https://doi.org/10.3311/PPci.37929

Periodica Polytechnica Civil Engineering, 69(1), pp. 347–362, 2025

Crack Identification and Quantification of Bridge Concrete Based on YOLOX and Image Processing Techniques

Bixiong Li[1*], Yunjun Liu[1], Guixing Kuang[1]

[1] Department of Architecture and Environment, Faculty of Civil Engineering, Sichuan University, Jiuyanqiao Wangjiang Road 29., 610065 Chengdu, China
[*] Corresponding author, e-mail: libix@126.com

Received: 20 July 2024, Accepted: 27 November 2024, Published online: 05 December 2024

Abstract

The investigation and analysis of bridge distress are critical for the assessment and maintenance of bridge safety, necessitating precise information regarding the condition of the bridge surface. In this study, a deep learning framework for automatically identifying bridge concrete cracks is proposed based on comparing the detection performance of YOLOX, SSD, and Faster R-CNN. The deep learning model YOLOX_s is initially trained and employed to identify bridge concrete cracks, and the detection results demonstrate that the bridge concrete crack identification accuracy rate of the YOLOX_s is 91.77% and much higher than that of SSD and Faster R-CNN, which are 88.09% and 86.57% separately. To perform bridge concrete crack quantification, several image processing techniques are applied. The process begins with the cropping of the identified cracks obtained by YOLOX_s followed by binarization using Otsu's method. Subsequently, the Zhang-Suen thinning algorithm is applied to extract the crack skeleton, while the Canny edge detection algorithm outlines the crack boundaries. Finally, a pixel accumulation-based method is implemented to calculate the crack dimensions. The findings indicate that the proposed method for measuring crack length and the maximum width achieves high accuracy levels of 96.6% and 95.86%, respectively.

Keywords

bridge, crack identification, YOLOX, crack quantification, image processing

1 Introduction

Reinforced concrete bridges are vulnerable to structural aging, damage, and even destruction due to several factors, such as overloading, impacts, environmental erosion, and material degradation [1]. To evaluate the service condition of bridges, the timely identification and accurate quantification of bridge concrete cracks are essentially required, because the emergence of cracks is the most direct manifestation of bridge distress and the quantity and dimensions of cracks are important metrics to evaluate bridge integrity and durability [2, 3]. On the other hand, the quantity and dimensions of cracks can provide valuable guidance for the prevention and control of major bridge distress by implementing associated countermeasures.

Bridge inspections are traditionally conducted by manual visual inspection. However, manual visual inspection is laborious, time-consuming, subjective, and even dangerous, particularly in the case of large-scale or complex bridges [4–6]. Moreover, manual inspections depend to a large extent on the experience and knowledge of inspectors, which may lead to varying judgement results. In pursuit of improving detection efficiency and objectivity, as well as achieving reliable detection results, the field of bridge crack detection is increasingly leveraging deep learning techniques to facilitate automated and cost-effective inspections. State-of-the-art object recognition networks such as Single Shot Multi-Box Detector (SSD) [7], Faster Region-based Convolutional Network (R-CNN) [8], and You Only Look Once (YOLO) [9] have been successfully deployed for concrete crack identification. For instance, Yan et al. [10] developed a novel SSD-based model for the automated detection of asphalt pavement cracks, while Huyan et al. [11] utilized Faster R-CNN for discerning complex crack patterns in road surfaces. Park et al. [12] employed the YOLO algorithm for real-time crack detection and quantification. Moreover, significant advancements have been achieved in the automatic crack inspection of other structural infrastructure, including

Cite this article as: Li, B., Liu, Y., Kuang, G. "Crack Identification and Quantification of Bridge Concrete Based on YOLOX and Image Processing Techniques", Periodica Polytechnica Civil Engineering, 69(1), pp. 347–362, 2025. https://doi.org/10.3311/PPci.37929

tunnels [13, 14], bridges [15–17], and dams [18–20]. These advancements also demonstrate the strong ability of deep learning methods in crack localization and determination.

In the deep learning domain, convolutional neural networks (CNNs) stand at the forefront, and serve as an important role [21–23]. Their prowess lies in the automatic extraction of salient features from image data, coupled with robust accuracy in image classification, objection detection, and image segmentation [24]. Models like Faster R-CNN, SSD, and YOLO are CNN-based and have shown remarkable practicability in object detection challenges [25]. Faster R-CNN is known for its accuracy, introducing a region proposal network (RPN) for generating high-quality region proposals, followed by precise classification and bounding box regression [8, 26]. However, it is too computationally intensive for embedded systems and too slow for real-time applications [7]. To solve this problem, SSD introduces several optimization measures in detection principles and processes, such as simultaneously predicting bounding box positions and object categories in a single forward pass and using feature maps of various scales for detection [27]. While significantly enhancing detection speed, it still ensures satisfactory accuracy. As for the YOLO series, its developers have put much effort into striking an ideal equilibrium between detection speed and accuracy [28]. The latest YOLO model can outperform Faster R-CNN in detection accuracy and outperform SSD in detection speed.

Given its proficiency in both processing speed and detection accuracy, the YOLO series has become a popular choice for researchers tackling various crack detection tasks. For instance, Du et al. [29] successfully implemented the YOLOv3 model to discern various pavement distresses, including different crack patterns, culminating in an overall detection accuracy of 73.64%. Zhang et al. [30] refined the YOLOv3 model to instantly identify bridge surface cracks, achieving enhanced accuracy and processing speeds. Similarly, Yao et al. [31] augmented the YOLOv4 model to facilitate real-time detection of concrete surface cracks. Hu et al. [32] explored the application of YOLOv5 for pavement crack detection, reporting that the trained model exceeded an accuracy threshold of 85%. These endeavors also underscore the ongoing scholarly pursuit to harness deep learning for advancing crack detection capabilities through algorithmic improvements and optimizations.

The swift advancements in computer technology have given rise to numerous high-performing models in the field of deep learning. Despite these advancements, there is often a lag in the exploration and validation of these models' performance in concrete crack detection. In 2021, Ge et al. [33] introduced You Only Look Once X (YOLOX), which incorporated several advancements over its predecessors, including YOLOv3 through YOLOv5. While models like those in the YOLO series have proven effective for such tasks due to their accuracy, speed and adaptability, the full potential of newer models remains an open area for research and could represent a significant opportunity for advancements in automated crack detection systems. However, there are few studies on damage recognition using the YOLOX, and the application effect of the YOLOX in crack detection remains to be further investigated [34]. Drawing on the advantages and current application of YOLOX, this study employs YOLOX for crack classification and localization. At the same time, in order to examine its better detection performance, it is compared with Faster R-CNN and SSD.

Indeed, in the context of bridge safety and maintenance, the identification of cracks is just the first step. The precise quantification of crack geometry is also essential for evaluating the structural integrity and longevity of bridges and can provide reference index for whether to take repair measures and what kind of repair measures. Accurate crack data extraction is a pivotal step in this quantification process. Various image-processing techniques (IPT) have been documented in the literature as effective methods for extracting detailed information about cracks from images. These include thresholding for distinguishing cracks from the background, edge detection algorithms for identifying crack outlines, thinning algorithms for extracting the crack skeleton, and Canny algorithms to minimize edge detection omissions [23, 35–37]. Kim et al. [37] achieved over 89% accuracy in crack width estimation using an image binarization method. Furthermore, Kim et al. [38] implemented a hybrid binarization approach, which precisely estimated crack width with minimal crack length data loss, detecting cracks as thin as 0.1 mm with a maximum length estimation error of 7.3%. Despite these advancements, the lack of a standardized approach for threshold setting and parameter optimization for feature extraction persists, rendering crack quantification laborious, inefficient, and costly. To more easily and quickly separate the crack from the background, while ensuring the integrity of the crack information, Otsu's thresholding method emerges as a potent solution. Otsu's way, an adaptive thresholding technique, is adept at binarizing images to differentiate the foreground from the background. It does so by calculating the threshold that maximizes inter-class

variance – the variance between the pixel value distributions of the foreground and background [39]. By exhaustively evaluating all grayscale values in a crack image to find the maximal inter-class variance, Otsu's method autonomously ascertains the best segmentation threshold, thereby effectively distinguishing cracks from the background [40, 41]. This results in precise crack information crucial for estimating crack length and width.

By integrating YOLOX's detection prowess with the convenience of image processing techniques, this paper proposes an automated framework for identifying and quantifying bridge concrete cracks. The process initiates with training YOLOX to detect bridge concrete cracks, followed by a comparative analysis with SSD and Faster R-CNN to attest to the superior performance of YOLOX. Subsequently, original crack images are detected by YOLOX, and cracks within the bounding boxes are cropped out to undergo binarization by Otsu's method. Then, the Zhang-Suen thinning algorithm is employed to extract the crack skeleton, and the Canny edge detection is utilized to delineate crack boundaries. The final phase involves devising a pixel accumulation-based computational approach for measuring crack length and width.

2 Methodology and materials

The conduct of this research and the overarching procedure are graphically summarized in the flowchart presented in Fig. 1. The process is divided into four distinct modules:

1. Image acquisition and dataset preparation: this initial module is devoted to collecting images and preparing the dataset, which forms the foundation for subsequent analyses.
2. Model training and evaluation: the second module is centered around the training, verification, and assessment of the deep learning models, encompassing YOLOX, SSD, and Faster R-CNN. This stage is crucial for refining the models to ensure high accuracy and reliability in crack identification.
3. Crack identification and image processing: in the third module, the research advances with using YOLOX to detect cracks on bridge surfaces. Once identified, these cracks within the bounding boxes are cropped out to get crack images that contain much less background. These images are then processed by Otsu's method for threshold segmentation to obtain the binary results of cracks. A thinning algorithm is used to extract the skeletal representation of the binary results of cracks.
4. Crack quantification: the fourth and final module applies the processed images to quantify the cracks. This involves measuring and analyzing the cracks based on the extracted information, a procedure that will be elaborately discussed in Section 4 of this paper.

