Composites Part B 242 (2022) 110096

Contents lists available at ScienceDirect

Composites Part B

journal homepage: www.elsevier.com/locate/compositesb

Monitoring and automatic characterization of cracks in strain-hardening cementitious composite (SHCC) through intelligent interpretation of photos

Pengwei Guo, Xiangjun Meng, Weina Meng, Yi Bao *

Department of Civil, Environmental and Ocean Engineering, Stevens Institute of Technology, Hoboken, NJ, 07030, United States

ARTICLE INFO

Keywords:
Binocular stereo vision
Computer vision
Crack detection
Crack quantification
Deep learning
Strain-hardening cementitious composites (SHCC)

ABSTRACT

This paper presents an intelligent photo interpretation approach to automatically monitor and characterize dense interconnected microcracks in strain-hardening cementitious composite (SHCC) featuring unique crack patterns in terms of crack number and crack width. The presented approach employs a stereo vision system that integrates binocular and monocular cameras for automatic detection, ranging, and quantification of cracks as well as characterization of crack patterns. The presented approach was implemented into evaluation of SHCC in flexural tests and direct tension tests. Dense microcracks were detected and ranged by the stereo vision system, segmented by an encoder-decoder approach, and quantified by an efficient computer vision approach. Evolution of the cracks was traced throughout the loading process until failure, and a statistical analysis revealed that the crack width was retained while the crack number monotonically increased. The interpretation time was shorter than 0.4 s for each photo, making the approach promising for monitoring of SHCC. The proposed system can be deployed for automated assessment of cementitious composites with complex crack patterns in material research and engineering structures.

1. Introduction

Civil infrastructure in the United States is aging and compromises economic wealth and public safety. Many structures were built several decades ago and have shown poor conditions. According to the America's Infrastructure Report in 2021, the overall rate of civil infrastructure is C- [1]. The aging infrastructure poses a significant challenge and increases the financial burden for inspection, maintenance, and repair of infrastructure. It was projected that annual infrastructure maintenance need an additional $206 billion to close a funding gap of two trillion US dollars for 10 years [1]. The funding gap takes "a toll on families' disposable household income" and impacts "the quality and quantity of jobs in the U.S." – resulting in a loss of $4 trillion in GDP, 2.5 million jobs, and $3400 in annual disposable income for each household in the U.S [2]. The large funding gap drives the prioritization of critical structures for timely treatment so as to minimize catastrophic consequences associated with structural failure. Condition assessment and health monitoring of civil infrastructure play significant roles in decision making for asset management. Among various types of anomalies, crack is an important type of damage that may significantly compromise the load-carrying capacity and durability of engineering structures [3]. Timely and reliable assessment of cracks may effectively prevent catastrophic consequences and enable efficient structural repair at an early age of deterioration.

In current practices, visual inspection via bare eyes is still a widely used approach by many engineers in crack inspection, and a crack meter or scope is often employed to measure the crack opening width after a crack is located [4]. This approach is well accepted historically, but it has limitations: (1) The reliability is limited. The inspection results highly depend on the experience and performance of the inspector. There is lack of effective methods to judge reliability of different projects [5]. (2) Visual inspection is costly and time-consuming [6]. It often takes weeks for a crew of multiple structural engineers to inspect one structure, such as a bridge, a building, or a tunnel. (3) It is difficult to inspect structures in harsh environment and extreme weather, such as extreme temperature and precipitation, while extreme weather is projected to increase in frequency and magnitude due to climate change. Distributed fiber optic sensors were proposed to monitor cracks in cementitious composites [4,5] and automatically detect, locate, quantify, and visualize of cracks in prestressed concrete girder [7], pavement [4], and bridge deck [8].

Alternatively, computer vision approaches attracted increasing

* Corresponding author.
E-mail address: yi.bao@stevens.edu (Y. Bao).

https://doi.org/10.1016/j.compositesb.2022.110096
Received 3 February 2022; Received in revised form 22 May 2022; Accepted 29 June 2022
Available online 2 July 2022

Table 1
Summary of representative methods reported in existing publications.

Year	Reference	Method	Auto-ranging	Crack detection	Crack segmentation	Crack quantification	Human intervention
2017	[31]	CrackNet	No	No	Yes	No	Yes
2018	[27]	FCN	No	No	Yes	Yes	Yes
2019	[9]	U-net	No	No	Yes	No	Yes
2019	[6]	Mask-RCNN	No	Yes	Yes	Yes	Yes
2020	[32]	SCHNet	No	No	Yes	No	Yes
2020	[33]	FCN	No	No	Yes	No	Yes
2020	[34]	Faster-RCNN	No	Yes	No	No	Yes
2020	[35]	YOLOv3 + U-net	Yes	Yes	No	No	Yes
2020	[23]	U-net	No	No	Yes	No	Yes
2020	[36]	Faster-RCNN	No	Yes	Yes	No	Yes
2020	[37]	DeepLabv3+	No	No	Yes	Yes	Yes
2020	[11]	Faster-RCNN	No	Yes	No	No	Yes
2020	[25]	SegNet	No	No	Yes	No	Yes
2022	[30]	Mask-RCNN + U-net	No	Yes	Yes	Yes	Yes

interests in assessing cracks using photos or videos [9]. Different approaches have been proposed to detect, segment, and quantify cracks. Those approaches are categorized into: (1) two-stage detectors [10], and (2) one-stage detectors. Representative two-stage detectors include region-based convolutional neural network (R–CNN) [11], Faster-RCNN [12], Mask-RCNN [6,13], and Cascade-RCNN [14], which were mainly developed on convolutional neural network (CNN) [15]. Representative one-stage detectors include YOLO [16,17] and single shot detector (SSD) [18]. Two-stage detectors feature high accuracy, and one-stage detectors feature high efficiency [19]. YOLOv5 achieved a rate of 150 images per second [20]. In addition to the capability of detection, cracks were segmented and quantified using photos. Instance segmentation approaches were proposed to segment cracks [21–23]. Representative approaches include fully convolutional network (FCN) [24], U-shaped fully convolutional network (U-net) [9], and SegNet [25]. The segmentation accuracy and efficiency are dependent on the segmentation architecture [26]. A sophisticated architecture may improve the accuracy but compromise the efficiency. With segmented images, cracks were quantified by the pixels of cracks in photos [27]. Previous research demonstrated that ranging was essential for crack quantification. Laser radars or lidars are popular ranging devices [28], which are however costly. Binocular stereo vision approaches are promising to utilize cost-effective binocular cameras [29].

Table 1 compares representative approaches for crack assessment based on computer vision and deep learning. Limitations were identified from the previous research: (1) Automatic ranging was not incorporated into crack assessment. Ranging and crack assessment were often performed separately. In many cases, ranging was performed manually. (2) There is a tradeoff between the assessment accuracy and the efficiency. The time for crack identification and quantification was about 10 s or longer [30], which is too long in many applications. (3) The execution of existing methods involved human intervention and was not automatically conducted. (4) Most previous research focused on conventional concrete with relatively simple crack patterns in terms of crack interconnectivity and spacing. There is lack of research on densely distributed microcracks, which are typical in strain-hardening cementitious composite (SHCC).

SHCC are a family of advanced concrete featuring high mechanical properties such as high tensile strengths, ductility, and toughness [38–40]. After cracks are produced in SHCC, SHCC is able to carry higher loads, similar to the strain-hardening behavior of low-carbon steel [41,42]. An interesting observation of SHCC has unique crack patterns that are different from conventional concrete [43,44]. In cracked SHCC, the crack widths are limited while the crack number increases with the external load until failure when localized cracks are widened [45,46]. The unique crack patterns of SHCC pose challenges to the previous approaches developed based on conventional concrete [47]. A recent research has shown feasibility of assessing dense microcracks in SHCC using computer vision approaches [30,48]. However, the crack assessment efficiency was limited. Assessment of cracks took 11.2 s per photo, which is too long in many applications.

This research presents an intelligent interpretation system to achieve real-time monitoring and automatic characterization of cracks in SHCC. Compared with existing approaches used to detect cracks, the proposed system tackles unique cracks in SHCC by utilizing a stereo vision system that integrates binocular and monocular cameras for automatic detection, ranging, and quantification of cracks as well as characterization of crack patterns for SHCC. The presented approach was implemented into automatic evaluation of SHCC plates in direct tension tests. The proposed system will be deployed for automated assessment of cementitious composites with complex crack patterns in material research and engineering structures.

2. Methodology

2.1. Overview

The stereo vision system has a binocular camera and a high-

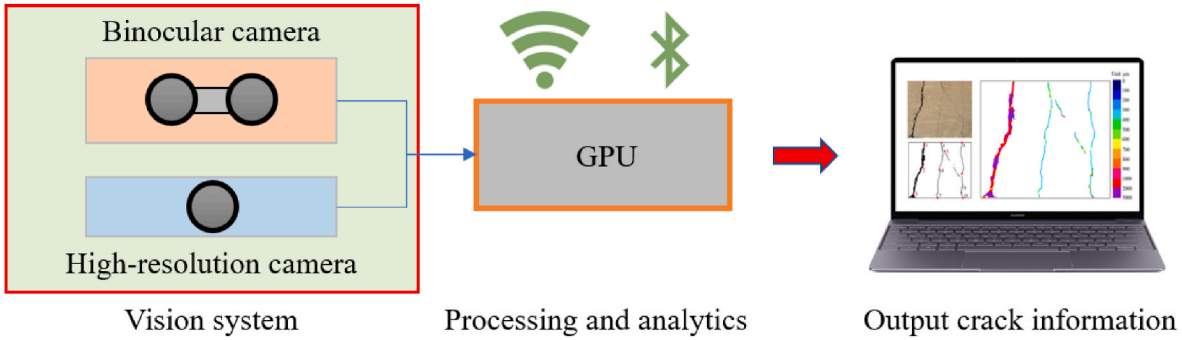


Fig. 1. The proposed smart vision system with a binocular camera and a high-resolution camera.

Table 2
Difference between four YOLOv5 models.

	YOLO5s	YOLO5m	YOLO5l	YOLO5x
Depth ratio	0.33	0.67	1.00	1.33
Width ratio	0.50	0.75	1.00	1.25

resolution monocular camera, as depicted in Fig. 1. Both the binocular camera (resolution: 960 × 720 pixels) and the high-resolution monocular camera (resolution: 2436 × 1125 pixels) were used to take photos of inspected objects. In this research, the high-resolution monocular camera and the binocular camera were connected to a laptop via USB cables or Bluetooth. Python codes were written to synchronize the cameras in acquiring data which was a video. The video was uploaded to a laptop, and Python codes were executed to analyze the frames of the videos. The data from binocular camera were used to measure the distance between the concrete surface and the camera lens via stereo vision. The high-resolution images were used to detect and quantify the cracks based on the deep learning approach. Fig. 1 shows the layout of the stereo vision system and the concept of using the vision system for assessing cracks in SHCC.

