ISSN: 2710-0588

Deep Learning-Based Crack Detection in Concrete Bridge Decks and Pavements: A Experimental Review

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

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

*Corresponding Author

Abstract

As transportation infrastructure gradually ages and maintenance costs continue to rise, efficient and precise defect detection of bridges, road surfaces, and concrete structures has become a critical factor in ensuring operational safety. Traditional manual crack identification methods are not only inefficient and subjective but also pose safety risks. The rapid advancement of computer vision and artificial intelligence technologies, particularly deep learning, offers effective solutions to address these challenges. This paper systematically reviews research on the application of deep learning models in typical tasks such as crack image classification, object localization, and pixel-level segmentation. It provides an in-depth analysis of the recognition mechanisms, technical advantages, and applicable scenarios of mainstream models, including convolutional neural networks and generative adversarial networks. Furthermore, it summarizes experimental findings and presents personal insights, aiming to provide reference and facilitate further research and application of deep learning in the intelligent maintenance and management of civil engineering infrastructure.

Keywords

Deep Learning, Attention Mechanism, Convolutional Neural Network (CNN).

1. Introduction

Crack identification methods have traditionally relied on manual inspection and semi-automatic detection based on conventional image processing. However, these methods increasingly reveal their limitations in complex real-world engineering environments, failing to meet the demands for high precision and low cost. For instance, in scenarios such as large bridge cables, tunnel interior structures, high-altitude domes, or areas with complex background interference, traditional approaches are not only inefficient and costly but also struggle to ensure the accuracy of identification results, thus falling short of the urgent need for intelligent and high-precision maintenance in modern engineering. To address this industry pain point, researchers worldwide have continuously pursued in-depth studies. Over time, deep learning-based crack identification methods have emerged accordingly and have permeated the field of civil engineering health monitoring.

At its core, crack identification technology based on deep learning is a computer vision approach that utilizes model architectures such as convolutional neural networks to automatically learn and extract multi-level features from massive datasets of crack images, ultimately enabling end-to-end intelligent recognition. This technology not only effectively overcomes interference from complex backgrounds, varying illumination conditions, and fine cracks but also demonstrates high detection efficiency and repeatability. As a result, it enables large-scale, automated, and high-precision crack identification, significantly advancing the intelligence level of structural health monitoring. Therefore, conducting a systematic review and research on deep learning-based crack identification methods is of critical importance for

ISSN: 2710-0588

promoting both the theoretical development and practical engineering application of this technology.

2. Research on Concrete Crack Identification Methods

Scholars worldwide have conducted in-depth analysis and optimization of key performance aspects in crack identification-such as classification accuracy, segmentation performance, model lightweighting, and anti-interference capabilities-by constructing diverse datasets and deep learning models. Various methods have demonstrated distinct advantages in different application scenarios, collectively advancing the development of this field.

To address the challenge of scarce high-quality crack datasets, Pan Yuan et al.[1]proposed a deep learning-based object detection framework that integrates textural features with concrete crack data. This method fuses textural features-such as entropy and contrast extracted from the gray-level co-occurrence matrix-with preprocessed image data to increase the number of feature channels. The results demonstrate that on a self-constructed steel-fiber reinforced concrete crack dataset, this approach effectively reduces the model's dependency on a large number of training samples and accelerates the training process. Under limited sample conditions, the proposed method not only improves detection accuracy compared to models without textural feature fusion but also significantly decreases the required training time. This study offers a practical solution for achieving efficient crack identification in data-scarce scenarios.

In addressing the balance between model lightweighting and accuracy, Chen Ruiqi et al.[2]investigated the application of lightweight convolutional neural networks in crack identification and segmentation to tackle the issues of high computational parameters and hardware demands in traditional models. The authors proposed an improved MobileNetV3-C/Sim classification model and an ES-DeepLab segmentation model, respectively. The results demonstrated that the classification model, by incorporating a C/Sim parallel attention mechanism, reduced computational parameters by 30.17% while maintaining high classification accuracy (97.00% for binary classification, 98.62% for multi-class classification). The segmentation model, achieved through backbone network replacement, multi-scale feature fusion, and the introduction of a Sobel edge attention mechanism, reduced parameters by 84.88% while increasing the crack Intersection over Union (IoU) by 3.51%. Its segmentation performance for complex backgrounds and fine cracks significantly outperformed comparative models. This study achieved an effective balance between accuracy and lightweight design, providing important insights for the practical deployment of models on edge devices. However, the automated workflow of its cascaded algorithm still has room for optimization in terms of real-time performance.