2.1 Deep learning models

2.1.1 Faster R-CNN

Faster R-CNN is an advanced object detection framework developed by Ren et al. [8] in 2016. It achieved a significant improvement over previous models by incorporating the innovative RPN, which is designed to efficiently

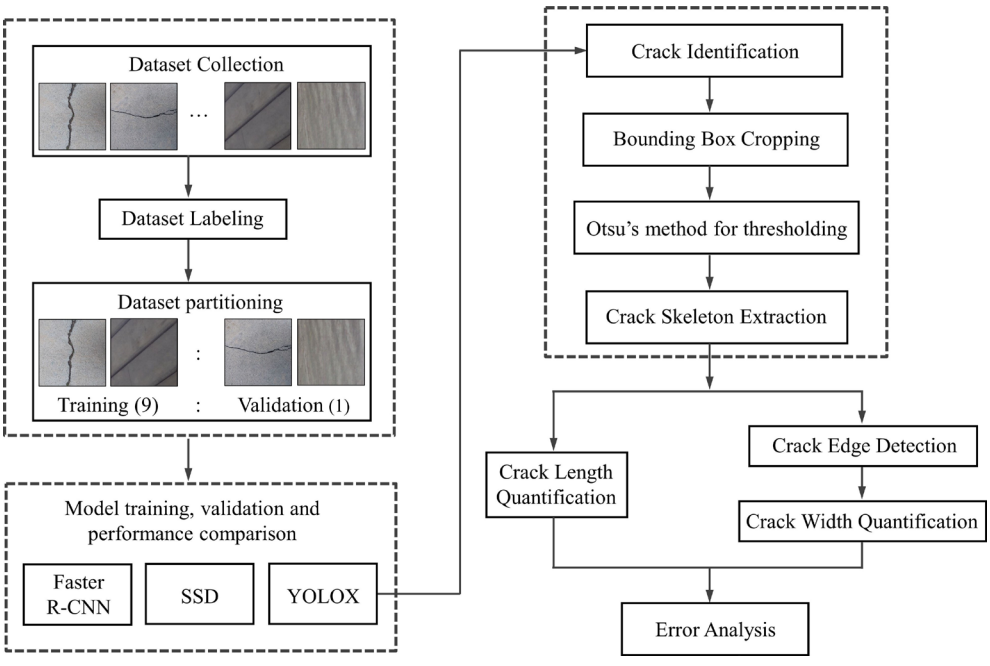


Fig. 1 Overview of the research process

generate high-quality candidate regions (proposals) for object detection. RPN operates in conjunction with the Fast R-CNN detection network, enabling the sharing of convolutional features between the RPN and the detection network. This facilitates efficient end-to-end training and feature utilization. To tackle the challenges of detecting objects at various scales, Faster R-CNN introduces a novel anchor mechanism. The model uses a set of predefined boxes of various sizes and ratios, known as anchors, as references for predicting the presence and location of objects. Furthermore, Faster R-CNN utilizes a Region of Interest (ROI) pooling layer to transform the irregularly shaped candidate regions into a feature map of a fixed size, which is suitable for processing by a fully connected network. These fixed-size feature maps are essential for the subsequent tasks of object classification and bounding box regression. By integrating all of these components, Faster R-CNN achieves a balance between detection accuracy and speed, making it a powerful tool for object detection tasks. However, it has limitations. The framework requires significant computational power, and despite its advancements, the detection speed still needs improvement for applications that require real-time performance.

2.1.2 SSD

Single Shot Multi-Box Detector (SSD) is an efficient one-stage object detection model that cleverly combines the fast detection capability of YOLO with the impressive detection accuracy of Faster R-CNN. Similar to YOLO, SSD directly generates object categories and bounding box locations in a single forward pass, which significantly improves detection speed. Inspired by Faster R-CNN, SSD also utilizes a region-based approach. In SSD, each cell on every feature map predicts a set of predefined bounding box (known as default boxes) as well as the probabilities of objects in these boxes belonging to specific categories. This enhances the model's ability to adapt to scale variations. In terms of performance, SSD not only outperforms YOLO in speed but also rivals Faster R-CNN in accuracy. Due to its exceptional performance, SSD is widely used in various computer vision tasks, particularly excelling in real-time detection scenarios.

2.1.3 YOLOX

Introduced by Ge et al. [33] in 2021, YOLOX emerged as a cutting-edge detection model, setting a new benchmark in object detection. As an advancement over its predecessors, YOLOv3, YOLOv4, and YOLOv5, YOLOX is distinguished by several key improvements. Foremost among these enhancements is the restructured detection head, which is adeptly partitioned into dual sectors for classification and regression. This partition streamlines the learning process and integrates seamlessly into the final prediction, yielding a synthesis of accuracy and efficiency. Moreover, the Mosaic data augmentation methodology is employed to substantially augment the diversity of the training dataset, amplifying the object detection capabilities of YOLOX. In addition, the innovative SimOTA dynamic label-matching strategy is a pivotal inclusion that intuitively matches positive samples with targets across a spectrum of sizes, thereby circumventing the need for laborious parameter optimization. The eschewal of conventional anchor boxes marks yet another stride in the YOLOX design, leading to a marked reduction in parameter complexity and an acceleration in the algorithm's processing cadence.

The tripartite constitution of the YOLOX model comprises the CSPDarknet backbone, tasked with initial feature extraction, followed by the FPN that bolsters feature discernment and concludes with the YOLO Head classifier regressor. This architecture, illustrated in Fig. 2, encapsulates a sophisticated operational principle that ensures the celerity and preciseness of detection.

2.2 Otsu's method for image thresholding

Thresholding is a fundamental technique in image processing aimed at partitioning an image into distinct regions by categorizing pixels according to their grayscale intensities. Generally, concrete cracks manifest as an abrupt change in the gray level between two juxtaposed regions possessing variant gray levels [42]. Employing the mean of these regions as a baseline, an optimized threshold can be established. For this study, Otsu's method, a nonparametric and unsupervised method of automatic threshold selection [43], is adopted for image segmentation.

The core principle of Otsu's method is the selection of a threshold that maximizes the inter-class variance of the two segments, namely, the foreground and the background. This inter-class variance, denoted by σ_ω^2 and calculated by Eq. (1), is a measure of the squared difference between the weighted mean grayscale intensities of the foreground and the background that are distinguished by a given threshold *t*. The optimal threshold is the one that yields the highest variance value:

$$\sigma_\omega^2 = \omega_0(t)\omega_1(t)\left[\mu_0(t) - \mu_1(t)\right]^2, \tag{1}$$

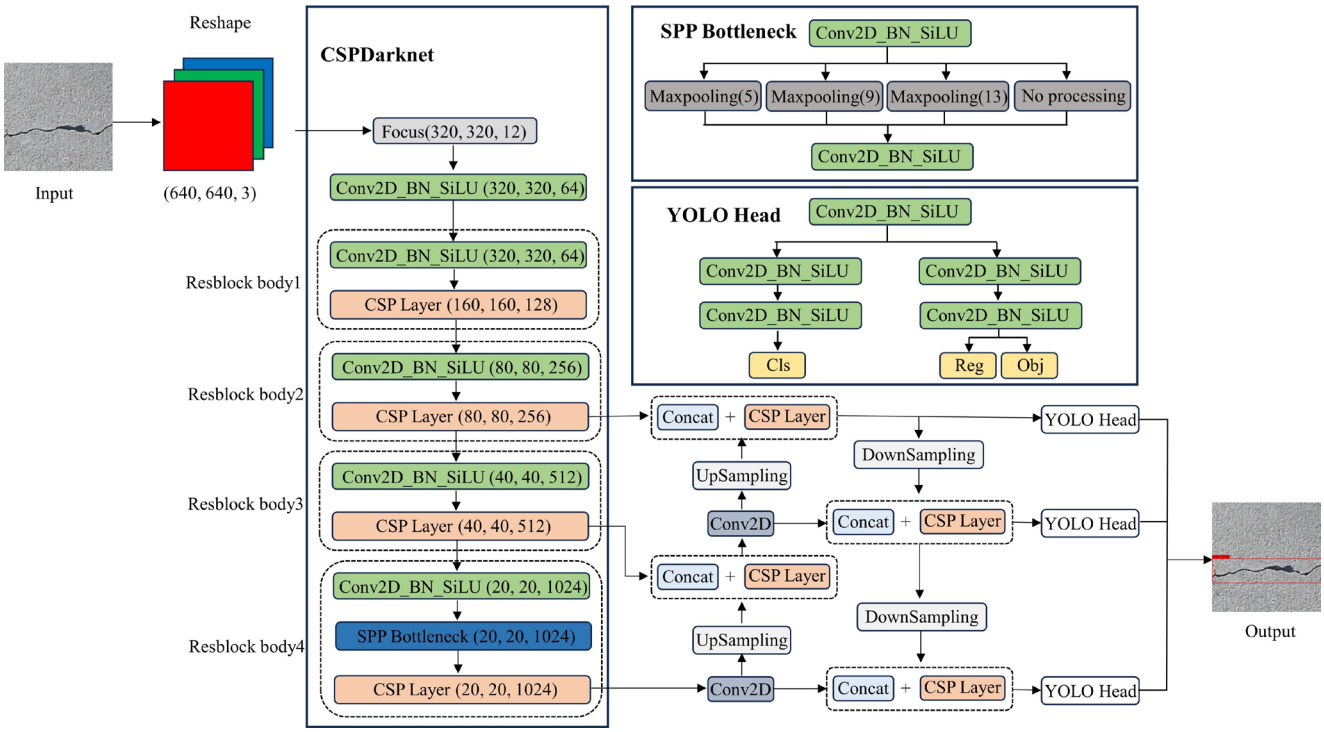


Fig. 2 Principle of the YOLOX bridge crack detection model

where, $\omega_0(t)$ represents the probability of the class of intensities below or at the threshold t, and $\omega_1(t)$ represents the probability of the class of intensities above the threshold t. μ_0, μ_1 are the mean grayscale intensities of the two classes. The probabilities $\omega_0(t)$ and $\omega_1(t)$ are derived from the histogram's L bins (each bin corresponds to a distinct range of intensity values within the image) as Eq. (2) and (3).