2.2. Crack detection

YOLOv5 is the latest generation of YOLO object detection model, and it is known to have the highest detection efficiency while retaining high accuracy. A YOLOv5 model has three main modules: backbone, neck, and prediction [49]. The backbone module consists of a focus layer, convolutional (conv) blocks, C3 layers, and a spatial pyramid pooling (SPP) layer. The focus layer is used to preprocess the image and accelerate the first convolution operations. The convolutional blocks are used to extract features (e.g., color and texture of cracks) from images. A convolutional block is composed of convolutional neural network, batch normalization, and SiLU activation function [50], and batch normalization is used to improve the accuracy and efficiency [51]. The C3 layer is a type of CSPNet (Cross Stage Partial Network), which enhances the learning capability, maintains detection accuracy, and reduces the detection time of CNN [52]. The SPP layer is used to integrate pooling features of different receptive fields and improve the detection accuracy of CNN [53]. The neck module is a combination of feature pyramid network and path aggregation network, and consists of multiple convolutional blocks and C3 layers, in order to utilize the features extracted by the backbone. The prediction module utilizes three different sizes of feature maps generated from the neck module to detect cracks and generate detection boxes on cracks.

This research investigates four YOLOv5 models, which are YOLO5s, YOLO5m, YOLO5l, and YOLO5x. The four models have different depths of C3 layers and different numbers of convolutional kernels, which characterize the sizes of the models. YOLO5s has the smallest size, YOLO5m has a medium size, YOLO5l has a large size, and YOLO5x has the largest size. The depth ratio and width ratio of the four different YOLOv5 models are shown in Table 2.

2.3. Crack segmentation

An encoder-decoder structure is used to generate binary images, as depicted in Fig. 2. In the down-sampling process, the input image is passed through the encoder, which extracts features of cracks and compress the size of feature map. In the up-sampling process, the decoder recovers the size of the feature map. Finally, the original image is converted into a binary image, where the pixels of cracks and concrete are shown in white color and black color, respectively.

Different encoders and decoders were investigated to optimize the segmentation efficiency and accuracy. The investigated encoders include ResNet 18 [54], ResNet 50 [54], ResNet 152 [54], DenseNet [55], VGG 19 [26], MobileNetV2 [56], and InceptionV4 [26]. The investigated decoders include U-shaped fully convolutional network (UNet) [57], Pyramid Scene Parsing Network (PSPNet) [58],

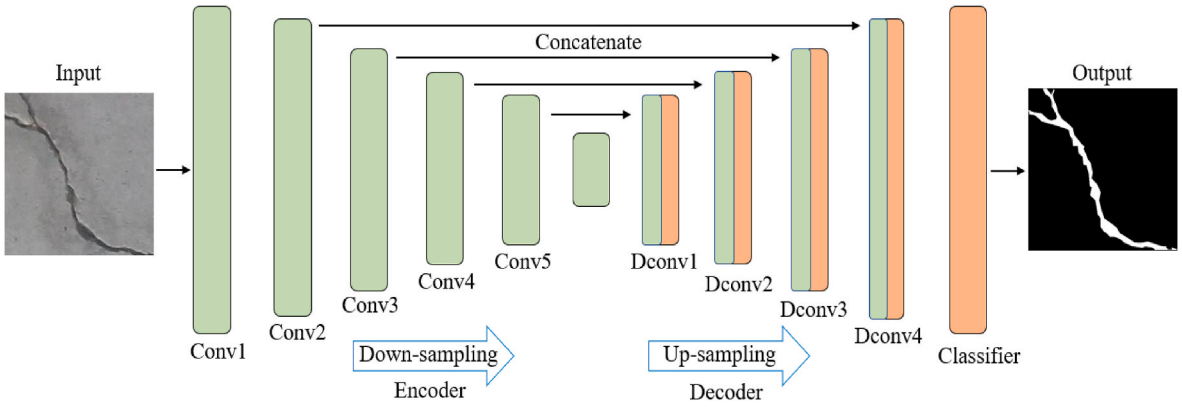


Fig. 2. Architecture of the encoder-decoder structure [26]. Conv is convolutional block used to extract features and compress data. Dconv is deconvolutional block used to recover the image.

Fig. 3. Calibration of binocular camera using a checkboard with black and white squares. "Left camera" and "right camera" refer to the left and the right telescopes of the binocular camera.

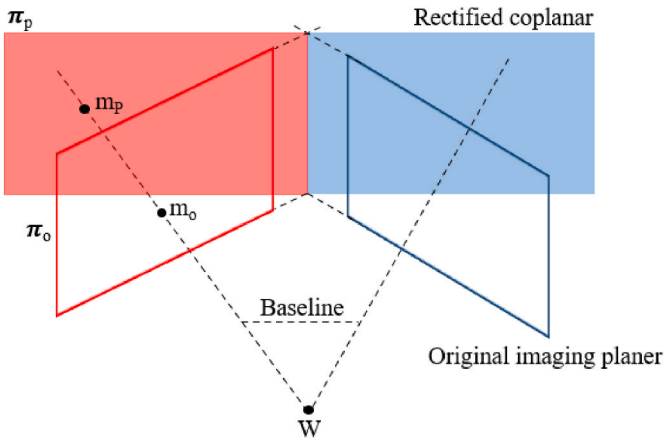


Fig. 4. Illustration of distortion correction. π_P is the rectified coplanar plane; π_o is the imaging plane; m_p is projected point on the rectified plane; m_o is projected point on imaging plane; w is point in the real world; and baseline is the spacing between the left and right cameras.

DeepLabV3+ [59], LinkNet [60], and Pyramid Attention Network (PAN) [61].

2.4. Ranging through binocular stereo vision

The proposed ranging method includes the calibration of cameras, distortion correction of lens, stereo correction, stereo matching and disparity calculation, denoising using filters, and determination of the distance. These steps are elaborated in the following subsections.

2.4.1. Camera calibration

Camera calibration was performed using the stereo camera calibration toolbox in MATLAB to determine the intrinsic parameters of the cameras, the external parameters, the radial distortion coefficient, and the tangent distortion coefficient for the binocular camera [62]. This research proposes to use a 7 × 10 checkerboard with intermediate black and white squares, as shown in Fig. 3. Each of the square area measures 25 mm by 25 mm in area.

2.4.2. Distortion correction

Light is distorted when it propagates in the optical system of a camera. There are two forms of distortion, which are radial distortion and tangent distortion. Radial distortion is related to the shapes of lens. Tangent distortion is caused when the lens and image surface are not parallel. The radial distortion and tangent distortion are corrected to improve accuracy. In this research, radial distortion coefficient and tangent distortion coefficient are obtained through the camera calibration process and used to eliminate the distortion effects [63].

2.4.3. Stereo correction

Photos obtained from binocular cameras are generated by light projected on imaging planes, different from the real planes. Stereo correction was performed to transform the photos from the imaging planes to rectified coplanar planes (see Fig. 4). The transformation was performed based on the triangulation principle, which was used to calculate the distance [63].

2.4.4. Stereo matching and disparity calculation

The stereo images obtained from the left and right cameras of the binocular camera are correlated but have binocular disparity. Stereo matching was performed to correlate the pixel points of the images obtained from the left and the right cameras, and disparity calculation was performed to eliminate the binocular disparity. In this research, stereo matching and disparity calculation were conducted using semi-global-block-matching (SGBM) package of OpenCV to achieve high-speed stereo matching. A disparity map was plotted based on the disparity values.

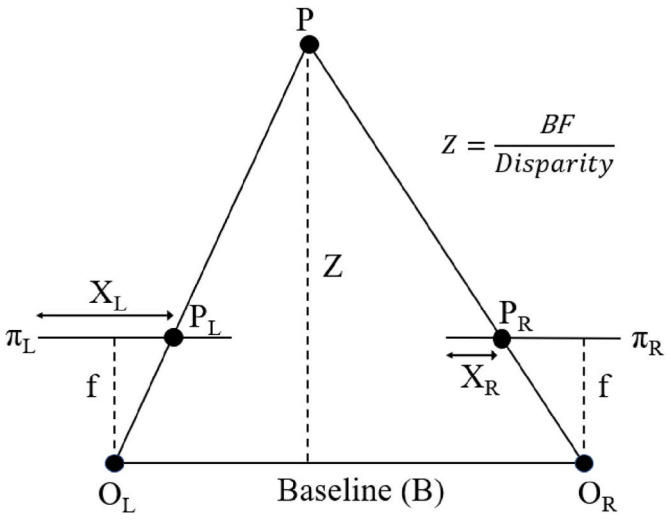


Fig. 5. Calculation of the distance [64]. P is a point in real world; f is the focal length; O_L and O_R are respectively the optical centers of the left and right camera; B is the baseline, which is the spacing between O_L and O_R; Z is the distance between the object and the optical center of the camera. π_L and π_R are the left and right projection planes; P_L and P_R are the projected points on the left and right projection planes; X_L is the distance between P_L and π_L; and X_R is the distance between P_R and π_R. The disparity value is equal to X_L - X_R.

2.4.5. Filtering and determination of distance

Typically, the disparity map has many noise points that compromise the accuracy. This research employed a median filter to remove the noise points from the disparity map. The distance (Z) between the object and the optical center of the camera were calculated using the disparity values, focal length, and baseline, as shown in Fig. 5.

2.5. Focal length of monocular high-resolution camera

The focal length of the monocular high-resolution camera was calculated based on the similar triangle theorem. This research used a coin measuring 24 mm in diameter to obtain the focal length, as illustrated in Fig. 6. The distance between the monocular camera and the coin was the same as the distance between the binocular camera and coin. The pixel number of the coin was determined by Hough circle transform [65]. The focal length was calculated using the actual diameter 24 mm of coin, distance, and pixel number (red color) of the coin.

2.6. Crack quantification

With the binary images obtained from segmentation, the crack was quantified based on the similar triangle theorem, as shown in Fig. 7. The pixels of a crack was turned into crack dimensions, given the focal length of camera and the distance between the object and camera.

Fig. 8 depicts the method for determining the pixels corresponding to cracks. In general, a crack has an arbitrary geometry in the binary image. Both the crack width and orientation may change along the crack. This research proposes two methods to handle the change of crack width and orientation, respectively. For the change of crack width, the image is segmented into thin strips, and, within each strip, the crack width is considered to be unchanged.

When the strips are thin, each strip has a constant width and orientation, so a parallelogram is used to characterize the crack in the strip. According to the pixel value, which is either 0 or 255 in the segmented images, the coordinates of the critical points of the parallelogram are

Fig. 6. Focal length calculation: (a) a photo of a coin; (b) an image after Hough circle detection.

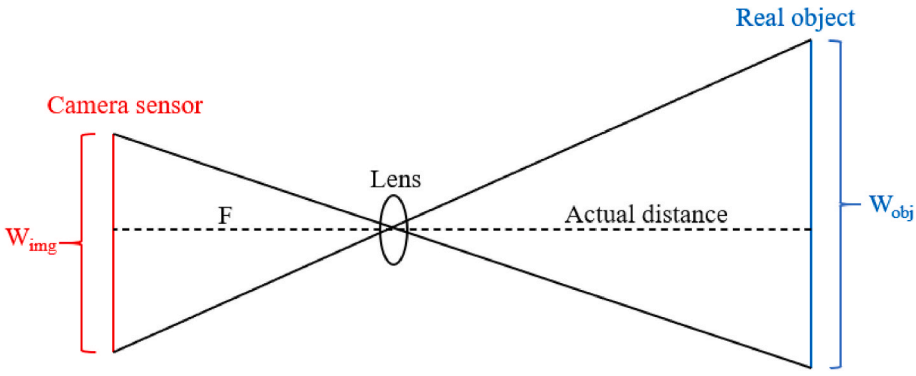


Fig. 7. The similar triangle theorem for crack quantification. F is the focal length of camera; W_{img} is the number of pixels for object in image; and W_{obj} is the actual width of object.

