

基于YOLOv8实例分割的桥梁裂纹识别

杨 明

安顺开放大学, 贵州 安顺

收稿日期: 2025年10月16日; 录用日期: 2025年11月14日; 发布日期: 2025年11月24日

摘 要

桥梁结构的安全性是工程领域的重要关注点, 裂纹检测是保障桥梁安全的关键环节。本文提出了一种基于YOLOv8实例分割的桥梁裂纹识别方法, 旨在解决传统检测方法在准确性上不足的问题。首先, 我们改进了数据增强方法, 增强了模型对不同环境下裂纹图像的鲁棒性和泛化能力, 使得模型在复杂实际应用中表现更加稳定。其次, 在C2F模块中引入了Coordinate Attention注意力机制。该机制通过对特征图的通道关系进行建模, 提升了模型对关键信息的关注度, 从而提高了裂纹检测的准确性。Coordinate Attention在空间和通道维度上融合信息, 使模型能够更精准地捕捉到裂纹的细节特征。实验结果表明, 本文提出的方法在本文构造的数据集上表现出色, 在精度上提高了3.5个百分点, 并且召回率和mAP均有提升, 对实际工程应用具有重要指导意义。

关键词

桥梁裂纹检测, YOLOv8, 数据增强, 注意力机制

Bridge Crack Recognition Based on YOLOv8 Instance Segmentation

Ming Yang

Anshun Open University, Anshun Guizhou

Received: October 16, 2025; accepted: November 14, 2025; published: November 24, 2025

Abstract

The safety of bridge structures is an important concern in the engineering field, and crack detection is a key link to ensure the safety of bridges. In this paper, we propose a bridge crack recognition method based on YOLOv8 instance segmentation, aiming at solving the problem of insufficient accuracy of traditional detection methods. First, we improve the data enhancement method to enhance the robustness and generalization ability of the model to crack images in different environments,

which makes the model perform more stably in complex practical applications. Second, the Coordinate Attention attention mechanism is introduced in the C2F module. This mechanism improves the accuracy of crack detection by modeling the channel relationship of the feature map, which enhances the model's attention to the key information. Coordinate Attention fuses the information in both spatial and channel dimensions, enabling the model to capture the detailed features of the cracks more accurately. The experimental results show that the method proposed in this paper performs well on several bridge crack datasets, with an improvement of 3.5 percentage points in precision and both recall and mAP, which is of great significance as a guide for practical engineering applications.

Keywords

Bridge Crack Detection, Yolov8, Data Enhancement, Attention Mechanism

Copyright © 2025 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

桥梁在现代社会中扮演着至关重要的角色，然而，随着桥梁的使用年限增长，桥梁可能出现裂纹等结构缺陷，这些缺陷对其安全性和可靠性构成了潜在威胁。因此，对桥梁结构进行实时的裂纹检测至关重要。近年来，基于深度学习的目标检测技术在图像处理领域取得了巨大的进展，然而，在桥梁裂纹的检测方面，仍然存在一定的挑战。

本文提出了一种基于 YOLOv8 实例分割的桥梁裂纹检测算法。为了提高模型的准确性和效率，本文主要从两个方面进行了创新。首先，我们改进了数据增强方法，通过在训练集中引入更多的变化和多样性，增强了模型的泛化能力。其次，我们引入了 Coordinate Attention 注意力机制[1]，通过对特征图中的位置信息进行建模，提升了模型对裂纹区域的关注度，从而改善了检测性能。

2. 相关研究

目前，在裂纹检测领域，深度学习算法已成为主导技术。目标检测模型主要分为两种类型：两阶段模型和一阶段模型。两阶段模型以 SPP-NET [2]、Faster-RCNN [3] 等代表。在训练过程中，首先对区域建议网络进行训练，随后训练针对目标区域的检测网络。因此，虽然这类模型具有较高的精度，但处理速度较慢。以 YOLO 系列和 SSD [4] 为代表的二阶段模型将目标检测视为一个回归问题。该模型利用初始锚点定位目标区域，直接预测类别，在不使用 RPN 的情况下完成整个检测过程，实现端到端目标检测。相应地，这种模型的速度比两阶段模型要快得多，但精度较低。在速度方面，单阶段模型比两阶段模型更适合工业应用。