$$\omega_0(t) = \sum_{i=0}^{t-1} p(i) \tag{2}$$

$$\omega_1(t) = \sum_{i=t}^{L-1} p(i) \tag{3}$$

2.3 Dataset

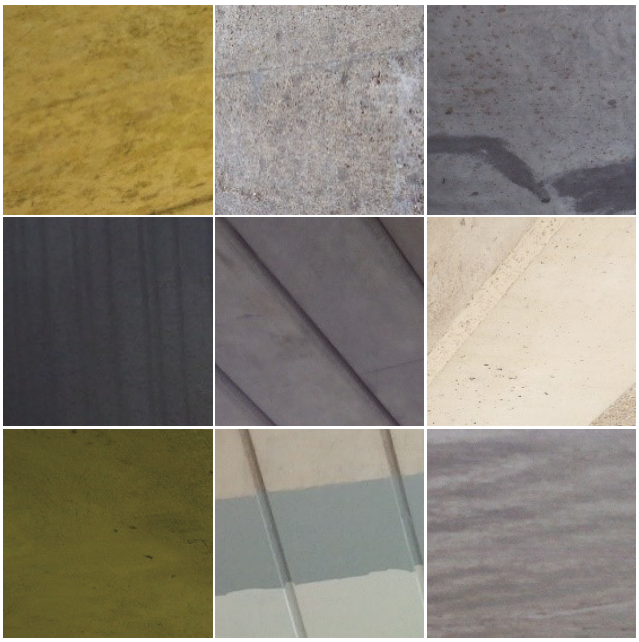
The dataset in this paper was constructed from two publicly available image datasets obtained from GitHub and crack images of main beams and piers captured by drones. Specifically, the dataset consisted of 5742 RGB images, including 2223 images with crack and 3519 images of non-defective backgrounds. The images were captured under varying weather and lighting conditions, and flashlights were used to illuminate dim sections of the bridge that revealed imperfections. It is important to note that the original image has not undergone any processing operations except for manual cropping. Additionally, the constructed dataset included concrete surface conditions of various colors and textures to account for variations in the appearance, extent, and severity of defects under actual bridge conditions. Fig. 3 presents sample images from the proposed dataset.

3 Crack identification

The study presents a three-stage approach to identify bridge concrete cracks, which is essential for developing an accurate and reliable detection system. The Dataset Construction stage forms the foundation of this approach, involving the collection of a diverse set of images from publicly accessible datasets and high-resolution drone captures of bridges. This variety ensures a realistic portrayal of diverse conditions. Meticulous labeling follows, with each image being annotated to mark the presence of cracks, providing high-quality ground-truth data for the models to learn from. Subsequently, the dataset is divided into training and validation sets to facilitate the development of an effective deep learning model and ensure impartial evaluation. For every ten data, nine are allocated to the training set, while one is assigned to the validation set.

During the next phase, Model Training, the advanced object detection architecture YOLOX is deployed. To validate its better performance in detecting bridge concrete cracks, YOLOX is compared to two state-of-the-art target detection models, namely, Faster R-CNN and SSD. The evaluation is grounded on standard evaluation metrics, including precision rate, recall rate, F1 score, and Average Precision (AP) value, ensuring a comprehensive

(a)



(b)

Fig. 3 Sample images of the constructed dataset; (a) bridge cracks; (b) non-defective backgrounds

and multifaceted comparison. All models undergo rigorous training on uniformly partitioned datasets, ensuring a fair and controlled training environment.

In the final stage, known as Model Verification, the trained model with the best overall performance undergoes thorough evaluation by testing on new datasets that were not included in the training data. This is done to assess its generalizability and robustness. Fig. 4 depicts a detailed and clear visual representation of the methodology for detecting cracks in bridge concrete.

3.1 Algorithm selection and training strategies

The YOLOX model is available in six variants – from the smallest nano to the largest x – catering to a spectrum of computational and accuracy requirements. The larger variants, such as l and x, boast enhanced feature extraction capabilities at the cost of increased computational parameters. Conversely, the smaller models, including nano and tiny, are characterized by their reduced parameter count and expedited processing times, though potentially compromising accuracy. Considering the available hardware and software infrastructure, storage capacities, and the trade-off between model performance and resource utilization, the YOLOX_s model is selected for this study. We employ the provided pre-training weight file in the training phase and train the YOLOX_s model with 300 epochs for our constructed dataset since SGD requires more time to converge. The backbone feature extraction network remains frozen for the initial 50 epochs, allowing only fine-tuning of the parameters. This process helps to improve training efficiency. For the next 250 epochs, the backbone feature extraction network is unfrozen to train and update all the network parameters.

To verify the superiority of the YOLOX_s model used in this paper, we purposefully compare it with Faster R-CNN and SSD model. To ensure a rigorous and fair performance comparison, the object detection models YOLOX, SSD, and Faster R-CNN undergo identical training, validation, and testing procedures using a consistent dataset, platform, and training methodologies, as detailed in Table 1.

3.2 Evaluation metrics

A suite of standard evaluation metrics – Precision, Recall, and F1 score – are employed to assess the efficacy of various crack detection models. Precision and Recall are quantified as Eq. (4) and (5). High values of Precision and Recall are indicative of superior model detection capabilities. However, Precision and Recall often exhibit an inverse relationship. An increase in Precision frequently coincides with a decrease in Recall and vice versa, as noted by [44]. To reconcile this trade-off and provide a more holistic measure of model performance, the F1 score is utilized, which is computed using Eq. (6). The F1 score harmonizes the balance between Precision and Recall, thereby serving as a singular measure of overall model effectiveness. It is imperative to clarify the terminology used: true positives (TP) refer to the count of crack samples accurately identified as cracks, while true negatives (TN) refer to the count of non-crack samples correctly recognized as background. Conversely, false positives (FP) refer to the count

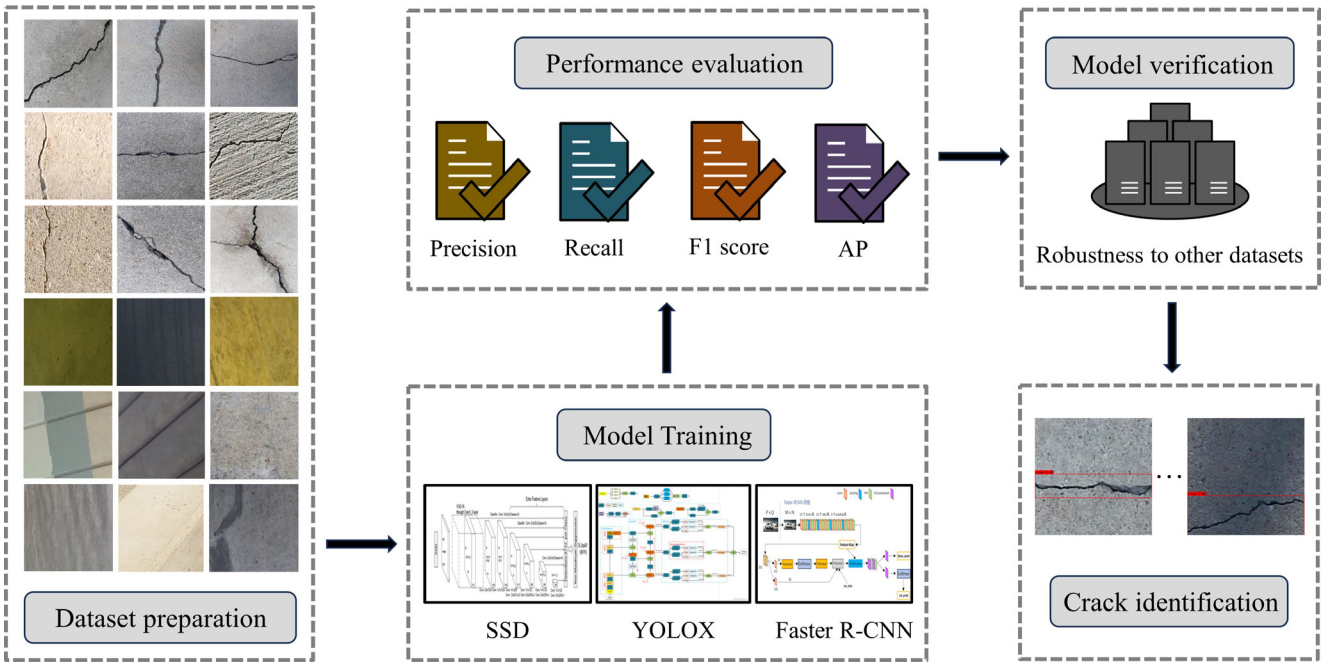


Fig. 4 The identification process of bridge concrete cracks

Table 1 Hardware/software environment and training strategies

Hardware	CPU	Intel(R) Core (TM) i7-8700
	RAM	16G
	GPU	NVIDIA GeForce GTX1060
Software	Os	Windows10
	CUDA	10.0
	CUDNN	7.4.1
	Python	3.6
	Torch	1.2.0
Training strategy	Optimizer	SGD
	Freeze_Batch size	16
	Unfreeze_Batch size	8
	Ir_decay_type	Cos
	Init_Ir	1×10^{-2}
	Min_Ir	1×10^{-4}
	Weight_decay	5×10^{-4}

of non-crack samples erroneously labeled as cracks, and false negatives (FN) refer to the count of crack samples mistakenly classified as background, as defined by [45].

$$\text{Precision} = \frac{TP}{TP + FP} \tag{4}$$

$$\text{Recall} = \frac{TP}{TP + FN} \tag{5}$$

$$\text{F1 score} = \frac{2 \times \text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \tag{6}$$

AP serves as an additional metric for quantifying the performance of different crack identification models. It is determined by calculating the area under the precision-recall (PR) curve, which graphically delineates the trade-off between precision, depicted on the y-axis, and recall, shown on the x-axis. The PR curve is thus a plot that illustrates the relationship between precision and recall across different thresholds. A higher AP value signifies a model's enhanced ability to classify cracks accurately, reflecting a combination of both high Recall and high Precision across varying levels of classification threshold. This metric is particularly valuable as it encapsulates the model's performance across the entire spectrum of possible detection decisions.