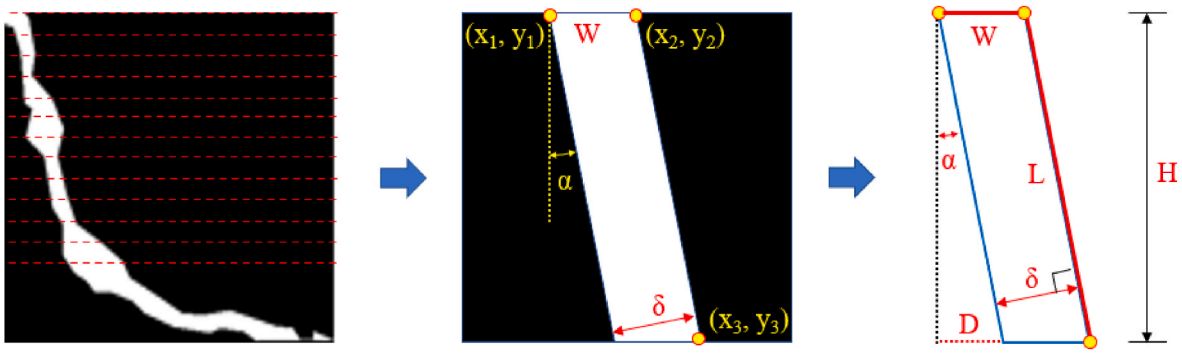


Fig. 8. Quantification of an arbitrary crack with changing width and orientation along the crack.

determined: (x_1, y_1), (x_2, y_2), and (x_3, y_3). Then, the crack width denoted by δ is described by the pixels. For the varying crack orientation, the slope (k) of the crack in the strip is utilized to consider the orientation. The slope is defined as:

$$k = \left|\frac{y_2 - y_3}{x_2 - x_3}\right| \tag{1}$$

Based on the crack slope, the strip orientation is adaptive to the crack orientation that is changing along the length of the crack. The strip is applied to ensure the angle between the strip and crack is not less than 45°. The angle ensures an adequate parallelogram to improve the calculation accuracy of crack width.

Then, the following formulae are used to calculate the crack width (δ) in each of the strips:

$$L = \sqrt[2]{(x_2 - x_3)^2 + (y_2 - y_3)^2} \tag{2}$$

$$cos\alpha = H/L \tag{3}$$

$$W = \begin{cases} |y_1 - y_2|, & k \le 1 \\ |x_1 - x_2|, & k > 1 \end{cases} \tag{4}$$

$$\delta = W \cdot cos\alpha = (H \cdot W)/L \tag{5}$$

where L, D, and H are the side lengths of the triangle, and W is the side length of the parallelogram along the edge of the strip.

The length ratio of an image to the real specimen is known as the scale factor, which indicates that the number of pixels of the image is proportional to the length of the real specimen. Therefore, the crack width is calculated by:

$$\Delta = \left(\frac{\delta}{W}\right)\Delta_{Picture} \tag{6}$$

where Δ is the crack width; and $\Delta_{Picture}$ is the width of the object shown in the picture.

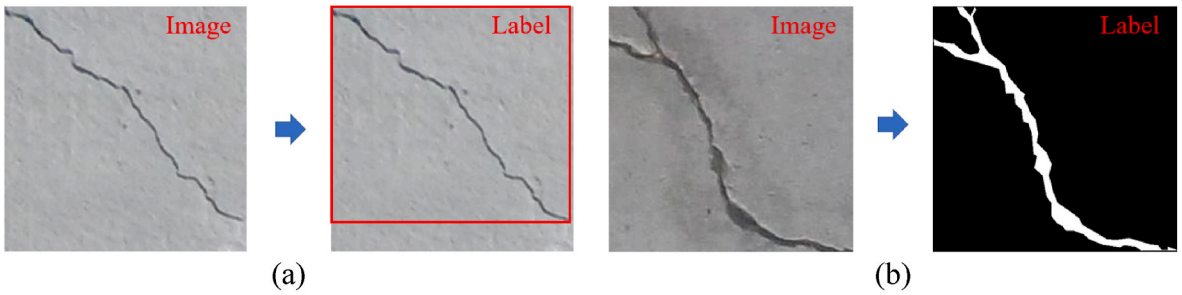
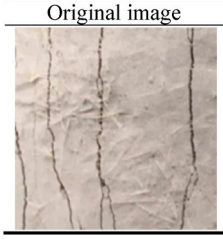
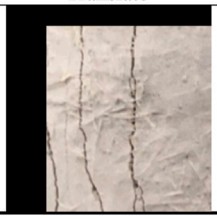
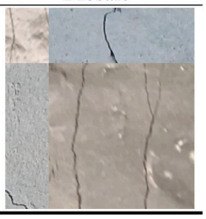


Fig. 9. Labels of the representative images for: (a) crack detection, and (b) crack segmentation. **3.2. Data augmentation**.

Table 3
Illustration of the data augmentation methods.

Original image	Translate	Flip	Mosaic
			

3. Implementation and performance evaluation

3.1. Dataset preparation

Two different datasets were built to train the deep learning models for crack detection and segmentation, respectively. The dataset for crack detection had 1500 images, and the dataset for crack segmentation had 1000 images. All images were resized into 512 × 512 pixels before they were labelled. The images in the two datasets were labelled in different ways. Fig. 9 shows the labels of two representative images. Regarding crack detection, the labels of each image were saved in a text file that only had four coordinates describing a rectangular box. Regarding crack segmentation, the labels of each image were saved in a text file that had many coordinates describing crack contours plotted as a binary image. Each of the two datasets were divided into training (80% data) and validation (20% data) sets.

Multiple strategies were applied to increase the sizes of the dataset because the sizes of the datasets significantly influence the accuracy and the generalizability of the trained models. The adopted data augmentation methods included flipping, translation, shearing, scaling, and mosaic, as illustrated in Table 3. The different data augmentation methods were randomly executed to increase the dataset sizes.

3.2. Crack measurement

The width of each crack was measured using a high-precision crack scope (model: CS-100; magnification: 25X; precision: 50 μm). The measurement of crack width is used as the ground truth that is utilized to evaluate the accuracy of the presented computer vision method based on deep learning.

3.3. Performance metrics

The performance of deep learning models can be evaluated using different metrics, such as precision, recall, dice coefficient (F1 score), intersection over union (IOU), mean average precision (mAP), and coefficient of determination (R^2). The precision, recall, and dice coefficient are defined using four parameters, which are the true positive (TP), true negative (TN), false positive (FP), and false negative (FN). TP is an outcome when the model correctly recognizes the positive class (i.e., cracks). TN is an outcome when the model correctly recognizes the negative class (i.e., uncracked concrete). FP is an outcome when the model recognizes uncracked concrete as crack. FN is an outcome when the model recognizes crack as uncracked concrete.

The precision is the proportion of correctly classified cracks to the total recognized cracks, as shown in Eq. (7):

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}} \tag{7}$$

The recall is the proportion of correctly classified cracks to the total cracks:

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}} \tag{8}$$

Dice coefficient is defined in Eq. (9):

$$\text{F1} = \frac{2\text{TP}}{2\text{TP} + \text{FP} + \text{FN}} \tag{9}$$

Precision and recall are used to evaluate crack detection accuracy. Dice coefficient is used to evaluate crack segmentation accuracy. Their values are in the range of 0–1, and 1 means the highest accuracy. In addition to dice coefficient, IOU is also used to evaluate the segmentation accuracy, as defined in Eq. (10). Typically, IOU higher than 0.5 represents a high accuracy.

$$\text{IOU} = \frac{\text{Predicted results} \cap \text{Ground truth}}{\text{Predicted results} \cup \text{Ground truth}} \tag{10}$$

The area under a precision-recall curve represents the average precision (AP), and mAP is equal to the AP value for each class divided by the number of classes. For example, mAP@0.5 represents the mAP value under the IOU higher than 0.5, as shown in Eq. (11).

$$\text{mAP@0.5} = \frac{\sum_{i=1}^{n} \text{AP}_i}{n} (\text{IOU} > 0.5) \tag{11}$$

where n is the number of class; and i is a certain class. mAP@0.5:0.95 denotes the average of mAP values corresponding to IOU from 0.5 to

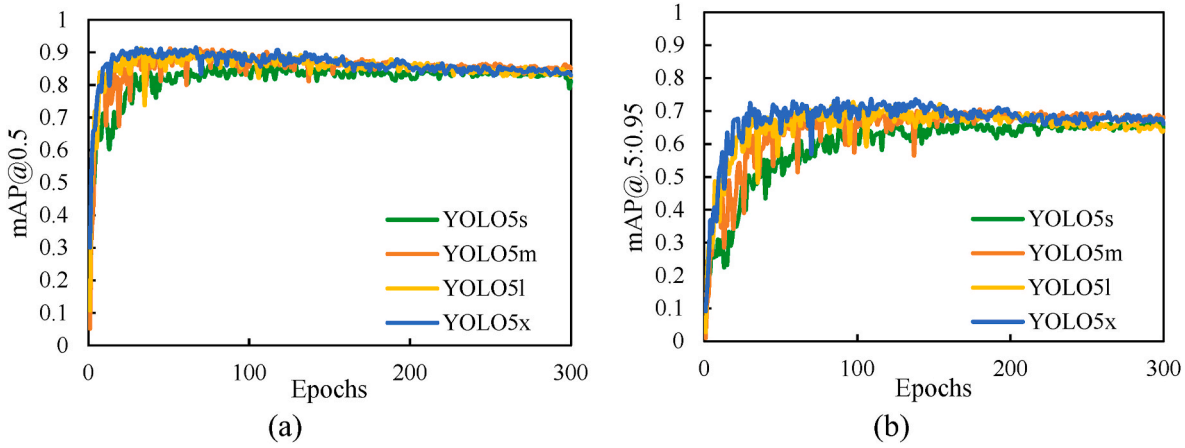


Fig. 10. Validation accuracy of the YOLOv5 models: (a) mAP@0.5, and (b) mAP@0.5:0.95.

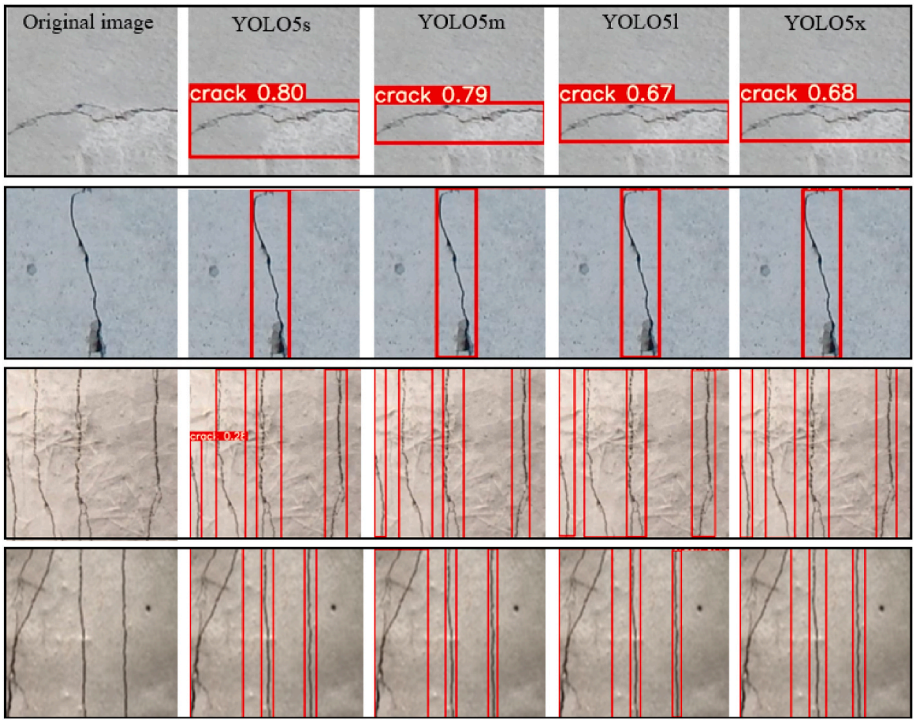


Fig. 11. Representative results of crack detection using the four trained YOLOv5 models.

