

# 基于改进YOLO v5的混凝土桥梁裂缝轮廓提取研究

侯珂<sup>1</sup>, 王翔<sup>2</sup>, 刘文江<sup>3</sup>, 程文党<sup>1</sup>

<sup>1</sup>山东交通学院, 轨道交通学院, 山东 济南

<sup>2</sup>济南城建集团有限公司, 山东 济南

<sup>3</sup>山东交通学院, 航空学院, 山东 济南

收稿日期: 2023年12月19日; 录用日期: 2024年1月16日; 发布日期: 2024年1月24日

## 摘要

裂缝是混凝土桥梁的常见病害之一, 当前主要依靠人工实施检测, 存在效率低和准确率低等问题, 研究自动化、智能化的裂缝检测方法一直是行业的热点问题。文章提出了一种改进YOLO v5的深度学习方法可以解决当前桥梁裂缝检测误检率高实时性差的问题。首先, 我们将YOLO v5中主干网络替换为残差网络, 进而加深网络的深度, 通过残差连接使网络学习更复杂的特征; 然后, 在颈部网络结构的最后一层使用GCNet网络结构, 通过全局相关层, 将全局信息编码到每个位置的特征中, 可以更好地捕捉裂缝的轮廓和细节。实验表明, 该方法能够快速准确地检测出裂缝, 且误检率较低, 精确率可达到67.1%, 比原始网络提高了9.2%, 在混凝土桥梁裂缝轮廓提取任务中展现了出色的性能。

## 关键词

桥梁裂缝, 实例分割, YOLO v5s, 残差网络, GCNet

# Research on Concrete Bridge Crack Outline Extraction Based on Improved YOLO v5

Ke Hou<sup>1</sup>, Xiang Wang<sup>2</sup>, Wenjiang Liu<sup>3</sup>, Wendang Cheng<sup>1</sup>

<sup>1</sup>School of Rail Transportation, Shandong Jiaotong University, Jinan Shandong

<sup>2</sup>Jinan Urban Construction Group Co. Ltd., Jinan Shandong

<sup>3</sup>School of Aeronautics, Shandong Jiaotong University, Jinan Shandong

Received: Dec. 19<sup>th</sup>, 2023; accepted: Jan. 16<sup>th</sup>, 2024; published: Jan. 24<sup>th</sup>, 2024

文章引用: 侯珂, 王翔, 刘文江, 程文党. 基于改进 YOLO v5 的混凝土桥梁裂缝轮廓提取研究[J]. 交通技术, 2024, 13(1): 67-72. DOI: 10.12677/ojtt.2024.131008

## Abstract

Cracks are a common disease in concrete bridges. Currently, crack detection mainly relies on manual inspection, which has issues such as low efficiency and accuracy. The research on automated and intelligent crack detection methods has always been a hot topic in the industry. This paper proposes an improved deep learning method based on YOLO v5 to address the high false detection rate and poor real-time performance in current bridge crack detection. Firstly, we replace the backbone network in YOLO v5 with a residual network to deepen the network and make it learn more complex features through residual connections. Secondly, we use the GCNet network structure in the last layer of the neck network structure. Through the global contextual layer, global information is encoded into the features of each position, which helps better capture the contours and details of the cracks. Experiments show that this method can quickly and accurately detect cracks with a low false detection rate. The precision rate can reach 67.1%, which is 9.2% higher than the original network, demonstrating excellent performance in the task of concrete bridge crack outline extraction.