3.3 Experimental results and analysis

In our comprehensive evaluation of object detection models, we juxtaposed YOLOX_s with two established models: SSD and Faster R-CNN. We compared the precision and recall for each model to gain insights into their detection capabilities. Furthermore, we plotted the PR curves, F1 score curves, and Loss function curves to visualize their performances.

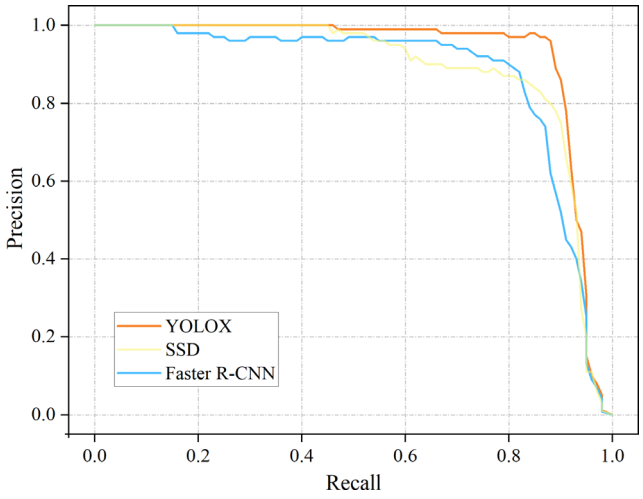
Table 2 presents the precision and recall metrics for all three models. Notably, YOLOX_s outperformed its counterparts in Precision, achieving a remarkable 95.38%, representing a substantial improvement over SSD's 89.19% and Faster R-CNN's 60.57%. However, it is worth mentioning that while YOLOX_s's Recall stands at 87.63%,

Table 2 Precision and recall of the three models

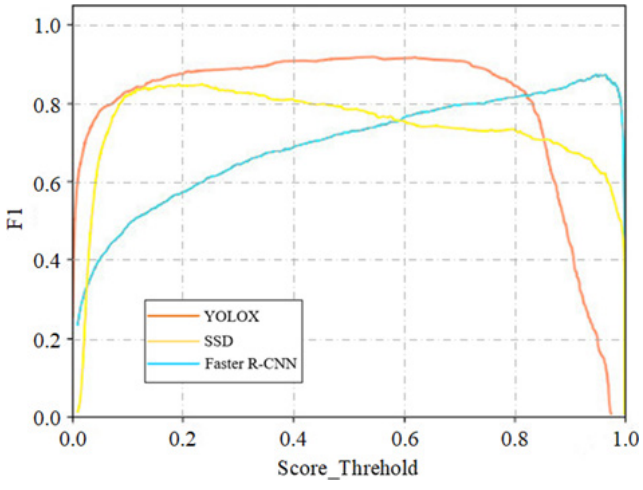
Model	Precision	Recall
YOLOX_s	95.38%	87.63%
SSD	89.19%	69.96%
Faster R-CNN	60.57%	90.11%

it is marginally outperformed by Faster R-CNN, suggesting that the latter may be slightly more effective at reducing FN. Collectively, these results highlight the potential of YOLOX_s as an object detection model with high accuracy, albeit with a slight trade-off in Recall.

As delineated in Fig. 5, the PR curve for YOLOX_s outstrips that of SSD and Faster R-CNN, exemplifying YOLOX_s's superior performance in object detection tasks. The AP metric, as reflected by the area under the PR curve, further corroborates this finding. YOLOX_s boasts an AP value of 0.9177, surpassing SSD's 0.8809 and Faster R-CNN's 0.8657. Additionally, when the decision threshold for classifying detections is established at 0.5 (a juncture that assigns equal importance to both Precision and Recall), YOLOX_s demonstrated the highest detection efficacy with an F1 score of 0.91. This score, a balanced harmonic mean of Precision and Recall, reinforces the robustness of YOLOX_s as an object detection model, particularly in scenarios where an equitable trade-off between Precision and Recall is desirable.



(a)



(b)

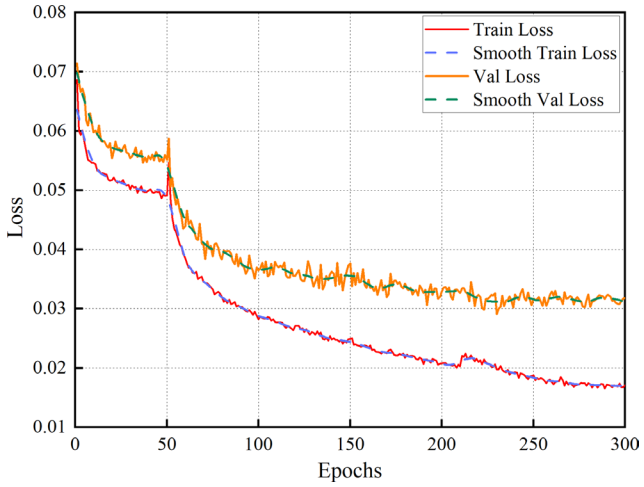
Fig. 5 The PR curve and F1 score curve of the three models; (a) PR curve; (b) F1 score curve

The loss curves of the YOLOX_s, SSD, and Faster R-CNN model are shown in Fig. 6. As the number of training sessions increases, the loss values gradually reduce and remain almost unchanged within the last few periods. Notably, the YOLOX_s model demonstrates a significant reduction in training loss, decreasing from 0.0686 to 0.0171, and a corresponding decline in validation loss from 0.0714 to 0.0318 by the conclusion of the 300th epoch. The loss curves suggest a stabilization in model learning post the 270th epoch despite intermittent fluctuations, indicative of the model's robust learning capabilities and potential for convergence without succumbing to overfitting.

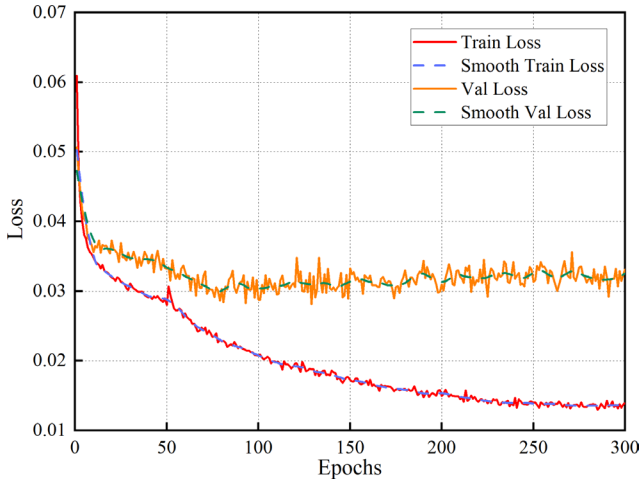
The loss curves of the YOLOX_s, SSD, and Faster R-CNN model are shown in Fig. 6. As the number of training sessions increase, the loss values gradually reduce and remain almost unchanged within the last few periods. Notably, the YOLOX_s model demonstrates a significant reduction in training loss, decreasing from 0.0686 to 0.0171, and a corresponding decline in validation loss from 0.0714 to 0.0318 by the conclusion of the 300th epoch. The loss curves suggest a stabilization in model learning post the 270th epoch despite intermittent fluctuations, indicative of the model's robust learning capabilities and potential for convergence without succumbing to overfitting.

In contrast, while SSD and Faster R-CNN exhibit substantial training loss reductions after the 150th epoch, the uptick in validation loss for these models implies a divergence between training and generalization performance. This trend suggests an inclination towards overfitting, where the models may learn idiosyncrasies of the training data – such as noise irrelevant to crack features – rather than extracting generalizable patterns.

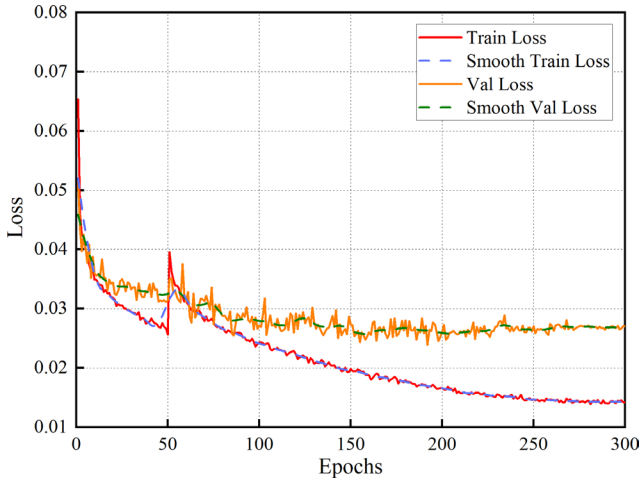
Through meticulous analysis of these training dynamics, we infer that the YOLOX_s model converges efficiently and generalizes better compared to SSD and Faster R-CNN. This indicates YOLOX's superior feature extraction capabilities, further cementing its position as the more adept model for crack detection tasks.



