Automated Crack Detection on Concrete Surfaces: An Evaluation of Deep Learning Approaches Using YOLOv8

Pshtiwan Shakor[1*], *Ilham* I. Mohammed[2], and *Shadyar* Radha[2]

[1]Department of Civil Engineering, Al-Qalam University, Kirkuk, 36001, Iraq
[2]Department of Civil Engineering, Faculty of Engineering, Tishk International University, Sulaymaniyah, Iraq

Abstract. Detecting structural cracks is vital for ensuring safety and preventing potential failures. However, manual inspection is time-consuming and subjective. As a result, researchers have turned to machine learning to automate the crack detection process. In this study, a deep learning-based approach is proposed to improve and boost the accuracy and efficiency of crack detection and health monitoring. The methodology involves creating a dataset of crack images and labelling them accordingly. Deep learning models, specifically YOLOv8, were trained on this dataset to effectively detect and pinpoint the location of cracks. Various preprocessing techniques such as denoising and color correction are applied to improve the quality of the images. Additionally, data augmentation techniques are used to diversify the dataset. Model performance was evaluated using Precision, Recall, and mean Average Precision (mAP). This research delves into investigating the advantages, challenges, and performance of machine learning algorithms (YOLOV8) for crack detection. Furthermore, it examines directional crack detection while comparing various instance segmentation models based on mAP scores. The study also discusses training results and presents graphs illustrating model performance and addresses dataset quality checks. Overall, this research contributes significantly towards evaluating object detection and instance segmentation methods in computer vision applications related to crack detection. The proposed deep learning approach shows promise in detecting cracks and analyzing them—an advancement that holds immense potential, for improving infrastructure integrity management systems.

1 Introduction

In order to ensure the safety and integrity of buildings, especially concrete ones, checking

---

[*] Corresponding author: pshtiwan.shakor@alqalam.edu.iq

and inspecting cracks is a pivotal step. Presently, the conventional way to do that is through manual crack detection which is laborious and requires a substantial amount of time, furthermore, this method is susceptible to error [1]. However, Advancements in computer programs have enabled the application of Artificial Intelligence (AI) to be used in order to detect cracks on concrete surfaces, and these new techniques with incorporating algorithms can detect cracks accurately [2]. In a 2022 study conducted by Ali and his colleagues, they utilized Machine learning in crack detection and achieved high accuracy. Moreover, other studies have also reported high detection accuracy [3, 4].

Cracks in concrete can occur due to factors like damage, temperature changes, moisture levels, concrete shrinkage, and natural ageing [5]. Different types of cracks can be found in slabs and beams, including plastic shrinkage cracks, expansion cracks, overloading cracks, shear cracks, flexural cracks, and torsion cracks [6,7]. Non-structural cracks can be caused by humidity, temperature fluctuations, alkaline-aggregate reactions, and steel corrosion [8]. According to the study by [9], the building assessment can be conducted with visual inspection, thermography and laser inspection. These techniques could be split up into three main parts: optical and infrared methods to assume monoscopic operation for two-dimensional images, and radar and laser techniques used for more detailed analysis in image processing. Ultrasonic techniques can also be employed for structural assessment and may be integrated with advanced inspection platforms to enhance field applicability and operational efficiency.

Cracks were inspected using colour as an indicator, where a dark colour or black line was considered to represent a crack [9]. It usually depends on two methods: a. This method analyses the threshold value of the photo pixel, so basically, this method only works with white and black colours and only detects them when it has those two available colours. The edge displayed in black and the rest showed in pure white.

The way of collecting the data could be crucial to have an enrich data resource for the civil infrastructures, these values will be highly potential for analyzing the data and critical in terms of project management that could protect structures from different type of hazardous such as natural hazardous and human-made hazardous. These hazards can be monitored and mitigated to reduce risks to human life by monitoring through automated cameras and sensors. As obvious, most common issues with the natural hazardous in structural engineered is a seismic activities where the tectonic plates meet. Automating detection is an important step toward the monitoring infrastructures in a high-quality standard with respect to analyzing and monitoring any defects of structural concrete.