Wang [5] 将双目立体视觉测量技术应用于无人机桥式裂缝系统。Kim 等 [6] 介绍了一种基于 R-CNN 的无人机平台的桥式裂缝识别算法。无人机正被越来越多地用于裂缝检测。然而，这些研究也有一定的局限性，特别是：1) 大多数研究集中在单调背景下的裂缝检测，如实验室中的混凝土构件或无杂物的桥面和路面。当无人机在某些结构的狭窄部分采集数据时，可能会出现 GPS 信号微弱的问题，这使得裂纹的位置难以复原。2) 大多数研究都在一些大型数据集上进行了迁移学习，这些数据集集中的图像大小通常很小，如 256×256 或 500×350 像素，因此迁移学习后的模型对小图像有很好的检测结果，但在大图像上会带来特征的消失。

YOLO 算法已广泛应用于无人驾驶汽车、机器人以及视频监控等领域,成为这些应用中的关键物体检测技术。Terven 等[7]人全面分析了 YOLO 的发展历程,详细介绍了从 YOLO 到 YOLOv8 的创新。雷源毅等[8]提出一种改进 YOLOv8n 的 RDCA-YOLO 安全帽检测算法,为解决工地场景下安全帽检测因背景复杂、目标密集等导致 YOLOv8n 算法识别精度低、易出现漏检及误检等问题。曹士杰[9]提出了一种基于改进 YOLOv5 的实例分割网络模型 Rock-YOLOv5,分别从通道注意力机制与空间注意力机制两方面对原始 YOLOv5 算法进行改进。通过对采集到的图像进行预处理,产生训练样本,构建出石块图像数据集。目标检测着重于识别物体的类别与位置,而实例分割则进一步对每个物体的实例进行区分,不仅识别类别和位置,还执行了分割,可以提供每个物体实例的精确轮廓,本文将基于改进 YOLOv8s 算法的实例分割实现对桥梁裂纹的识别。

3. 方法

3.1. 数据增强

为了获得更好的训练效果,Yolov4 [10]开始使用 mosaic 数据增强方法来达到丰富数据集特征的目的。yolov8 的默认数据增强策略是在除最多 10 个 epoch 外的所有数据中使用马赛克数据增强,同时对一半数据进行左右反转,并在数据反转后立即对部分数据进行颜色调整等。本文针对火焰和烟雾的数据特征,对 Yolov8 的默认马赛克策略进行了修改,同时增加了图像曝光处理[11]和图像模糊处理[12]如图 1 所示。

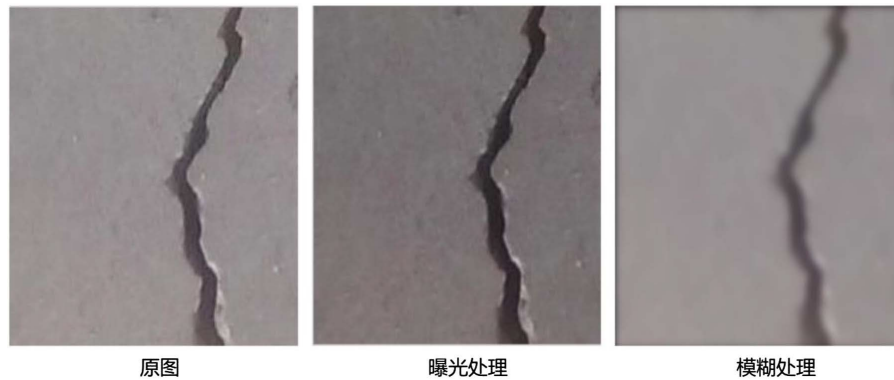


Figure 1. Data augmentation effect diagram
图 1. 数据增强效果图

图像曝光调整可以帮助模型适应不同的照明和摄像条件。它包括修改图像的伽玛曝光,使其更亮或更暗。对图像亮度或三基色值进行非线性运算或反运算的数学方法如下:

$$V_{out} = AV_{in}^{\gamma} \quad (1)$$

在方程中, A 代表常数,输入值和输出值都属于非负实数集。通常,在 A 设为 1 的情况下,输入和输出值的范围为 0 至 1。

为了增强模型对相机焦点变化的适应性,本文引入了随机高斯模糊技术。高斯模糊,作为一种图像模糊滤波器,通过对图像的每个像素点应用正态分布进行转换,其在 N 维空间的正态分布方程为:

$$G(r) = \frac{1}{\sqrt{2\pi\sigma^2}^N} e^{-r^2/(2\sigma^2)} \quad (2)$$

其中,模糊半径用 r 表示,正态分布的标准偏差用 σ 表示。

3.2. 注意力机制

传统的卷积神经网络主要捕获局部特征，难以建模对视觉任务至关重要的长距离依赖。因此，本文采用了 Coordinate Attention 机制，这种机制通过融合位置信息到通道注意力中，为移动网络设计了一种创新的注意力机制。Coordinate Attention 通过将通道注意力拆分为两个一维特征编码步骤，沿两个空间维度聚合特征，从而能够在一个方向上捕获长距离的依赖关系，同时在另一个方向上保持精确的位置信息。得到的特征图分别编码成一对具有方向感知和位置敏感的注意力图，这些图可以互补地用于增强输入特征图中感兴趣对象的表征。

