Tiako P.F. (ed) *Intelligent and Sustainable Solutions*. Chronicle of Computing. OkIP. CAIF25#9

https://doi.org/ 10.55432/978-1-6692-0011-6_4

Real-Time Crack Detection and Segmentation Using HMI Integrated YOLO v11-Seg Model for Industrial Applications

Shahmir Qureshi[a], Asfia Fatima[a], Jawaid Daudpoto[a,*], Aamir Shaikh[a], Abdul Qayoom[a]

[a]*Department of Mechatronic Engineering, Mehran University of Engineering & Technology, Jamshoro-76062, Sindh, Pakistan*

ABSTRACT

Cracks in metallic surfaces fail industrial systems, compromising systems' safety and productivity. It is therefore imperative to detect such cracks in their initial stage, so preventive measures can be taken to avoid downtime and associated risks. In this paper, we report the novel adaptation and application of the YOLO v11-seg model for precise crack segmentation in metallic surfaces. The specialized configuration optimizes both detection accuracy and inference speed, making it suitable for industrial applications. The deep learning model is integrated into a Human-Machine Interface (HMI) using the Open Neural Network Exchange (ONNX) format, to enable seamless real-time visualization and interaction. This enhances usability for industrial operators, bridging the gap between advanced AI models and practical deployment. We report a fully functional and versatile inspection tool by combining the deep learning model with hardware and real-time video processing via Python. A custom dataset comprising 1,111 training and 246 test images was curated, annotated with segmentation masks, and augmented using the Albumentations library to improve generalization. The model showed detection and segmentation precisions of 96% and 94%, respectively.

Keywords: Crack detection, YOLO v11, Crack segmentation, Mask Selection, Human Machine Interface.

I. INTRODUCTION

Due to the increase in the development of industries, it is necessary to be attentive to the durability and storage of metals used in building construction, pipelines, mechanical tools, aerospace, automobile industries, and many other sectors. As the demand for metal materials such as iron, aluminum, and copper increases globally, set to rise by 2 to 6 times by 2100 due to infrastructure development (Watari et al., 2020), the risks associated with fatigue, stress concentration, and the impact of the environment on structures with microcracks are increasing. Cracks can make machines and structures unsafe, less efficient, and expensive to fix, or even cause serious damage.

Traditional crack detection methods include ultrasonic testing, radiographic testing, and Eddy Current Testing. However, these methods have some drawbacks. They are frequently time-consuming, require skilled use, and may miss minor or deep cracks (Shen et al., 2024). This demonstrates the importance of improved cracking detection methods for various sectors. This is the outcome of substantial advances in computer vision and machine learning algorithms for material damage diagnosis, which aim to reduce hazards and extend the overall durability of metal structures while incurring lower maintenance costs. It remains a significant problem to create models that can handle many materials, detect minute flaws, perform rapidly, and be simple to use.

This project aims to solve these problems by a deep learning model for finding and outlining cracks, with a simple interface for users. The goal is to make crack identification more accurate, the process more efficient, and adaptable to different types and materials. This method is crucial for monitoring the health of structures, offering both improvement in detection and an easy-to-use interface for real inspections.

The paper is structured as follows: Section II of this paper presents other related literature in the field, while Section III presents the methodology, training process, experiment, and result. Lastly, Section IV provides the conclusion, the result and suggestion for further research. Finally, the acknowledgment section follows.

II. RELATED WORK

Crack detection has been addressed through various approaches, which can be categorized based on their underlying methodologies, including semantic segmentation, threshold segmentation, edge detection, and others.

Semantic segmentation is widely used for crack detection due to its ability to perform pixel-level classification. Several methods have been proposed to improve performance in terms of precision, recall, speed, and robustness to different crack patterns. For instance, KTCAM-Net combines classification and segmentation, achieving an F1-score of 88.6%, precision of 88.7%, recall of 88.2%, and 28 FPS (Al-Huda, Peng, Algburi, Al-antari, AL-Jarazi, & Zhai, 2023). ADDU-Net, an asymmetric dual-decoder U-Net, improves crack detection accuracy with an F1-score of 78.1%, precision of 84.6%, recall of 72.4%, and 35 FPS (Al-Huda, 2023). A deep learning approach using U-Net and YOLOv7 for bridge deck crack analysis reached an F1-score of 78.1% and a crack length detection accuracy of 92.38% (Tran, 2023). For tiny cracks in steel beams, the FCN-SFW fusion algorithm achieved an F1 score of 68.28% with a 1.58-second inference time (Wang, 2020). AFFNet, utilizing ResNet101 with attention modules, attained an MIoU of 84.49% and 52 ms inference time for concrete crack detection (Hang, 2023). PCSN, using SegNet, reported a recall of 50% and mAP of 83% (Chen, 2020), while DEHF-Net, with dual-path encoding, achieved 86.3% precision and 92.4% recall (Bai, 2024). The Student+Teacher Model with EfficientUNet demonstrated

