据,可以增加数据集的多样性,使网络模型能够更好地学习和适应不同场景和背景条件下的目标检测任务。

YOLO v8 网络采用超参数调节机制,通过调整深度和宽度关键参数实现对网络结构深度和宽度的精确控制,它提供了 n、s、m、l、x 共 5 个版本,其中 n 版本是最小的网络模型,而 x 版本则是性能最强的网络模型,但网络模型太大,对试验条件要求过高,s、m、l 版本检测性能相差较小,但模型大小相差较大,因此,本文综合考虑检测性能和模型大小的因素,选择 s 版本,与原版模型相比,其深度比例为 0.33,宽度比例为 0.50,在保持相对较小的模型的同时尽可能提高检测性能。

目前,尽管 YOLO v8s 算法在目标检测任务中展现了出色的性能,但在复杂背景下进行桥梁裂缝检测任务时,仍然存在一些瓶颈和改进空间:(1)针对桥梁裂缝细小模糊等问题,特征提取能力较差;(2)由于完全交并比(Complete Intersection over Union, CIoU)^[19]损失函数引入的纵横比增量在处理复杂背景下桥梁裂缝检测时可能存在一定的模糊性,尤其当裂缝与周围复杂背景相似或存在干扰物时,这种模糊性可能会影响到裂缝边界框的精确定位和准确性,且当数据标注存在误差时,裂缝位置的预测精度也会下降;(3)图片背景噪声干扰过大,存在大量无效信息,导致细小裂缝在检测任务中竞争力不足,从而影响裂缝检测效果。

2 基于改进 YOLO v8s 的裂缝检测算法

本文提出一种基于改进 YOLO v8s 的混凝土桥梁裂缝检测算法,该算法主要采用计算速度更快

的焦距交并比 (Wise Intersection over Union, WIoU) 损失函数、全维度动态卷积 (Omni-Dimensional Dynamic Convolution, ODConv) 和级

联双层特征改进注意力 (Concatenated Two-Layer Feature-Modified Attention, C2f-MA) 融合模块, 算法结构如图 1 所示。

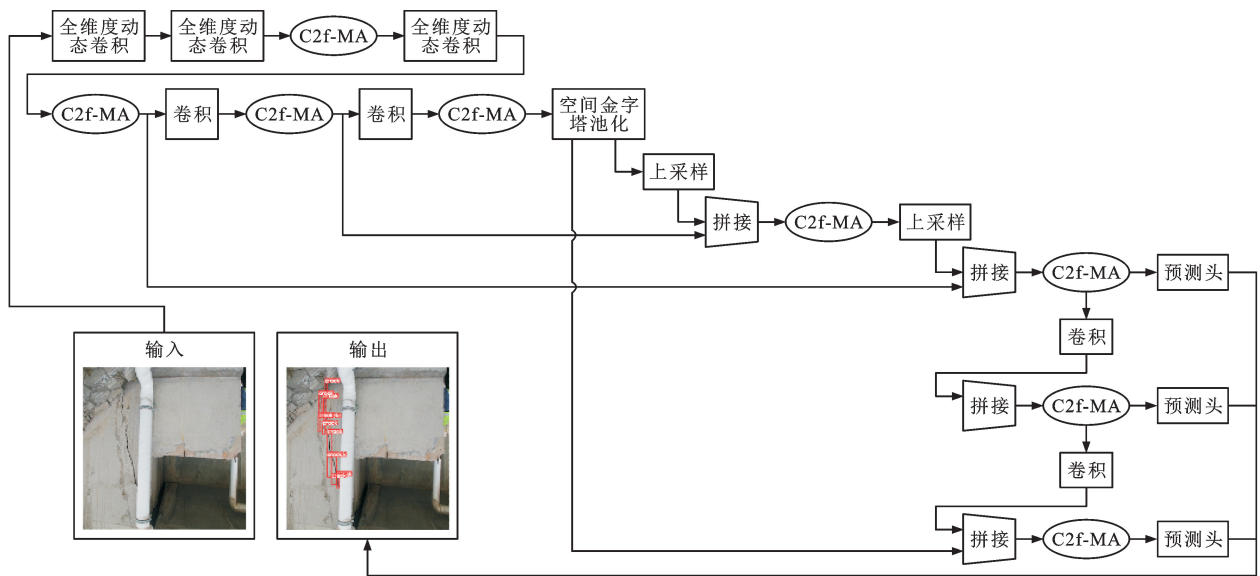


图 1 改进 YOLO v8s 网络结构

Fig.1 Improved YOLO v8s network structure

2.1 损失函数

边界框回归 (Bounding-Box Regression, BBR) 的损失函数对于目标检测任务来说至关重要, YOLO v8s 网络采用 CIoU^[20] 损失函数作为目标定位函数, CIoU 损失函数引入了边界框回归的重叠面积、中心点距离和纵横比这 3 个几何因素, 但纵横比作为一种相对值, 无法确定裂缝尺寸或形状信息, 且未考虑难易样本的平衡问题, 因此, 本文引入 WIoU^[21] 损失函数, 该损失函数提出了动态非单调的聚焦机制, 用离群度替代 IoU 对锚框进行质量评估。由于本文数据集标注过程中存在一定误差, 导

致出现低质量示例, 而 WIoU 损失函数提供了明智的梯度增益分配策略, 可以降低高质量锚框的竞争力, 减小低质量示例产生的有害梯度, 且 WIoU 没有计算纵横比, 计算速度优于 CIoU, 故 WIoU 相对于 CIoU 表现较好, 提升了目标神经网络的检测性能。CIoU 和 WIoU 损失函数如图 2 所示, 其中: B_p 为桥梁裂缝预测框; B_t 为桥梁裂缝真实标注框; W_g 为真实值与预测值最小封闭框的高; H_g 为真实值与预测值最小封闭框的宽; d 为真实标注框与预测框的中心点距离; c 为同时包含 2 个框的最小区域的对角线长度。