YOLOv8 中的 C2f 模块承担了一个关键任务：融合来自高层的语义特征和来自低层的细节特征。C2f 模块采用 Split 进行特征分层，并在特征提取过程中使用跳层连接以确保轻量级并获得更丰富的梯度流信息。C2f 的结构如图 2 所示。

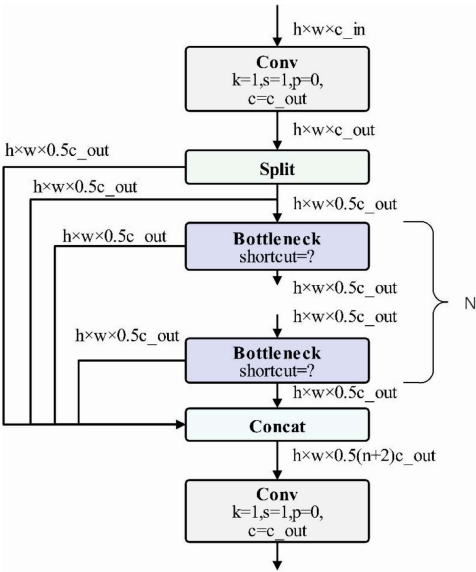


Figure 2. Structure of the C2f Module
图 2. C2f 模块结构

本文将 Coordinate Attention 引入 yolov8 的 c2f 模块中，构建一个 c2fca 模块，结合 c2f 和 Coordinate Attention 的优点，增强远程依赖的特征提取，模块结构如图 3。

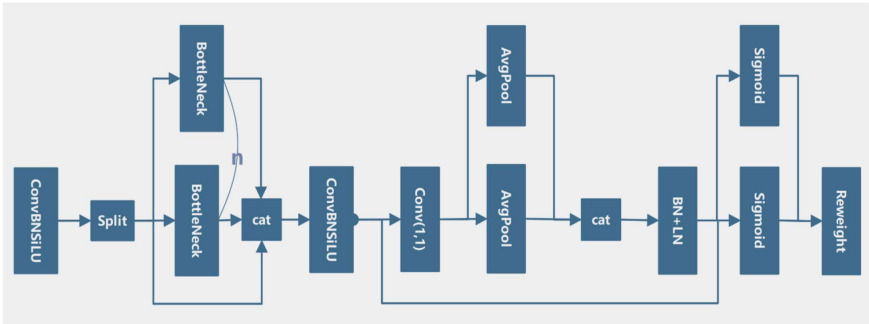


Figure 3. The constructed C2fca Module
图 3. 构建的 C2fca 模块

3.3. 改进后的网络模型

首先，在 input 阶段先进行 exposed 和 blur 的数据增长，再进行 mosaic 增强，在训练过程中后 50 epoch 不进行 mosaic 增强。图片处理阶段完成之后，在主干网络上进行特征提取，和原算法的主要区别是在 P3、P4、P5 部分进行模块替换，将原算法中的 c2f 模型替换为 c2fca 模块，改进后的算法如图 4 所示。

