ISSN: 2710-0588

A Review on Deep Learning-Based Segmentation Methods for Concrete Cracks in Bridges and Pavements

Chunlei Zhao, Ziqin Ye, Zhan Zhang, Yuanyuan Cao, and Liwei Wu*

North China University of Science and Technology, Tangshan 063210, China

*Corresponding Author

Abstract

As critical transportation infrastructure, the structural integrity of bridges and pavements directly impacts the safety and service life of transportation systems. Traditional manual crack detection methods suffer from inefficiency, subjectivity, and difficulties in result quantification. In recent years, the rapid advancement of computer vision technologies and artificial intelligence theories has provided effective solutions to these challenges, particularly through intelligent crack segmentation methods based on deep learning, which have demonstrated significant advantages.This paper systematically reviews the application of mainstream segmentation models-represented by encoder-decoder architectures, multi-scale feature fusion, and attention mechanisms-in this field. It provides an in-depth analysis of the technical characteristics, performance strengths and limitations, and applicable scenarios of various algorithms. By summarizing experimental findings and offering insights, this study aims to provide theoretical reference and forward-looking guidance for promoting the practical development and intelligent advancement of structural health monitoring technologies.

Keywords

Deep Learning, Segmentation, Multi-scale Feature Fusion, Attention Mechanism.

1. Introduction

The detection of surface defects in bridges and pavements has traditionally relied on manual visual inspection and simple tools. This approach exhibits clear limitations in modern engineering contexts, including low efficiency, subjectivity, and an inability to provide precise, quantifiable data. These shortcomings are particularly evident in scenarios such as inspecting the interior of large cross-sea bridge girders, conducting routine assessments of extensive highway networks, and performing rapid post-storm evaluations. In recent years, advancements in computational power and the adoption of data-driven methods have facilitated the emergence of intelligent crack segmentation based on deep learning. This technology can automatically identify and quantify crack characteristics-even detecting fine cracks invisible to the naked eye-thereby significantly improving the objectivity and accuracy of inspections.

Deep learning-based crack segmentation is a sophisticated application of computer vision for pixel-level image analysis in engineering. This method utilizes deep neural networks to autonomously learn and extract multi-level crack features, enabling precise pixel-wise identification and contour delineation in images. It reliably distinguishes cracks from challenging backgrounds like shadows, oil stains, and surface textures, and accurately quantifies key parameters such as width, length, and distribution patterns. Consequently, it provides reliable data for assessing structural health, predicting residual service life, and formulating scientific maintenance strategies. A systematic review of deep learning advances

ISSN: 2710-0588

in this domain is therefore crucial, holding significant theoretical and practical value for advancing intelligent and automated civil infrastructure inspection.

2. Research on Concrete Crack Segmentation Methods

Scholars worldwide have significantly enhanced the precision, efficiency, and utility of crack segmentation through model development, dataset creation, image processing, and algorithm optimization.

Wang Lu et al.[1]systematically compared the performance of 14 segmentation models from the YOLO v8, v9, and v11 series on a self-built dataset (BACD 2025), investigating the impact of different network architectures on concrete crack segmentation. The results indicate that the YOLO v9c-seg model achieved the best overall performance, with a mAP@0.5 of 98.1% and an F1-score of 94.99%, striking an optimal balance between accuracy and speed. The study also revealed that the number of model parameters and computational complexity are not positively correlated with performance. Lightweight models, such as YOLO v8n-seg, maintained high accuracy while achieving an inference speed of up to 312.5 FPS, making them more suitable for real-time detection scenarios. This research provides direct guidance for model selection in engineering practice.

Feng Jingyi et al.[2]proposed an R-PANNet model integrating multi-scale features and an attention mechanism, employing ResNet-50 as the backbone network combined with a Path Aggregation Network (PANet) and a dual attention mechanism (Channel Attention Module and Position Attention Module). Experimental results on a self-constructed dam surface crack dataset showed that the model achieved an Intersection over Union (IoU) of 82.02% and an F1-score of 90.12% for crack segmentation. The study revealed that deploying the attention mechanism at deeper network layers (P5 level) most effectively enhances semantic information and significantly improves segmentation accuracy. Additionally, the segmentation results enabled precise quantification of geometric features such as crack area, length, and width. This method demonstrates outstanding capability in capturing crack characteristics under complex backgrounds.