The data acquisition is a most crucial part in this work, it should considered precisely without fault or miscalculation, therefore, many drone types could be used in collection of these such as unmanned aerial vehicles (UAV) and micro aerial vehicles (MAV), these are could be used as a sensor system to capture required data. According to the study of used ultra-sonic sensors incorporated with UAV to identify the obstacles in front of the aerial vehicle, they also used low cost 2D laser scanner which be able to work to the Time of flight principle and has an accuracy of 30 mm up to 50 mm in a distance range of 0.1 m up to 30 m. This camera has a low weight about 210 gram and size (62×62×87.5) mm.

Various methods are available for crack detection in concrete. Pixel intensity thresholding separates pixels based on their intensity values [10] proposed a bi-thresholding method for crack identification in concrete images. The minimal path finding approach classifies cracks by representing them as a graph and finding the shortest path. developed an approach for crack detection in roadway pavement photographs [11]. Percolation uses liquid permeation to identify and analyze clusters of crack pixels in concrete surface images. Introduced the percolation model for crack detection. Feature engineering involves extracting relevant features from raw data for machine learning. proposed a crack detection technique based on the percolation model [2]. Deep learning utilizes deep neural networks to learn complex patterns. proposed the use of a convolutional deep neural network (DNN) for crack detection in concrete [12].

The lack of effective and efficient methods for automatically detecting cracks in concrete is a problem. Current methods rely on manual inspection, which is time-consuming and error-prone [13, 14]. The aim is to be to able to develop a machine learning model using deep learning that can detect concrete cracks [15]. Dataset of annotated images on concrete surfaces will be used for training and evaluation. The main aim for this research is to find whether a model can be created that not only can detect the crack, but as well as its direction either being horizontal, vertical, or diagonal.

1.1 Research Gap and Technical Contribution

Although there have been many improvements made using deep learning techniques to improve the accuracy of automatic crack detection, there are still some significant issues or shortcomings in the current use of YOLO-, CNN- based frameworks for crack detection. Most previous studies have treated the issue of crack detection primarily as a binary localization problem, without including the directional awareness of the structure involved in crack detection. As an example, while crack orientation can directly relate to the stress distribution in reinforced concrete members (i.e., flexural, shear, and bond cracking), most previous studies have ignored the orientation of cracks. Furthermore, limited research has systematically compared the performance trade-offs, robustness, and deployment implications of bounding box detection versus instance segmentation, on an identical dataset. Also, few researchers have systematically evaluated the effects of environmental sensitivity, particularly illumination variability, even though illumination variability will directly impact the ability to monitor structures in the field. Therefore, this paper advances the conventional research in crack detection, by introducing directional crack classification into a YOLOv8 framework, structurally comparing object detection and instance segmentation approaches, and assessing the illumination robustness of each approach, to provide enhanced structural diagnostic relevance and field applicability to automated crack detection systems.

2 Methodology

Cracks in concrete have different shapes and pattern on the surface or deeper inside the concrete, some of them directly affect the structural properties of the materials, however, some of them have non-structural affect but might increase or enlarge in the shape which could structural element into danger. For example, some previous work type the cracks according to their shape. Therefore, Most observed cracks are classified according to their shape or way of the crack propagation such as a. horizontal, b. diagonal, c. vertical, as shown in Figure 1.