0.95 at an interval of 0.05, as shown in Eq. (12):

$$\mathrm{mAP@0.5:0.95} = \frac{\{\mathrm{mAP(IOU > 0.5):\ \ mAP\ \ (IOU > 0.95)\}@0.05}}{10} \tag{12}$$

Coefficient of determination (R^2) is used to evaluate the crack quantification accuracy:

$$R^2 = \frac{\sum_{i=1}^{n}(\Delta_A - \overline{\Delta_M})}{\sum_{i=1}^{n}(\Delta_M - \overline{\Delta_M})} \tag{13}$$

where n is the number of data instances; Δ_A is the analysis result of crack width; Δ_M is the measurement of crack width from the crack scope; and $\overline{\Delta_M}$ is the mean value of the measurement of crack width from the crack scope.

4. Results and discussion

4.1. Crack detection

Fig. 10 shows the validation accuracy of the YOLOv5 models trained in 300 epochs. In the first 20 epochs, the accuracy rapidly increases with the epoch number. As the epoch number is further increased, the accuracy is stabilized. In the training of each model, the parameters which provide the highest mAP@0.5 are applied to the model. The YOLO5s model has the lowest mAP@0.5 and mAP@0.5:0.95, indicating the lowest validation accuracy of detection, which is attributed to the simple architecture of YOLO5s. Overall, comparable accuracy is shown by the YOLO5m, YOLO5l, and YOLO5x models. This is because crack is the only class of object to be detected. Thus, the accuracy is not significantly increased by using a large architecture.

The trained YOLOv5 models are then used to detect cracks in 60 photos of cracked SHCC specimens. The 60 photos are not used in

Table 4
Comparison of performance for four YOLO models.

Model	mAP @0.5	mAP @0.5:0.95	Total time (s)	Average time (s)
YOLO5s	0.89	0.67	0.52	0.008
YOLO5m	0.91	0.71	0.65	0.010
YOLO5l	0.91	0.73	0.86	0.013
YOLO5x	0.92	0.74	1.08	0.016

Table 5
Comparison of performance for different decoders tested on 60 images.

Encoder	Decoder	Total time (s)	Speed (s/image)	IOU (0.5)	F1 score
ResNet50	PAN	2.58	0.042	0.976	0.989
ResNet50	UNet	2.41	0.040	0.986	0.993
ResNet50	PSPNet	1.71	0.028	0.971	0.988
ResNet50	LinkNet	2.18	0.035	**0.987**	**0.993**
ResNet50	DeeplabV3+	2.38	0.039	0.984	0.992

training the models, so they represent new unseen data for the trained models. Fig. 11 shows some representative crack detection results. The results showed that all cracks in the photos are recognized by the four models, regardless of the different crack patterns in conventional concrete and SHCC.

Table 4 lists the statistics of accuracy based on validation data, total time, and average time for the YOLO models. The adopted computer had a single RTX 3090 GPU, Intel i9-11900F CPU, and 64 GB memory. The machine learning models were coded based on Python 3.8.8 and Pytorch 1.7.1. The YOLO5x model has the highest mAP@0.5 and mAP @0.5:0.95, indicating the highest accuracy. The implementation of time analysis is based on 60 unlabeled photos. The total time is the time used to detect cracks in the 60 images. The results of the total time of the YOLO5s, YOLO5m, YOLO5l, and YOLO5x models are 0.52 s, 0.65 s, 0.86 s, and 1.08 s, respectively. The results of average time per image for the YOLO5s, YOLO5m, YOLO5l, and YOLO5x models are 0.008 s, 0.010 s, 0.013 s, and 0.016 s, respectively. The different models do not show a significant difference in terms of the assessment efficiency. The YOLO5x model is selected for further tests.

4.2. Crack segmentation

Each segmentation model is trained for 30 epochs. The model with the highest IOU for validation set is used to segment 60 images that are not used in the training. The total time and average time used to segment 60 images are analyzed. Table 5 shows the total time, average speed, and accuracy of different decoders. With the same encoder (ResNet50), LinkNet has the highest accuracy and the second shortest running time. The IOU and F1 score of LinkNet are 0.987 and 0.993, respectively.

Fig. 12 shows representative segmentation results from the different decoders. Compared with the ground truth obtained from visual inspection, LinkNet provides reasonable segmentation results for different types of concrete with different crack patterns. All the five models showed good segmentation accuracy for the photo with a single wide crack, but the models based on DeeplabV3+, PSPNet, UNet, and PAN show errors for segmentation of complex crack patterns. For example, the models based on DeeplabV3+, PSPNet, and PAN fail to segment intersections of cracks at some spots. LinkNet is selected as the decoder for crack segmentation.

Table 6 shows the running time, speed per image, and validation accuracy of segmentation using different encoders. With a same decoder structure (LinkNet), ResNet50 and Inception v4 have the highest segmentation accuracy. MobileNet v2 model has the shortest time. The IOU and F1 score of ResNet50 and Inception v4 are 0.987 and 0.993, respectively.

Fig. 13 shows representative segmentation results from the different

Table 6
Comparison of performance for different encoders.

Decoder	Encoder	Run time (s)	Speed (s/image)	IOU (0.5)	F1 score
LinkNet	ResNet18	1.80	0.030	0.979	0.990
LinkNet	ResNet50	2.20	0.036	**0.987**	**0.993**
LinkNet	ResNet152	3.04	0.051	0.985	0.992
LinkNet	VGG19	4.16	0.069	0.985	0.993
LinkNet	Inception v4	4.03	0.072	**0.987**	**0.993**
LinkNet	MobileNet v2	**1.74**	**0.029**	0.974	0.989
LinkNet	DenseNet161	4.44	0.074	0.985	0.993

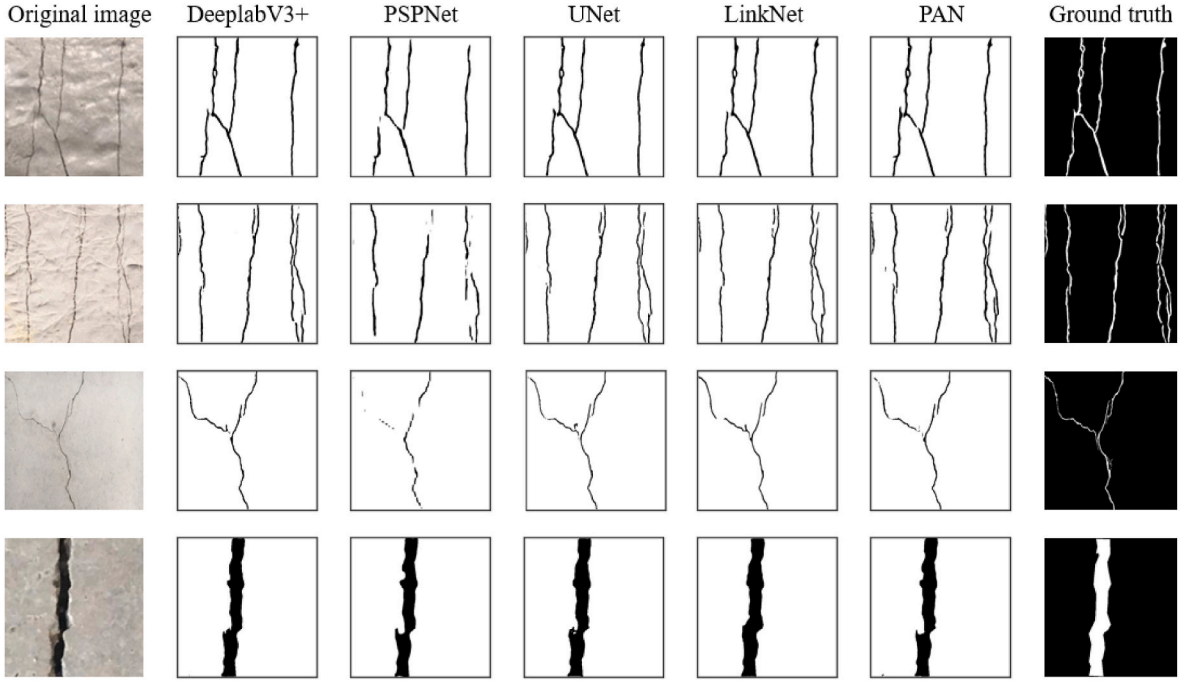


Fig. 12. Representative results of crack segmentation using the different decoder structures.

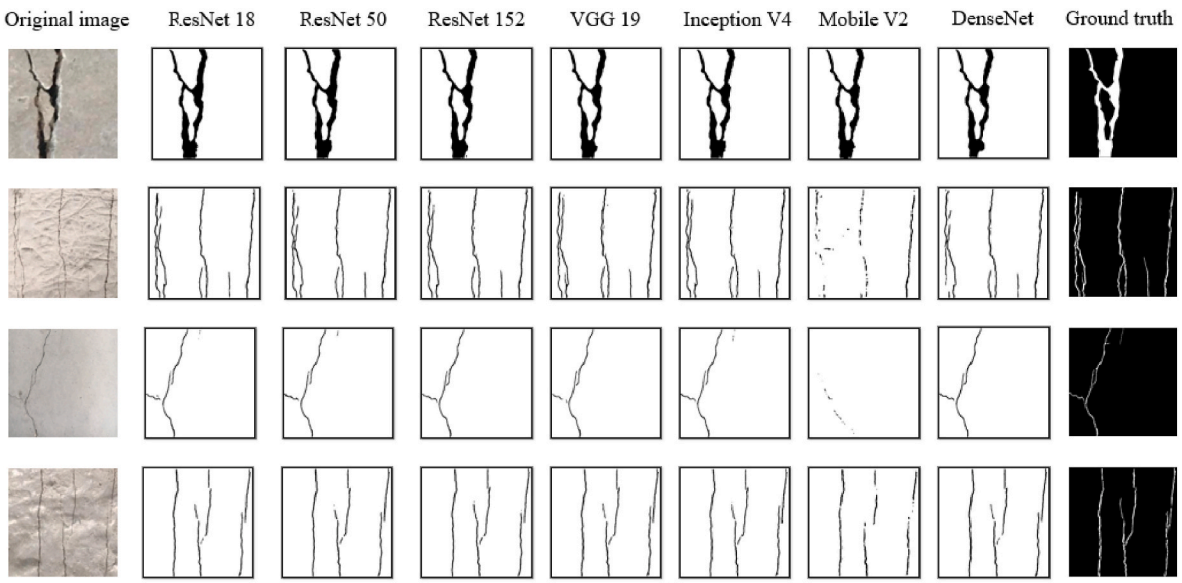


Fig. 13. Representative results of crack segmentation using the different encoder structures.