* Corresponding author E-mail: jawaid.daudpoto@faculty.muet.edu.pk

Presented at the 2025 OkIP International Conference on Artificial Intelligence Frontiers (CAIF) in Oklahoma City, OK, USA, and Online, on April 2, 2025

strong semi-supervised performance with an F1-score of 83.21% (W. Wang & Su, 2021). Mask R-CNN achieved an F1-score of 78.47% for crack detection in concrete structures (Yamane & Chun, 2020). These studies show the progress and challenges in using semantic segmentation for crack detection. They also point out that there are still many opportunities for more research and improvement in this area.

When threshold segmentation methods are combined with image processing, they are often used for crack detection. For instance, the Threshold-based WSIS framework (T. He et al., 2024; H. Zhang et al., 2022) delivered solid results, with an F1 score of 66%, precision of 52%, recall of 90%, and accuracy of 98%. It effectively bridges the gap between unsupervised and supervised methods for detecting pavement cracks. Another study (Z. He & Xu, 2024) improved Otsu thresholding by integrating it with YOLOv7, reaching an impressive 98.4% accuracy for detecting crack repair traces, with an inference time of just 8.9 seconds, significantly better than the original method. Likewise, Local thresholding with a DCNN (Su et al., 2022) accomplished an F1 score of 91%, precision of 92%, and recall of 91% for bridge crack detection. Although these methods are efficient, they confront difficulties in adapting to varying lighting conditions and complex crack patterns.

Sobel and Canny edge detection methods when integrated with deep learning algorithms can identify cracks more efficiently. For instance, a study (Luo et al., 2023) combined a version of Canny with DeepLabV3+ for better feature integration. The performance outcome resulted in a 6.5% increase in MIoU and an F1 score of 63%. (Ranyal et al., 2024)utilized a vehicle-based approach with GPS-tagged images, and an attention-enhanced RetinaNet was trained, achieving an F1 score of 85.21%, precision of 85.96%, and recall of 84.48%. Additionally, (K. Liu & Chen, 2023)presented Crack-DA which is an unsupervised domain-adaptive framework. The study applied methods based on depth and edge information reached an F1 score of 74.7% on UAV data.

When it comes to crack recognition and segmentation, region-based approaches such as Mask R-CNN and YOLO place an emphasis on spatial features. For tunnel defect images, (Xu et al., 2021) used a Mask R-CNN with Path Augmentation Feature Pyramid Network (PAFPN) and achieved 92.03% precision and 96.26% recall. While YOLOv4 (J. Zhang et al., 2023) obtained 93.96% accuracy and 92% F1-score with minimal processing load. A Mask R-CNN model (Z. Liu et al., 2023) for non-destructive testing of asphalt cracks using GPR images had an average precision of 83.3%, F1-score of 82.4%, and mIoU of 70.1%.

Morphological operations, when used with deep learning, improve crack segmentation and classification. For example, a Mask R-CNN model that used morphological closing (Huang et al., 2022) reached an F1 score of 68.68% for tunnel lining segmentation. Then, there's the Parallel ResNet method (Fan et al., 2022), which did even better with an F1 score of 93.08% on the CrackTree200 dataset for pavement crack detection. Finally, when U-Net models combined with morphological operations (Dong et al., 2021), enhanced the segmentation of steel fatigue cracks, achieving an mF1 score of 42.79%. InceptionV3 (Nguyen et al., 2023), applied to ASR cracks, achieved an F1 score of 93.7% and an accuracy of 94.07%, providing an efficient AI solution for structural health monitoring.

Despite advancements in Crack detection, some challenges remain including limited data, imbalanced classes, and generalization across materials and types. High computational cost and real-time industrial integration are also some major limitations. To address these issues, this paper proposes an improved model using deep learning techniques to enhance accuracy, improve efficiency, and increase adaptability across different industrial applications. The developed model showed higher performance, speed, and better generalization compared to the models discussed in related works, even with a relatively small dataset. We successfully integrated this model into a WPF application using C# for backend operations. Reliance on Python libraries is eliminated, showcasing the feasibility of deploying models trained in one library across different platforms, simplifying the process, and providing users with an intuitive interface for inspection.