Peng Yaopan et al.[3]addressed the requirements for model lightweighting and real-time detection by proposing an improved DeepLabV3+ crack segmentation algorithm. Using MobileNetV3 as the backbone network, the method incorporates an Intra-Scale Feature Interaction module (AIFI), Normalized Attention Module (NAM), hybrid attention mechanism (ACmix), and a C2f-SCConv decoding module. On the Concrete3k and Asphalt3k datasets, the model achieved an mIoU of 86.21%, reduced parameter count by 88.1%, and reached an average frame rate of 47.91 fps, significantly improving operational efficiency while maintaining high accuracy. This study offers a viable solution for deployment on edge devices, though its generalization capability in complex backgrounds requires further validation.

Yang Hanlong et al.[4]proposed an automated annotation strategy integrating Otsu binarization and Canny edge detection, constructing a Swin-Unet model by combining Swin-Transformer with Unet. On a self-constructed dataset of 19,101 crack images, the model achieved 100% classification accuracy and 93.1% IoU segmentation precision. The results demonstrate that the hierarchical attention mechanism of Swin-Transformer effectively enhances the model's ability to correlate global and local features, significantly improving crack segmentation performance. The study also analyzed the aggregation characteristics of crack orientation and shape using minimum enclosing rectangles, providing insights for crack propagation prediction.

In summary, the collective findings of researchers demonstrate strong consistency, revealing a clear consensus on multiple key aspects of concrete crack identification technology. Innovations in model architecture-particularly the incorporation of attention mechanisms,

ISSN: 2710-0588

multi-scale feature fusion, and Transformer structures-have been shown to significantly enhance a model's ability to capture complex crack patterns and global contextual information, thereby effectively improving segmentation accuracy. Meanwhile, lightweight design strategies substantially increase detection speed while maintaining high precision.

3. Research on Pavement Crack Segmentation Methods

Scholars worldwide have significantly enhanced the precision, robustness, and practicality of pavement crack segmentation through the development of deep learning models, incorporation of frequency-domain feature enhancement mechanisms, and optimization of network architectures and post-processing workflows.

Wang Jianxin et al.[5]proposed a pavement crack segmentation method that integrates frequency-domain Mamba with a recursive gated Transformer. This approach adopts a U-shaped encoder-decoder architecture. The encoder incorporates a large-kernel recursive attention mechanism and a gated edge feed-forward network to enhance local and global feature interaction. A frequency-domain Mamba enhancement module is embedded in the skip connections, which uses wavelet transform to separate image content from noise, thereby improving the model's perception of high-frequency details. Experiments on three public datasets-DeepCrack537, CrackLS315, and YCD-demonstrate that this method outperforms mainstream segmentation models in metrics such as mIoU and F1-score, exhibiting stronger robustness especially in scenarios with complex backgrounds and weak edges. The study provides a new perspective of cross-domain feature fusion for the fine-grained segmentation of pavement cracks.

Dai Shaosheng et al.[6]proposed an image processing-based method for pavement crack feature extraction. This approach first applies morphological processing and connected component analysis to deep learning segmentation results, filtering out interference regions by combining grayscale, area, and shape features. Subsequently, an improved fast parallel thinning algorithm is used to extract the crack skeleton, followed by skeleton connection and growth based on endpoint direction and spacing to address discontinuities and shortening issues. Results demonstrate that this method effectively produces continuous, complete, single-pixel crack skeletons without burrs, achieving precision and recall rates of 92.97% and 93.29% respectively, outperforming traditional Zhang-Suen and medial axis transformation methods. The study establishes a complete workflow from pixel-level segmentation to morphological feature extraction, providing a reliable foundation for quantitative crack assessment.