Table 7
Results of camera calibration.

	Left camera	Right camera
Intrinsic parameter	$I_L = \begin{pmatrix} 703.97 & 0 & 647.29 \\ 0 & 703.16 & 339.35 \\ 0 & 0 & 1 \end{pmatrix}$	$I_R = \begin{pmatrix} 705.74 & 0 & 706.02 \\ 0 & 704.64 & 343.83 \\ 0 & 0 & 1 \end{pmatrix}$
Distortion vector	D_L = (0.060, 0.016, 0.002, −0.002, −1.017)	D_R = (0.053, −0.097, −0.003, 0.002, 0.095)
Rotation matrix	$R = \begin{pmatrix} 0.999 & 0.002 & -0.008 \\ -0.002 & 1.000 & 0.004 \\ 0.009 & -0.004 & 1.000 \end{pmatrix}$	
Translation vector	T = (−60.895, 0.081, −1.423)	

encoders. Compared with the ground truth, ResNet50 and Inception v4 provided reasonable segmentation results, but Inception v4 had longer running time. Therefore, ResNet50 was selected as the decoder for crack segmentation in this research.

4.3. Distance measurement

4.3.1. Camera calibration for binocular camera

Camera calibration is performed to determine the intrinsic matrices, external parameters (rotation matrix and translation vector), and distortion coefficients of the binocular camara. The calibration results are listed in Table 7.

The focal length and baseline used for calculation of distance are 703.97 (from I_L) and 60.895 (from T), respectively. The distortion vector consists of three radial distortion coefficients and two tangential distortion coefficients. The rotation matrix represents the rotation of the right camera relative to the left camera. Translation vector is the translation relationship of the right camera relative to the left camera.

4.3.2. Distortion and stereo correction of binocular images

The images obtained from the left and the right cameras are shown in Fig. 14(a) and **14(b)**, respectively. The rectified images for the left and the right cameras are shown in Fig. 14(c) and **(d)**, respectively. The images after distortion and stereo correction are coplanar. The points of the left and the right images are located on the same line, as shown in Fig. 14(e).

4.3.3. Disparity map and denoising

Stereo matching and disparity calculation are performed and generate a gray scale disparity map, as shown in Fig. 15(a). To identify disparity, the gray scale image is converted into a RGB colored map, as shown in Fig. 15(b). To note, the disparity calculation is more accurate for objects near the center of an image. The mean filter is applied to reduce the noise in the disparity map. After denoising, the voids of both gray and colored images are reduced, as shown in Fig. 15(c) and (d).

4.3.4. Calculation of distance

Table 8 lists some representative results of distance measured using the binocular stereo vision method in comparison with the distance measured using a ruler and denoted as actual distance. The statistics of all results show that the R^2 value is 0.99, the largest error is 1.1%, and the total time is up to 0.24 s, indicating that the ranging method can provide reasonable accuracy and desired efficiency.

4.4. Crack quantification

With the segmented images and distance, the cracks can be quantified. Fig. 16 shows a representative result of crack quantification. It can be seen that the presented method can provide reasonable detection and quantification of distributed cracks with a realistic crack pattern in civil engineering structures. The cracks have varying orientation and width, and one crack may have multiple branches along the crack. In the segmentation image, 10 points at the detected cracks are arbitrarily selected and used to evaluate the crack quantification accuracy. The crack widths at the 10 points were also measured using the crack scope.

Table 9 compares the results of crack width measured from the proposed method and crack scope. The largest absolute discrepancy is generated at point 2, which is 365 μm and 9% of the measurement from the crack scope. The second largest absolute discrepancy is generated at point 3, which is 137 μm and 8% of the measurement from the crack scope.

While the relatively large errors of the wide cracks are related to many factors, a main cause is related to the complex crack pattern of wide cracks in the crack depth direction, as illustrated in Fig. 17. When the crack is vertical to the surface, the edges of the crack can be clearly identified. However, when the crack is inclined, because the wide opening width allows more light in the crack, some light can be reflected by the material inside the crack, which makes it difficult to clearly identify the edges of the crack. Such effect can be aggravated when the

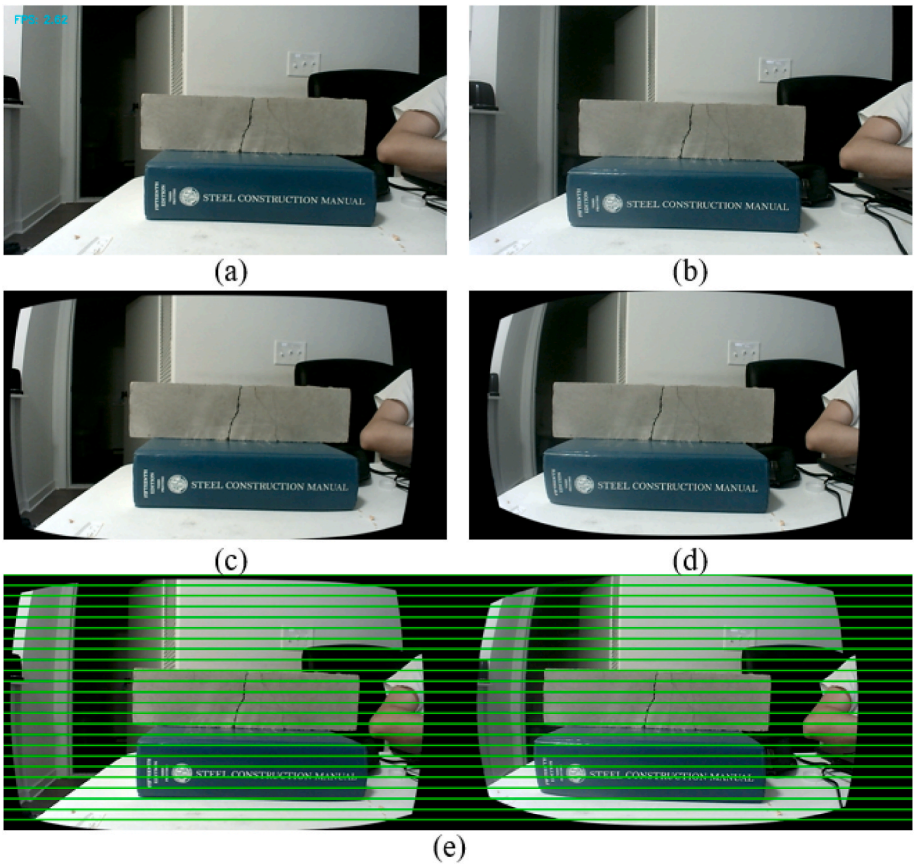


Fig. 14. Distortion and stereo correction: (a) original left image, (b) original right image, (c) rectified left image, (d) rectified right image, and (e) correlation of the left and right images.

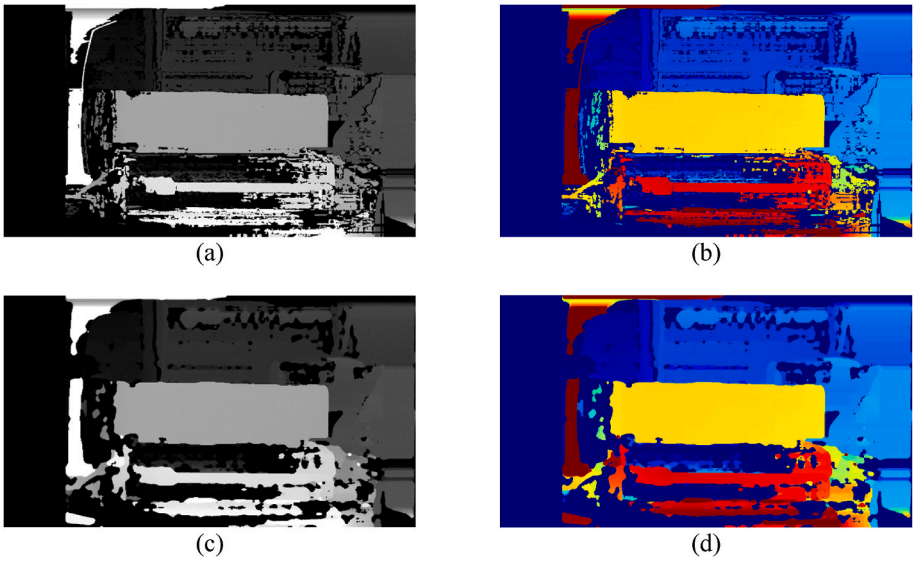


Fig. 15. Disparity maps from stereo matching and disparity calculation: (a) gray scale disparity map, (b) colored disparity map, (c) gray scale disparity map after denoising, and (d) colored disparity map after denoising.

crack is near the edge of the image. Therefore, point 10 also have a relatively high discrepancy for the crack width, and the relative error of point 10 is the largest (10%). The statistics of more results show that the R^2 value of the comparison of the crack width results is 0.98. It was found that the execution of the crack quantification codes took about 0.32 s, and the whole process from photo acquisition to crack quantification was completed within 0.4 s.

4.5. Characterization of crack pattern

Based on the capabilities of crack detection and quantification, the proposed approach was employed to characterize the crack patterns of

Table 8
Result of distance measurement.

Range (mm)	Actual distance (mm)	Measured distance (mm)	Error (mm)	Error (%)
0–200	–	–	–	–
200–300	240	242.1	2.1	0.9
300–400	319	321.3	2.3	0.7
400–500	406	403.3	2.7	0.7
500-1000	980	973.1	6.9	0.7
1000–3000	2450	2422.6	27.4	1.1

SHCC specimens loaded under direct tension. The test setup is shown in Fig. 18(a). A dog-bone specimen made of SHCC was loaded using a load frame. The test was performed under displacement control with a constant displacement rate at 0.05 mm/min. The applied load was measured by a load cell embedded in the load frame, and the elongation of the specimen was measured by an extensometer deployed over the gauge length of the specimen designed in accordance with reference [66]. Fig. 18(b) shows a set of test results of the stress-strain curves. After the specimens were cracked, they carried higher tensile loads and exhibited dense microcracks, as shown in Fig. 18(c). During the tests, it was found that the crack patterns changed with the increase of the elongation of the specimens. Three points denoted by A, B, and C are used to represent three different damage levels. Point A represents minor damage, which occurred shortly after the specimen was cracked; point B represents moderate damage; point C represents severe damage; and point D represents failure. At each level, the photo of a cracked specimen can be processed by the proposed approach. The cracks are visualized and quantified, as shown in Fig. 18(d). The colored legend represents the crack width.