III. METHODOLOGY

A. *Dataset Preparation*

The method of detecting cracks using convolutional neural networks (CNNs) needs a variety of datasets to learn important features. To achieve this, we collected small open-source datasets focused on metal cracks of different sizes and shapes, such as those in metal sheets, welds, pipes, and machine parts. Figure 1 shows examples of cracks in metal parts, including welded joints, pipes, gear teeth, and bicycle frames. These cracks are difficult to find because they have different shapes and appear in random locations. Finding them is important to keep structures and machines safe and working properly.

To prepare the dataset, we removed low-quality or irrelevant images and made sure the labels were accurate. While some datasets already included basic augmentations, we added more enhancements such as rotation, flipping, translation, noise, and adjustments to brightness and contrast. These were based on existing methods (Golding et al., 2022; Z. Wang et al., 2020) and implemented using the Albumentations library (ÖNLER & Eray, 2018). Figure 2 shows how these data augmentation techniques and annotations improve crack detection. It highlights different crack patterns, orientations, and surfaces, marked with bounding boxes and outlines. These changes make the model more reliable by helping it handle different lighting conditions, angles, and backgrounds.

Fig. 1. Dataset Sample

The images were manually labeled with bounding boxes and segmentation masks using CVAT (Computer Vision Annotation Tool), which is a tool for labeling images (Guillermo et al., 2020). After augmentation, the dataset contained 1357 images, 1111 for training, and 246 for testing. YOLOv11 was used to train and evaluate the model.

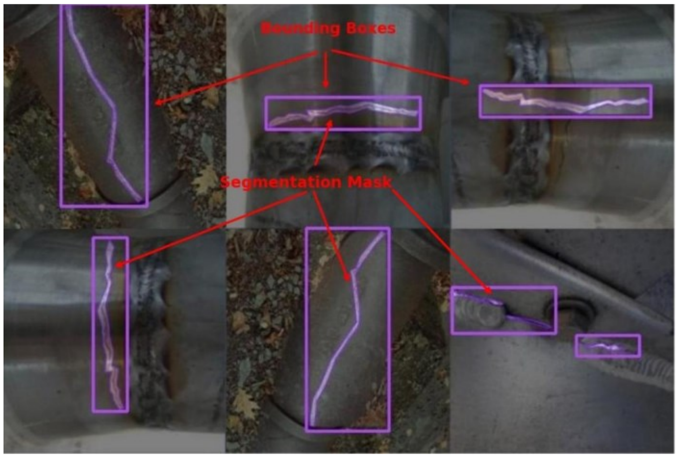

Fig. 2. Augmented and annotated data samples showing metallic cracks

B. *Model Architecture*

YOLOv11 is an improved version of YOLO; it has a shorter response time and increased accuracy. Building on YOLOv8, it introduces features that improve performance in tasks such as posture estimation and instance segmentation (Khanam & Hussain, 2024). In the model's backbone, the researchers employ smaller convolutions to enhance speed. The image processing consists of blocks, which increases the focus on priorities, improving the identification of both tiny and hidden objects. It further enhances image recognition, in which knowledge from one phase of the visuals is combined with knowledge from other phases. Extra layers are added to make more precise estimations. In summary, YOLOv11 is faster, more efficient, and better suited for real-time object recognition across diverse computer vision applications.

C. *Experimental Environment*

The experiment environment of this article is: Google Colab with GPU acceleration. NVIDIA Tesla T4 GPU (16 GB VRAM), an Intel Xeon processor, 13 GB RAM, and Ubuntu 20.04 with Python 3.10.12.