In the context of integrating traditional image processing with deep learning, Xu Jiaqi et al.[3]developed a concrete crack identification framework (CCI) based on edge extraction and the VGG16 network. This method begins with a series of preprocessing steps-including grayscale conversion, median filtering, and morphological operations-to enhance features, followed by edge extraction using the Sobel operator. The processed image is then fed into the VGG16 network for classification. Results show that the framework achieves 96% identification accuracy on the test set. With GPU acceleration, the identification time for a single engineering image is reduced to less than 1 second. In practical applications on roads and concrete specimens, it maintains approximately 90% accuracy, demonstrating both efficiency and practicality. The study highlights the effectiveness of combining traditional image processing techniques with classical deep learning models. However, the preprocessing pipeline is relatively cumbersome, and the model's false positive rate may increase when confronted with complex interference patterns resembling crack morphology.

In addressing the balance between model lightweighting and accuracy, Wu Xionghua et al.[4]focused on the issues of low recognition accuracy and insufficient model generalization for concrete crack identification in complex backgrounds. They investigated the application of an improved residual structure-based ResNet+ network model for classifying genuine and pseudo cracks. By optimizing the residual connections and introducing quasi-residual connections to replace traditional 3×3 convolution kernels, the network's capability for extracting fine-grained crack features was enhanced. Experimental results demonstrated that the ResNet+ model achieved a classification accuracy of 98.40%, a recall rate of 99.20%, and an F1-score of 98.80% on their self-constructed dataset. In cross-dataset testing, it attained accuracy rates exceeding 94% on public datasets like CRACK500 and CFD, showing strong generalization capability. Furthermore, the model has only 16.87 million parameters, which is significantly reduced compared to the original ResNet model, and it also outperforms classical networks like VGGNet in terms of memory usage and computational efficiency. This study achieved model lightweighting while maintaining high accuracy through structural optimization, providing a feasible solution for deployment on embedded devices.

For automated inspection of large-scale infrastructure, Xu Liang et al.[5]addressed the limitations of traditional manual inspection-such as low efficiency and safety concerns-by investigating crack detection and quantification methods in UAV-based intelligent inspection systems. The authors employed the YOLOv5 algorithm for crack detection and integrated the U^2-net model to achieve pixel-level crack segmentation and parameter measurement. By analyzing the relationship between flight altitude, camera zoom ratio, and Ground Sampling Distance (GSD), the optimal inspection parameters were determined. Experimental results showed that at 2.0x zoom and 8m flight altitude, the accuracy of crack width identification reached 95.90%–99.10%, with the average relative error controlled within 6.90%–8.30%. By maintaining GSD ≤ 1.0 mm, the system effectively identified millimeter-wide cracks and consistently achieved high recall rates and low false alarm rates across various inspection scenarios. The study established a complete workflow from image acquisition and crack extraction to parameter calculation, offering a practical approach for automated dam crack detection and intelligent early warning. However, the system’s path planning and obstacle avoidance capabilities in complex vertical structures and multi-obstacle environments still require further improvement.

In summary, the experimental results from various researchers are closely aligned, demonstrating that by integrating textural features, introducing attention mechanisms, optimizing network architectures, and combining traditional image processing with UAV (drone) technology, significant improvements can be achieved in the classification accuracy, segmentation performance, and model lightweighting of concrete crack identification. Furthermore, consistent progress has been made in reducing sample requirements, lowering computational costs, and enhancing training speed and generalization capabilities.

3. Research on Bridge Crack Identification Methods

Scholars worldwide have continuously optimized and enhanced key performance metrics in crack identification tasks-such as classification accuracy, segmentation precision, model lightweighting, and anti-interference capability in complex scenarios-by constructing diverse bridge crack datasets and designing various deep learning models. These efforts have steadily advanced the technology toward practical implementation.