(a)



(b)



(c)

Fig. 6 Loss curves of the three models: (a) YOLOX; (b) SSD; (c) Faster R-CNN

3.4 Detection performance

Given the commendable overall performance of the trained YOLOX_s model compared to SSD and Faster R-CNN, this study proceeds to adopt YOLOX_s for subsequent crack identification. To further substantiate the model's capabilities, we conducted preliminary detection tasks on the SDNET2018 dataset using the trained YOLOX_s model. As depicted in Fig. 7, the results offer empirical evidence of the model's robust generalization ability when applied to new datasets. Specifically, the trained YOLOX_s demonstrated proficiency in accurately identifying cracks on the SDNET2018 dataset, discerning relevant features while successfully disregarding extraneous areas. Moreover, the model exhibited high positional accuracy, with detected crack locations closely aligning with the defects' actual positions, highlighting the model's practical utility.

4 Crack quantification

Accurate characterization of surface concrete crack geometry is critical for evaluating the condition of bridges. Consequently, subsequent to the precise identification of bridge concrete cracks through a deep learning model, it is essential to develop corresponding quantitative methodologies that can measure the crack dimensions, particularly the length and width of the cracks. Fig. 8 presents a comprehensive and systematic strategy for measuring and analyzing bridge cracks. Initially, an advanced detection and identification process is implemented using the previously trained and validated YOLOX_s model. The bounding box coordinates provided by the YOLOX_s model are used as a reference for subsequent image cropping, highlighting the crack zones for further examination. Following segmentation, Otsu's method is applied to

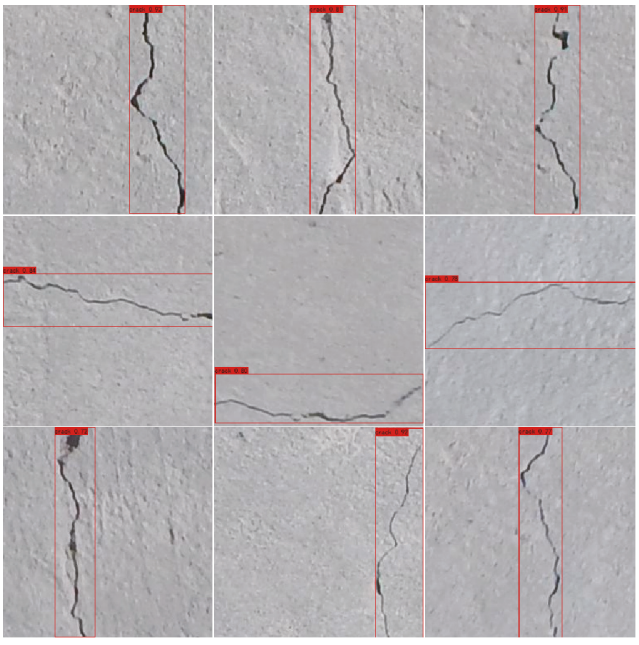


Fig. 7 Samples results of the crack identification for SDNET2018

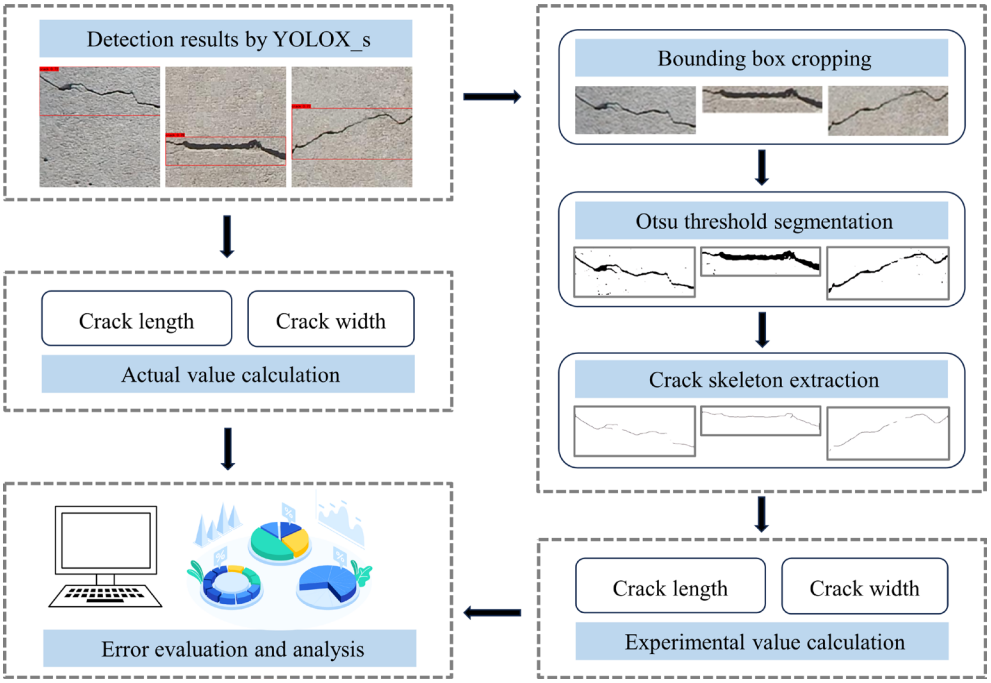


Fig. 8 The quantification framework for cracks

perform thresholding, which adeptly transforms the targeted images into binary form. Subsequently, the Zhang-Suen thinning algorithm is used to extract the crack skeleton, while the Canny edge detection algorithm is employed for crack edge extraction. Finally, the measurements of the cracks are determined by carefully counting the number of pixels corresponding to the crack features and multiplying by the known pixel cell width. This process produces the measured values, which are then compared to the actual dimensions of the cracks to assess the accuracy of the quantification method through error analysis.

4.1 Crack extraction

4.1.1 Bounding box cropping

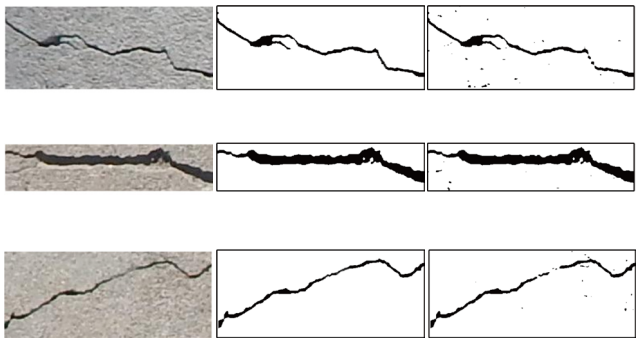
After identifying cracks using YOLOX_s, the model uses bounding boxes to accurately pinpoint the locations of the detected cracks. Based on the coordinates of the bounding boxes, the images within the bounding boxes are cropped out. This step significantly reduces the presence of unnecessary background elements in the crack image, thereby reducing the computational requirements of subsequent processing stages. Critically, in scenarios where multiple cracks are present within a single image, this approach enables the individual isolation of each crack. Isolating the cracks is crucial for reducing the risk of dimensional quantification errors caused by the interactive effects of adjacent cracks, thus improving the overall accuracy of the quantification process.

4.1.2 Thresholding

Otsu's thresholding method is used to accurately separate cracks from their background, aiming to achieve comprehensive crack characterization. Fig. 9 displays a selection of sample results that enable a comparative analysis between the actual ground truth and the outcomes obtained using Otsu's method. It is evident that the binary crack images obtained using Otsu's method are complete, precise, and accurately reflect the actual state of the cracks. This indicates that Otsu's method delivers exceptional performance in crack segmentation, demonstrating significant adaptability and robustness.

4.1.3 Skeleton extraction

The thinning algorithm is a well-established technique that can be used to extract the skeletons of cracks by



(a) (b) (c)

Fig. 9 Sample results of thresholding: (a) crop results; (b) ground truth; (c) Otsu's method

reducing cracks into single-pixel lines. After binarizing crack images using Otsu's method, the Zhang-Suen thinning algorithm is applied to extract the skeletal representations of the cracks.

To evaluate the validity and precision of the Zhang-Suen thinning algorithm, a congruence analysis is conducted to compare the skeleton extracted by the algorithm with the original crack depiction. Fig. 10 illustrates the results, showing the original crack and its binary result, the skeleton obtained through the Zhang-Suen thinning algorithm, and the fit of the skeleton to the original crack. The results reveal a high degree of correspondence between the extracted skeleton and the median line of the original crack, effectively preserving the crack's complete topological structure. These observations confirm that the Zhang-Suen algorithm is suitable for the precise extraction of bridge crack skeletons.