D. *Training Process*

This research employed a supervised learning approach with labeled training data, annotated in YOLO format via CVAT. The data included class IDs, coordinates, width, height, and key points for masks. A YAML file organized the image and label details. Training and evaluation were conducted Google Colab with GPU support, using the YOLOv11-seg model for 200 epochs, a batch size of 16, and an input size of 640x640 pixels. The learning rate started at 0.002 and was optimized, with Automatic Mixed Precision (AMP) enabled for faster processing. The model's performance plateaued after 160 epochs, leading to early termination. After training, the model was converted to ONNX format, enabling smooth deployment across different platforms. Figure 3 shows the metal crack detection system in the WPF application. On the left, it displays a welded joint image with a crack highlighted in red. On the right, it shows a video frame with a detected crack marked as "Crack 0.82" inside a blue box. Users can select folders, detect cracks, navigate through images and videos, and save results. They can also browse pre-recorded or live video, select a camera, and adjust settings like the confidence threshold and frame interval. Auto-processing can be enabled to automatically process all images from the selected folder, saving results, including bounding boxes and segmentation masks, to the desktop folder. The system uses the YOLOv11-seg model for accurate crack detection. The model's architecture, Input format, and Output analysis, all were verified by using Netron. The values obtained are (1, 37, 8400) for bounding box predictions and (1, 32, 160, 160) for segmentation masks. The first output includes bounding box coordinates, confidence scores, and class probabilities, while the second represents 32 predicted masks. Mask at index 11 was selected as it gave better results in comparison to others. Subsequently, the ONNX model was integrated into a WPF application with an easy-to-use interface for image processing. To improve detection, Non-Maximum Suppression (NMS) was used to remove extra bounding boxes. This setup made the application standalone, so external tools like Python environment were not required. Video processing was also added to the app, using Python scripts to handle both live and pre-recorded videos with adjustable frame rates to save on computing power. Both methods have their advantages. The integrated WPF approach provides a smooth, low-resource experience, making it easy to deploy a standalone application. On the other hand, the Python-based method offers real-time video processing, greater flexibility, and quicker model updates.

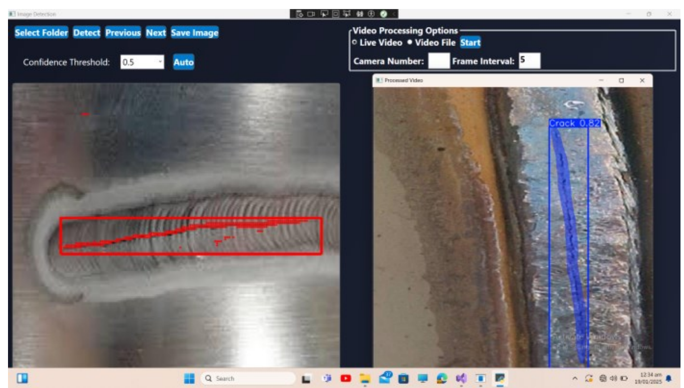

Fig 3. Metal Crack Detection in Image and Video via WPF App

E. *Experiments and Results*

This experiment focused on crack segmentation, where the YOLOv11 model identified cracks pixel-wise. Evaluated on a test dataset, it showed strong precision and recall for both bounding box detection (precision: 96.58%, recall: 93.43%) and segmentation (precision: 94.69%, recall: 91.61%). The model achieved excellent mAP50 scores of 96.88% for detection and 93.53% for segmentation at an IoU threshold of 50%, demonstrating its ability to detect cracks of various sizes and types. Figure 4 highlights the predicted results on the test dataset. YOLOv11 effectively detects and segments both thin and thick cracks in materials like metal sheets, pipes, and weld joints, providing bounding boxes with confidence scores. It identifies fine cracks (e.g., img5.png, 0.89 confidence) and thicker cracks (img6.png, 0.93 confidence). However, occasional misses and low-confidence detections suggest limitations due to a small training set and less distinct features. With an inference speed of 12.71 milliseconds per image, YOLOv11 is well-suited for real-time crack detection applications.