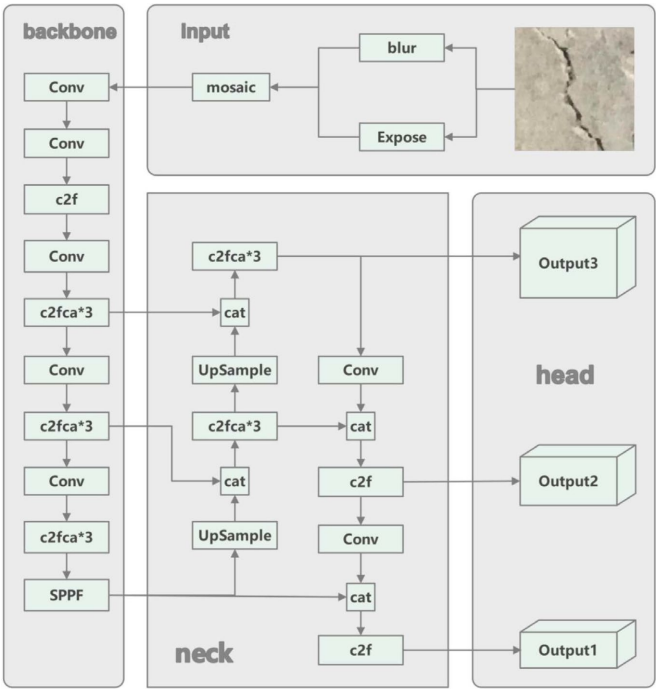


Figure 4. Improved network model
图 4. 改进后的网络模型

4. 实验

4.1. 数据集准备

本文实验使用的数据集包含 2400 张桥梁裂纹的图像，部分来自网络公开数据集，部分来自现场采集。我们按 8:1:1 的比例将数据集分为训练集、验证集和测试集。其中，验证集加入了部分现场采集到桥梁裂纹图片，以便更好评估模型在生产环境中性能。在目标检测中，通常物体检测任务中需要为每个物体标注边界框及其对应的类别标签。相对地，在实例分割任务中，除了标注物体的类别外，还需对每个像素进行标注，明确其属于哪一个具体实例，实现对每个实例的像素级分割，裂纹标注如图 5 所示。

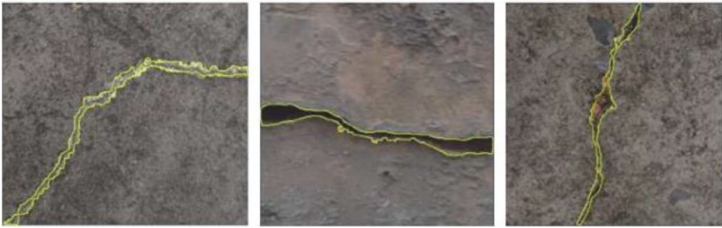


Figure 5. Annotation example
图 5. 标注示例

4.2. 实验评估

在本文的物体检测任务中，我们使用三个主要评价指标来评估模型，分别是精度、召回率和平均精度(AP)。精度作为一个指标，量化了模型在预测过程中做出准确预测的比例。它的计算方法是，在全部样本中，预测正确的样本所占的比例。本文将精度作为评估模型准确性的重要性能指标。另一方面，召回率衡量的是模型正确检测到的正样本占正确检测正样本与错误检测负样本总和的比例，用以评估模型的检测能力和覆盖范围。

在本文中，我们特别关注召回率，因为在物体检测任务中，错误检测会导致严重后果。此外，平均精度(AP)也是评估物体检测模型整体性能的重要指标。平均精度通过两种方法确定：AP@0.5 和 AP@0.5:0.95。AP@0.5 具体表现为 0.5 交集大于 IOU 时的平均精度，衡量预测边界框与实际边框之间的重叠程度。同时，AP@0.5:0.95 考虑了 IOU 值从 0.5 到 0.95 之间的平均精度。这两种 AP 计算方法都能有效评估模型在不同 IOU 临界值下的性能。需要强调的是，我们只考虑真实 IOU 大于 50%、小于 95%的类与预测边界框之间的匹配。这种方法有助于排除低质量的预测，确保评估模型对目标对象的精确定位和识别能力。

$$\text{Precision} = \frac{TP}{TP + FP} \tag{3}$$

$$\text{Recall} = \frac{TP}{TP + FN} \tag{4}$$

$$AP = \int_0^1 P(R) dR \tag{5}$$

$$mAP = \frac{\sum_{i=1}^N AP_i}{N} \tag{6}$$

4.3. 实验结果分析

基于制作的数据集，相较原始 YOLOv8s，改进模型的精度(P)由 0.865 提升至 0.900 (+3.5 个百分点)，召回率(R)由 0.781 提升至 0.793，mAP@0.5 由 0.865 提升至 0.888，mAP@0.5:0.95 由 0.567 提升至 0.578；模型大小仅由 11.1 MB 增至 12.2 MB，结果如表 1 所示。

Table 1. Comparison of experimental results
表 1. 实验结果对比表

模型	P	R	mAP 0.5	mAP 0.5:0.95	大小 (MB)
Yolov5s	0.852	0.786	0.867	0.563	7.2
Yolov7	0.859	0.785	0.868	0.56	36.4
Yolov8s	0.865	0.781	0.865	0.567	11.1
改进的 Yolov8s	0.90	0.793	0.888	0.578	12.2

为了验证我们算法的优越性，我们在相同的输入规模和相似的模型大小下进行了多次实验，并与当前物体检测中流行的 YOLOv5s [13]、YOLOv7s [14]等算法进行了比较。在相同环境下，我们的算法表现更优。尽管改进算法模型的参数略有增加，但运行效率并未受到影响。这归功于改进后的方法降低了模型的计算复杂度，同时增强了模型理解输入数据的能力。