In Fig. 18(d), the width of the specimen is divided into 20 equal strips by 19 lines in red color. Along each line, the crack number and crack width are evaluated. Based on statistics of the results from the 19 lines, the crack pattern of the specimen at the specific load level can be characterized, as shown in Fig. 19(a). Each column represents the average number of cracks with a specific range of crack widths. The widths of most cracks are in the range of 61 μm–100 μm when the specimen fails with localized cracks. Overall, the distribution of cracks satisfies the normal distribution shown by the grey dash line. With the same method, the crack pattern of a specimen can be characterized at an arbitrary load level in real time. Fig. 19(b) illustrates the evolution of the crack pattern. As the load level increased from point A to point B, the crack number and crack width increased. As the load level increased from point B to point C, the crack number increased while the crack width was relatively sustained. As the load level increased from point C to point D, the crack number slightly increased, and the crack width was significantly increased. Characterization of crack patterns is relevant to the mechanical properties and the durability of SHCC which are closely related to the crack number and width [67]. It is desired to have narrow cracks because presence of microcracks do not highly compromises the mechanical properties nor the durability when the crack width is lower than 50 μm [66]. More images are shown in Appendix.

5. Conclusions and prospects

This research proposes an intelligent interpretation approach for automatic assessment and characterization of unique crack patterns in SHCC in real time. The proposed approach integrates stereo vision and deep leaning for automatic detection, ranging, localization, quantification, and visualization of cracks. The approach is implemented to monitor and characterize the evolution of crack patterns for SHCC. Based on the investigations, the following conclusions are drawn:

- The proposed approach provides detailed information about the characteristics of unique cracks in SHCC throughout the loading process. The detailed information includes crack number and crack width that reflect the damage condition of the SHCC specimen, despite the complex crack patterns different from conventional concrete. The evolution of the crack patterns can be monitored to assess the condition of SHCC. The total assessment time is less than 0.4 s per photo.
- The YOLOv5 models show desired accuracy and efficiency in crack detection. The mAP@0.5 score is 0.89 for YOLO5s, 0.91 YOLO5m, 0.91 for YOLO5l, and 0.92 for YOLO5x. The time needed to detect cracks in one image is 0.008 s for YOLO5s, 0.010 s for YOLO5m, 0.013 s for YOLO5l, and 0.016 s for YOLO5x. YOLOv5 is promising for real-time automatic crack detection. In the investigated encoders and decoders, ResNet50 and LinkNet demonstrated the highest segmentation accuracy and high efficiency. With combination of

Table 9
Results of crack width quantification.

Point	Proposed method (μm)	Crack scope (μm)	Discrepancy (μm)	Error (%)
1	936	970	−34	−4%
2	4465	4100	365	9%
3	1513	1650	−137	−8%
4	490	520	−30	−6%
5	267	280	−13	−5%
6	444	410	34	8%
7	222	210	12	6%
8	271	250	21	8%
9	333	350	−17	−5%
10	913	830	83	10%

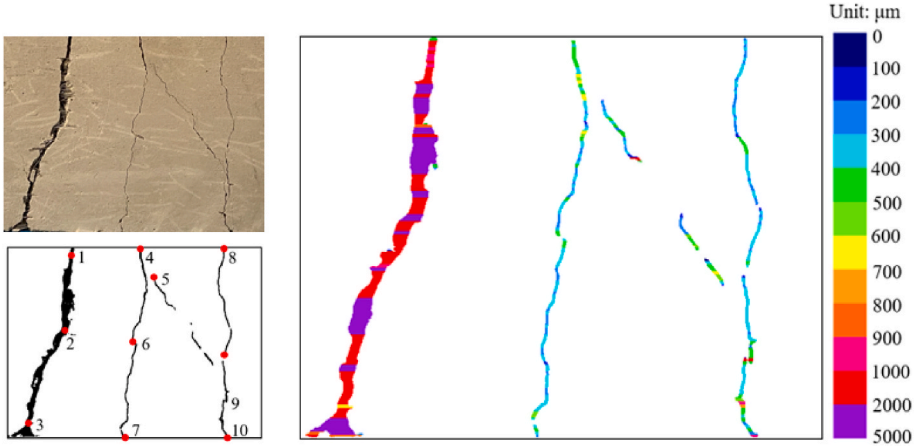


Fig. 16. Visualization of distributed cracks detected and quantified using the presented method.

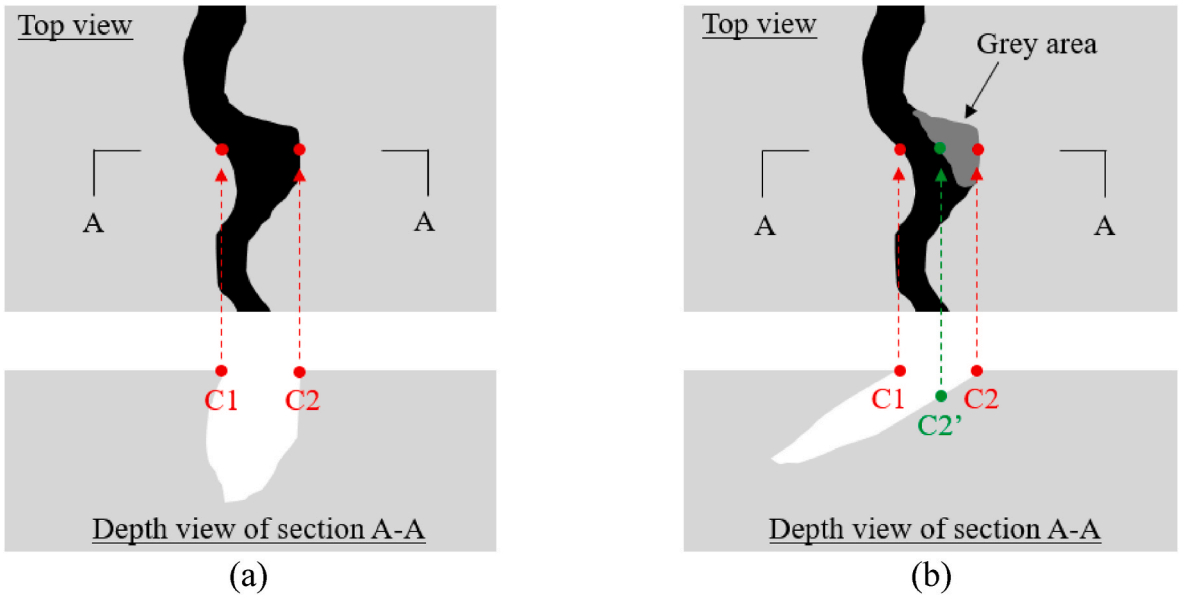


Fig. 17. Effect of crack pattern in the depth direction on the accuracy of detection of cracks: (a) a vertical crack; and (b) an inclined crack.

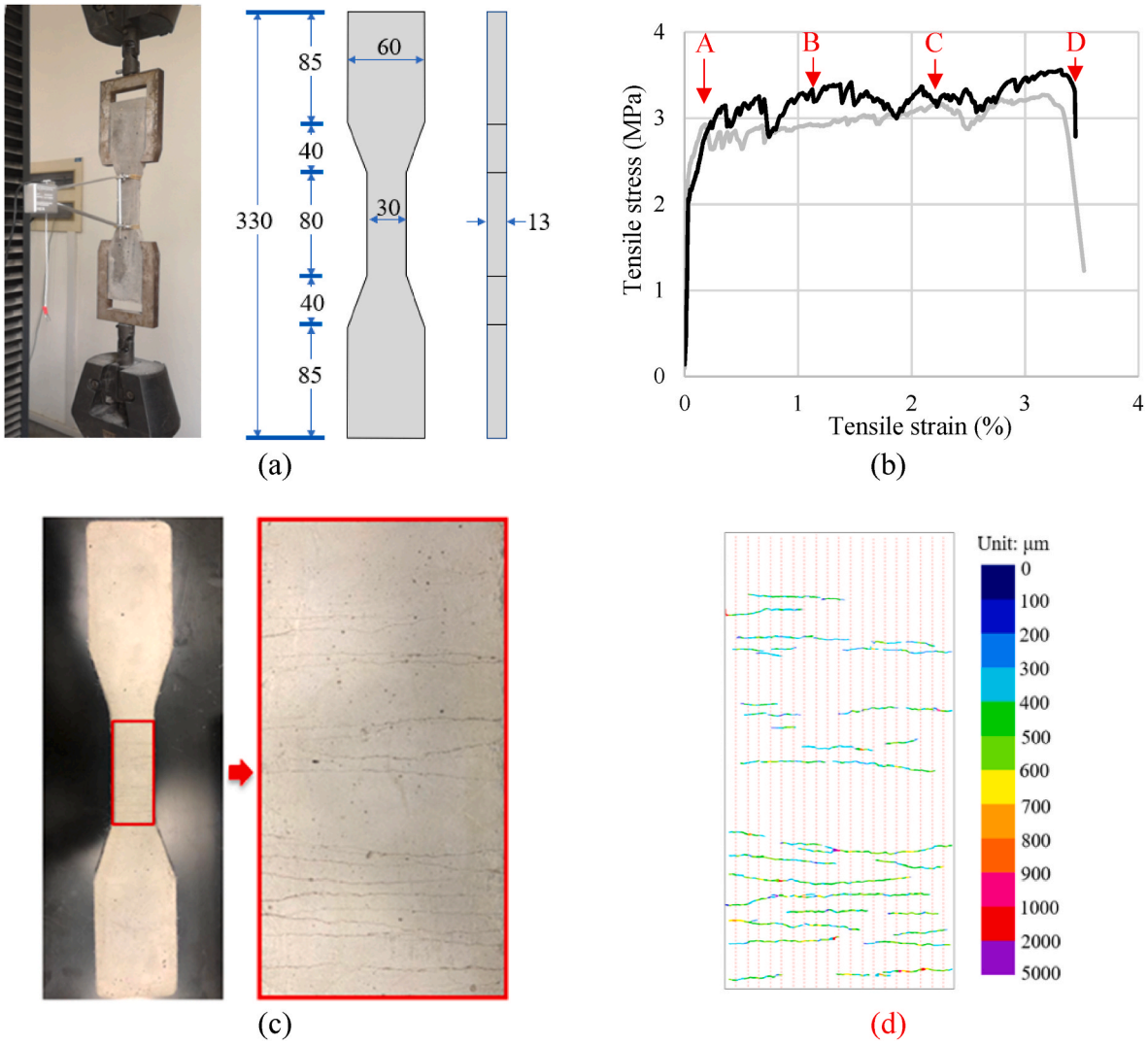


Fig. 18. Crack evaluation of SHCC specimens: (a) specimen and test setup (unit: mm), (b) stress-strain curves, (c) photos of a cracked specimen, and (d) crack visualization and quantification.

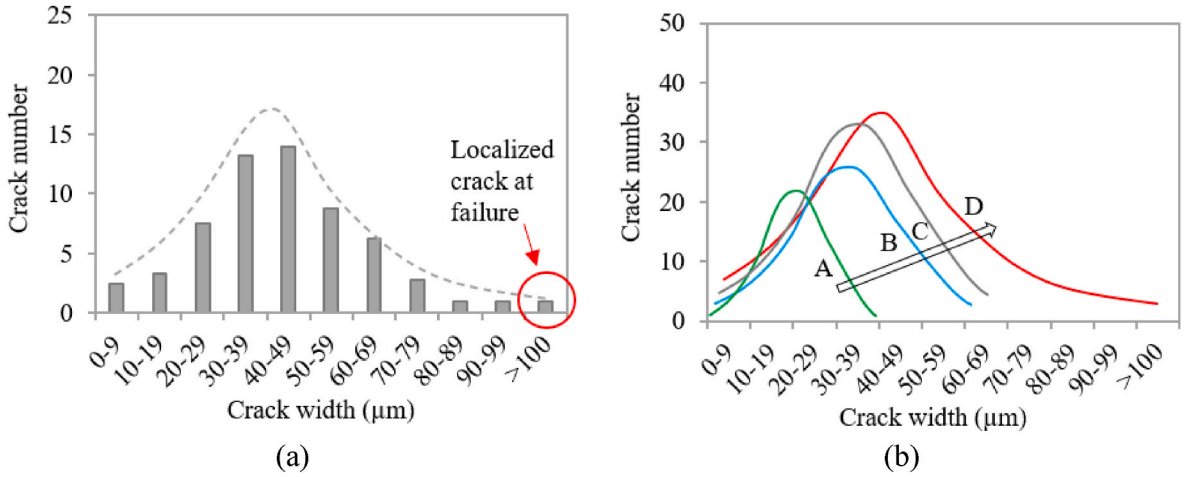


Fig. 19. Statistics of crack patterns of a SHCC specimen: (a) crack number versus crack width; and (b) evolution of the crack pattern in the loading process.

