

Crack and Noncrack Classification from Concrete Surface Images Using Machine Learning

Hyunjun Kim, Eunjong Ahn, Myoungsu Shin and Sung-Han Sim^{ID}

Abstract

In concrete structures, surface cracks are important indicators of structural durability and serviceability. Generally, concrete cracks are visually monitored by inspectors who record crack information such as the existence, location, and width. Manual visual inspection is often considered ineffective in terms of cost, safety, assessment accuracy, and reliability. Digital image processing has been introduced to more accurately obtain crack information from images. A critical challenge is to automatically identify cracks from an image containing actual cracks and crack-like noise patterns (e.g. dark shadows, stains, lumps, and holes), which are often seen in concrete structures. This article presents a methodology for identifying concrete cracks using machine learning. The method helps in determining the existence and location of cracks from surface images. The proposed approach is particularly designed for classifying cracks and noncrack noise patterns that are otherwise difficult to distinguish using existing image processing algorithms. In the training stage of the proposed approach, image binarization is used to extract crack candidate regions; subsequently, classification models are constructed based on speeded-up robust features and convolutional neural network. The obtained crack identification methods are quantitatively and qualitatively compared using new concrete surface images containing cracks and noncracks.

Keywords

Concrete crack identification, convolutional neural network, digital image processing, machine learning, speeded-up robust features

Introduction

Cracks in concrete structures are primary indicators of possible structural damage and durability.^{1–9} Most of the developed countries conduct regular crack assessment of civil engineering structures as part of infrastructure maintenance. Manual visual inspection is the most commonly employed method in practice for obtaining crack information such as the existence, location, and width, which can be used to prepare maintenance plans. Although crack information can be obtained from a manual visual inspection, it is labor-intensive, costly, time-consuming, and often unreliable because the results depend on the experience and skill of the inspector.

To overcome the drawbacks of manual visual inspection, digital image processing has been introduced as a promising alternative for crack monitoring. Generally, the surface images of concrete structures are used for image processing, from which crack information such

as existence, location, and width is determined. Widely used image processing algorithms for crack identification are based on image binarization, edge detection, and mathematical morphology. Image binarization, which helps convert the pixels in a grayscale image to either black or white, can be used for crack detection, because dark cracks are generally categorized as black whereas relatively lighter backgrounds appear white in the binarized image.^{10–12} In edge detection, concrete cracks are detected by localizing the borders of the crack pixels.^{13,14} Mathematical morphology is used as

School of Urban and Environmental Engineering, Ulsan National Institute of Science and Technology (UNIST), Ulsan, Republic of Korea

Corresponding author:

Sung-Han Sim, School of Urban and Environmental Engineering, Ulsan National Institute of Science and Technology (UNIST), 50 UNIST-gil, Ulju-gu, Ulsan 44919, Republic of Korea.
Email: ssim@unist.ac.kr

an additional process to modify crack shapes and thereby improve the identification performance.^{15,16} Jahanshahi et al.¹⁷ and Koch et al.¹⁸ summarized the image processing methods used for the crack detection of concrete structures.

Although previous studies on the use of image processing for crack identification have shown enormous potential, the underlying common assumption that the given images contain actual cracks critically limits full automation. For instance, the surface images of the entire exterior of a concrete structure captured manually using a digital camera or with the aid of an unmanned aerial vehicle (UAV) taken for structural maintenance may contain cracks and/or noncracks such as dark stains, shades, dust, lumps, and holes, which are difficult to distinguish in the aspect of image processing.^{19,20} Moreover, image binarization may categorize a dark stain as black (i.e. a crack), resulting in a false positive detection. Therefore, the process of distinguishing cracks from surface images containing actual cracks and/or crack-like noncracks is essential for a fully automated crack monitoring.