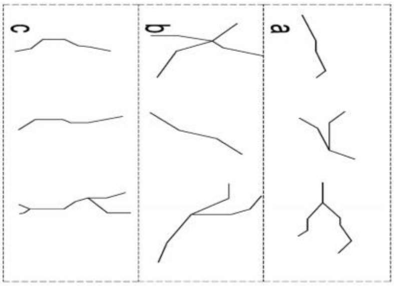

Fig. 1. Illustrated different types of cracks, a. longitudinal cracks, b. diagonal cracks, and c. vertical crack

You Only Look Once (YOLO) is an advanced object detection model, the first iteration of this algorithm was introduced in 2015. The latest iteration of this algorithm is called YOLOV8, which stands for (you only look once version eight). This algorithm is very powerful, it is developed as to not only detect objects in images, furthermore it can also be used to detect object on videos in real time, which makes it very effective for applications that needs instantaneous outcomes. YOLOv8 uses a CNN (Convolutional Neural Network) architecture. This structure is designed to extract features from the input image. One difference between the YOLOv8 with other iterations and other algorithms is that it conducts the detection process in a single pass, in contrast to multiple stages in other models. Predetermined shapes that aid to predict in different dimensions better are called Anchor boxes, and YOLOv8 uses this mechanism. Furthermore, the loss function in this algorithm factors localization loss for bounding box, classification loss. Finally, confidence loss which is used for objectless score. When the predictions are made, the algorithm uses Non-Maximum Suppression to remove and eliminate excessive and redundant bounding boxes, as shown in Figure 2. Three models have been created, two of them use bounding box for detection and the third one uses instance segmentation (see Table 1 for the differences between these two detection types).

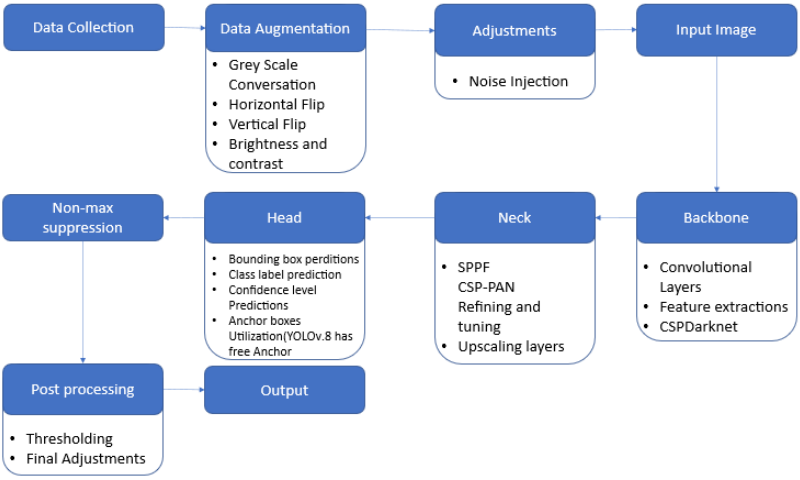

Fig.2. Research methodology flow diagram

The methodology for crack detection using YOLOv8 algorithm includes first to collect data, which can be done by taking photos of concrete surfaces, the next step which is which is one of the most significant one is Annotation, annotation is the process of labeling images with the location and identity of objects. This information is then used to train the model to detect objects in new images. In order to obtain a diverse set of images for our study, cracks on concrete surfaces were captured from several buildings and the photos were taken in different distances and angles. Moreover, some of the photos were taken with tele-photo lens in order to provide a detailed view of the cracks. The photos were taken at different times of day so as to have a more diverse dataset in different lighting conditions. There exist two types of annotation, one is to draw a bounding box around the object we want to detect and the other one is instance segmentation as shown in Figure 3 which includes drawing a mask around each object.

Fig.3. Annotated vertical cracks on the surface of concrete

In this research, the latter was used as we created all the models from the same dataset. The next step is preprocessing and augmentation, both are used to make model robust, we applied resizing to images, as the images were taken in different formats and sizes, we had to convert all of them into one size, the size we chose was 640×640 pixels. Then we had to generate more data as the models work better with more data, thus we applied augmentation layers, we applied vertical and horizontal flips, the next augmentation step that was applied was to greyscale 17% of generated images were grayscale and finally in last augmentation step ±28% (shown in Figure 4) brightness was applied to simulate dark and bright environments. The final step in methodology was training. The algorithm had to be trained on the dataset, this process is rather automatic, its duration is ultimately on the number of images, parameters, the architecture of the algorithm as well as other factors.