4.2 Crack quantification

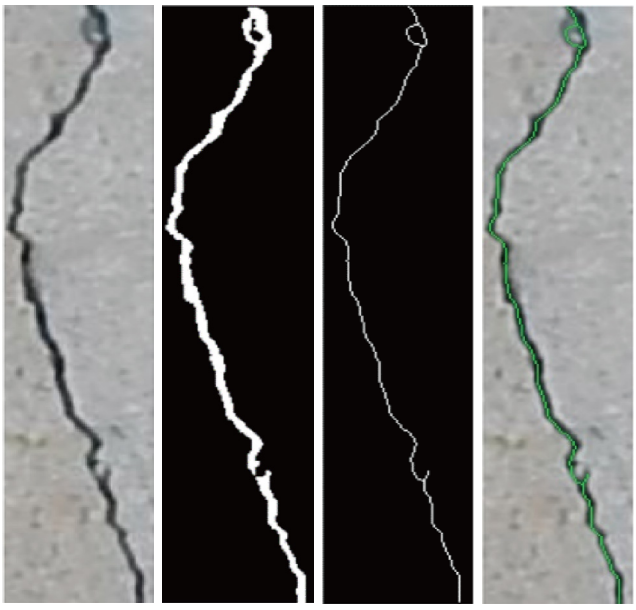
Building upon the successfully extracted crack skeleton, this study introduces an automated crack quantification method that accurately measures both the length and width of cracks.

4.2.1 Length of crack

The total length of a crack is ascertained by summing the micro-crack lengths between two adjacent pixel cells along the crack skeleton. As demonstrated in Fig. 11, three distinct distribution patterns exist among two adjacent pixel cells, characterized by the angle formed between the line connecting their center points and the horizontal line. These angles are 0°, 45°, and 90°, respectively. When the angle is 0° or 90°, the corresponding micro-crack length measures 1 pixel. When the angle is 45°, the length of this micro-crack measures $\sqrt{2}$ pixels. Fig. 11 illustrates a schematic that clarifies the spatial relationship between two adjacent pixel cells and the corresponding methodology for computing the micro-crack length. When tallying the crack length, each pixel cell on the skeleton is compared against its preceding cell, counting the occurrences of 0° (or 90°) and 45° angles. The aggregate crack length is derived using Eq. (7):

$$L = \left(n_1 + \sqrt{2}n_2\right)d, \tag{7}$$

where n_1 is the count of pixel cells forming 0° (or 90°) angles with the preceding cell, n_2 is the count of pixel cells at 45° angles with the preceding cell, and d represents the actual width of each pixel cell.



(a) (b) (c) (d)

Fig. 10 A congruence analysis between the extracted skeleton and the original crack (a) original crack image (b) binary image (c) crack central skeleton (d) the result of the congruence analysis between the extracted skeleton and the original crack

4.2.2 Width of crack

Accurate identification of crack edges constitutes a pivotal step in precisely measuring crack width. This study employs the Canny edge detection algorithm to delineate these boundaries effectively. A critical aspect of crack width computation involves pinpointing the tangent direction at the measurement point along the crack skeleton, which can be intricate. To address this, we introduce the notion of local tangent direction to describe the direction in which the crack width is measured.

The local tangent direction is established by considering the measurement point and the two adjacent pixel cells within its eight-neighborhood. We categorized the local tangent direction into eight distinct cases, as demonstrated in Fig. 12. These cases are differentiated by the angles formed between a line connecting the centers of the two adjacent pixel cells and the horizontal axis, with the specific angles being 0°, 26.6°, 45°, 63.4°, 90°, 116.6°, 135°, and 153.4°. Once the local tangent direction is ascertained, a perpendicular bisector to the tangent line is drawn through the center of the width measurement point. This bisector intersects the identified crack edges at two points, denoted as P and Q. The width of the crack at this juncture is determined by the distance between points P and Q, which can be calculated as follows in Eq. (8):

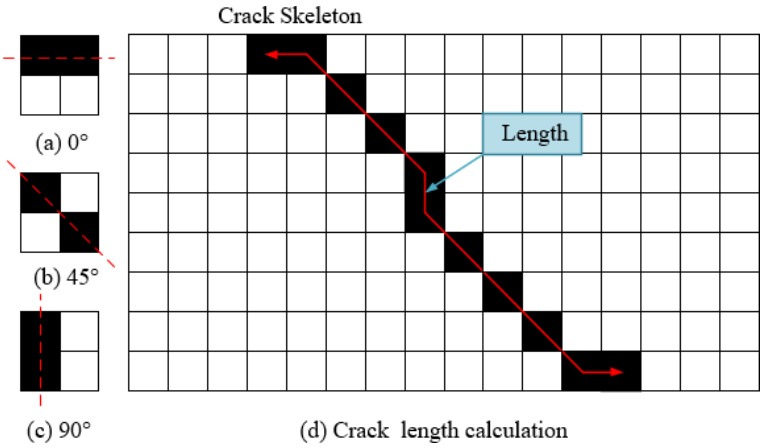


Fig. 11 Illustration of the distributions of two adjacent pixel cells and crack length calculation; the distribution of two adjacent pixel cells at an angle of: (a) 0°; (b) 45°; (c) 90°, respectively; (d) micro-crack length accumulation and cracklength calculation

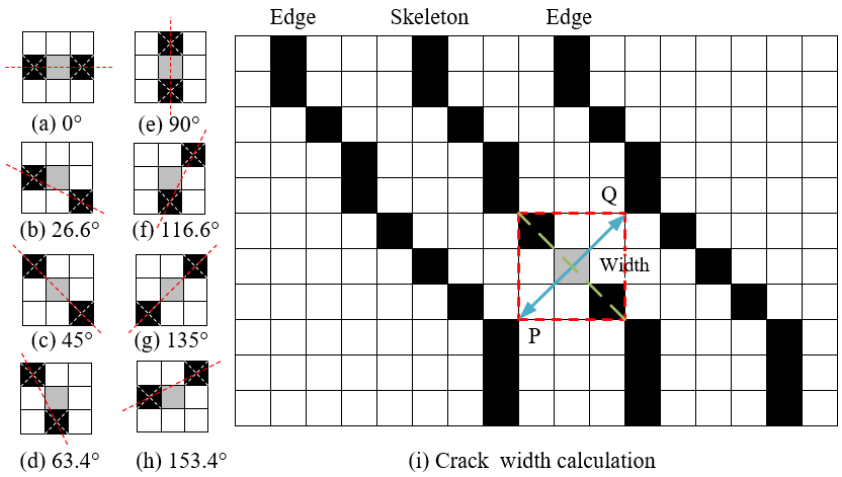


Fig. 12 Illustration of the 8 cases of the local tangent direction and crack width calculation; the distribution of a measurement point and its two adjacent pixel cells at an angle of: (a) 0°; (b) 26.6°; (c) 45°; (d) 63.4°; (e) 90°; (f) 116.6°; (g) 135°; (h) 153.4°, respecively; (i) crack width calculation at a measurement point

$$W_{(P,Q)} = \sqrt{\left(x_P - x_Q\right)^2 + \left(y_P - y_Q\right)^2}, \tag{8}$$

where (x_P, y_P) is the coordinate of P, (x_Q, y_Q) is the coordinate of Q.

4.3 Error analysis

The dataset employed for error analysis is compiled using authentic images of structural cracks. It encompasses 15 high-resolution photographs, which were taken with a SONY camera mounted on a tripod to secure a perpendicular angle to the photographed surface, thereby mitigating angular distortions. To aid in the accurate measurement of the cracks, a ruler was included in each image as a scale reference, assisting with the calibration of pixel dimensions to actual lengths. The specifics of the calibration procedure are depicted in Fig. 13. The calibration target is a square with sides measuring 15 mm, represented on a 240 × 240 pixels grid. Consequently, this translates to each pixel representing a real-world length of 0.0625 mm.

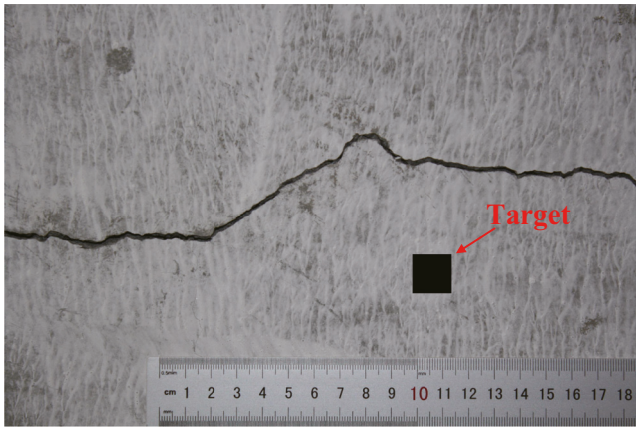


Fig. 13 Target calibration

To validate the validity and precision of the proposed method for calculating crack dimensions, the relative and

absolute errors between the measured and actual values are calculated and analyzed. The relative error offers insights into the accuracy of the measured value in relation to the actual value, while the absolute error provides a direct measure of the deviation from the actual measurements.