5. 结语

本文提出了一种基于 YOLOv8 实例分割的桥梁裂纹识别方法,旨在解决桥梁结构安全性的重要问题。针对传统方法在实时性和准确性上的不足,本文提出了一系列创新性的改进。首先,在数据增强方面,通过改进马赛克策略并引入图像曝光处理和图像模糊处理,增强了模型对裂纹图像的鲁棒性和泛化能力。其次,在 C2F 模块中加入 Coordinate Attention 注意力机制,提升了模型对关键信息的关注度,从而提高了裂纹检测的准确性。实验结果表明,改进后的算法在精度、召回率和 mAP 等指标上均取得了显著提升,尽管模型的大小略有增加,但运行效率并未受到影响。本文提出的基于 YOLOv8 实例分割的桥梁裂纹识别方法在解决桥梁结构安全性问题方面具有重要意义。

基金项目

2025 年安顺市理论创新课题(联合课题)“基于大语言模型的开放教育学习辅助智能体研究”(ASLCLH-2025-15)。

参考文献

- [1] Hou, Q., Zhou, D. and Feng, J. (2021) Coordinate Attention for Efficient Mobile Network Design. 2021 *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, Nashville, 20-25 June 2021, 13708-13717. <https://doi.org/10.1109/cvpr46437.2021.01350>
- [2] He, K., Zhang, X., Ren, S. and Sun, J. (2015) Spatial Pyramid Pooling in Deep Convolutional Networks for Visual Recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **37**, 1904-1916. <https://doi.org/10.1109/tpami.2015.2389824>
- [3] Ren, S., He, K., Girshick, R. and Sun, J. (2017) Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **39**, 1137-1149. <https://doi.org/10.1109/tpami.2016.2577031>
- [4] Liu, W., Anguelov, D., Erhan, D., Szegedy, C., Reed, S., Fu, C., et al. (2016) SSD: Single Shot MultiBox Detector. In: Leibe, B., Matas, J., Sebe, N. and Welling, M., Eds., *Computer Vision—ECCV 2016*, Springer, 21-37. https://doi.org/10.1007/978-3-319-46448-0_2
- [5] Wang, J., Dai, H., Chen, T., Liu, H., Zhang, X., Zhong, Q., et al. (2023) Toward Surface Defect Detection in Electronics Manufacturing by an Accurate and Lightweight Yolo-Style Object Detector. *Scientific Reports*, **13**, Article No. 7062. <https://doi.org/10.1038/s41598-023-33804-w>
- [6] Kim, I., Jeon, H., Baek, S., Hong, W. and Jung, H. (2018) Application of Crack Identification Techniques for an Aging Concrete Bridge Inspection Using an Unmanned Aerial Vehicle. *Sensors*, **18**, Article 1881. <https://doi.org/10.3390/s18061881>
- [7] Terven, J. and Cordova-Esparza, D. (2023) A Comprehensive Review of YOLO: From YOLOv1 to YOLOv8 and Beyond. arXiv: 2304.00501.
- [8] 雷源毅, 朱文球, 廖欢. 复杂场景下的改进 YOLOv8n 安全帽佩戴检测算法[J]. 软件工程, 2023, 26(12): 46-51.
- [9] 曹士杰, 张竹林. 基于改进 YOLOv5 的矿山石块实例分割算法[J]. 兰州工业学院学报, 2023, 30(6): 19-25.
- [10] Bochkovskiy, A., Wang, C.Y. and Liao, H.Y.M. (2020) Yolov4: Optimal Speed and Accuracy of Object Detection. arXiv: 2004.10934.
- [11] 梁锦秀. 低光照条件下的图像增强和识别关键技术研究[D]: [博士学位论文]. 广州: 华南理工大学, 2021.
- [12] Zheng, C., Jia, W., Wu, S. and Li, Z. (2023) Neural Augmented Exposure Interpolation for Two Large-Exposure-Ratio Images. *IEEE Transactions on Consumer Electronics*, **69**, 87-97. <https://doi.org/10.1109/tce.2022.3214382>
- [13] 贾世娜. 基于改进 YOLOv5 的小目标检测算法研究[D]: [硕士学位论文]. 南昌: 南昌大学, 2022.
- [14] Wang, C., Bochkovskiy, A. and Liao, H.M. (2023) YOLOv7: Trainable Bag-Of-Freebies Sets New State-Of-The-Art for Real-Time Object Detectors. 2023 *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, Vancouver, 17-24 June 2023, 7464-7475. <https://doi.org/10.1109/cvpr52729.2023.00721>