To address the detection challenges of complex backgrounds and easily lost crack details in images of ancient architectural bridges, Zhu Qiankun et al.[6]proposed a crack identification and measurement method based on an improved YOLO11 and SegFormer. To reduce computational costs, the authors introduced the StarNet lightweight backbone network. They combined it with an HSANet neck network to enhance edge detail retention and designed an OSCD detection head to optimize multi-scale detection efficiency. In the segmentation stage, a Feature Fusion Module (FFM) and a High-Low Frequency Decomposition Block (HLFDB) were incorporated to mitigate information loss during downsampling. Experimental results demonstrated that the YOLO-CD model achieved an F1-score of 67.8%, mAP50 of 71.5%, and mAP50-95 of 46.4%, with floating point operations reduced by 47.6% compared to the original YOLO11. The SegFormer-HF model achieved an mIoU of 90.51%, mPA of 85.15%, and an F1-score of 91.50%, outperforming mainstream segmentation models. This study establishes an integrated workflow from detection and segmentation to geometric parameter calculation, providing an effective means for non-contact automated crack detection in ancient bridges. However, the control of false detections under extremely complex textures still requires further improvement.

In the domain of model lightweighting and UAV deployment applications, Yang Guojun et al.[7] addressed the challenges of large parameter sizes in bridge crack detection models and unstable image quality in UAV-captured images by proposing a lightweight identification and measurement method based on improved YOLOv7 and SeaFormer. The authors introduced an EVC module into YOLOv7-tiny to enhance long-range dependency modeling and local feature extraction capabilities, combined with the SeaFormer segmentation model for efficient semantic segmentation. Additionally, they proposed a crack length and width calculation method based on an inscribed circle approach. Experimental results demonstrated that the improved YOLOv7-tiny-EVC model achieved an AP value of 91% with only 4.6 million parameters. The SeaFormer model attained IoU and PA metrics of 90.21% and 94.16% respectively, significantly outperforming comparative models. In practical UAV inspections, the relative errors for crack length and width measurements were controlled within 10% and 18% respectively. This study achieved an effective balance between accuracy and lightweight design, providing a feasible solution for UAV-based deployment. However, the precision of width measurements under conditions of blurred crack boundaries still requires further improvement.

In addressing the engineering practicality of automated inspection systems, Tang Yong et al.[8] developed an automatic bridge crack identification system based on UAV images and Support Vector Machine (SVM). The method first employs a LOG filter for image preprocessing and edge enhancement, followed by an SVM one-vs-rest multiclass classifier to categorize cracks into transverse, longitudinal, and other types. Parameters such as crack length, maximum width, and area proportion are then extracted through pixel calculation and morphological erosion operations, ultimately enabling automated assessment of bridge health conditions. System testing results demonstrated an overall classification accuracy of 96% on 100 crack images, with crack parameters automatically computed and exported as structured reports, highlighting its strong practical applicability. The study established a complete detection workflow from image input to evaluation output, providing a valuable reference for the application of traditional machine learning methods in automated crack identification. However, the system's robustness in scenarios with shadows and stain coverage requires further improvement.

To enhance the anti-interference capability of algorithms in complex scenarios, Feng Haolong et al.[9]addressed the issues of complex bridge surface image conditions and numerous interfering objects by proposing a two-stage identification algorithm termed "classification-screening and crack-identification." The authors improved the MobileNetV3 network architecture to construct an image screening model for automatically filtering out interference images such as backgrounds, ancillary structures, and severe contamination. Subsequently, they employed transfer learning to train a box-grid fusion crack identification model, achieving effective distinction between cracks and interfering objects. Experimental results

ISSN: 2710-0588

demonstrated that the image classification model achieved a maximum accuracy of 97.9%, while the crack identification model reached an AP of 65.6% for grid prediction, significantly outperforming single-stage identification models. This study notably enhances the algorithm's anti-interference capability in complex scenarios through a two-stage processing mechanism, providing a reliable method for automated screening and identification in practical bridge inspection work, though there remains room for improvement in the localization accuracy of fine and elongated cracks.

In summary, the experimental results from various scholars align well, demonstrating that optimizing detection model architectures, integrating traditional image processing methods, or adopting two-stage identification strategies can effectively enhance the accuracy, lightweight design, and anti-interference capability of bridge crack identification. These approaches exhibit promising generalization performance and practical utility in real-world engineering applications.

4. Research on Road Crack Identification Methods

Scholars worldwide have continuously worked on optimizing and improving key performance metrics in pavement crack identification, including detection accuracy, model efficiency, and adaptability to complex scenarios, by constructing diverse datasets and developing advanced deep learning models .