## Keywords

Bridge Crack, Instance Segmentation, YOLO v5s, Residual Network, GCNet

Copyright © 2024 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

在交通运输领域中，桥梁是重要的基础设施之一，对于确保交通安全和促进经济发展具有至关重要的意义。然而，混凝土桥梁长期使用和环境因素可能导致裂缝问题，数据显示裂缝是桥梁损坏的主要因素之一[1][2]。这将严重影响桥梁的结构稳定性和使用寿命。依靠人眼传统的混凝土桥梁裂缝检测方法主要依赖于人工巡查和目视检测，然而这种方法存在人力成本高、效率低和主观性强的问题。因此，提出一种能够快速准确地检测和识别混凝土桥梁裂缝的方法对于未来大量混凝土桥梁的检测任务至关重要。

近年来，基于深度学习的图像处理技术在众多领域取得了重要突破，为解决实际问题提供了全新的解决方案。由于相对于传统算法的优越性，深度学习相关算法在工业领域得到了广泛应用[3]。深度学习的概念由 G.E.Hinton 等人[4]于 2006 年提出，它包含多隐含层和感知层，能够更好地描述抽象和深层次的信息。将深度学习算法应用于桥梁裂缝检测领域可以实现桥梁裂缝的自动化和实时检测，从而减轻桥梁养护部门的工作负担。由于桥梁裂缝检测对实时性的要求，需要使用检测速度更快的目标检测算法。YOLO (You Only Look Once)是代表单阶段检测算法的一种方法，其速度更快[5]。随着 YOLO v3 [6]、YOLO v4 [7]和 YOLO v5 的先后提出，YOLO 系列算法在检测速率和精度方面表现出较大的优势，被广泛应用于结构裂缝的智能识别。Zhang 等人[8]提出了一种改进的 YOLO v3 模型，用于桥梁病害检测。他们通过引入焦点损失函数(Focal Loss)和新的迁移学习方法，成功地提高了病害识别的准确率。Jang 等人[9]提出了一种基于混合扩展卷积块的深度学习网络 HDCB-Net，并结合 YOLO v4 实现像素级的裂缝检测。相较于之前的版本，YOLO v5 系列算法在识别精度和检测效率方面都得到了显著提升，可以实现桥梁裂缝的实时精确检测[10]。

本文基于 YOLOv5 V7.0 实例分割模型进行改进, 并采用 Python 3.8 编写模块程序。我们使用了 2164 张和 538 张带有标记信息的图像分别作为改进后模型和原有模型的训练集和验证集, 进行了检测对比实验。对比结果显示, 改进后的模型在轮廓提取效果上有明显提升, 具有更广泛的应用价值。

## 2. YOLO v5 模型介绍

YOLO v5 是基于深度学习的目标检测模型框架, YOLO (You Only Look Once) 系列模型的核心思想是将目标检测任务转化为一个单次前向传播的问题, YOLO v5 根据网络深度和宽度的大小不同分为 s、m、l 和 x 四个版本, 其中 YOLO v5s 是 YOLO 系列算法中网络深度最小的一个, 也是检测速度最快的, 可以满足检测实时性的需求。

YOLO v5 v7.0 的网络模型结构可以分为主干网络(Backbone)、特征金字塔网络(FPN)和分割头(Segmentation Head)三个主要部分。

1) 主干网络(Backbone): 主干网络是 YOLO v5 的特征提取器, 负责从输入图像中提取特征。YOLO v5 采用了 CSPDarknet53 作为主干网络, 它是一种深度卷积神经网络结构。由一系列的卷积层、批量归一化层和激活函数层组成。该结构使用了 CSP (Cross Stage Partial) 连接, 能够减少参数量同时提高特征的代表能力。CSPDarknet53 可以从输入图像中提取出丰富的语义特征, 用于后续的裂缝分割任务。

2) 特征金字塔网络(FPN): 特征金字塔网络是为了解决不同尺度目标问题而引入的。它通过在主干网络中添加额外的卷积层和上采样操作, 将不同层级的特征进行融合, 从而得到具有多尺度信息的特征图。YOLO v5 使用的特征金字塔网络是 PANet (Path Aggregation Network)。PANet 通过自底向上和自顶向下的路径来进行特征融合。自底向上路径从主干网络的底层特征图中提取语义信息, 并进行上采样操作。自顶向下路径则将上一层的特征图和自底向上路径的特征图进行融合。通过这种融合操作, PANet 能够生成具有多尺度信息的特征图, 用于检测不同大小的目标。