ResNet50 and LinkNet, the IOU and F1 score are 0.987 and 0.993, respectively, and the assessment time is 0.036 s per photo.

- The binocular stereo vision approach provides high accuracy and efficiency for ranging cracks. The proposed processes integrate the camera calibration, distortion and stereo correction, stereo matching and disparity calculation, and distance calculation. Regarding the ranging accuracy, the R^2 value is higher than 0.99, and relative error is less than 0.9%. With segmented images and determined distance, the widths of cracks can be quantified according to the pixels of cracks. The assessment time is 0.32 s per photo.

Based on the established capabilities of crack assessment, there are opportunities to perform further research in the following aspects:

- The proposed approach has been deployed on a land robotic car to automatically perform condition assessment of structures such as bridge decks and girders. It is interesting to incorporate the approach into drones with stronger accessing capabilities. It is envisioned that utilization of the developed approach will help save cost and minimize human intervention in crack assessment. Further efforts are needed to evaluate the performance of the proposed method in different use cases.
- It is interesting to further improve the efficiency by reducing the total running time from 0.4 s to millisecond level for real-time condition assessment of engineering structures. It is speculated that multi-threaded parallel and distributed computing will further reduce the computation time.
- This research shows that the proposed method provides desired performance in assessing cracks in different types of concrete. It is interesting to evaluate the performance of the proposed method for assessing cracks in other materials such as steel and masonry blocks with different textures and crack patterns. It is important to test the proposed method in scenarios involving different brightness of photos.

Author statement

Pengwei Guo: Investigation; data collection; methodology; software; coding; writing—
Original draft preparation.
Xiangjun Meng: Investigation; data collection; methodology; coding.
Weina Meng: Conceptualization; supervision; resources; funding acquisition; writing—review and editing.
Yi Bao: Conceptualization; supervision; visualization; funding acquisition; writing—review and editing; project administration.
All authors have read and agreed to the published version of the manuscript.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgement

This research was funded by National Science Foundation of the United States [grant number CMMI-2046407].

Appendix

Fig. A1 shows the change of cracks in a SHCC dog-bone specimen with the load time. These images were extracted from the different frames of a video produced using the proposed approach. The change of cracks in the SHCC reflects the fracturing process, as elaborated in Section 4.5.

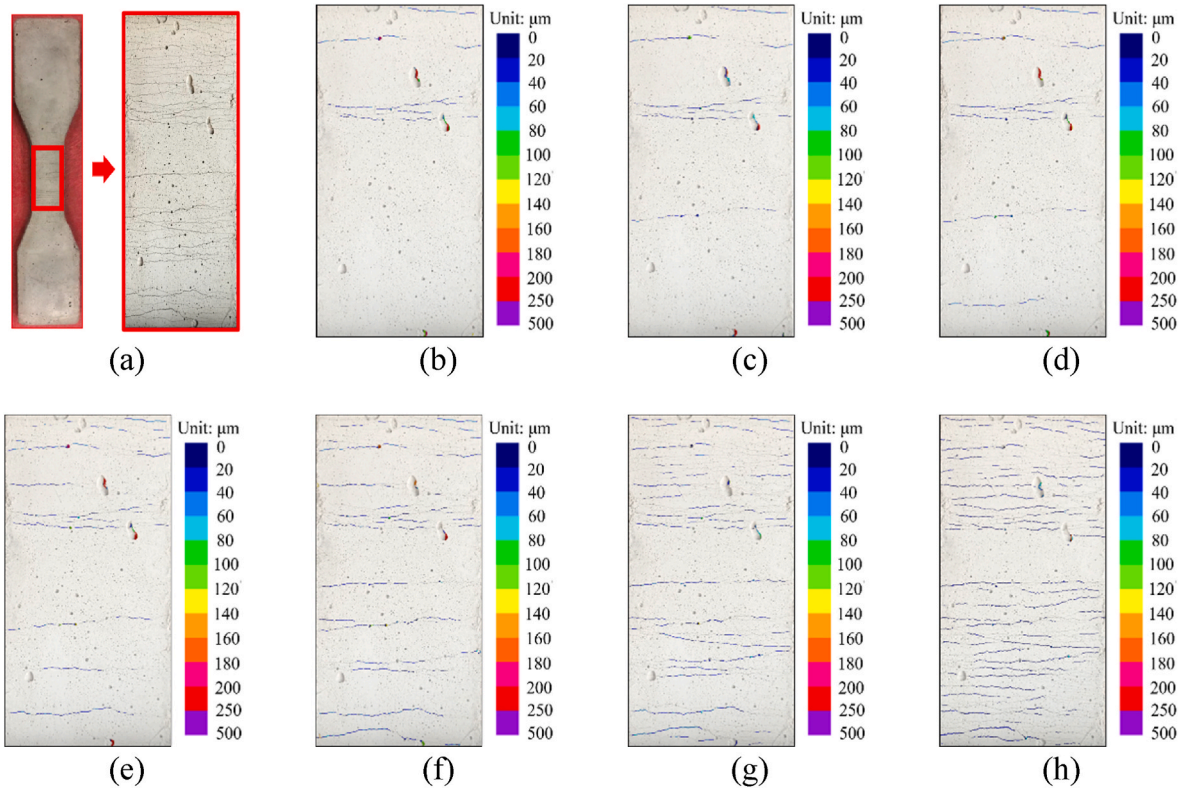


Fig. A1. Monitoring of the change of cracks in a SHCC specimen under direct tension: (a) the tested specimen; (b) 1 min; (c) 2 min; (d) 3 min; (e) 4 min; (f) 5 min; (g) 6 min; and (h) 7 min.

References

[1] Report ASCE. Card for America's Infrastructure. cited 2021 4/10]; Available from: http://www.infrastructurereportcard.org/; 2017.

[2] ASCE. Failure to act: closing the infrastructure investment gap for America's economic future [cited 2021 4/10]; Available from: https://www.asce.org/failuretoact/; 2016.

[3] Faber MH, Sorensen JD. Indicators for inspection and maintenance planning of concrete structures. Structural Safety 2002;24(2–4):377–96. https://doi.org/10.1016/S0167-4730(02)00033-4.

[4] Bao Y, Tang F, Chen Y, Meng W, Huang Y, Chen G. Concrete pavement monitoring with PPP-BOTDA distributed strain and crack sensors. Smart Structures and Systems 2016;18(3):405–23. https://doi.org/10.12989/sss.2016.18.3.405.

[5] Tan X, Bao Y. Measuring crack width using a distributed fiber optic sensor based on optical frequency domain reflectometry. Measurement 2021;172:108945. https://doi.org/10.1016/j.measurement.2020.108945.

[6] Kim B, Cho S. Image-based concrete crack assessment using mask and region-based convolutional neural network. Structural Control and Health Monitoring 2019;26(8):e2381. https://doi.org/10.1002/stc.2381.

[7] Ye C, Butler LJ, Elshafie MZ, Middleton CR. Evaluating prestress losses in a prestressed concrete girder railway bridge using distributed and discrete fibre optic sensors. Construction and Building Materials 2020;247:118518. https://doi.org/10.1016/j.conbuildmat.2020.118518.

[8] Yao Y, Yan M, Bao Y. Measurement of cable forces for automated monitoring of engineering structures using fiber optic sensors: a review. Automation in Construction 2021;126:103687. https://doi.org/10.1016/j.autcon.2021.103687.

[9] Liu Z, Cao Y, Wang Y, Wang W. Computer vision-based concrete crack detection using U-net fully convolutional networks. Automation in Construction 2019;104:129–39. https://doi.org/10.1016/j.autcon.2019.04.005.

[10] Li C, Luo B, Hong H, Su X, Wang Y, Liu J, Wang C, Zhang J, Wei L. Object detection based on global-local saliency constraint in aerial images. Remote Sensing 2020;12(9):1435. https://doi.org/10.3390/rs12091435.

[11] Deng J, Lu Y, Lee VCS. Concrete crack detection with handwriting script interferences using faster region-based convolutional neural network. Computer-Aided Civil and Infrastructure Engineering 2020;35(4):373–88. https://doi.org/10.1111/mice.12497.

[12] Ren S, He K, Girshick R, Sun J. Faster r-cnn: towards real-time object detection with region proposal networks. IEEE Transactions on Pattern Analysis and Machine Intelligence; 2017. p. 1137–49. https://doi.org/10.1109/TPAMI.2016.2577031.

[13] He K, Gkioxari G, Dollár P, Girshick R. Mask r-cnn. In: Proceedings of the IEEE international conference on computer vision; 2017. p. 2961–9. https://doi.org/10.1109/ICCV.2017.322.

[14] Cai Z, Vasconcelos N. Cascade r-cnn: delving into high quality object detection. In: Proceedings of the IEEE conference on computer vision and pattern recognition; 2018. p. 6154–62. https://doi.org/10.1109/CVPR.2018.00644.

[15] Cha Y-J, Choi W. Vision-based concrete crack detection using a convolutional neural network. Dynamics of Civil Structures; 2017. p. 71–3. https://doi.org/10.1007/978-3-319-54777-0_9.

[16] Li Y, Han Z, Xu H, Liu L, Li X, Zhang K. YOLOv3-lite: a lightweight crack detection network for aircraft structure based on depthwise separable convolutions. Applied Sciences 2019;9(18):3781. https://doi.org/10.3390/app9183781.

[17] Huang R, Pedoeem J, Chen C. YOLO-LITE: a real-time object detection algorithm optimized for non-GPU computers. In: 2018 IEEE international conference on big data; 2018. p. 2503–10. https://doi.org/10.1109/BigData.2018.8621865.

[18] Liu W, Anguelov D, Erhan D, Szegedy C, Reed S, Fu C-Y, Berg AC. Ssd: single shot multibox detector. In: European conference on computer vision; 2016. p. 21–37. https://doi.org/10.1007/978-3-319-46448-0_2.

[19] Li Y, Lu Y, Chen J. A deep learning approach for real-time rebar counting on the construction site based on YOLOv3 detector. Automation in Construction 2021;124:103602. https://doi.org/10.1016/j.autcon.2021.103602.

[20] Ultralytics YOLOv5. Available from: https://github.com/ultralytics/yolov5; 2021.

[21] Kalfarisi R, Wu ZY, Soh K. Crack detection and segmentation using deep learning with 3D reality mesh model for quantitative assessment and integrated visualization. Journal of Computing in Civil Engineering 2020;34(3):04020010. https://doi.org/10.1061/(ASCE)CP.1943-5487.0000890.