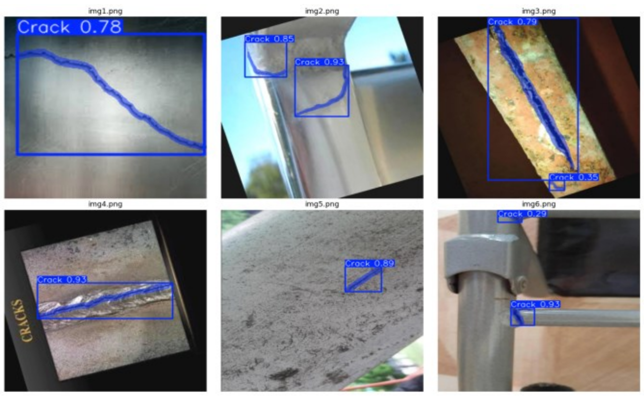

Fig 4. Predicted results on test dataset

F. *Comparison with other Models*

The crack detection model of this study was compared with several popular methods, including models that use semantic segmentation, edge detection, region-based methods, and more. Figure 5 shows the performance (Precision, Recall, and F1 Score) of different crack detection models, including KTCAM-Net, ADDU-Net, YOLOv7, U-Net, CGTr-Net, Efficient U-Net, Mask R-CNN, and the new YOLOv11. YOLOv11 gives the best results with the highest Precision (94.69%) and Recall (91.61%) and has a fast inference time of 12.71 ms. It processes images in 0.027 seconds on average, achieving 37.11 FPS. YOLOv11 works well on different crack shapes, even small and unbalanced datasets, making it a good choice for real-world use. The above results suggest that YOLOv11 is one of the best models for real-time crack detection in the industry. Our model, built on the new YOLOv11, shows clear improvements in precision, recall, and speed, solving many problems found in earlier models.

TABLE 1. PERFORMANCE COMPARISON

Model	Prec (%)	Rec (%)	F1 (%)	Inf Time (ms)
KTCAM-Net (Al-Huda, Peng, et al., 2023)	88.7	88.2	88.6	42
ADDU-Net (Al-Huda, et al., 2023)	84.6	72.4	78.1	28
U-Net (Tran et al., 2023)	83.5	72.9	77.8	-
CGTr-Net (Wang, Leng, & Zhang, 2024)	88.8	88.3	88.7	-
Efficient U-Net (W. Wang & Su, 2021)	83.11	86.06	83.2	-
Mask R-CNN (Z. Liu et al., 2023)	83.3	-	82.4	4.2FPS
Our Work	94.69	91.61	93.1	12.71

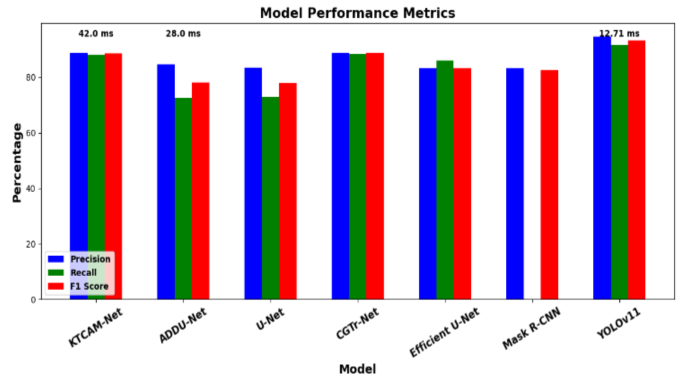

Fig 5. Performance of Different Crack Segmentation Models

IV. CONCLUSION

In this study, real-time crack detection and segmentation on metallic surfaces is introduced using the YOLOv11-seg model. The model demonstrated high precision (96.58%) and recall (93.43%) in detecting cracks, along with strong segmentation performance (precision: 94.69%, recall: 91.61%). This makes it ideal for use in industrial inspection, in such areas where quick and efficient inspection is needed. Through the implementation of the model in the WPF application using the ONNX format, we developed an effective and adaptive approach for industrial operators to leverage AI in the industry.

In comparison with other models like KTCAM-Net, ADDU-Net, and Mask R-CNN, the developed model outperforms others in precision and speed, requiring only 12.71 ms on average to generate predictions for images. This makes YOLOv11 suitable for addressing a variety of crack types where the shapes can be complex and the sample datasets may be small. Compared to traditional systems that require highly skilled workers and significant time investment, the proposed system offers an efficient solution for monitoring the health of metallic structures.

The results reaffirm the necessity of incorporating deep learning with other HMI tools to respond to critical difficulties in industrial crack identification. The real-time image and video processing capabilities expand the deployment of the technology into aerospace, construction, and manufacturing industries, reducing risks and minimizing time wasted on equipment and structures.

However, our experiment also showed some drawbacks: the precision of crack detection was insufficient for very small cracks, and low illumination conditions affected the performance of the algorithm. These limitations point out important areas for future improvement. Expanding the dataset to include more crack types and structural conditions, along with refining the model, could significantly enhance its accuracy and robustness.

ACKNOWLEDGMENT

We would like to thank Mehran University of Engineering and Technology (MUET) and the Mechatronics Engineering Department for providing the lab and computer facilities needed for this research. We also thank Ignite for funding this project under grant number NGIRI-2024-27094.