3) 分割头(Segmentation Head): 分割头用于预测每个目标的掩码。首先, 分割头接收来自检测头的特征图和预测框, 然后对特征图进行上采样, 使其与输入图像的分辨率相同。接着, 分割头根据预测框的位置, 对特征图进行裁剪, 得到每个目标的局部特征图。然后, 分割头对每个局部特征图进行卷积操作, 得到每个目标的掩码预测, 掩码的形状为  $h \times w \times 1$ , 其中  $h$  和  $w$  是预测框的高度和宽度。最后, 分割头将每个掩码预测映射到输入图像的坐标系中, 得到每个目标的最终掩码输出。

## 3. 算法改进

### 3.1. 主干网络结构改进

ResNet 的核心原理包括残差块、堆叠残差块和全局平均池化。残差块引入了残差学习的概念, 通过跳跃连接将输入直接添加到输出, 有助于梯度更轻松地传播, 缓解了梯度消失和训练困难的问题。通过堆叠多个残差块, ResNet 构建了深层网络, 使网络能够学习更深层次的特征表示。最后, 采用全局平均池化层而非全连接层, 以减少参数量、降低过拟合风险, 并有助于提取整体图像的特征。ResNet 网络结构图如图 1 所示。

将 YOLOv5 中的主干网络替换为 ResNet 对于裂缝分割任务将带来多方面的好处。通过残差学习实现对输入和目标之间的残差的学习, 有助于更好地捕捉图像中的复杂特征, 尤其是对于裂缝分割任务中的边缘、纹理等细节信息。其深层结构使得网络能够学习层次化的特征表示, 更好地理解图像的不同层次信息。残差连接的引入有助于梯度在网络中更顺畅地传播, 有助于缓解梯度消失的问题, 对于训练深层网络和处理裂缝分割等需要高级特征的任务至关重要。

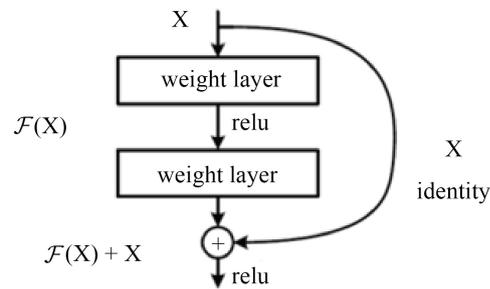


Figure 1. ResNet network architecture diagram

图 1. ResNet 网络结构图

### 3.2. GCNet 注意力机制

GCNet 是一种基于注意力机制的网络结构，它通过引入全局相关层来增强特征图的表达能力和鲁棒性。对于裂缝轮廓提取这样一个细粒度的分割任务，在颈部网络最后一层使用 GCNet 注意力机制，可以通过全局相关层，将全局信息编码到每个位置的特征中，更好地捕捉裂缝的轮廓和细节。GCNet 网络如图 2 所示。

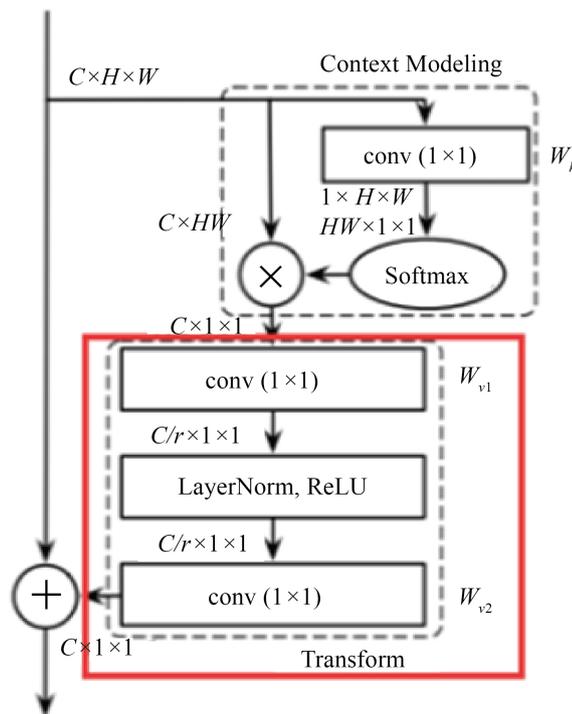


Figure 2. GCNet network architecture diagram

图 2. GCNet 网络结构图

## 4. 实验

### 4.1. 数据集来源及模型训练

实验数据集为在 AI 图像识别模型训练数据标注平台 Roboflow 使用的公开数据集，包含混凝土桥梁裂缝图像 2702 张，按照 8:2 的比例随机将其分为训练集和验证集。