Machine learning has been recognized as an innovative tool in various civil engineering applications. In particular, supervised learning, which is a type of machine learning, can be used to resolve crack recognition problems in conjunction with computer vision. This combined approach typically involves identifying the unique characteristics of cracks and noncracks from training images, which are used in classification methods such as support vector machines (SVMs)²¹ and random forests.²² The trained classification model is subsequently applied to new images in which surface cracks are to be detected. The geometric patterns (e.g. eccentricity and number of pixels in each pixel group) and statistical properties of pixel intensities (e.g. mean and standard deviation) have been selected as features to distinguish cracks and noncracks and thereby generate a classification model.^{23–26} Although user-defined empirical thresholds are unnecessary in these methods, crack-like noncracks that share similar geometry and colors with cracks still remain undistinguishable. For an effective classification, advanced features need to be extracted from cracks and noncracks to generate a robust classification model.

Modern feature detection algorithms used in the computer vision field can be employed to recognize the salient features of cracks to enable accurate identification.^{27–29} In particular, speeded-up robust features (SURF),²⁹ which is one of the most widely employed local feature detectors, has a proven performance in terms of computational time.³⁰ SURF can be used to efficiently select interest points as features from a similarity-invariant representation; these features can

collectively represent a characteristic descriptor of a specific object. Although SURF has a strong potential for automated crack monitoring, its use for crack identification has not been reported in the literature.

However, deep learning, which is a cascade of multiple layers, has recently been introduced as a powerful method for crack identification.^{31–34} Concrete surface images labeled as either a cracked surface or as an intact surface have been used for training a classification model using convolutional neural network (CNN).³⁵ In the validation stage, the trained classification model is used to test new concrete surface images. Previous studies that employed deep learning have successfully detected cracked regions; however, the classification in the presence of crack-like noncracks, which are unavoidable in real-world applications, was not fully studied. It is important to accurately detect and filter possible noncrack objects in concrete surface images. However, this problem has rarely been discussed in the literature.

This article presents a framework for concrete crack identification using machine learning. The framework can help determine the existence and location of cracks from concrete surface images. The proposed approach is designed to perform accurately, particularly when the images contain noncracks that are difficult to be distinguished from cracks using existing image processing algorithms. The main contribution of this study can be summarized as follows: (1) an efficient classification framework based on a crack candidate region (CCR) is proposed to effectively categorize cracks and noncracks, (2) comparative analysis between SURF-based and CNN-based methods is conducted to evaluate the classification performances, and (3) a comprehensive crack identification in the presence of crack-like noncracks is conducted for practical applications.

Background

To automatically categorize crack and noncrack objects from concrete surface images, two types of classification models are considered in this study: (1) SURF-based classification and (2) CNN-based classification. In general, local features are used in the SURF-based method, whereas global features are extracted in the CNN-based method to obtain the classification model. The overall processes of each method are briefly explained in this section.

SURF-based classification

Csurka et al.³⁶ proposed a bag-of-words (BoW) model for the natural image classification of objects such as

trees, cars, phones, and books. This process consists of three stages: (1) feature extraction, (2) visual vocabulary construction, and (3) classification. Because the crack identification method used in this study is based on the categorization process proposed by Csurka et al.,³⁶ the image processing and machine learning algorithms used in the three stages are briefly discussed.

Feature extraction: SURF. Feature extraction, which is a process of determining the unique characteristics of an image, is a vital part of object identification using image processing. In contrast to Csurka et al.,³⁶ who used scale-invariant feature transform (SIFT)²⁸ for feature extraction, we selected SURF owing to its high performance and computational efficiency. SURF, which is designed to obtain distinctive features from digital images, consists of two main procedures: (1) interest point detection and (2) interest point description. To detect the interest points on elements such as blobs, corners, and edges, the determinant of the Hessian matrix is used as a measure for evaluating the local change around each pixel. After the interest points are obtained, Haar wavelet responses are calculated within a circular neighborhood; an orientation is then assigned to each point using these responses. A square region is subsequently generated along the obtained orientation to address the image rotations. A feature vector with 64 elements is finally computed using the Haar wavelet responses in both the horizontal and vertical directions in 4×4 sub-regions.

Visual vocabulary construction: k-means clustering. The feature vectors of all the interest points are used to generate a visual word that serves as a representative, small image segment to demonstrate features such as color, shape, and surface texture. An image contains various interest points and corresponding feature vectors; therefore, it is necessary to determine the characteristic features of cracks and noncracks to efficiently handle the large volume of images in the training stage. *k*-means clustering,³⁷ which is a popular method for cluster analysis, is introduced to determine the representative clusters, in which the mean values of the feature vectors are the visual words. The results of the *k*-means clustering (i.e. visual words) are then grouped, and this group is called visual vocabulary or the bag of features.