The dataset had 1041 images that were labeled as being related to cracks in concrete structures. Images were taken under different light sources and environmental conditions. Each image contained one of three types of crack orientations: horizontal, vertical or diagonal; each type of crack was represented almost equally in the images, resulting in approximately 347 images per orientation to minimize the effects of classification bias. The images were labeled using standard label procedures based on both bounding boxes and pixel-level masks and the accuracy of the labels was confirmed through manual verification by cross checking all of the labels. A stratified random split was used to divide the images into training sets, validation sets, and test sets such that the training set was made up of 819 (78.7%), the validation set was made up of 175 (16.8%), and the test set was made up of 47 (4.5%). In doing so, we ensured that the number of images with each of the three orientations were maintained across the sets and prevented data leakage due to the fact that there was no intersection of images across the subsets. We assessed model performance using multiple statistical measures, including Precision, Recall, F1-score, mAP@0.5 and mAP@0.5:0.95, as well as confusion matrices to assess whether or not models demonstrated sufficient classification balance and generalization ability to support the validity of our results while reducing the potential for overfitting. All reported mean Average Precision values correspond to mAP@0.5 (Intersection over Union threshold = 0.5), unless otherwise stated.

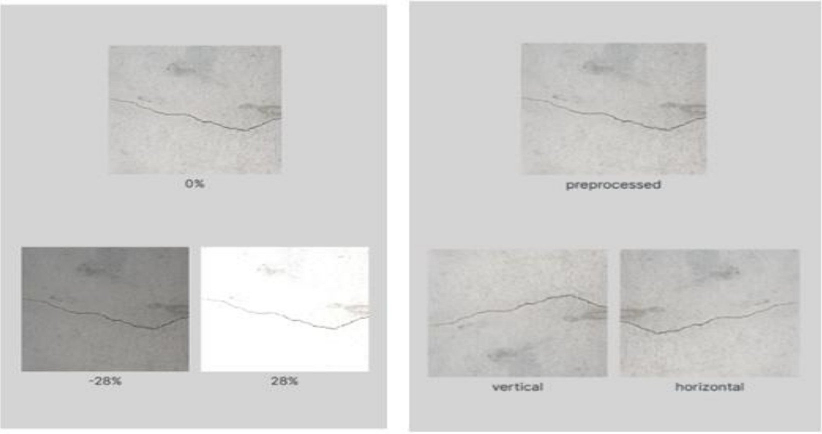

Fig.4. Showing flip and brightness in augmentation step

As stated earlier, two different versions of YOlOv8 were used, bounding box and instance segmentation, in Table 1 they are compared and Table 2 shows YOLOv8s specifications and in Table 3, hyper parameters are showed.

Table 1. Introduction Segmentation

	Instance Segmentation	Object Detection (Bounding Box)
Definition	Assigns a pixel-level label to each object instance.	Identifies and localizes objects with bounding boxes.
Output	Segmented masks for each object instance.	Bounding box coordinates and class labels.
Granularity	Fine-grained segmentation at the pixel level.	Coarse-grained localization at the object level
Level of Detail	Provides detailed information about object boundaries	Limited to rectangular bounding boxes
Use Cases	Precise object detection, instance-level analysis.	General object detection, localization tasks.
Complexity	More computationally expensive and resource intensive.	Less computationally expensive and faster.
Training Data	Requires pixel-level annotations for each object instance.	Requires bounding box annotations for objects.