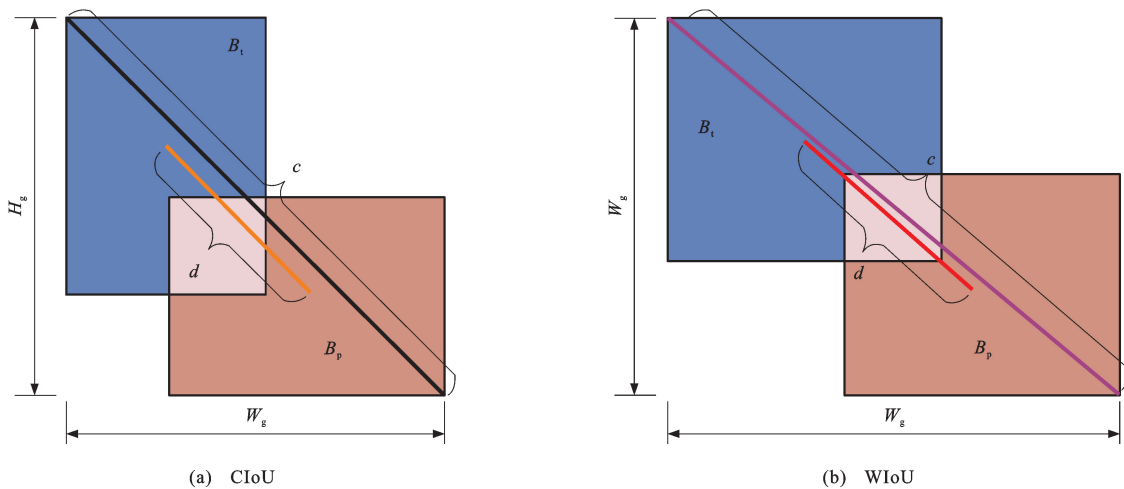


图 2 CIoU 和 WIoU 损失函数

Fig.2 Loss functions of CIoU and WIoU

CIoU 和 WIoU 的计算式分别为

$$C = I - \frac{d^2}{c^2} - B \quad (1)$$

$$W = \exp\left[\frac{d^2}{(c^2)^*}\right] \quad (2)$$

式中: C 为 CIoU; W 为 WIoU; I 为真实标注框与预测框重叠面积和总面积的比值; B 为权衡参数与 2 个框长宽比一致性的乘积; $*$ 表示将 W_g 和 H_g 从计算图中分离出来。

IoU 计算式为

$$I = \frac{|B_p \cap B_t|}{|B_p \cup B_t|} \quad (3)$$

2.2 全维度动态卷积

为解决桥梁裂缝检测中裂缝细小模糊的问题,本文将 YOLO v8s 主干网络中前 3 个常规卷积替换为 ODConv,ODConv 是由 Li 等^[22]提出的一种通过并行策略引入的多维注意力机制,以增强卷积核在空间的 4 个维度上学习互补的注意力。通过逐步递

进地对卷积操作在位置、通道、滤波器和核等维度上应用不同的注意力权重,可以使卷积操作更好地处理输入数据中各个维度之间的差异性,捕获更加丰富的上下文信息,提升卷积的特征提取能力,输出表达能力更强的裂缝特征,其结构如图 3 所示,图中 Sigmoid 函数用于归一化卷积核权,使其值在 $0 \sim 1$ 。对于动态卷积来说,是通过多个卷积核进行线性加权来进行卷积操作,而这些权重则与输入有关,因此,动态卷积具有输入依赖性,而 ODConv 延续了动态卷积的定义,即

$$y = (\alpha_{w1}W_1 + \dots + \alpha_{wi}W_i + \dots + \alpha_{wn}W_n)x \quad (4)$$

$$y = (\alpha_{w1}\alpha_{f1}\alpha_{c1}\alpha_{s1}W_1 + \dots + \alpha_{wi}\alpha_{fi}\alpha_{ci}\alpha_{si}W_i + \dots + \alpha_{wn}\alpha_{fn}\alpha_{cn}\alpha_{sn}W_n)x \quad (5)$$

式中: x 为输入特征; y 为输出特征; W_i 为第 i ($i = 1, 2, \dots, n, n$ 为卷积核总数)个卷积核的权重; α_{wi} 为第 i 个卷积核的注意力权重; α_{fi} 、 α_{ci} 和 α_{si} 分别为第 i 个卷积核的输出通道、输入通道和空间维度的可学习权重。

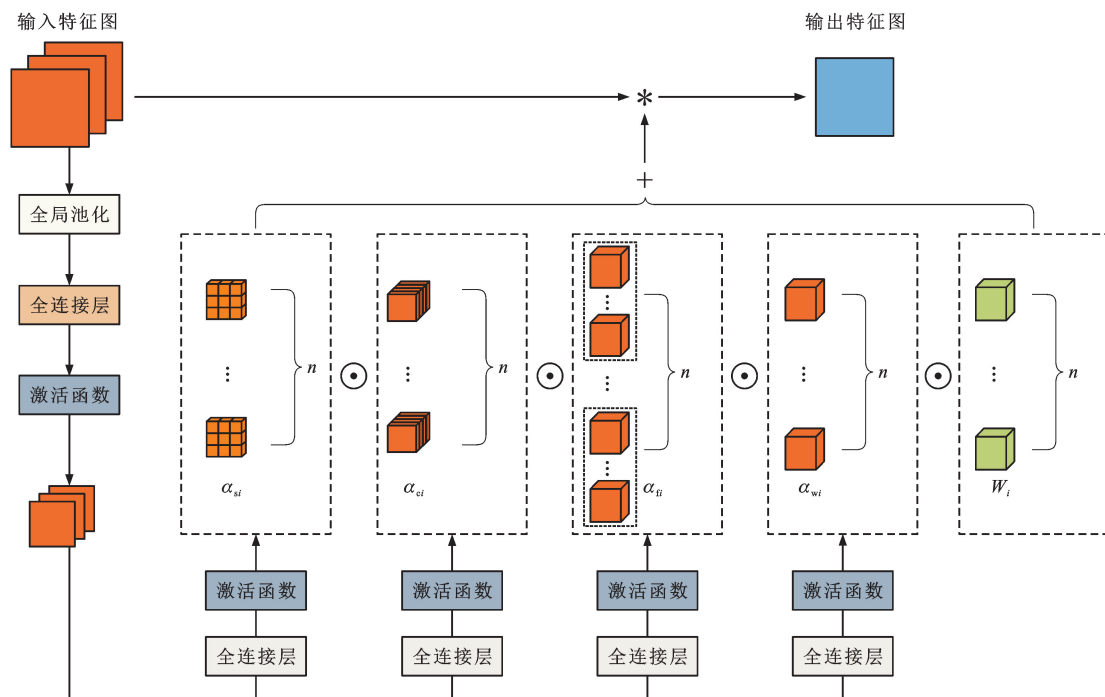


图 3 ODConv 结构

Fig. 3 ODConv structure

如图 3 所示,ODConv 模块能获得更丰富的上下文信息,较好地提升卷积特征提取能力,获得更有效的细小裂缝特征信息,从而提高对细小裂缝的检测能力,优化复杂背景下细小模糊特征提取困难的问题。

2.3 C2f-MA 模块

C2f-MA 模块由 C2f 模块与改进的 MA-ECA

模块串行组成,结构如图 4 所示。高效通道注意力 (Efficient Channel Attention, ECA) 模块^[23]是基于压缩与激发 (Squeeze-and-Excitation, SE^[24]) 模块开发的一种轻量级高效通道注意力模块,ECA 模块保持输入特征的通道维度不变,在局部范围内利用一维卷积操作来捕捉通道间的关联性,并通过自适应调整卷积核大小来适应不同的上下文信息。原始