Classification: SVM. To categorize the visual vocabulary through *k*-means clustering, Csurka et al.³⁶ used SVM, which is one of the most common classification algorithms owing to its robustness, computational efficiency, and resistance to over-fitting. When two different sets (i.e. cracks and noncracks) of images are

trained for the classification, a visual vocabulary should be first generated from all the images using *k*-means clustering. Subsequently, the frequency of occurrence of the visual words in the vocabulary is calculated for each category. The obtained feature histograms are then inputted to the SVM to construct the classification model. Among the various SVM classifiers (e.g. linear, quadratic, cubic, and Gaussian), the linear SVM classifier, which is the most widely used, is selected in this work.

CNN-based classification

The CNN is a feed-forward artificial neural network, which has been demonstrated as a powerful tool for image classification. Krizhevsky et al.³⁸ presented AlexNet, by implementing CNN, to classify natural images into 1000 categories. In contrast to the SURF-based classification, the architecture of AlexNet is a hierarchical structure, having five convolutional layers and three fully connected layers. Each convolutional layer handles an input image having different kernels and corresponding sizes. Furthermore, AlexNet is equipped with rectified linear units (ReLUs) and max pooling between the convolutional layers to enhance the classification performance in terms of the computational time and accuracy. After passing through the convolutional layers, the output will go through three fully connected layers with the softmax activation function to identify the class of the image, such as animal, car, fruit, or vegetable. Figure 1 shows the overall process of the CNN-based and SURF-based classifications, modified from the study by Zheng et al.³⁹ Note that the CNN-based method directly uses global features for the classification, whereas the SURF-based method uses visual words clustered from local features.

For training the classification model, a set of surface images needs to be prepared. A typical method of applying CNN is to employ a scanning window, in which the input images are divided into a number of sub-images with a fixed resolution, as shown in Figure 1.^{31,33} The sub-images are manually categorized as either a cracked surface or as an intact surface to build the classification model, which is used to determine the existence and locations of the cracks. Although the CNN shows strong potential, the scanning window was found to be inefficient in that the intact surface, which takes up a majority of an image, has the highest influence in the training. As an alternative to the scanning window, Faster R-CNN,⁴⁰ which can be used to automatically detect important objects, has been used for classifying concrete crack and steel delamination and corrosion.⁴¹ However, crack identification from images that contain crack-like noncrack

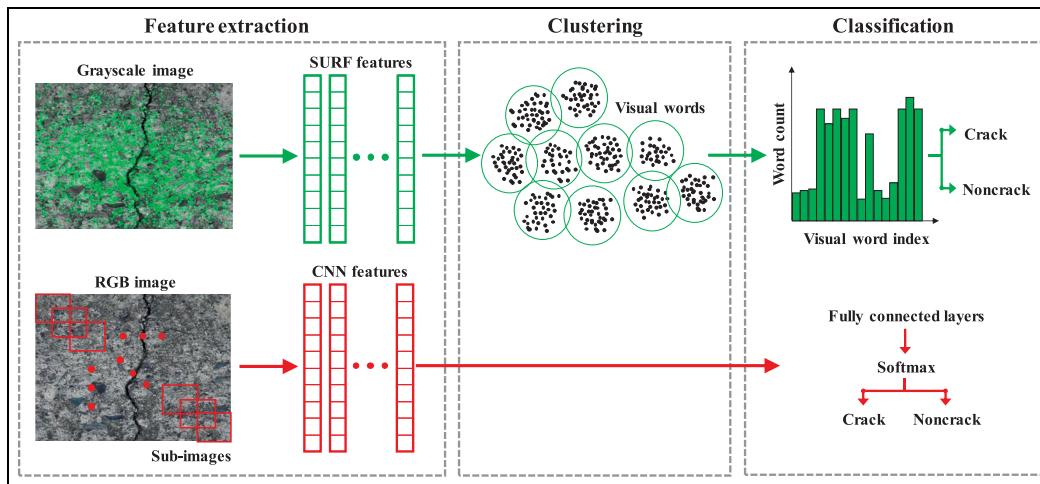


Figure 1. Schematic of SURF-based and CNN-based methods.
Source: Modified from Zheng et al.³⁹

objects have received little attention, despite this case being quite common in practice.