Table 2. Summary of YOLOv8s-seg Model

	YOLOv8s-seg Summary
Layers	261

Parameters	11,791,644
Gradients	11,791,628
GFLOPs	42.7

Table 3. Hyperparameters for Object Detection

Hyperparameter	Default Value
--image -size	640
--batch-size	16
--epochs	300
--lr	0.01
--weight-decay	0.0005
--iou-thres	0.6
--conf-thres	0.25

Table 4. Hyperparameters for Instance Segmentation

Hyperparameter	Default Value
--model	yolov8s
--mask-thres	0.5

3 Results

The first test was to test the detection capability of the bounding box detection and the instance segmentation models with a high-resolution image. Figure 5 illustrates the comparative detection results of the bounding box and instance segmentation models. The bounding box model correctly located cracks and determined crack direction, while the instance segmentation model also classified crack direction, but with additional pixel-level crack boundary delineation.

Fig.5. Comparative detection results of bounding box (left) and instance segmentation (right) models for crack localization and directional classification.

In the second evaluation, the models were assessed for robustness when exposed to various levels of illumination (low and high brightness) as shown in Figure 6. The instance segmentation model experienced a large decrease in confidence values under varying light conditions, while the bounding box model saw a smaller decrease in confidence from 71% to 69%, showing better resistance to brightness variations than the instance segmentation model.

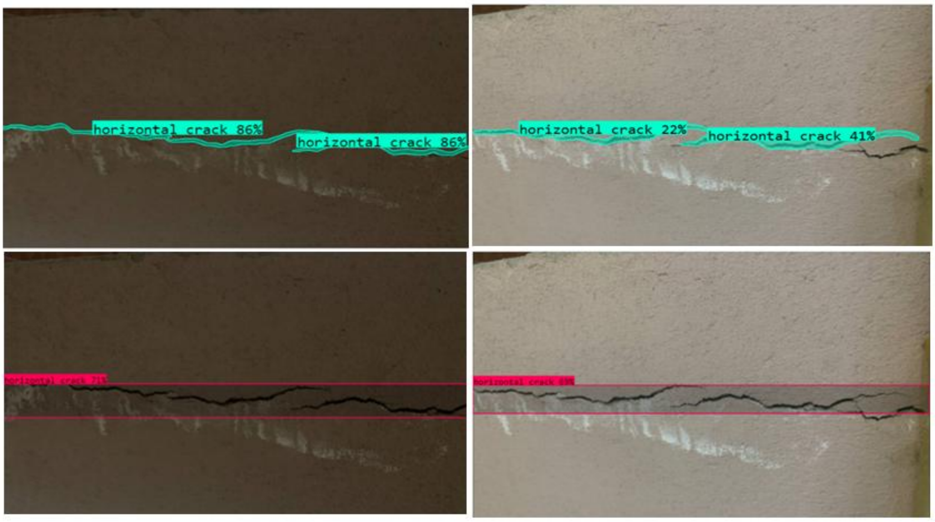

Fig.6. Showing the performance of the models in different light conditions

The ability of each model to identify and classify multiple crack orientations within one image was evaluated. Both models were able to find cracks oriented horizontally and diagonally at the same time in the same structural surface as shown in Figure 7.