Table 3 presents the quantified measurements of crack length across various images, and Table 4 details the calculated measurements for the maximum crack width obtained from various images. The data suggests that the maximum absolute error for crack length measurements is limited to 1.8 mm when compared to the actual values. Furthermore, the relative error is kept below 5.3%, with the mean relative error being 3.400%. Furthermore, the relative error is kept below 5.3%, with the mean relative error being 3.400%. This indicates a high level of detection precision for determining crack length, achieving an accuracy rate of 96.600%. In the measurement of the maximum crack width, the maximum absolute error recorded is 0.04 mm, with the maximum relative error reaching 9.76%, and the data reveals an average relative

Table 3 The measured results of crack length

Crack number	Number of pixel cells		Real value (mm)	Measured value (mm)	Absolute error (mm)	Relative error (%)
	0° (90°)	45°				
1	136	124	20.4	19.5	0.9	4.2
2	343	76	27.1	28.3	−1.2	4.2
3	331	169	34.5	35.8	−1.3	3.6
4	326	180	35.6	36.4	−0.8	2.2
5	333	178	38.5	36.7	1.8	4.9
6	388	142	38.7	37.0	1.7	4.6
7	418	113	34.9	36.3	−1.4	3.9
8	396	119	34.9	35.4	−0.5	1.4
9	347	151	34.1	35.2	−1.1	3.1
10	336	182	36.9	37.2	−0.3	0.9
11	243	127	27.9	26.5	1.4	**5.3**
12	338	173	35.8	36.7	−0.9	2.5
13	340	170	35.4	36.6	−1.2	3.3
14	382	116	35.2	34.4	0.8	2.3
15	359	183	37.1	38.9	**−1.8**	4.6
Average error					1.140	3.400

Table 4 The measured results of the maximum crack width

Crack number	Local tangent direction	Real value (mm)	Measured value (mm)	Absolute error (mm)	Relative error (%)
1	26.6°	0.63	0.60	−0.03	4.76
2	26.6°	0.45	0.48	0.03	6.67
3	45°	0.38	0.36	−0.02	5.26
4	26.6°	0.55	0.54	−0.01	1.82
5	45°	0.41	0.45	**0.04**	**9.76**
6	0°	0.74	0.76	0.02	2.70
7	153.4°	0.73	0.72	−0.01	1.37
8	0°	0.60	0.63	0.03	5.00
9	0°	0.54	0.57	0.03	5.56
10	116.6°	1.18	1.21	0.03	2.54
11	135°	0.64	0.60	0.04	6.25
12	45°	0.35	0.36	0.01	2.86
13	153.4°	0.58	0.60	0.02	3.45
14	26.6°	0.72	0.70	−0.02	2.78
15	63.4°	0.77	0.78	0.01	1.30
Average error				0.023	4.139

error of 4.139%. In other words, the overall detection accuracy for the maximum crack width is calculated to be 95.86%, which falls within the acceptable range for precision, affirming the method's suitability for practical engineering applications where precision is critical.

A more intuitional examination of the method's quantitative capabilities has been conducted. Fig. 14 provides a comparative visualization of the actual versus measured values for crack lengths. Similarly, Fig. 15 maps out the relationship between the actual and measured values for the maximum crack widths. Inspection of Figs. 14 and 15 reveals a noteworthy degree of accuracy in the measurements yielded by the proposed quantification technique, as evidenced by the proximity of these values to the actual observed statistics.

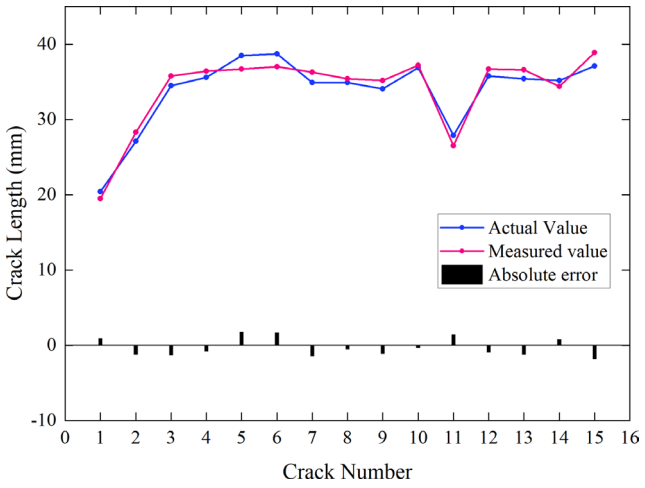


Fig. 14 Comparison of actual and measured values of crack length

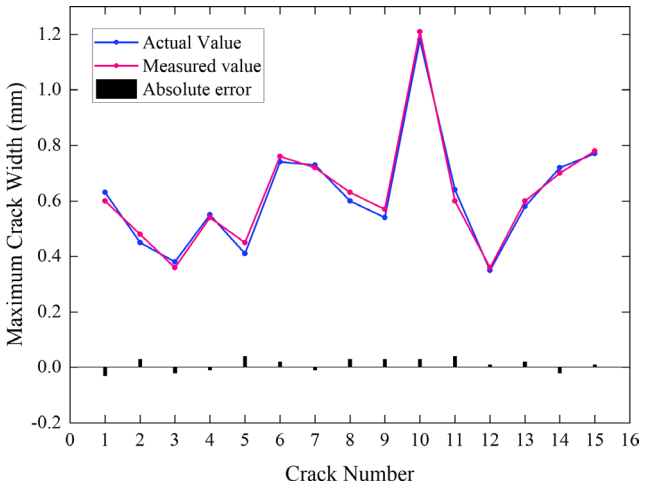


Fig. 15 Comparison of actual and measured values of the maximum crack width

5 Conclusions

Combining the benefits of deep learning and image processing techniques, this paper introduces an automated framework for identifying and quantifying bridge concrete cracks. Three deep learning models, Faster R-CNN, SSD, and YOLOX_s, are trained on a dataset consisting of 5742 RGB images. A comprehensive comparative analysis demonstrates that YOLOX_s outperforms its counterparts in terms of reliability and accuracy. Therefore, YOLOX_s is chosen as the superior model for bridge concrete crack identification. For crack quantification, the detection results of YOLOX_s are used to crop out cracks within the bounding box, which then undergo a thresholding process using Otsu's method. Upon successful threshold segmentation, we employ the Zhang-Suen thinning algorithm and the Canny edge detection algorithm to extract the skeleton and outlines of cracks. Subsequently, we develop a method to calculate the geometric properties of these cracks, with a focus on their lengths and widths. An error analysis is conducted to compare the measured and actual values. Based on the experimental findings, the conclusions can be summarized as follows:

1. YOLOX, coupled with image processing techniques, can offer a valuable tool for crack identification, extraction, and quantification.
2. Comprehensive evaluations confirm that YOLOX_s outperforms SSD and Faster R-CNN in identifying bridge concrete cracks, achieving a Precision of 95.38%, a Recall of 87.63%, and an AP of 0.9177.
3. The automatic crack quantification method based on image processing techniques is simple and efficient, which yields high-precision measurements. The assessment of crack length shows a maximum relative error of only 5.3% with an average relative error of 3.4%. When it comes to the maximum crack width measurements, the relative error remains below 9.76%, with an average relative error of 4.139%.
4. With UAVs to capture high-quality images for training the YOLOX model, YOLOX will achieve better performance of bridge concrete crack identification. Furthermore, the complete and accurate extraction of crack information is critical to ensure the accurate quantification of cracks. By employing more advanced threshold segmentation methods, the accuracy of quantification can be further enhanced. It is believed that with further development, the developed method can provide an automated platform for inspecting bridges.

Acknowledgement

The authors would like to graciously thank the National Natural Science Foundation of China (Grant No. 51678379) and Key Research and Development Program of Sichuan province (2023YFQ0047) for the financial support of this work.

References

[1] Laxman, K. C., Ross, A., Ai, L., Henderson, A., Elbatanouny, E., Bayat, M., Ziehl, P. "Determination of vehicle loads on bridges by acoustic emission and an improved ensemble artificial neural network", Construction and Building Materials, 364, 129844, 2023. https://doi.org/10.1016/j.conbuildmat.2022.129844

[2] Ai, D., Jiang, G., Kei, L. S., Li, C. "Automatic Pixel-Level Pavement Crack Detection Using Information of Multi-Scale Neighborhoods", IEEE Access, 6, pp. 24452–24463, 2018. https://doi.org/10.1109/ACCESS.2018.2829347

[3] Peng, X., Zhong, X., Zhao, C., Chen, A., Zhang, T. "A UAV-based machine vision method for bridge crack recognition and width quantification through hybrid feature learning", Construction and Building Materials, 299, 123896, 2021. https://doi.org/10.1016/j.conbuildmat.2021.123896

[4] Laxman, K. C., Tabassum, N., Ai, L., Cole, C., Ziehl, P. "Automated crack detection and crack depth prediction for reinforced concrete structures using deep learning", Construction and Building Materials, 370, 130709, 2023. https://doi.org/10.1016/j.conbuildmat.2023.130709

[5] Koshti, A. M. "X-ray ray tracing simulation and flaw parameters for crack detection", In: Health Monitoring of Structural and Biological Systems XII, Denver, CO, USA, 2018, 106000R. ISBN 9781510616967 https://doi.org/10.1117/12.2286784

[6] Li, K., Li, T., Ma, M., Wang, D., Deng, W., Lu, H. "Laser cladding state recognition and crack defect diagnosis by acoustic emission signal and neural network", Optics & Laser Technology, 142, 107161, 2021. https://doi.org/10.1016/j.optlastec.2021.107161

[7] Liu, W., Anguelov, D., Erhan, D., Szegedy, C., Reed, S., Fu, C.-Y., Berg, A. C. "SSD: Single Shot MultiBox Detector", In: Computer Vision – ECCV 2016, Amsterdam, Netherlands, 2016, pp. 21–37. ISBN 978-3-319-46447-3 https://doi.org/10.1007/978-3-319-46448-0_2

[8] Ren, S., He, K., Girshick, R., Sun, J. "Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks", [preprint] arXiv, arXiv:1506.01497, 04 June 2015. https://doi.org/10.48550/arXiv.1506.01497

[9] Redmon, J., Divvala, S., Girshick, R., Farhadi, A. "You Only Look Once: Unified, Real-Time Object Detection", In: 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Las Vegas, NV, USA, 2016, pp. 779–788. ISBN 978-1-4673-8852-8 https://doi.org/10.1109/CVPR.2016.91

[10] Yan, K., Zhang, Z. "Automated Asphalt Highway Pavement Crack Detection Based on Deformable Single Shot Multi-Box Detector Under a Complex Environment", IEEE Access, 9, pp. 150925–150938, 2021. https://doi.org/10.1109/ACCESS.2021.3125703