Concrete crack identification using machine learning

Based on the categorization process described in the previous section, a concrete crack identification approach is developed, consisting of two main processes: (1) generation of CCRs and (2) SURF-based and CNN-based classifications. Unlike the natural image classification process, the proposed approach can handle concrete surface images containing multiple cracks and noncracks that generally cover small portions of the entire image area. To enable this, crack candidates, which can be actual cracks or crack-like noncracks, are initially extracted using image binarization and then manually categorized as either a crack or as a noncrack in the training stage. Subsequently, SURF and CNN features are obtained from the CCRs, from which the classification models are constructed. The trained models are finally applied to new images to evaluate the classification performances.

CCR

The proposed approach is employed for identifying cracks in concrete surface images that may contain crack and/or crack-like noncrack objects. The proposed framework is designed to initially select crack candidates from surface images that may contain either a crack or a noncrack. The selected crack candidates constitute the CCRs, which are further used in building and applying the classification model.

The crack candidates, which represent both actual crack and crack-like noncrack objects, are selected from a concrete surface image for effective classification. The crack elements are typically represented by dark colors, which can be simply extracted using image binarization methods. In the image binarization approach, all the pixels are converted into zero (black) or one (white) based on a threshold calculated using the statistical properties, such as pixel intensities and user-defined parameters such as sensitivity and window size. Among the various image binarization methods^{42–44} available for detecting the CCRs, Sauvola's binarization is used in this study owing to its high performance in noisy and high-contrast images,⁴³ as shown in equation (1)

$$T = m \times \left\{ 1 - k \times \left(1 - \frac{s}{R} \right) \right\} \quad (1)$$

where R is a factor for normalizing the standard deviation, k is the sensitivity, and m and s are the mean and standard deviation of pixel intensities, respectively. Note that the sensitivity controls the contribution of the statistical properties, and the window represents a rectangular box in which the threshold of each pixel is calculated. In contrast to other methods that directly employ the standard deviation, Sauvola's binarization makes it possible to amplify the contribution of the standard deviation in an adaptive manner by a factor of R , making it effective with noisy and high-contrast images. The image binarization finally returns the crack and noncrack objects marked as black in the binary images. Most of the obtained objects appear to be clearly noncracks because of noisy surface textures, which can be removed based on their geometric patterns such as the eccentricity and the number of pixels in each pixel group, as shown in equation (2)

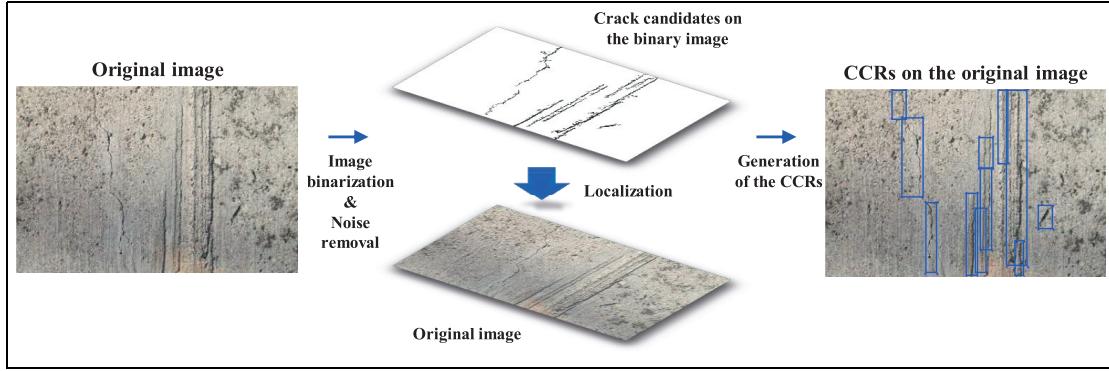


Figure 2. Generation of the CCRs in the entire image.

$$\begin{aligned} e &> e_{threshold} \\ A &> A_{threshold} \end{aligned} \quad (2)$$

where e and A are the eccentricity and the number of pixels of a pixel group in the binary image, respectively. The computational efficiency can be improved by filtering the unnecessary noisy objects. Finally, the smallest rectangles containing crack candidates are marked in the original image, as shown in Figure 2. Note that the CCR may contain either a true crack or a crack-like noncrack object. This implies that if only Sauvola's binarization is applied to an input image without further machine learning-based classification, all the CCRs are considered as cracks, even if some of them are noncracks (0% accuracy for true negative).