Zhou Shuangxi et al.[10]addressed the limitations of traditional road crack detection methods-such as poor real-time performance and low accuracy-by proposing an improved YOLOv5-based road crack detection algorithm (YOLOv5s-attention) incorporating an attention mechanism. The method integrates a channel attention module (SENet) into the backbone feature extraction network of YOLOv5s and embeds the CBAM module, which combines channel and spatial attention, into the C3 module to enhance the model's crack feature extraction capability and anti-interference performance. Experimental results demonstrated that the improved YOLOv5s-attention model achieved a detection precision of 56.4% on a self-constructed road crack dataset, representing a 1.0% improvement over the original YOLOv5s model. The comprehensive evaluation metric F1-score increased by 0.9%, and the mean average precision (mAP) improved by 1.8%. The model exhibited enhanced detection capability for fine targets such as 30° angled cracks. This study validates the effectiveness of attention mechanisms in improving the crack detection performance of YOLO-series models, providing a practical approach for automated and rapid road crack detection. However, further optimization is needed to reduce false positives and missed detections in extremely complex backgrounds.

In summary, the findings of Zhou Shuangxi et al. confirm that integrating the SENet channel attention mechanism and the CBAM spatial-channel hybrid attention mechanism into the YOLOv5 backbone effectively enhances the model's capability for extracting pavement crack features and improves its noise immunity, leading to notable improvements across all detection metrics. This study validates the effectiveness of attention mechanisms in boosting the performance of lightweight detection models, offering a practical approach for automated and rapid pavement crack identification.

5. Conclusion

Through a systematic review and analysis of relevant domestic and international literature, it is evident that while significant progress has been made in intelligent road distress detection technologies, distinct limitations persist in research methodologies and practical applications, warranting further in-depth exploration. In terms of adaptability to complex environments, the robustness of existing algorithms under challenging scenarios such as extreme lighting

ISSN: 2710-0588

conditions, severe surface contamination, and dense occlusions still requires enhancement. Regarding model performance optimization, the balance between lightweight design and detection accuracy needs further refinement. Future research should focus on improving the stability and practicality of algorithms in real-world environments, thereby advancing distress detection technologies toward greater precision, robustness, and practical applicability.

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] Pan Y, Zhou S X, Huang X S, et al. A deep learning-based concrete crack identification method incorporating textural features[J]. Concrete, 2024, (02): 45-51, 57.

[2] CHEN Ruiqi. Research on Concrete Crack Identification Method Based on Lightweight Convolutional Neural Network[D]. Sichuan Normal University, 2025. DOI:10.27347/d.cnki.gssdu.2025.001868.

[3] XU Jiaqi. Research on Concrete Crack Identification Method Based on Edge Extraction and VGG16 Deep Convolutional Neural Network[J]. Construction Technology(Chinese and English), 2023, 52(09): 11-15.

[4] Wu X H, Zhou X L, Wang Y Q. Research on concrete crack identification method based on ResNet+ network model[J]. Highway, 2025, 70(10): 381-388.

[5] XU Liang, HE Wei, YEERDA·Yeerdingdala, et al. Research on Methods and Mechanisms of Concrete Crack Detection Using UAV Intelligent Inspection[J]. Water Resources and Hydropower Engineering(Chinese and English), 2024, 55(S1): 249-256. DOI:10.13928/j.cnki.wrahe.2024.S1.039.

[6] ZHU Qiankun, XIE Chenhui, ZHANG Qiong, et al. Ancient Bridge Crack Identification Method Based on Computer Vision and Deep Learning[J/OL]. Journal of Southwest Jiaotong University, 1-12[2025-11-15]. Link.

[7] YANG G J, QI Y H, DU Y F, et al. Bridge crack identification and measurement using improved YOLOv7 and SeaFormer[J]. Journal of Railway Science and Engineering, 2025, 22(01): 429-442. DOI:10.19713/j.cnki.43-1423/u.T20240458.

[8] TANG Yong, YAO Xuechun, WANG Chen, et al. Automatic Bridge Crack Recognition System Based on UAV Image Technology and Support Vector Machine (SVM)[J]. Highway Engineering, 2024, 49(06): 49-56. DOI:10.19782/j.cnki.1674-0610.2024.06.008.

[9] FENG H L, LIU Y F, LIU X G, et al. Research on two-stage identification algorithm for bridge cracks based on convolutional neural network[J/OL]. Engineering Mechanics, 1-13[2025-11-15]. https://link.cnki.net/urlid/11.2595.O3.20250113.1014.002.

[10] ZHOU Shuangxi, YANG Dan, PAN Yuan, et al. Pavement Crack Detection and Recognition Using YOLOv5 with Attention Mechanism[J]. Journal of East China Jiaotong University, 2024, 41(02): 56-63. DOI:10.16749/j.cnki.jecjtu.20240307.002.