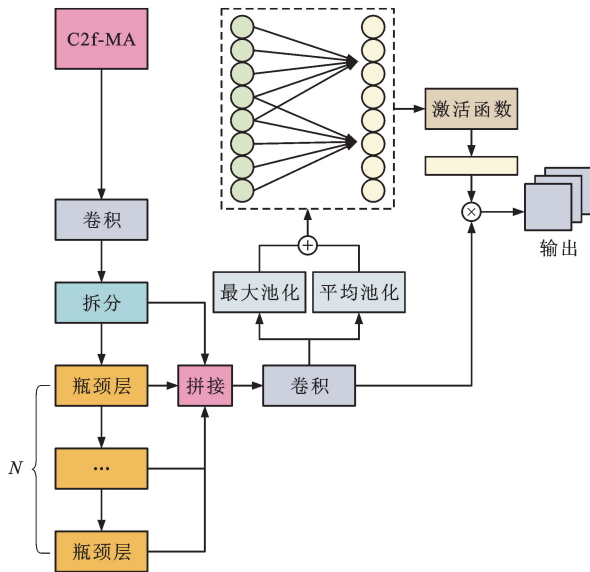


图 4 C2f-MA 结构

Fig. 4 C2f-MA structure

ECA 模块使用平均池化提取特征图,能够较好地保留背景信息,但复杂背景下局部信息的丢失和过多无用背景信息均会影响细小模糊裂缝特征的提取效果,因此,本文在ECA模块中引入最大池化来增强

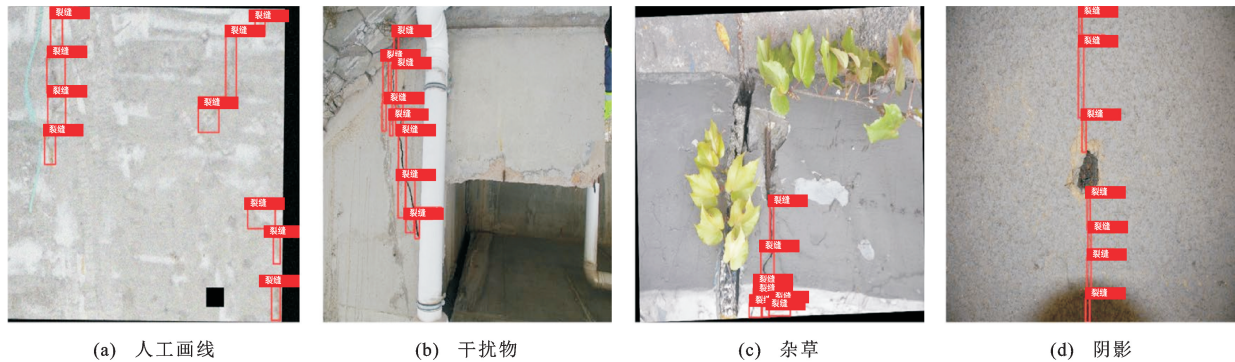


图 5 桥梁裂缝标注

Fig. 5 Bridge crack annotations

3.2 试验环境

本文使用 Pytorch 网络框架,基于 Python 代码运行算法,为保证试验顺利进行,对试验环境进行测试,具体如表 1 所示。

3.3 试验参数设置

网络训练采用的优化器为动量随机梯度下降 (Stochastic Gradient Descent with Momentum, SGD)算法,动量因子设置为 0.937,初始学习率设置为 0.01,最终学习率设置为 0.01,学习衰减策略为余弦退火策略^[25],当前训练步 t 的学习率 η_t 为

$$\eta_t = \frac{1}{2} \left[1 + \cos\left(\frac{t\pi}{T}\right) \right] \eta \quad (6)$$

式中: T 为学习率衰减的周期; η 为初始学习率。

对纹理特征信息的敏感性,更好地保留图像的显著特征,减少无效信息的影响。首先将输入特征经过卷积和拆分后,使用 N 个瓶颈层进行特征提取(本文中 N 取值为 3),然后将拆分特征图和每个瓶颈层的特征图分别拼接后进行最大池化与平均池化操作,并将池化后的特征图相加,生成语义信息更丰富的特征图,使之更关注通道内重要特征信息,抑制无关的特征信息,对提取复杂背景下细小模糊的裂缝特征有更好的效果。针对复杂背景下的细小模糊裂缝检测,C2f-MA 模块能使该模型关注更重要的特征信息,降低背景噪声对模型的影响,有效提高了复杂背景下桥梁裂缝的检测精度。

3 桥梁裂缝数据集与超参数设置

3.1 试验数据集构建

从桥梁检测报告中筛选出存在裂缝细小模糊、阴影、人工画线、杂草等具有复杂背景的裂缝照片,如图 5 所示,图像分辨率为 640 像素 \times 640 像素,对拍摄角度、拍摄距离、焦距、对比度等参数无特定要求,并采用开源 LabelImg 工具标注每张图像的裂缝。

表 1 试验环境

Table 1 Experimental environment

名称	型号
显卡(GPU)	NVIDIA Quadro RTX 3090
显存/GB	24
操作系统	Ubuntu 20.04
Pytorch	1.10.0
Python	3.8.16
CUDA	11.3
Torchvision	0.11.1

模型输入图像尺寸为 640 像素 \times 640 像素,分别采用旋转、平移、改变亮度、加噪声等方式进行数据增强,提升模型的泛化能力,增强过后共计 2 140 个图

像,并按照 4 : 1 划分为训练集与验证集。

4 试验结果

4.1 模型评价指标

本文选取目标检测任务中的评价指标评估模型性能,包括查准率、查全率、裂缝的平均准确率。其检测结果可分为 4 种:真正例、假正例、真负例、假负例。准确率 P 及召回率 R 分别为

$$P = \frac{T_P}{T_P + F_P} \quad (7)$$

$$R = \frac{T_P}{T_P + F_N} \quad (8)$$

精度均值 A 为

$$A = \int_0^1 P dR \quad (9)$$

式中: T_P 为真正例,即模型正确识别出存在裂缝的区域的比例; F_P 为假正例,即模型错误地将没有裂缝的区域预测为有裂缝区域的比例; F_N 为假负例,即模型漏掉了实际存在裂缝的区域的比例。

本文主要采用 R 、 $P_{0.5}$ 、 $P_{0.5:0.95}$ 、帧率和模型存储容量来评价模型,其中: $P_{0.5}$ 为 IoU 阈值为 0.5 时的平均准确率,当检测框与真实框的 IoU 值超过此阈值时,才会被认为是正确检测,强调在较高置信度下模型的准确性; $P_{0.5:0.95}$ 为考虑 0.50~0.95 的一系列 IoU 阈值的平均精度,反映模型在不同置信度下的综合性能。 $P-R$ 曲线则是根据不同的 IoU 阈值绘制而成,展示在不同召回率水平下的准确率。交并比 IoU 是衡量目标检测算法预测框与真实标注框重叠程度的关键指标,通过计算两者的交集与并集面积之比实现。在目标检测任务中,准确率是衡量模型性能的重要标准。本文同时采用 $P_{0.5:0.95}$ 和 $P_{0.5}$,旨在从不同角度全面反映模型的检测能力,确保结果的可靠性和模型的泛化能力。