The advantages of the CCRs in the proposed framework can be summarized as follows:

1. The application of the CCRs is tailored to the classification of actual cracks and crack-like noncrack objects. Previous studies utilizing the scanning window focused on detecting cracks on intact surfaces.^{31,33} However, the CCRs enable constructing a classification model trained with cracks and crack-like noncracks.
2. The computational efficiency can be enhanced because only the selected CCRs are used in the training and validation stages. Considering that the image background, which does not contain possible crack or noncrack objects, occupies a major portion of the concrete surface image, excluding the background can significantly reduce the computational burden.
3. A robust classification model can be constructed from the CCRs. Previous studies that have used the scanning window have an issue that classification accuracy can be degraded when a crack or a noncrack is located at the edges of an image.³¹ In contrast to the scanning window, as a crack or a

noncrack in the CCRs is generally located at the center of an image, the proposed CCR-based framework is optimized for the classification.

SURF-based and CNN-based classification models

To construct the classification models, SURF and CNN features are obtained from the CCRs. In the SURF-based method, a grayscale image is used to extract the local features. A concrete surface image typically contains a large number of local features because of the noisy surface texture, thus affecting the classification of the cracks and noncracks. Because the important features are largely located on crack-like shapes (either actual cracks or noncracks), the binary information of the CCRs is used to preferentially select the SURF features on the crack segments, whereas most of the noisy SURF features on the concrete surface are filtered out, as shown in Figure 3. In contrast to the SURF-based method, the CNN-based method resizes the RGB image to a fixed resolution of $227 \times 227 \times 3$ for the input image in the employed CNN architecture. Note that the input size of AlexNet is introduced in the proposed approach.

The classification models of the SURF-based and CNN-based methods are constructed using the CCRs obtained from the concrete surface images. From the features obtained using SURF, the visual words that contain representative, small image segments are generated using k -means clustering. Subsequently, the obtained visual words are grouped to create a visual vocabulary. Here, the frequency of occurrence of the visual words in each category (i.e. cracks and noncracks) is calculated, from which the classification model is obtained using the linear SVM classifier. The trained model can be used to categorize new CCRs. Note that the clustering and classification processes used in this work follow the procedures described in section "SURF-based classification." In the