Tang Yuan et al.[7]proposed an improved SegFormer-based algorithm for pavement crack segmentation. This method discards the original multilayer perceptron decoder of the SegFormer network and designs a feature fusion network that integrates multi-scale features. It introduces a channel attention module to suppress redundant information, and constructs a spatially efficient convolutional pooling module incorporating partial convolutions to enhance high-level semantic information extraction. Additionally, a random replacement data augmentation strategy is proposed to improve the model's generalization performance. Experiments on the Crack500 dataset show that the improved model achieves an F1-score of 90.01% and an mIoU of 82.17%, representing improvements of 1.03% and 1.32% over the original SegFormer, and demonstrates superior performance in segmenting thin and indistinct cracks. This research effectively enhances the model's adaptability to multi-scale cracks while maintaining a lightweight design.

In summary, the research findings of various scholars demonstrate strong consistency, revealing a clear consensus on multiple key aspects of automated pavement crack segmentation and feature extraction technology. The incorporation of frequency-domain information effectively enhances the model's perception of crack edges and global context, significantly improving segmentation accuracy and anti-interference capability. Lightweight network designs and efficient decoding architectures reduce computational complexity while maintaining high performance, facilitating practical deployment. Additionally, post-processing optimization through traditional image processing methods further improves the continuity and completeness of segmentation results, providing reliable support for crack morphology analysis and quantitative evaluation.

4. Research on Bridge Crack Segmentation Methods

Scholars worldwide have conducted in-depth research and optimization on key performance metrics of concrete bridge crack segmentation-including segmentation accuracy, detection efficiency, model lightweighting, and measurement precision-by introducing methods such as multi-scale feature fusion, lightweight network design, Fully Convolutional Neural Networks, and crack measurement algorithms.

Song Fengquan et al.[8]proposed a semantic segmentation method for concrete bridge cracks based on multi-scale features. The approach comprises a dual-branch multi-scale feature fusion network (DMFFNet) and a lightweight Strip-DeepLabV3+ network. DMFFNet employs a dual-encoder branch to extract global semantic features and local detail features separately. It incorporates a context information perception module, a hybrid-scale feature fusion module, and a multi-scale wavelet fusion module to enhance feature representation capability and mitigate information loss during downsampling. Strip-DeepLabV3+ reduces computational complexity by replacing the backbone network with MobileNetV2, and introduces dynamic strip convolution, a dynamic strip pyramid pooling module, and a strip attention gate module to improve perception of slender cracks. Experimental results show that DMFFNet achieved mIoUs of 86.12% and 85.07% on the Crack Dataset and Bridge_Crack_Image_Data dataset, respectively, with the highest recall reaching 93.94%. Meanwhile, Strip-DeepLabV3+, while reducing parameters by 51.75M and computational load by 201.49G, increased mIoU by 3.46%, effectively balancing segmentation accuracy with model lightweighting.

Yao Yukai et al.[9]addressed the issues of high computational cost and low segmentation speed of the Deeplabv3+ model in bridge crack segmentation by proposing an improved Deeplabv3+ algorithm. The enhanced approach incorporates an RFB multi-branch convolutional module after the ASPP module to expand the receptive field, replaces the backbone network with MobileNetv2 to improve speed, employs depthwise separable convolution instead of standard convolution to reduce parameter count, and adds an additional low-level feature fusion step to enhance detail segmentation. Experimental results demonstrate that the improved model achieves a detection accuracy of 90.15% on a self-constructed dataset, representing a 4.07% increase over the original model, while the inference speed (FPS) improves from 36.23 to 53.38. Both segmentation accuracy and speed are significantly enhanced, highlighting the practical value of this approach for crack detection tasks.

To address the issues of insufficient denoising effects and poor crack continuity in traditional image segmentation methods, Hu Wenkui et al.[10]proposed an automatic bridge crack segmentation model (BCI-AS) based on a fully convolutional neural network (FCN), along with a crack width measurement algorithm utilizing projection technology and least squares fitting of the centerline. The BCI-AS model achieves end-to-end pixel-level segmentation through an encoder-decoder (contraction path and expansion path) architecture. It incorporates image preprocessing techniques-such as grayscale conversion, piecewise linear transformation, and median filtering-collectively referred to as the GPM method, to enhance image quality by improving contrast and reducing noise. For width measurement, the algorithm obtains the coordinates of the crack's upper and lower edges via projection technology, fits the centerline using the least squares method, and calculates the perpendicular width relative to this

ISSN: 2710-0588

centerline. Experimental results demonstrated that the BCI-AS model achieved a segmentation accuracy of 94.45%, produced well-connected cracks after segmentation, and kept the relative error of width measurement below 7% with fast computational speed. This validates the feasibility and accuracy of the proposed algorithm for crack segmentation and width calculation.