4.2 试验对比分析

4.2.1 改进 YOLO v8s 与 YOLO v8s 对比

YOLO v8s 与改进 YOLO v8s 算法性能对比如表 2 所示,提出的算法针对桥梁裂缝检测取得了更佳的性能,其 $P_{0.5}$ 提升了 3.6%, $P_{0.5:0.95}$ 提升了 4.5%。

表 2 YOLO v8s 与改进 YOLO v8s 性能对比

Table 2 Performance comparison between YOLO v8s and improved YOLO v8s

方法	R	$P_{0.5}$	$P_{0.5:0.95}$	帧率/ (帧·s ⁻¹)	模型存储 容量/MB
YOLO v8s	0.789	0.857	0.586	139	11.13
改进 YOLO v8s	0.829	0.893	0.631	123	11.14

图 6 为 2 种算法的 $P-R$ 曲线,实际检测效果如图 7 所示,图中数字为裂缝检测置信度,可知:所提算法的 $P-R$ 曲线面积大于 YOLO v8s 的曲线面积,在不同召回率下均能保持较高的准确率,表明其在裂缝检测任务中具有更强的特征提取能力,更加适用于复杂背景下的桥梁裂缝检测场景;相比之下,YOLO v8s 算法可能受到复杂背景和细小模糊裂缝等因素的影响,其检测效果较差,因此,改进 YOLO v8s 算法在处理裂缝处于复杂背景时表现出更高的鲁棒性和准确性。

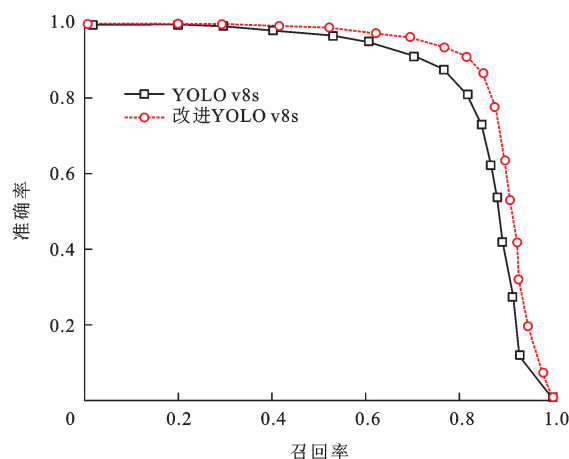


图 6 两种算法的 $P-R$ 曲线对比

Fig. 6 Comparison of $P-R$ curves between two algorithms

通过图 7 中可视化结果对比可知:在阴影、杂草、裂缝细小模糊、人工画线等干扰下,所提算法对裂缝检测具有更高的识别精度,能有效改善因背景复杂、裂缝细小模糊等因素造成的裂缝错检、漏检情况。

4.2.2 消融试验结果分析

为深入分析算法的有效性,对不同模块对网络的影响,对提出的改进 YOLO v8s 算法进行了消融试验,将其划分为 6 组不同网络模型,结果如表 3 所示,其中:×表示不使用该模块;√表示使用了该模块。

表 3 中,试验 1 为基准网络测试结果,试验 2~4 分别为 WIoU 模块、ODConv 模块、C2f-MA 模块对基准网络性能的影响,试验 5 在试验 2 的基础上引入 ODConv 模块,试验 6 在试验 5 的基础上引入 C2f-MA 模块,即提出的改进 YOLO v8s 算法。

从表 3 可知:试验 2 中引入了 WIoU 损失函数,能更好地识别低质量样本,同时由于减少了纵横比计算步骤,进一步加快了模型的计算速度,使得对桥梁裂缝的检测更加精准且更快速, $P_{0.5}$ 相比于试验 1 提升了 1.4%;试验 3 中引入了 ODConv 模块,该模块中卷积核在空域、输入通道、输出通道等部分赋予不同的注意力,使之能够捕获丰富的上下文信息,