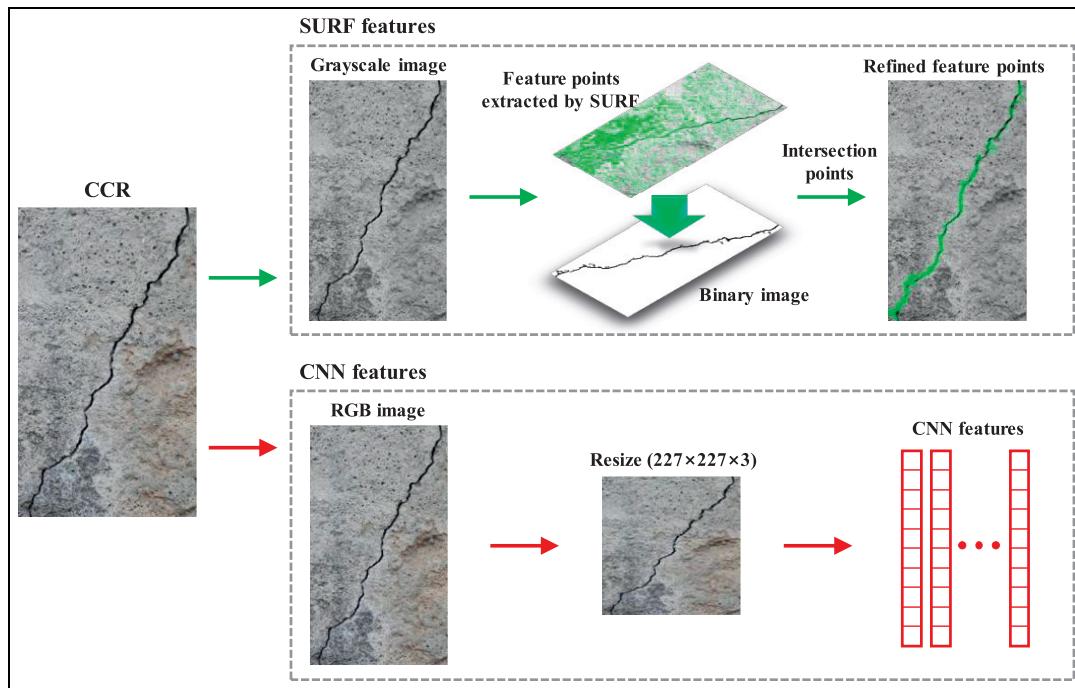


Figure 3. Feature extraction process of SURF and CNN.

CNN-based method, the obtained CNN features pass through the fully connected layers and then through the output layer to categorize the label, as described in section “CNN-based classification.” Figure 4 shows the schematic of the overall process of the proposed approach.

Experimental validation

Experimental setup

The proposed crack identification approach is evaluated to demonstrate its performance using surface images obtained from concrete structures. The image binarization is applied to 487 images captured using digital cameras (see Table 1) to extract the CCRs including cracks and noncracks. The user-defined parameters of the image binarization are selected as 0.07 and 131 for the sensitivity and the window size, respectively.¹¹ In addition, the thresholds of the noise object removal are selected as 0.9 and 5000 for the eccentricity and the number of pixels in each pixel group, respectively. Finally, 3186 CCRs are generated, which consist of 527 actual cracks and 2659 noncracks. To obtain a robust classification model, the image set is collected from various concrete surfaces under different working distances between the camera and the concrete surface, and under different illuminance conditions. Figure 5 shows typical sample images taken from the set. The images contain noncracks such as dark shadows, stains flowing down from the top, dust, and protruding lumps generated from the casts, which are generally found in

concrete structures. Furthermore, these kinds of crack-like noncracks are found to be similar to cracks in terms of geometry (e.g. long and thin) and color (both are dark). Note that the image database also includes branched cracks, spalling, and various orientations of cracks. All images can be downloaded at: http://shm.unist.ac.kr/files/Image_Pool.zip.

Classification performance comparison between SURF and CNN

The classification models of the SURF-based and CNN-based methods are implemented using MATLAB.⁴⁵ To evaluate the classification performances with respect to the size of CCRs, six sets (i.e. 100, 200, 500, 1000, 2000, and 3000) of CCRs are constructed from 3186 CCRs. In the feature extraction stage, SURF and CNN features are obtained by following the procedure of the proposed approach, as shown in Figure 3. To generate the classification model of the SURF-based method, three cases with different sizes of visual words (i.e. 100, 500, and 1000) are considered in the *k*-means clustering. Three cases with different minibatch sizes (i.e. 50, 100, and 200) are selected for the CNN-based method. With regard to the computational environment, a PC with an Intel Core i7-7700 processor clocked at 3.60 GHz and with 16 GB of RAM was employed. Moreover, a dedicated GPU (NVIDIA GeForce GTX 1080) was used.