[11] Huyan, J., Li, W., Tighe, S., Zhai, J., Xu, Z, Chen, Y. "Detection of sealed and unsealed cracks with complex backgrounds using deep convolutional neural network", Automation in Construction, 107, 102946, 2019. https://doi.org/10.1016/j.autcon.2019.102946

[12] Park, S. E., Eem, S.-H., Jeon, H. "Concrete crack detection and quantification using deep learning and structured light", Construction and Building Materials, 252, 119096, 2020. https://doi.org/10.1016/j.conbuildmat.2020.119096

[13] Kuang, G., Li, B., Mo, S., Hu, X., Li, L. "Review on Machine Learning-based Defect Detection of Shield Tunnel Lining", Periodica Polytechnica Civil Engineering, 66(3), pp. 943–957, 2022. https://doi.org/10.3311/PPci.19859

[14] Zhou, Z., Zhang, J., Gong, C. "Automatic detection method of tunnel lining multi-defects via an enhanced You Only Look Once network", Computer-Aided Civil and Infrastructure Engineering, 37(6), pp. 762–780, 2022. https://doi.org/10.1111/mice.12836

[15] Yu, Z., Shen, Y., Shen, C. "A real-time detection approach for bridge cracks based on YOLOv4-FPM", Automation in Construction, 122, 103514, 2021. https://doi.org/10.1016/j.autcon.2020.103514

[16] Zhang, J., Qian, S., Tan, C. "Automated bridge surface crack detection and segmentation using computer vision-based deep learning model", Engineering Applications of Artificial Intelligence, 115, 105225, 2022. https://doi.org/10.1016/j.engappai.2022.105225

[17] Wan, C., Xiong, X., Wen, B., Gao, S., Fang, D., Yang, C., Xue, S. "Crack detection for concrete bridges with imaged based deep learning", Science Progress, 105(4), 00368504221128487, 2022. https://doi.org/10.1177/00368504221128487

[18] Shi, P., Fan, X., Ni, J., Wang, G. "A detection and classification approach for underwater dam cracks", Structural Health Monitoring, 15(5), pp. 541–554, 2016. https://doi.org/10.1177/1475921716651039

[19] Feng, C., Zhang, H., Wang, H., Wang, S., Li, Y. "Automatic Pixel-Level Crack Detection on Dam Surface Using Deep Convolutional Network", Sensors, 20(7), 2069, 2020. https://doi.org/10.3390/s20072069

[20] Fan, X., Cao, P., Shi, P., Chen, X., Zhou, X., Gong, Q. "An underwater dam crack image segmentation method based on multi-level adversarial transfer learning", Neurocomputing, 505, pp. 19–29, 2022. https://doi.org/10.1016/j.neucom.2022.07.036

[21] Zhao, Z.-Q., Zheng, P., Xu, S.-T., Wu, X. "Object Detection With Deep Learning: A Review", IEEE Transactions on Neural Networks and Learning Systems, 30(11), pp. 3212–3232, 2019. https://doi.org/10.1109/TNNLS.2018.2876865

[22] Wang, W., Su, C. "Semi-supervised semantic segmentation network for surface crack detection", Automation in Construction, 128, 103786, 2021.
https://doi.org/10.1016/j.autcon.2021.103786

[23] Kao, S.-P., Chang, Y.-C., Wang, F.-L. "Combining the YOLOv4 Deep Learning Model with UAV Imagery Processing Technology in the Extraction and Quantization of Cracks in Bridges", Sensors, 23(5), 2572, 2023.
https://doi.org/10.3390/s23052572

[24] Ali, R., Chuah, J. H., Abu Talip, M. S., Mokhtar, N., Shoaib, M. A. "Structural crack detection using deep convolutional neural networks", Automation in Construction, 133, 103989, 2022.
https://doi.org/10.1016/j.autcon.2021.103989

[25] Srivastava, S., Divekar, A. V., Anilkumar, C., Naik, I., Kulkarni, V., Pattabiraman, V. "Comparative analysis of deep learning image detection algorithms", Journal of Big Data, 8(1), 66, 2021.
https://doi.org/10.1186/s40537-021-00434-w

[26] Zhou, S., Qiu, J. "Enhanced SSD with interactive multi-scale attention features for object detection", Multimedia Tools and Applications, 80(8), pp. 11539–11556, 2021.
https://doi.org/10.1007/s11042-020-10191-2

[27] Sultana, F., Sufian, A., Dutta, P. "A Review of Object Detection Models Based on Convolutional Neural Network", In: Mandal, J. K., Banerjee, S. (eds.) Intelligent Computing: Image Processing Based Applications, 2020, pp. 1–16. ISBN 978-981-15-4287-9
https://doi.org/10.1007/978-981-15-4288-6_1

[28] Tulbure, A.-A., Tulbure, A.-A., Dulf, E.-H. "A review on modern defect detection models using DCNNs – Deep convolutional neural networks", Journal of Advanced Research, 35, pp. 33–48, 2022.
https://doi.org/10.1016/j.jare.2021.03.015

[29] Du, Y., Pan, N., Xu, Z., Deng, F., Shen, Y., Kang, H. "Pavement distress detection and classification based on YOLO network", International Journal of Pavement Engineering, 22(13), pp. 1659–1672, 2021.
https://doi.org/10.1080/10298436.2020.1714047

[30] Zhang, Y., Huang, J., Cai, F. "On Bridge Surface Crack Detection Based on an Improved YOLO v3 Algorithm", IFAC-PapersOnLine, 53(2), pp. 8205–8210, 2020.
https://doi.org/10.1016/j.ifacol.2020.12.1994

[31] Yao, G., Sun, Y., Wong, M., Lv, X. "A Real-Time Detection Method for Concrete Surface Cracks Based on Improved YOLOv4", Symmetry, 13(9), 1716, 2021.
https://doi.org/10.3390/sym13091716

[32] Hu, G. X., Hu, B. L., Yang, Z., Huang, L., Li, P. "Pavement Crack Detection Method Based on Deep Learning Models", Wireless Communications and Mobile Computing, 2021, 573590, 2021.
https://doi.org/10.1155/2021/5573590

[33] Ge, Z., Liu, S., Wang, F., Li, Z., Sun, J. "YOLOX: Exceeding YOLO Series in 2021", [preprint] arXiv, arXiv:2107.08430, 18 July 2021.
https://doi.org/10.48550/arXiv.2107.08430

[34] Wang, L., Li, J., Kang, F. "Crack Location and Degree Detection Method Based on YOLOX Model", Applied Sciences, 12(24), 12572, 2022.
https://doi.org/10.3390/app122412572

[35] Liu, Y.-F., Cho, S., Spencer, B. F., Fan, J.-S. "Concrete Crack Assessment Using Digital Image Processing and 3D Scene Reconstruction", Journal of Computing in Civil Engineering, 30(1), 04014124, 2016.
https://doi.org/10.1061/(ASCE)CP.1943-5487.0000446

[36] Abdel-Qader, L., Abudayyeh, O., Kelly, M. E. "Analysis of Edge-Detection Techniques for Crack Identification in Bridges", Journal of Computing in Civil Engineering, 17(4), pp. 255–263, 2003.
https://doi.org/10.1061/(ASCE)0887-3801(2003)17:4(255)

[37] Kim, H., Ahn, E., Cho, S., Shin, M., Sim, S.-H. "Comparative analysis of image binarization methods for crack identification in concrete structures", Cement and Concrete Research, 99, pp. 53–61, 2017.
https://doi.org/10.1016/j.cemconres.2017.04.018

[38] Kim, H., Lee, J., Ahn, E., Cho, S., Shin, M., Sim, S.-H. "Concrete Crack Identification Using a UAV Incorporating Hybrid Image Processing", Sensors, 17(9), 2052, 2017.
https://doi.org/10.3390/s17092052

[39] Otsu, N. "A Threshold Selection Method from Gray-Level Histograms", IEEE Transactions on Systems, Man, and Cybernetics, 9(1), pp. 62–66, 1979.
https://doi.org/10.1109/TSMC.1979.4310076

[40] Sun, Y., Salari, E., Chou, E. "Automated pavement distress detection using advanced image processing techniques", In: 2009 IEEE International Conference on Electro/Information Technology, Windsor, ON, Canada, 2009, pp. 373–377. ISBN 978-1-4244-3354-4
https://doi.org/10.1109/EIT.2009.5189645

[41] Hoang, N.-D. "Detection of Surface Crack in Building Structures Using Image Processing Technique with an Improved Otsu Method for Image Thresholding", Advances in Civil Engineering, 2018, 3924120, 2018.
https://doi.org/10.1155/2018/3924120

[42] Flah, M., Suleiman, A. R., Nehdi, M. L. "Classification and quantification of cracks in concrete structures using deep learning image-based techniques", Cement and Concrete Composites, 114, 103781, 2020.
https://doi.org/10.1016/j.cemconcomp.2020.103781

[43] Yang, P., Song, W., Zhao, X., Zheng, R., Qingge, L. "An improved Otsu threshold segmentation algorithm", International Journal of Computational Science and Engineering, 22(1), pp. 146–153, 2020.
https://doi.org/10.1504/IJCSE.2020.107266

[44] Lu, G., He, X., Wang, Q., Shao, F., Wang, J., Jiang, Q. "Bridge crack detection based on improved single shot multi-box detector", PLoS ONE, 17(10), e0275538, 2022.
https://doi.org/10.1371/journal.pone.0275538

[45] Zoubir, H., Rguig, M., El Aroussi, M., Chehri, A., Saadane, R., Jeon, G. "Concrete Bridge Defects Identification and Localization Based on Classification Deep Convolutional Neural Networks and Transfer Learning", Remote Sensing, 14(19), 4882, 2022.
https://doi.org/10.3390/rs14194882