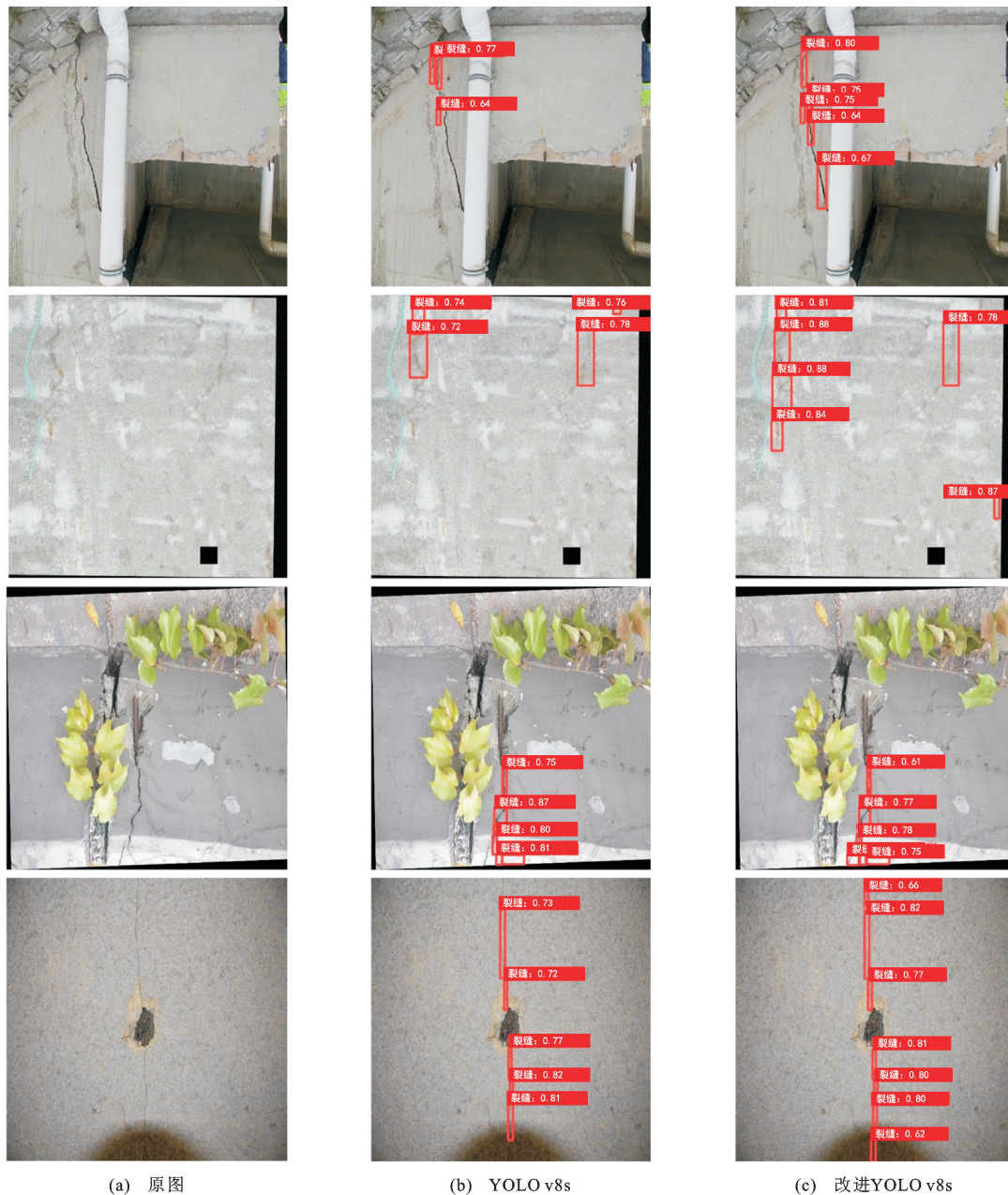


图 7 裂缝识别结果对比

Fig. 7 Comparison of crack recognition results

表 3 消融试验结果

Table 3 Results of ablation experiments

试验编号	WIoU	ODConv	C2f-MA	R	$P_{0.5}$	$P_{0.5:0.95}$	帧率/(帧·s ⁻¹)	模型存储容量/MB
1	×	×	×	0.789	0.857	0.586	139	11.13
2	√	×	×	0.821	0.871	0.595	143	11.13
3	×	√	×	0.823	0.872	0.611	135	11.14
4	×	×	√	0.808	0.878	0.600	128	11.13
5	√	√	×	0.827	0.879	0.613	141	11.14
6	√	√	√	0.829	0.893	0.631	123	11.14

提升卷积的特征提取能力,从而进一步解决了桥梁裂缝细小模糊的问题,相比于试验 2 在 $P_{0.5}$ 上提升

了 1.5%,但检测速度受网络深度的影响有所下降,并增加了一定的参数量;试验 4 中引入改进的 C2f-

MA 模块,相比于试验 3 进一步增强了裂缝特征的语义信息,弱化了背景信息,故 $P_{0.5}$ 提升了 2.1%;试验 5 同时引入了 WIoU 与 ODConv 模块,准确率相比于试验 4 提升了 2.2%;试验 6 的 $P_{0.5}$ 整体提升了 3.6%,但网络深度与复杂度发生了改变,最终检测速度有所下降。

为验证 C2f-MA 模块的有效性,在确保其他模块不变的基础上,对该模块引入不同位置的消融试验,如表 4 所示。试验 1 为仅在头部检测网络中加入 C2f-MA 模块;试验 2 为仅在主干网络部分引入 C2f-MA 模块;试验 3 为在全局网络中引入 C2f-MA 模块。通过对比可知,试验 3 对裂缝的检测准确率优于试验 1、2,其 $P_{0.5}$ 分别较试验 1、2 高 0.6% 和 0.9%,但总体检测速度略低。综上所述,C2f-MA 模块在主干网络、检测头部和全局网络中的不同位置都对模型性能有一定贡献,且在全局网络中的影响最显著。

表 4 C2f-MA 模块不同位置消融对比结果

Table 4 Comparison of ablation experiment results of C2f-MA module at different locations

试验编号	R	$P_{0.5}$	$P_{0.5:0.95}$	帧率/ (帧·s ⁻¹)	模型存储 容量/MB
1	0.823	0.884	0.619	130	11.13
2	0.823	0.887	0.623	132	11.13
3	0.829	0.893	0.631	123	11.14

对损失函数来说,其超参数取值对目标检测任务的影响也至关重要,因此,本文在保证其他模型不变的基础上,通过调节超参数进行几组消融试验。如表 5 所示,通过调节损失函数的超参数取值,可发现当置信度损失平衡权重 $\alpha=1.6$,边界框回归平衡权重 $\delta=4$ 时在召回率和准确率上得到了较高的指标。

表 5 WIoU 超参数对试验结果的影响

Table 5 Effects of WIoU hyperparameters on experimental results

α	δ	R	$P_{0.5}$	$P_{0.5:0.95}$
2.3	2	0.833	0.887	0.620
1.9	3	0.810	0.877	0.605
1.6	4	0.829	0.893	0.631
1.4	5	0.825	0.890	0.627
1.3	6	0.830	0.886	0.625

4.2.3 多种目标检测算法结果对比

为了验证提出的改进 YOLO v8s 算法的有效性,本文使用经典的两阶段算法 Faster R-CNN、可变形检测变换器(Deformable Detection Transformer, Deformable DETR)^[26]算法、改进型除噪锚框的检

测变换器(DETR with Improved Denoising Anchor Boxes, DINO)^[27]算法与 YOLO v8-GAM-Wise-IoU^[28]进行对比试验,结果具体如表 6 所示,改进 YOLO v8s 算法在检测精度方面取得了显著提升,且模型存储容量仅为 11.14 MB,远小于其他目标检测模型,因此,改进 YOLO v8s 算法在综合性能方面具备显著优势,对复杂背景下混凝土桥梁裂缝图像的检测能力更强。

表 6 多种目标检测算法结果对比

Table 6 Comparison of results from multiple target detection algorithms

方法	R	$P_{0.5}$	$P_{0.5:0.95}$	模型存储 容量/MB
Faster R-CNN	0.302	0.449	0.232	41.35
Deformable DETR	0.582	0.810	0.464	40.10
YOLO v8-GAM-Wise-IoU	0.763	0.832	0.544	12.87
DINO	0.687	0.858	0.585	47.54
改进 YOLO v8s	0.829	0.893	0.631	11.14

图 8 展示了桥梁裂缝在其他目标检测网络上的可视化结果对比,可以看出:改进 YOLO v8s 算法在实际检测中取得了较好的效果,Faster R-CNN 算法在实际检测过程中存在误检,将图片背景中的杂草、人工画线等作为裂缝,Deformable DETR 算法和 DINO 算法的检测效果略低于提出的算法,且存在误检现象,且 Deformable DETR 算法和 DINO 算法模型过大,检测速度较低,难以应用,而 YOLO v8-GAM-Wise-IoU 算法的检测效果较差,存在漏检问题,因此,所提算法对于复杂背景下的桥梁裂缝检测有更好的应用价值。

4.3 裂缝长宽测量

本文研究裂缝目标检测算法的目的是对图像中的裂缝进行识别与定位,可将裂缝与复杂背景分离,但无法量化分析图像特征,因此,在本文方法研究基础上,还需结合相关语义分割方法获取裂缝的长宽信息,如图 9 所示,实现过程如下。