模型训练在本人电脑上进行,使用 CPU 为 AMD Ryzen7 5800H, RAM 为 16 GB 随机存取内存, GPU 为 RTX 3060 Laptop, 6 GB 显示内存。深度学习框架为 pytorch, 整个训练过程设置 epoch 为 156, 根据数据集的大小和模型大小情况将 Batch\_size 设置为 4, 训练时使用 SGD 优化算法进行参数优化, 初始学习率为 0.01, 输入图片分辨率为  $640 \times 640$ 。

## 4.2. 实验结果与分析

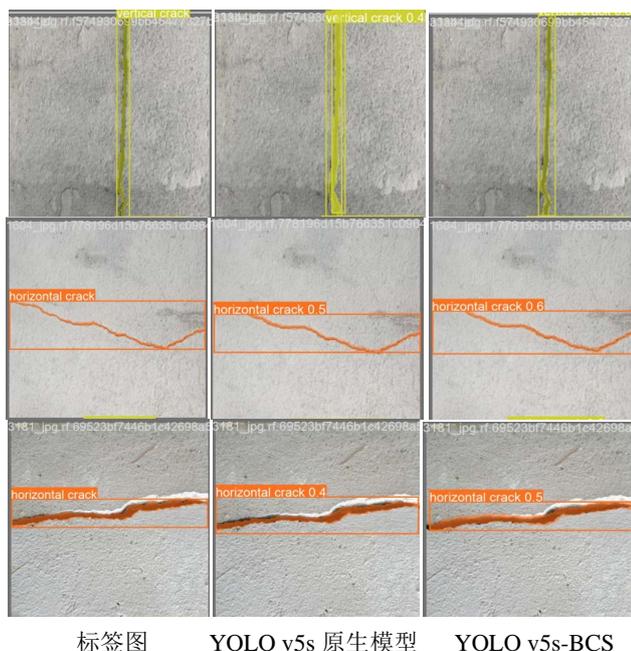
为了验证本文所提出模型方法的有效性, 考虑混凝土桥梁裂缝轮廓提取实时性的要求, 使用 YOLO v5s 模型作为原始模型, 分别与原生的 YOLO v5s、使用 Resnet 作为主干网络的 YOLO v5s-RS、加入 GCNet 注意力机制并将主干网络替换为残差网络的 YOLO v5s-GERS 和本文提出的添加 GCNet 注意力机制并将模型中主干网络替换为残差网络的 YOLO v5s-BCS 通过在相同数据集上进行实验分析。

通过表 1 可以看出, 改进后的 YOLO v5s-BCS 相较于 YOLO v5s 的原生模型效果有明显提升。通过将主干网络替换为残差网络, YOLOv5s-RS 相较于原生模型精确率提升了 3.8%; 由于在添加 GCNet 注意力机制的基础上将主干网络替换为残差网络, 召回率提升了 0.7%; 本文提出的模型在 YOLO v5s-RS 的基础上在颈部网络最后一层使用 GCNet 注意力机制, 精确率提升了 9.2%。

**Table 1.** Experimental results of YOLO v5s model and its improved version

**表 1.** YOLO v5s 模型与其改进后模型实验结果

Model	mAP50	Map50-95	Precision	Recall
YOLOv5s	0.485	0.209	0.632	0.524
YOLOv5s-RS	0.482	0.216	0.67	0.529
YOLOv5s-GERS	0.443	0.2	0.619	0.531
YOLOv5s-BCS	0.431	0.199	0.671	0.474