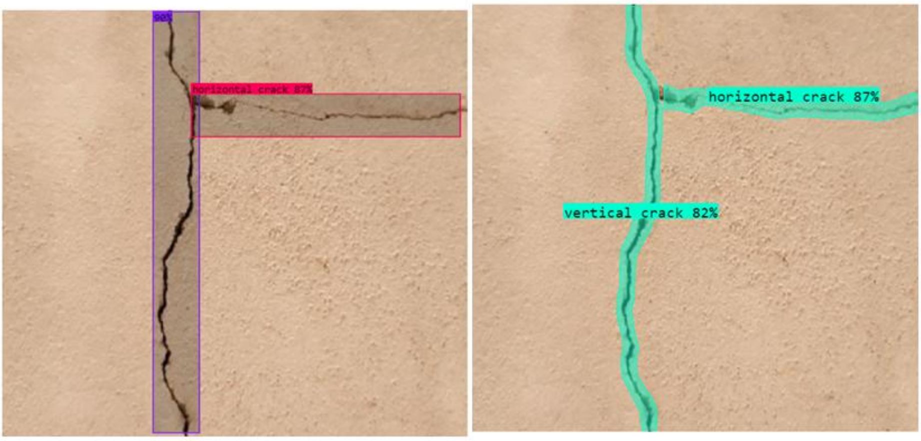

Fig.7. Performance of the models on an image that has two different directions

To examine further how dataset influence affects model performance, two instance segmentation models trained on different split datasets were compared to each other. The first model was trained on 819 training images and 175 validation images and tested on 47 test images. The second model was trained on 738 training images and 225 validation images and tested on 185 test images. The first configuration produced substantily higher detection accuracy than the second configuration, showcasing how much the volume of training data and the distribution of the training data will affect model accuracy.

Fig.8. Detection Results from the first instance segmentation model compared with the 2nd.

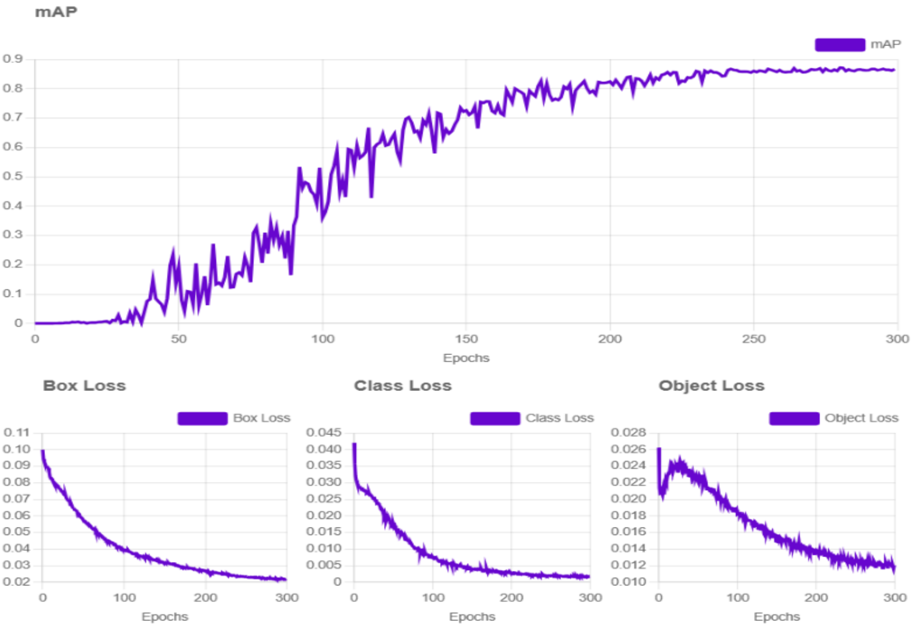

Fig.9. The graphs show the bounding box model was trained for 300 epochs, as the training progressed, box loss, class loss and objection loss decreased.

All reported map values correspond to mAP@0.5 (IoU threshold = 0.5). The object detection model achieved 87.1% while the best performing instance segmentation configuration achieved 88.5%. Analysis of the precision-recall curves demonstrated stable behavior for both sensitivity and specificity of the detector. Evaluation of confusion matrix also demonstrated that the classification results remained consistent between all crack categories (horizontal, vertical, diagonal), demonstrating satisfactory generalization. Figures 8 and 9 display the training curves for the bounding box model over 300 epochs. The loss components (box loss, classification loss, objectness loss) decreased over the 300 epochs with no significant oscillations or divergence.

4 Discussion

The comparative analysis demonstrated that both bounding box detection and instance segmentation were successful at localizing cracks and classifying direction. The bounding box model was able to detect a much smaller decrease in confidence (only a 2% drop from 71% to 69%) than the instance segmentation model when exposed to varying illuminations. This behavior can be attributed to the fact that instance segmentation models perform dense classification across every pixel, as such they are more sensitive to changes in intensity gradient and contrast. Bounding box detection models, while still capable of being sensitive to these factors, rely heavily on the aggregation of features in a region; therefore, exhibit comparatively lower sensitivity to localized changes in brightness. These comparisons demonstrate a significant tradeoff between the accuracy of detecting boundaries and the ability to operate effectively under various illumination conditions.