(1) 裂缝定位和识别检测。使用改进 YOLO v8s 算法对裂缝进行定位和识别检测,提取裂缝区域,确定裂缝的具体位置,使用目标框将裂缝从背景中分离出来,抑制其他复杂背景的干扰,提高裂缝的分割精度。

(2) 裂缝区域提取。基于第 1 步所提模型得到的裂缝边界框位置信息提取对应的裂缝矩形区域,即形成感兴趣区域(Region of Interest, ROI),将裂缝区域作为后续分割任务的输入,将注意力集中在

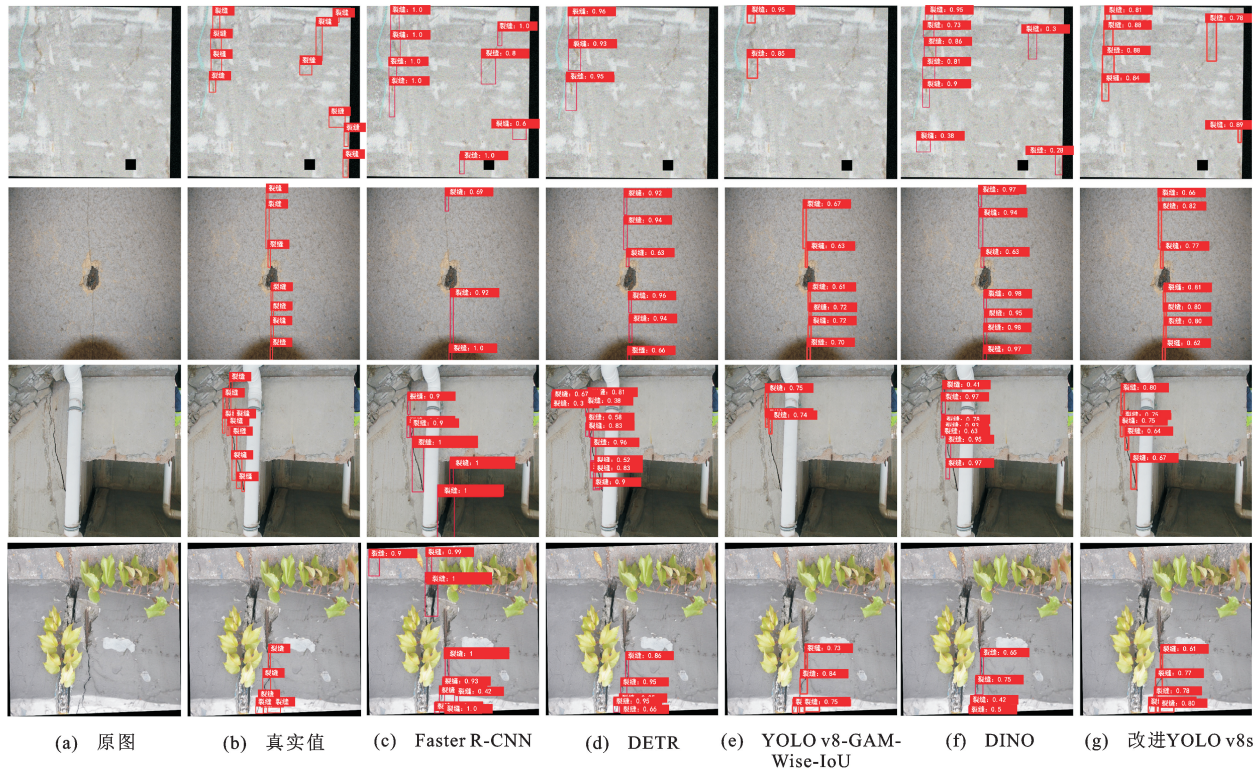


图 8 多种目标检测算法对比结果

Fig. 8 Comparison of results from multiple target detection algorithms

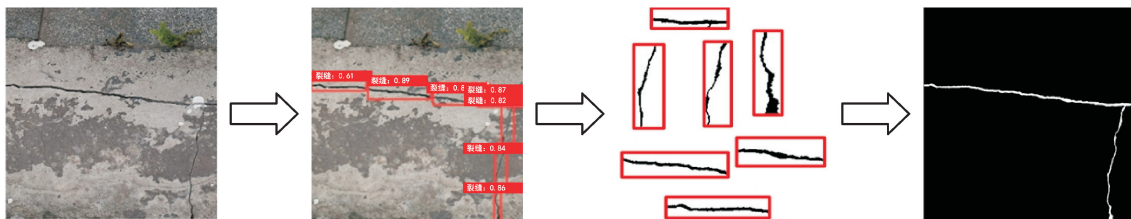


图 9 裂缝提取主骨架结果

Fig. 9 Result of crack extraction of main skeleton

裂缝区域上,提高分割的准确性和效率。

(3)裂缝分割。在第 2 步的基础上将提取的 ROI 作为输入传递给 Hang 等^[29]提出的裂缝语义分割算法,进行像素级别的裂缝分割,通过该步骤可获得裂缝的二值掩码图像。

(4)主骨架线提取。如图 10 所示,基于裂缝的二值掩码图像,应用基于像素点的八邻域计算方法^[30]提取裂缝的主骨架线,以准确表示裂缝的形态特征,为后续长宽测量提供准确输入。通过对二值图像中的每个像素点应用一系列基于八邻域的迭代条件判断,以删除边缘像素并保持图像连通性,从而将线条减至单个像素宽度,该算法的迭代步骤如下。

a. 初始化。将 Hang 等^[29]提出的裂缝语义分割算法得到的二值掩码图像作为输入,其中 0 值为背景,非 0 值为前景。

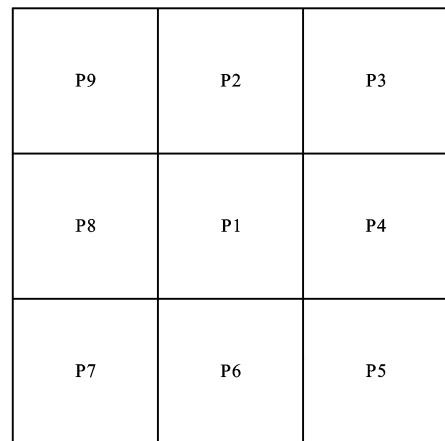


图 10 八邻域示意

Fig. 10 Schematic of eight-neighborhood

b. 循环遍历。对图像中的每个前景像素点进行循环遍历。

c. 条件检查。对每个前景像素点应用一系列条件检查,以确定是否符合删除条件,其条件包括:①相邻前景像素点的数量应在 2~6;②从 P2 到 P9 再到 P2 的像素中,0 和 1 出现的累计次数应为 1;③3 个连续的像素点 P2、P4、P6 和 P4、P6、P8 的值的乘积应至少有一个为 0。