REFERENCES

Al-Huda, Z., Peng, B., Algburi, R. N. A., Al-antari, M. A., AL-Jarazi, R., Al-maqtari, O., & Zhai, D. (2023). Asymmetric dual-decoder-U-Net for pavement crack semantic segmentation. *Automation in Construction*, *156*, 105138. https://doi.org/10.1016/j.autcon.2023.105138

Al-Huda, Z., Peng, B., Algburi, R. N. A., Al-antari, M. A., AL-Jarazi, R., & Zhai, D. (2023). A hybrid deep learning pavement crack semantic segmentation. *Engineering Applications of Artificial Intelligence*, *122*, 106142. https://doi.org/10.1016/j.engappai.2023.106142

Bai, S., Ma, M., Yang, L., & Liu, Y. (2024). Pixel-wise crack defect segmentation with dual-encoder fusion network. *Construction and Building Materials*, *426*, 136179. https://doi.org/10.1016/j.conbuildmat.2024.136179

Chen, T., Cai, Z., Zhao, X., Chen, C., Liang, X., Zou, T., & Wang, P. (2020). Pavement crack detection and recognition using the architecture of segNet. *Journal of Industrial Information Integration*, *18*, 100144. https://doi.org/10.1016/j.jii.2020.100144

Dong, C., Li, L., Yan, J., Zhang, Z., Pan, H., & Catbas, F. N. (2021). Pixel-level fatigue crack segmentation in large-scale images of steel structures using an encoder–decoder network. *Sensors*, *21*(12), 4135. https://doi.org/10.3390/s21124135

Fan, Z., Lin, H., Li, C., Su, J., Bruno, S., & Loprencipe, G. (2022). Use of Parallel ResNet for High-Performance Pavement Crack Detection and Measurement. *Sustainability (Switzerland)*, *14*(3), 1825. https://doi.org/10.3390/su14031825

Golding, V. P., Gharineiat, Z., Munawar, H. S., & Ullah, F. (2022). Crack Detection in Concrete Structures Using Deep Learning. *Sustainability (Switzerland)*, *14*(13), 8117. https://doi.org/10.3390/su14138117

Guillermo, M., Billones, R. K., Bandala, A., Vicerra, R. R., Sybingco, E., Dadios, E. P., & Fillone, A. (2020). Implementation of Automated Annotation through Mask RCNN Object Detection model in CVAT using AWS EC2 Instance. *IEEE Region 10 Annual International Conference, Proceedings/TENCON*, *2020-November*, 708–713. https://doi.org/10.1109/TENCON50793.2020.9293906

Hang, J., Wu, Y., Li, Y., Lai, T., Zhang, J., & Li, Y. (2023). A deep learning semantic segmentation network with attention mechanism for concrete crack detection. *Structural Health Monitoring*, *22*(5), 3006–3026. https://doi.org/10.1177/14759217221126170

He, T., Li, H., Qian, Z., Niu, C., & Huang, R. (2024). Research on Weakly Supervised Pavement Crack Segmentation Based on Defect Location by Generative Adversarial Network and Target Re-optimization. *Construction and Building Materials*, *411*, 134668. https://doi.org/10.1016/j.conbuildmat.2023.134668

He, Z., & Xu, W. (2024). Deep learning and image preprocessing-based crack repair trace and secondary crack classification detection method for concrete bridges. *Structure and Infrastructure Engineering*, 1–17. https://doi.org/10.1080/15732479.2024.2330702

Huang, H., Zhao, S., Zhang, D., & Chen, J. (2022). Deep learning-based instance segmentation of cracks from shield tunnel lining images. *Structure and Infrastructure Engineering*, *18*(2), 183–196. https://doi.org/10.1080/15732479.2020.1838559

Khanam, R., & Hussain, M. (2024). *YOLOv11: An Overview of the Key Architectural Enhancements*. *abs/2410.17725*. http://arxiv.org/abs/2410.17725

Liu, K., & Chen, B. M. (2023). Industrial UAV-Based Unsupervised Domain Adaptive Crack Recognitions: From Database Towards Real-Site Infrastructural Inspections. *IEEE Transactions on Industrial Electronics*, *70*(9), 9410–9420. https://doi.org/10.1109/TIE.2022.3204953