There is an apparent discrepancy between the aggregate mean Average Precision (mAP) performance and the confidence levels for each test image. It would appear that the discrepancy is caused by either sampling or stratification, as opposed to simply differing performance. Confidence is a measure of how confident the model is about detecting cracks on a given test image. Confidence will be heavily influenced by the way that features are distributed throughout the model, as well as variability from one test image to another. The lower number of test images used in the high performing configuration reduces the amount of variability represented by the data, which increases the likelihood that the confidence behavior will be sensitive to the specific type(s) of cracks present. Conversely, using a larger number of test images in a low performing configuration, although resulting in a lower level of overall accuracy, can result in confidence being stabilized. This illustrates why it is important to carefully consider dataset stratification, as well as the design of the validation process, when developing models to detect defects such as cracks, to ensure that the models generalize robustly, while minimizing the occurrence of overfitting. Beyond basic crack identification, direction of crack orientation is a very useful diagnostic tool for understanding how a structure has been subjected to stress. Diagonal (at an angle) cracks generally indicate shear (sliding) action; horizontal (side-to-side) cracks often indicate poor bonding or problems with the reinforcing steel; vertical cracks are usually indicative of flexure (bending). The inclusion of directional crack classification within the developed system enables engineers to make preliminary assessments of the general condition of structures under evaluation. In addition, this feature can be used to assist engineers in determining which structures have the highest priority when performing inspections/maintenance on multiple large-scale structures at one time.

The experiment using a random split of the dataset illustrates that performance of a model is heavily dependent upon the distribution of the training data and the composition of the training data. Models trained using a larger training subset demonstrated better performance, suggesting that there was better feature extraction and generalization ability. When compared to other studies utilizing deep learning to identify cracks in structures, this study has advanced the evaluation paradigm through incorporation of directional classification and analysis of illumination robustness. Through incorporation of orientation-aware detection, this study has enhanced the structural diagnostics of distinguishing among flexural, shear, and bond related cracking mechanisms in reinforced concrete members. Overall, the results provide evidence supporting the use of YOLOv8 based frameworks for performing automatic structural crack assessments in real world applications.

5 Conclusion

The study used image processing to detect cracks and their directions in concrete structures automatically. The study applied deep learning methods such as instance segmentation and object detection to locate and propagate cracks. The study achieved mAP of 87.1% for object detection and 88.5% for instance segmentation. The study showed the potential of image processing for structural health monitoring and computer vision applications in civil engineering. The study used YOLOv8 and image processing techniques for automatic crack detection in concrete structures. However, the results were affected by the lighting conditions and the size of the dataset. Therefore, some possible recommendations for future work are: To improve the robustness of the crack detection method, different lighting conditions and angles should be considered in the image processing steps, such as contrast enhancement, edge detection, and thresholding. To expand the scope and applicability of the crack detection method, more types and sizes of cracks should be included in the dataset, such as hairline cracks, longitudinal cracks, transverse cracks, diagonal cracks, etc. To evaluate the performance of the crack detection method in real-world scenarios, more field tests and experiments should be conducted on actual concrete structures with different environmental factors and structural conditions.

Funding Declaration

No funding was received to assist with the preparation of this manuscript.