[22] Piyathilaka L, Preethichandra D, Izhar U, Kahandawa G. Real-time concrete crack detection and instance segmentation using deep transfer learning. Engineering Proceedings 2020;2(1):91. https://doi.org/10.3390/ecsa-7-08260.

[23] Sun M. Semantic segmentation using modified U-net architecture for crack detection. Available from: https://openprairie.sdstate.edu/etd/3937/; 2020.

[24] Dung CV. Autonomous concrete crack detection using deep fully convolutional neural network. Automation in Construction 2019;99:52–8. https://doi.org/10.1016/j.autcon.2018.11.028.

[25] Chen T, Cai Z, Zhao X, Chen C, Liang X, Zou T, Wang P. Pavement crack detection and recognition using the architecture of SegNet. Journal of Industrial Information Integration 2020:100144. https://doi.org/10.1016/j.jii.2020.100144.

[26] Yakubovskiy P. Segmentation models. GitHub repository; 2019. Available from: https://github.com/qubvel/segmentation_models.

[27] Yang X, Li H, Yu Y, Luo X, Huang T, Yang X. Automatic pixel-level crack detection and measurement using fully convolutional network. Computer-Aided Civil and Infrastructure Engineering 2018;33(12):1090–109. https://doi.org/10.1111/mice.12412.

[28] Jung HG, Cho YH, Yoon PJ, Kim J. Scanning laser radar-based target position designation for parking aid system. IEEE Transactions on Intelligent Transportation Systems 2008;9(3):406–24. https://doi.org/10.1109/TITS.2008.922980.
[29] Liu Y, Wang H, Dong C, Chen Q. A car-following data collecting method based on binocular stereo vision. IEEE Access 2020;8:25350–63. https://doi.org/10.1109/access.2020.2965833.
[30] Guo P, Meng W, Bao Y. Automatic identification and quantification of dense microcracks in high-performance fiber-reinforced cementitious composites through deep learning-based computer vision. Cement and Concrete Research 2021;148:106532. https://doi.org/10.1016/j.cemconres.2021.106532.
[31] Zhang A, Wang KC, Li B, Yang E, Dai X, Peng Y, Fei Y, Liu Y, Li JQ, Chen C. Automated pixel-level pavement crack detection on 3D asphalt surfaces using a deep-learning network. Computer-Aided Civil and Infrastructure Engineering 2017;32(10):805–19. https://doi.org/10.1111/mice.12297.
[32] Pan Y, Zhang G, Zhang L. A spatial-channel hierarchical deep learning network for pixel-level automated crack detection. Automation in Construction 2020;119:103357. https://doi.org/10.1016/j.autcon.2020.103357.
[33] Ren Y, Huang J, Hong Z, Lu W, Yin J, Zou L, Shen X. Image-based concrete crack detection in tunnels using deep fully convolutional networks. Construction and Building Materials 2020;234:117367. https://doi.org/10.1016/j.conbuildmat.2019.117367.
[34] Fang F, Li L, Gu Y, Zhu H, Lim J-H. A novel hybrid approach for crack detection. Pattern Recognition 2020;107:107474. https://doi.org/10.1016/j.patcog.2020.107474.
[35] Liu J, Yang X, Lau S, Wang X, Luo S, Lee VCS, Ding L. Automated pavement crack detection and segmentation based on two-step convolutional neural network. Computer-Aided Civil and Infrastructure Engineering 2020;35(11):1291–305. https://doi.org/10.1111/mice.12622.
[36] Kang D, Benipal SS, Gopal DL, Cha Y-J. Hybrid pixel-level concrete crack segmentation and quantification across complex backgrounds using deep learning. Automation in Construction 2020;118:103291. https://doi.org/10.1016/j.autcon.2020.103291.
[37] Ji A, Xue X, Wang Y, Luo X, Xue W. An integrated approach to automatic pixel-level crack detection and quantification of asphalt pavement. Automation in Construction 2020;114:103176. https://doi.org/10.1016/j.autcon.2020.103176.
[38] Xie C, Cao M, Guan J, Liu Z, Khan M. Improvement of boundary effect model in multi-scale hybrid fibers reinforced cementitious composite and prediction of its structural failure behavior. Composites Part B: Engineering 2021;224:109219. https://doi.org/10.1016/j.compositesb.2021.109219.
[39] Zhang D, Yu J, Wu H, Jaworska B, Ellis BR, Li VC. Discontinuous micro-fibers as intrinsic reinforcement for ductile Engineered Cementitious Composites (ECC). Composites Part B: Engineering 2020;184:107741. https://doi.org/10.1016/j.compositesb.2020.107741.
[40] Zhang W, Zheng Q, Ashour A, Han B. Self-healing cement concrete composites for resilient infrastructures: a review. Composites Part B: Engineering 2020;189:107892. https://doi.org/10.1016/j.compositesb.2020.107892.
[41] Rezaie AB, Liebscher M, Ranjbarian M, Simon F, Zimmerer C, Drechsler A, Frenzel R, Synytska A, Mechtcherine V. Enhancing the interfacial bonding between PE fibers and cementitious matrices through polydopamine surface modification. Composites Part B: Engineering 2021;217:108817. https://doi.org/10.1016/j.compositesb.2021.108817.
[42] Khandelwal S, Rhee KY. Recent advances in basalt-fiber-reinforced composites: tailoring the fiber-matrix interface. Composites Part B: Engineering 2020;192:108011. https://doi.org/10.1016/j.compositesb.2020.108011.
[43] Du J, Meng W, Khayat KH, Bao Y, Guo P, Lyu Z, Abu-obeidah A, Nassif H, Wang H. New development of ultra-high-performance concrete (UHPC). Composites Part B: Engineering 2021;224:109220. https://doi.org/10.1016/j.compositesb.2021.109220.
[44] Akbar A, Liew K. Influence of elevated temperature on the microstructure and mechanical performance of cement composites reinforced with recycled carbon fibers. Composites Part B: Engineering 2020;198:108245. https://doi.org/10.1016/j.compositesb.2020.108245.
[45] Mazur K, Jakubowska P, Romańska P, Kuciel S. Green high density polyethylene (HDPE) reinforced with basalt fiber and agricultural fillers for technical applications. Composites Part B: Engineering 2020;202:108399. https://doi.org/10.1016/j.compositesb.2020.108399.
[46] Li M, Pu Y, Thomas VM, Yoo CG, Ozcan S, Deng Y, Nelson K, Ragauskas A. Recent advancements of plant-based natural fiber–reinforced composites and their applications. Composites Part B: Engineering 2020;200:108254. https://doi.org/10.1016/j.compositesb.2020.108254.
[47] Akbar A, Liew K. Multicriteria performance evaluation of fiber-reinforced cement composites: an environmental perspective. Composites Part B: Engineering 2021;218:108937. https://doi.org/10.1016/j.compositesb.2021.108937.
[48] Meister S, Wermes M, Stüve J, Groves RM. Investigations on Explainable Artificial Intelligence methods for the deep learning classification of fibre layup defect in the automated composite manufacturing. Composites Part B: Engineering 2021;224:109160. https://doi.org/10.1016/j.compositesb.2021.109160.
[49] *YOLO v5 structure*. Available from: https://github.com/ultralytics/yolov5/issues/280.
[50] Elfwing S, Uchibe E, Doya K. Sigmoid-weighted linear units for neural network function approximation in reinforcement learning. Neural Networks 2018;107:3–11. https://doi.org/10.1016/j.neunet.2017.12.012.
[51] Ioffe S, Szegedy C. Batch normalization: accelerating deep network training by reducing internal covariate shift. International Conference on Machine Learning 2015:448–56. Available from: http://proceedings.mlr.press/v37/ioffe15.html.
[52] Wang C-Y, Liao H-YM, Wu Y-H, Chen P-Y, Hsieh J-W, Yeh I-H. CSPNet: a new backbone that can enhance learning capability of CNN. Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition 2020:390–1. https://doi.org/10.1109/CVPRW50498.2020.00203.
[53] He K, Zhang X, Ren S, Sun J. Spatial pyramid pooling in deep convolutional networks for visual recognition. IEEE Transactions on Pattern Analysis and Machine Intelligence 2015;37(9):1904–16. https://doi.org/10.1109/TPAMI.2015.2389824.
[54] He K, Zhang X, Ren S, Sun J. Deep residual learning for image recognition. In: Proceedings of the IEEE conference on computer vision and pattern recognition; 2016. p. 770–8. https://doi.org/10.1109/CVPR.2016.90.
[55] Talo M. Automated classification of histopathology images using transfer learning. Artificial Intelligence in Medicine 2019;101:101743. https://doi.org/10.1016/j.artmed.2019.101743.
[56] Sandler M, Howard A, Zhu M, Zhmoginov A, Chen L-C. Mobilenetv2: inverted residuals and linear bottlenecks. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition 2018:4510–20. https://doi.org/10.1109/CVPR.2018.00474.
[57] Ronneberger O, Fischer P, Brox T. U-net: convolutional networks for biomedical image segmentation. In: International conference on medical image computing and computer-assisted intervention; 2015. p. 234–41. https://doi.org/10.1007/978-3-319-24574-4_28.
[58] Zhao H, Shi J, Qi X, Wang X, Jia J. Pyramid scene parsing network. In: Proceedings of the IEEE conference on computer vision and pattern recognition; 2017. p. 2881–90. https://doi.org/10.1109/CVPR.2017.660.
[59] Chen L-C, Zhu Y, Papandreou G, Schroff F, Adam H. Encoder-decoder with atrous separable convolution for semantic image segmentation. In: European conference on computer vision; 2018. p. 801–18. https://doi.org/10.1007/978-3-030-01234-2_49.
[60] Chaurasia A, Culurciello E. Linknet: exploiting encoder representations for efficient semantic segmentation. In: 2017 IEEE visual communications and image processing; 2017. p. 1–4. https://doi.org/10.1109/VCIP.2017.8305148.
[61] Li H, Xiong P, An J, Wang L. Pyramid attention network for semantic segmentation. arXiv preprint arXiv:180510180 2018. https://arxiv.org/abs/1805.10180.
[62] MATLAB. Stereo camera calibration. Available from: https://www.mathworks.com/discovery/stereo-vision.html.
[63] Szeliski R. Computer vision: algorithms and applications. Available from: https://szeliski.org/Book/; 2010.
[64] Depth Map from Stereo Images. Available from: https://docs.opencv.org/3.1.0/dd/d53/tutorial_py_depthmap.html.
[65] Hough Circle Transform. Available from: https://docs.opencv.org/3.4/da/d53/tutorial_py_houghcircles.html.
[66] Li X, Lv X, Zhou X, Meng W, Bao Y. Upcycling of waste concrete in eco-friendly strain-hardening cementitious composites: mixture design, structural performance, and life-cycle assessment. Journal of Cleaner Production 2022;330:129911. https://doi.org/10.1016/j.jclepro.2021.129911.
[67] Mahjoubi S, Barhemat R, Guo P, Meng W, Bao Y. Prediction and multi-objective optimization of mechanical, economical, and environmental properties for strain-hardening cementitious composites (SHCC) based on automated machine learning and metaheuristic algorithms. Journal of Cleaner Production 2021;329:129665. https://doi.org/10.1016/j.jclepro.2021.129665.