In summary, the experimental results of various authors align well, demonstrating that the integration of Fully Convolutional Neural Networks with traditional image processing techniques can effectively enhance the segmentation accuracy, detection efficiency, and measurement precision of concrete bridge cracks. Consistent progress has been achieved in reducing model parameters, lowering computational costs, and improving model generalization capability.

5. Conclusion

Through a systematic review and analysis of recent research on deep learning-based segmentation methods for concrete cracks in bridges and pavements, it has been observed that significant progress has been made in this technological field. However, the following issues still require further investigation: Currently, most models struggle with generalization capability and robustness when dealing with complex background interference, varying lighting conditions, and various pseudo-cracks in practical engineering scenarios. Additionally, the precise segmentation and continuity preservation of fine cracks, initial cracks, and crack tips remain challenging problems for current technologies.

Acknowledgments

Basic scientific research business cost project of provincial universities of North China University of Science and Technology(JJC2024067).

Study on the Behavior and Design Method of Grouted Sleeve Connections for Steel Tubes under Combined Compression, Bending, and Torsion(S202510081040).

References

[1] WANG Lu. Research on Semi-automatic Construction of Concrete Crack Semantic Segmentation Dataset and Model Performance[D]. Dalian Jiaotong University, 2025. DOI:10.26990/d.cnki.gsltc.2025.000826.

[2] FENG Jingyi, LIANG Hui, QI Zhiyong, et al. Concrete Crack Semantic Segmentation Model Integrating Multi-scale Features and Attention Mechanism[J]. Journal of Hydroelectric Engineering, 2025, 44(09): 114-124.

[3] PENG Y P, ZHANG R F, LIU Y H, et al. Lightweight concrete crack segmentation algorithm incorporating feature interaction and attention mechanisms[J]. Journal of Optoelectronics·Laser, 2025, 36(07): 722-732. DOI:10.16136/j.joel.2025.07.0070.

[4] YANG H L, CHEN J J, PAN Y. Research on automatic annotation and segmentation method for dam surface concrete cracks using Swin-Unet[J]. Journal of Hydroelectric Engineering, 2024, 43(12): 23-33.

[5] WANG Jianxin, LI Linyi, WANG Jin, et al. Pavement Crack Segmentation Method Integrating Frequency Domain Mamba and Recursively Gated Transformer[J/OL]. China Journal of Highway and Transport, 1-17[2025-11-15]. Link.

[6] DAI Shaosheng, MAO Xinghua, YU Zi'an. Image Processing-Based Method for Pavement Crack Feature Extraction[J]. Semiconductor Optoelectronics, 2024, 45(03): 508-514. DOI:10.16818/j.issn1001-5868.2024010304.

ISSN: 2710-0588

[7] TANG Y, DONG S J, LIU C, et al. A pavement crack segmentation algorithm based on improved SegFormer network[J]. Journal of Shaanxi University of Science & Technology, 2024, 42(03): 166-173. DOI:10.19481/j.cnki.issn2096-398x.2024.03.016.

[8] SONG Fengquan. Research on Semantic Segmentation Method for Concrete Bridge Cracks Based on Multi-scale Features[D]. Chongqing Jiaotong University, 2025. DOI:10.27671/d.cnki.gcjtc.2025.000729.

[9] YAO Yukai, GUO Baoyun, LI Cailin, et al. Research on Bridge Crack Segmentation Algorithm Based on Improved Deeplabv3+[J]. Journal of Shandong University of Technology(Natural Science Edition), 2024, 38(02): 21-26. DOI:10.13367/j.cnki.sdgc.2024.02.008.

[10] HU Wenkui, DENG Hui, FU Zhixu, et al. Bridge Crack Segmentation and Measurement Method Based on Fully Convolutional Neural Network[J]. Industrial Construction, 2022, 52(04): 192-201+218. DOI:10.13204/j.gyjzG21053111.