References

1. Y. Han, R. Jin, H. Wood, and T. Yang, Investigation of Demographic Factors in Construction Employees Safety Perceptions. KSCE J. Civ. Eng. **23**, 2815–2828 (2019). doi: https://doi.org/10.1007/s12205-019-2044-4
2. L. Lei, D. Haijuan, and C. Bing, Experimental Investigation on Concrete Using Corn Stalk and Magnesium Phosphate Cement under Compaction Forming Technology. J. Mater. Civ. Eng. **32**, 4020370 (2020). doi: 10.1061/(ASCE)MT.1943-5533.0003487.
3. R. Ali, J. H. Chuah, M. S. A. Talip, N. Mokhtar, and M. A. Shoaib, Structural crack detection using deep convolutional neural networks. Autom. Constr. **133**, 103989 (2022). doi: https://doi.org/10.1016/j.autcon.2021.103989.
4. R. Amhaz, S. Chambon, J. Idier, and V. Baltazart, Automatic Crack Detection on Two-Dimensional Pavement Images: An Algorithm Based on Minimal Path Selection. IEEE Trans. Intell. Transp. Syst. **17**, 2718–2729 (2016). doi: 10.1109/TITS.2015.2477675.
5. S. E. Park, S.-H. Eem, and H. Jeon, Concrete crack detection and quantification using deep learning and structured light. Constr. Build. Mater. **252**, 119096 (2020). doi: https://doi.org/10.1016/j.conbuildmat.2020.119096
6. M. Antoun, C. A. Issa, G. Aouad, and N. Gerges, Sustainable masonry blocks:

Olive wood waste as substitute for fine aggregates. Case Stud. Constr. Mater. **15**, e00590 (2021). doi: https://doi.org/10.1016/j.cscm.2021.e00590

7. N. Kumar and M. Barbato, Effects of sugarcane bagasse fibers on the properties of compressed and stabilized earth blocks. Constr. Build. Mater. **315**, 125552 (2022). doi: https://doi.org/10.1016/j.conbuildmat.2021.125552
8. C. V. Dung and L. D. Anh, Autonomous concrete crack detection using deep fully convolutional neural network. Autom. Constr. **99**, 52–58 (2019). doi: https://doi.org/10.1016/j.autcon.2018.11.028
9. S. S. Man, S. Alabdulkarim, A. H. S. Chan, and T. Zhang, The acceptance of personal protective equipment among Hong Kong construction workers: An integration of technology acceptance model and theory of planned behavior with risk perception and safety climate. J. Safety Res. **79**, 329–340 (2021). doi: https://doi.org/10.1016/j.jsr.2021.09.014
10. Z. Fan, S. Member, Y. Wu, J. Lu, and W. Li, Automatic Pavement Crack Detection Based on Structured Prediction with the Convolutional Neural Network. 1–9. https://doi.org/10.48550/arXiv.1802.02208
11. B.K.Biskaborn, L. Nazarova, T. Kröger, L.A. Pestryakova, L. Syrykh, G. Pfalz, U. Herzschuh and B. Diekmann, Late Quaternary Climate Reconstruction and Lead-Lag Relationships of Biotic and Sediment-Geochemical Indicators at Lake Bolshoe Toko, Siberia. Front. Earth Sci. **9**, 733533 (2021). doi: 10.3389/feart.2021.737353 .
12. M. A. Khan, S.-H. Kee, A.-S. K. Pathan, and A.-A. Nahid, Image Processing Techniques for Concrete Crack Detection: A Scientometrics Literature Review, Remote Sensing, **15,** 9 (2023). doi: 10.3390/rs15092400.
13. D. Ai, G. Jiang, S.-K. Lam, P. He, and C. Li, Computer vision framework for crack detection of civil infrastructure—A review, Eng. Appl. Artif. Intell. **117**, 105478(2023).doi: https://doi.org/10.1016/j.engappai.2022.105478 .
14. T. Jiang, Y. Hong, J. Zheng, L. Wang, and H. Gu, Crack Detection of FRP-Reinforced Concrete Beam Using Embedded Piezoceramic Smart Aggregates, sensors, **19**, 3390 (2019). doi: 10.3390/s19091979.
15. C. He, B. McCabe, and G. Jia, Effect of leader-member exchange on construction worker safety behavior: Safety climate and psychological capital as the mediators. Saf. Sci. **142**, 105401, (2021). doi: https://doi.org/10.1016/j.ssci.2021.105401