d. 标记删除。如果一个前景像素点满足上述所有条件,则将其标记为删除。

e. 交替执行。通常会有 2 个迭代步骤交替执行,直到不再发生任何改变。

f. 终止条件。当一个完整的迭代过程中没有像素被删除或标记时,算法终止。

(5) 宽度测量。基于裂缝主骨架线的拓扑结构计算裂缝宽度,具体步骤如下。

a. 初始化距离图像。创建一个与主骨架线图像大小相同的距离图像,并将所有像素点的初始距离设为无穷大,该图像用于存储每个主骨架线像素点与最近裂缝边缘像素点之间的距离。

b. 遍历主骨架线像素点。首先遍历所有裂缝边缘像素点;其次采用欧氏距离计算当前主骨架线像素点与每个裂缝边缘像素点之间的距离;最后比较并更新当前主骨架线像素点与所有裂缝边缘像素点之间的距离,选择最小距离作为该像素点与最近裂缝边缘像素点的距离。

c. 更新最小距离图像。遍历主骨架线上的每个像素点,将其与最近裂缝边缘像素点的最小距离更新到最小距离图像的相应位置。

d. 计算裂缝宽度。遍历最小距离图像的每个像素点,将每个像素点与最近裂缝边缘像素点之间的

最小距离乘以 2,即可得到裂缝的宽度。

本文对测量到的裂缝宽度数据进行了可视化分析,并通过绘制沿主骨架线方向的裂缝宽度分布热力图来展示结果,如图 11 所示,其中热力图的颜色深浅表示不同裂缝测量点的裂缝宽度。具体而言,颜色越深表示裂缝宽度越大,最大裂缝宽度为 20 像素。

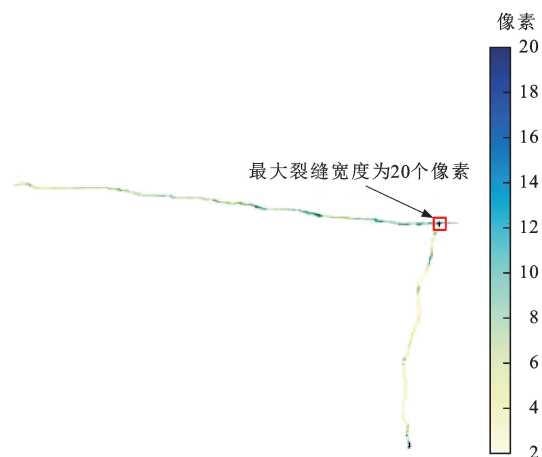


图 11 沿主骨架线方向的裂缝宽度热力图

Fig. 11 Crack width heatmap along main skeleton line direction

裂缝长度的计算是基于裂缝主骨架线实现的。首先对图像中的像素进行归一化处理。将像素范围缩放到 0~1,使得不同图像具有相似的像素范围;然后对归一化后的像素进行累加操作,获得裂缝的像素长度,即裂缝在图像中连续的像素数量,如图 12 所示,可计算得出裂缝长度为 1 006 个像素。由于缺少带有真实裂缝参数的裂缝图像,提出的裂缝量化方法只是在像素级别进行计算。未来考虑结合拍摄设备的相关参数和实际拍摄条件来建立数学模型,将像素级别的裂缝参数映射到裂缝的实际长宽中。

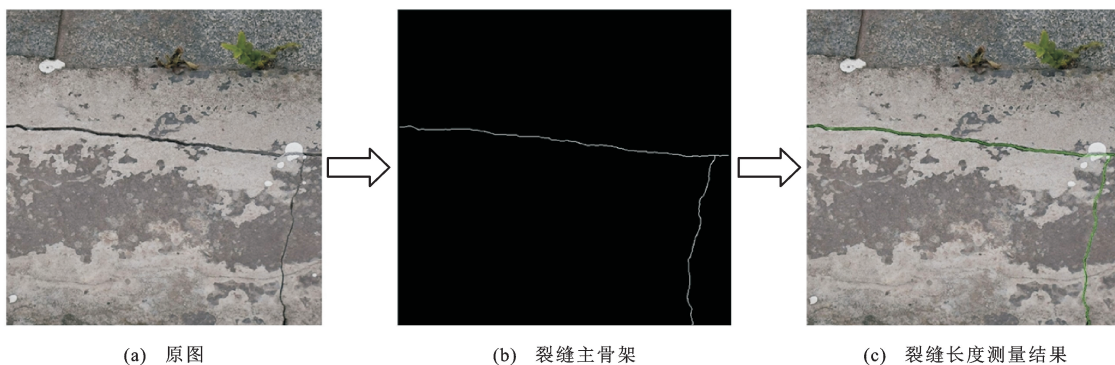


图 12 裂缝长度测量结果

Fig. 12 Crack length measurement results

5 结 语

(1)在裂缝长度和宽度测量过程中,背景颜色(如裂缝所在处的水泥颜色)及杂物等干扰信息往往

会影响裂缝像素的准确抓取,为有效解决这一问题,提出改进 YOLO v8s 算法,旨在精确定位桥梁裂缝位置信息,特别是在复杂背景下,仍能保持高效的检测性能。通过采用高精度裂缝定位和像素分割算

法,进一步提取裂缝主骨架线,从而实现对裂缝宽度和长度的可靠测量。通过试验验证,该算法的检测平均精度达到 89.3%。

(2)本文试验数据集均来自桥梁检测报告,改进 YOLO v8s 算法在桥梁裂缝图像的拍摄角度、焦距和对比度方面具有一定的鲁棒性,能够适应不同条件下的图像,为实际应用带来灵活性和便利性。

(3)尽管所提算法在桥梁裂缝检测方面取得了显著效果,但仍需在未来工作中进一步验证其稳健性和适应性。由于复杂背景和桥梁结构表面病害多样,仅对裂缝这一类病害进行检测难以满足工程实际需求,针对多种病害实现模型更小、更快速、精度更高的识别还有待研究。另外,本文在计算裂缝长度和宽度参数时仅实现了以像素为单位的长度和宽度的测量,未实现从像素单位映射到实际长度和宽度,由于裂缝的实际长度和宽度参数对于混凝土结构的健康评估至关重要,需进一步研究能够测量实际裂缝长度和宽度的方法,以获取更精确的尺寸数据。

参 考 文 献 :

References :