Liu, Z., Yeoh, J. K. W., Gu, X., Dong, Q., Chen, Y., Wu, W., Wang, L., & Wang, D. (2023). Automatic pixel-level detection of vertical cracks in asphalt pavement based on GPR investigation and improved mask R-CNN. *Automation in Construction*, *146*, 104689. https://doi.org/10.1016/j.autcon.2022.104689

Luo, J., Lin, H., Wei, X., & Wang, Y. (2023). Adaptive Canny and Semantic Segmentation Networks Based on Feature Fusion for Road Crack Detection. *IEEE Access*, *11*, 51740–51753. https://doi.org/10.1109/ACCESS.2023.3279888

Nguyen, A., Gharehbaghi, V., Le, N. T., Sterling, L., Chaudhry, U. I., & Crawford, S. (2023). ASR crack identification in bridges using deep learning and texture analysis. *Structures*, *50*, 494–507. https://doi.org/10.1016/j.istruc.2023.02.042

ÖNLER, & Eray. (2018). Image Augmentation In Agriculture Using The Albumentations Library. *New Trends in Agriculture, Forestry and Aquaculture Sciences*, *abs/1809.06839*, 89. https://www.researchgate.net/publication/366685124

Ranyal, E., Sadhu, A., & Jain, K. (2024). Enhancing pavement health assessment: An attention-based approach for accurate crack detection, measurement, and mapping. *Expert Systems with Applications*, *247*, 123314. https://doi.org/10.1016/j.eswa.2024.123314

Shen, X., Lu, X., Guo, J., Liu, Y., Qi, J., & Lv, Z. (2024). Nondestructive Testing of Metal Cracks: Contemporary Methods and Emerging Challenges. *Crystals*, *14*(1), 54. https://doi.org/10.3390/cryst14010054

Su, H., Wang, X., Han, T., Wang, Z., Zhao, Z., & Zhang, P. (2022). Research on a U-Net Bridge Crack Identification and Feature-Calculation Methods Based on a CBAM Attention Mechanism. *Buildings*, *12*(10), 1561. https://doi.org/10.3390/buildings12101561

Tran, T. S., Nguyen, S. D., Lee, H. J., & Tran, V. P. (2023). Advanced crack detection and segmentation on bridge decks using deep learning. *Construction and Building Materials*, *400*, 132839. https://doi.org/10.1016/j.conbuildmat.2023.132839

Wang, S., Pan, Y., Chen, M., Zhang, Y., & Wu, X. (2020). FCN-SFW: Steel Structure Crack Segmentation Using a Fully Convolutional Network and Structured Forests. *IEEE Access*, *8*, 214358–214373. https://doi.org/10.1109/ACCESS.2020.3040939

Wang, W., & Su, C. (2021). Semi-supervised semantic segmentation network for surface crack detection. *Automation in Construction*, *128*, 103786. https://doi.org/10.1016/j.autcon.2021.103786

Wang, Z., Yang, J., Jiang, H., & Fan, X. (2020). CNN training with twenty samples for crack detection via data augmentation. *Sensors (Switzerland)*, *20*(17), 1–17. https://doi.org/10.3390/s20174849

Watari, T., Nansai, K., Giurco, D., Nakajima, K., McLellan, B., & Helbig, C. (2020). Global Metal Use Targets in Line with Climate Goals. *Environmental Science and Technology*, *54*(19), 12476–12483. https://doi.org/10.1021/acs.est.0c02471

Xu, Y., Li, D., Xie, Q., Wu, Q., & Wang, J. (2021). Automatic defect detection and segmentation of tunnel surface using modified Mask R-CNN. *Measurement: Journal of the International Measurement Confederation*, *178*, 109316. https://doi.org/10.1016/j.measurement.2021.109316

Yamane, T., & Chun, P. (2020). Crack Detection from a Concrete Surface Image Based on Semantic Segmentation Using Deep Learning. *Journal of Advanced Concrete Technology*, *18*(9), 493–504. https://doi.org/10.3151/jact.18.493

Zhang, H., Qian, Z., Tan, Y., Xie, Y., & Li, M. (2022). Investigation of pavement crack detection based on deep learning method using weakly supervised instance segmentation framework. *Construction and Building Materials*, *358*, 129117. https://doi.org/10.1016/j.conbuildmat.2022.129117

Zhang, J., Qian, S., & Tan, C. (2023). Automated bridge crack detection method based on lightweight vision models. *Complex and Intelligent Systems*, *9*(2), 1639–1652. https://doi.org/10.1007/s40747-022-00876-6