**Figure 3.** Comparison chart of test results

**图 3.** 检测结果对比图

通过图 3 可以看出, 本文提出的模型在裂缝的定位识别方面相较于原生模型有提升, 原生模型存在一条裂缝识别出多个裂缝的情况, 而改进后的模型识别定位更为准确, 且在评价指标上表现更优, 使用改进后的模型能更完整的对裂缝轮廓进行提取, 可以满足日常混凝土桥梁裂缝轮廓提取的需求。

## 5. 结语

在本研究中, 以基于深度学习的混凝土桥梁裂缝检测为主题, 对 YOLOv5s 模型进行了改进和优化。通过将模型主干网络替换为残差网络, 并在颈部网络部分通过在网络最后一层使用 GCNet 注意力机制成功地提升了轮廓提取精确率和性能。在相同的数据集上, 对改进后的模型与 YOLOv5s 原生模型进行了对比实验。实验结果表明, 经过改进后的模型在精确率指标上取得了 9.2% 的提升。这证明了我们所引入的改进策略的有效性和优越性。通过优化网络结构和引入注意力机制, 提高了混凝土桥梁裂缝轮廓提取的准确性和效率, 为相关领域的研究和实际应用提供了有价值的参考。我们相信这些工作能够为交通运输领域的桥梁安全和维护提供有益的技术支持, 并为未来相关研究和应用提供了启示和借鉴。

## 课题来源

交通运输行业重点科技专项, 项目编号: 2022-ZD4-066。

## 参考文献

- [1] Bu, G.P., Chanda, S., Guan, H., *et al.* (2015) Crack Detection Using a Texture Analysis-Based Technique for Visual Bridge Inspection. *Electronic Journal of Structural Engineering*, **14**, 41-48. <https://doi.org/10.56748/ejse.141881>
- [2] 阮小丽, 王波, 荆国强, 等. 桥梁混凝土结构表面裂缝自动识别技术研究[J]. 世界桥梁, 2017, 45(6): 55-59.
- [3] 孙志军, 薛磊, 许阳明, 等. 深度学习研究综述[J]. 计算机应用研究, 2012, 29(8): 2806-2810.
- [4] Hinton, G.E. and Salakhutdinov, R.R. (2006) Reducing the Dimensionality of Data with Neural Networks. *Science*, **313**, 504-507. <https://doi.org/10.1126/science.1127647>
- [5] Du, Y.C., Pan, N., Xu, Z.H., *et al.* (2021) Pavement Distress Detection and Classification Based on Yolo Network. *International Journal of Pavement Engineering*, **22**, 1659-1672. <https://doi.org/10.1080/10298436.2020.1714047>
- [6] Redmon, J. and Farhadi, A. (2018) YOLOv3: An Incremental Improvement. <https://arxiv.org/abs/1804.02767>.
- [7] Bochkovskiy, A., Wang, C.Y. and Liao, H.Y.M. (2020) YOLOv4: Optimal Speed and Accuracy of Object Detection. <https://arxiv.org/abs/2004.10934>.
- [8] Zhang, C.B., Chang, C.C. and Jamshidi, M. (2020) Concrete Bridge Surface Damage Detection Using a Single-Stage Detector. *Computer-Aided Civil and Infrastructure Engineering*, **35**, 389-409. <https://doi.org/10.1111/mice.12500>
- [9] Jiang, W., Liu, M., Peng, Y., *et al.* (2020) HDCB-Net: A Neural Network with the Hybrid Dilated Convolution for Pixel-Level Crack Detection on Concrete Bridges. *IEEE Transactions on Industrial Informatics*, **17**, 5485-5494. <https://doi.org/10.1109/TII.2020.3033170>
- [10] Shu, Z.K., Yan, Z.Y., Xu, X.H., *et al.* (2021) Pavement Crack Detection Method of Street View Images Based on Deep Learning. *Journal of Physics: Conference Series*, **1952**, 22-43. <https://doi.org/10.1088/1742-6596/1952/2/022043>