- [1] 陈浩瀚. 混凝土路面图像的裂缝检测算法研究[D]. 东莞: 东莞理工学院, 2023.
CHEN Hao-han. Research on crack detection algorithm of concrete pavement images[D]. Dongguan: Dongguan University of Technology, 2023. (in Chinese)
- [2] 邓 露, 褚鸿鹤, 龙砺芝, 等. 基于深度学习的土木基础设施裂缝检测综述[J]. 中国公路学报, 2023, 36(2): 1-21.
DENG Lu, CHU Hong-hu, LONG Li-zhi, et al. Review of deep learning-based crack detection for civil infrastructure[J]. China Journal of Highways and Transport, 2023, 36(2): 1-21. (in Chinese)
- [3] 《中国公路学报》编辑部. 中国桥梁工程学术研究综述·2021[J]. 中国公路学报, 2021, 34(2): 1-97.
Editorial Department of *China Journal of Highway and Transport*. Review on China's bridge engineering research; 2021[J]. China Journal of Highway and Transport, 2021, 34(2): 1-97. (in Chinese)
- [4] HSIEH Y A, TSAI Y J. Machine learning for crack detection: review and model performance comparison[J]. Journal of Computing in Civil Engineering, 2020, 34(5): 04020038.
- [5] 刘宇飞, 樊健生, 聂建国, 等. 结构表面裂缝数字图像法识别研究综述与前景展望[J]. 土木工程学报, 2021, 54(6): 79-98.
LIU Yu-fei, FAN Jian-sheng, NIE Jian-guo, et al. Review and prospect of digital-image-based crack detection of structure surface[J]. China Civil Engineering Journal, 2021, 54(6): 79-98. (in Chinese)
- [6] MUNAWAR H S, HAMMAD A W A, HADDAD A, et al. Image-based crack detection methods: a review[J]. Infrastructures, 2021, 6(8): 115.
- [7] LECUN Y L, BENGIO Y, HINTON G. Deep learning[J]. Nature, 2015, 521(7553): 436-444.
- [8] 沙爱民, 童 峥, 高 杰. 基于卷积神经网络的路表病害识别与测量[J]. 中国公路学报, 2018, 31(1): 1-10.
SHA Ai-min, TONG Zheng, GAO Jie. Recognition and measurement of pavement diseases based on convolutional neural networks[J]. China Journal of Highway and Transport, 2018, 31(1): 1-10. (in Chinese)
- [9] DENG Li, CHU Han-han, SHI Peng, et al. Region-based CNN method with deformable modules for visually classifying concrete cracks[J]. Applied Sciences, 2020, 10(7): 2528.
- [10] 晏班夫, 徐观亚, 栾 健, 等. 基于 Faster R-CNN 与形态法的路面病害识别[J]. 中国公路学报, 2021, 34(9): 181-193.
YAN Ban-fu, XU Guan-ya, LUAN Jian, et al. Pavement distress detection based on faster R-CNN and morphological operations[J]. China Journal of Highway and Transport, 2021, 34(9): 181-193. (in Chinese)
- [11] 余加勇, 李 锋, 薛现凯, 等. 基于无人机及 Mask R-CNN 的桥梁结构裂缝智能识别[J]. 中国公路学报, 2021, 34(12): 80-90.
YU Jia-yong, LI Feng, XUE Xian-kai, et al. Intelligent identification of bridge structure cracks based on unmanned aerial vehicle and mask R-CNN[J]. China Journal of Highway and Transport, 2021, 34(12): 80-90. (in Chinese)
- [12] REN Shao-qing, HE Kai-ming, 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 (TPAMI), 2017, 39(6): 1137-1149.
- [13] LIU Wei, ANGUELOV D, ERHAN D, et al. SSD: single shot multi-box detector[C] // Springer. Proceedings of the 2016 European Conference on Computer Vision. Berlin: Springer, 2016: 21-37.
- [14] REDMON J, FARHADI A. YOLOv3: an incremental improvement[J]. 2018, DOI:10.48550/arXiv.1804.02767.
- [15] 彭雨诺, 刘 敏, 万 智, 等. 基于改进 YOLO 的双网络桥梁表面病害快速检测算法[J]. 自动化学报, 2022, 48(4): 1018-1032.
PENG Yu-nuo, LIU Min, WAN Zhi, et al. A dual deep network based on the improved YOLO for fast bridge surface defect detection[J]. Acta Automatica Sinica, 2022, 48(4): 1018-1032. (in Chinese)
- [16] HUSSAIN M. YOLO-v1 to YOLO-v8, the rise of YOLO and its complementary nature toward digital manufacturing and industrial defect detection[J]. Machines, 2023, 11(7): 677.
- [17] SWATHI Y, CHALLA M. YOLO v8: advancements and innovations in object detection[C] // Springer. International Conference on Smart Computing and Communication. Berlin: Springer, 2024: 1-13.
- [18] BOCHKOVSKIY A, WANG C Y, LIAO H Y M. YOLOv4: optimal speed and accuracy of object detection[J]. 2020, DOI:10.48550/arXiv.2004.10934.
- [19] ZHENG Zhao-hui, WANG Ping, REN Dong-wei, et al.

- Enhancing geometric factors in model learning and inference for object detection and instance segmentation[J]. IEEE Transactions on Cybernetics, 2022, 52(8): 8574-8586.
- [20] ZHENG Zhao-hui, WANG Peng, LIU Wei, et al. Distance-IoU loss: faster and better learning for bounding box regression[C]// AAAI. Proceedings of the AAAI Conference on Artificial Intelligence. New York: AAAI, 2020: 12993-13000.
- [21] TONG Zan-jia, CHEN Yu-hang, XU Ze-wei, et al. Wise-IoU: bounding box regression loss with dynamic focusing mechanism[J]. 2023, DOI:10.48550/arXiv.2301.10051.
- [22] LI Chao, ZHOU Ao-jun, YAO An-bang. Omni-dimensional dynamic convolution[J]. 2022, arXiv preprint arXiv: 2209.07947.
- [23] WANG Qi-long, WU Bang-gu, ZHU Peng-fei, et al. ECA-Net: efficient channel attention for deep convolutional neural networks[C]//IEEE. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. New York: IEEE, 2020: 11534-11542.
- [24] HU Jie, SHEN Li, ALBANIE S, et al. Squeeze-and-excitation networks[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2020, 42(8): 2011-2023.
- [25] HE Tong, ZHANG Zhi, ZHANG Hang, et al. Bag of tricks for image classification with convolutional neural networks[C]// IEEE. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. New York: IEEE, 2019: 558-567.
- [26] ZHU Xi-zhou, SU Wei-jie, LU Le-wei, et al. Deformable DETR: deformable transformers for end-to-end object detection[J]. 2020, DOI:10.48550/arXiv.2010.04159.
- [27] ZHANG Hao, LI Feng, LIU Shi-long, et al. DINO: DETR with improved denoising anchor boxes for end-to-end object detection[J]. 2022, DOI:10.48550/arXiv.2203.03605.
- [28] XONG Chen-qi, ZATED T, ABDELKADER E M. A novel YOLO v8-GAM-Wise-IoU model for automated detection of bridge surface cracks[J]. Construction and Building Materials, 2024, 414: 135025.
- [29] HANG Die, YANG Jian-xi, JIANG Shi-xin, et al. Lightweight mesh crack detection algorithm based on efficient attention mechanism[J]. International Journal of Robotics and Automation, 2023, 38(3): 170-179.
- [30] ZHANG T Y, SUEN C Y. A fast parallel algorithm for thinning digital patterns[J]. Communications of the ACM, 1984, 27(3): 236-239.