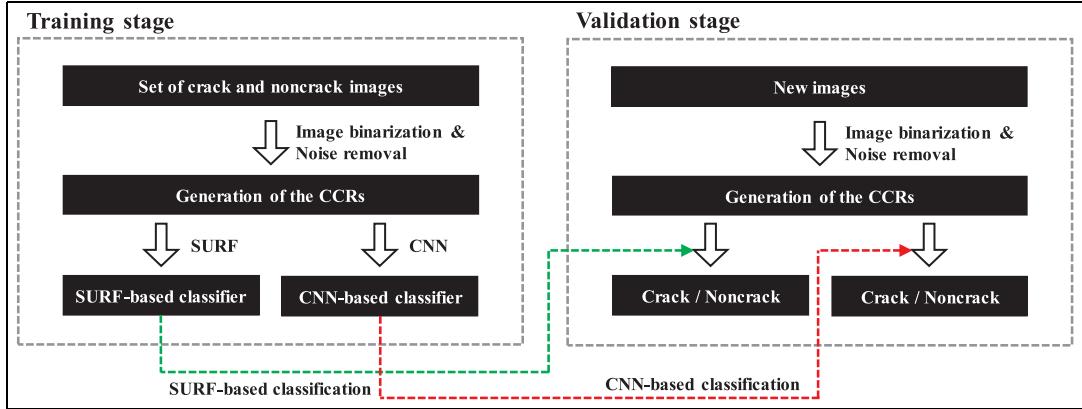


Figure 4. Flowchart of the proposed approach for concrete crack identification.

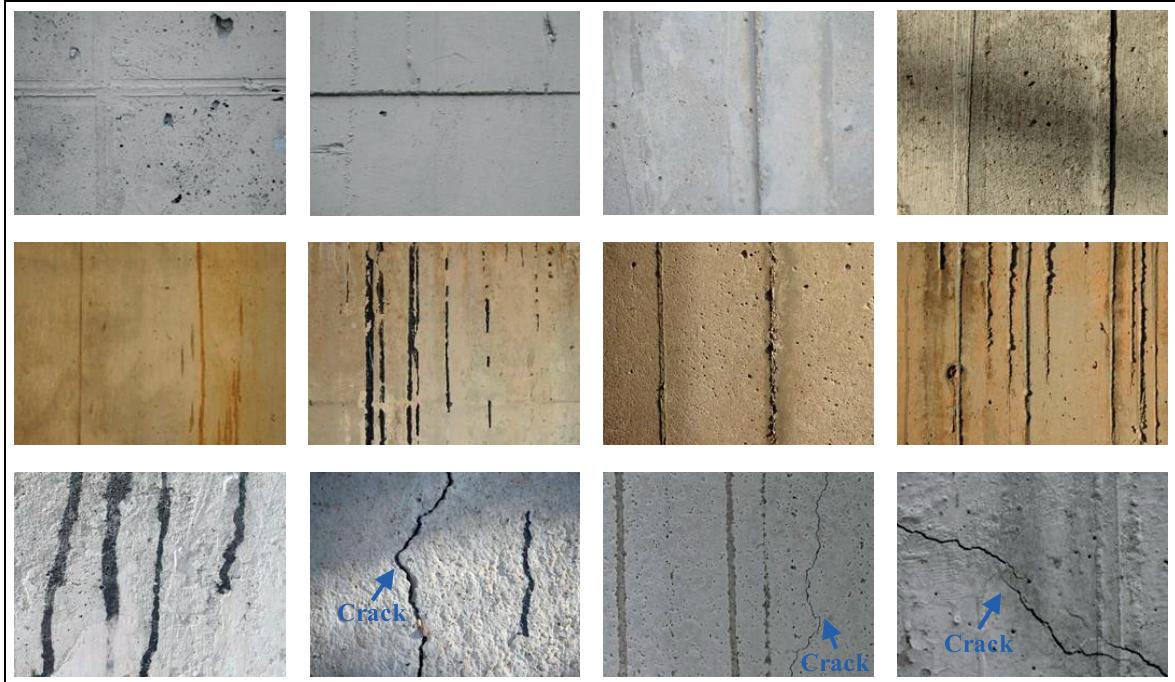


Figure 5. Sample images of concrete surfaces used for experimental validation.

Figure 6 shows the typical classification results. Both the SURF-based and CNN-based methods successfully categorize the CCRs in the sample images as either a

crack or as a noncrack, as indicated by the blue and red boxes. Note that only a few representative CCRs are shown for effective demonstration.

The trained classification models of the SURF-based and CNN-based methods are compared to quantitatively evaluate the identification performances. A 10-fold cross-validation is conducted for each CCR set (i.e. 100, 200, 500, 1000, 2000, and 3000). Figure 7 shows the results of the SURF-based method with three different visual words (i.e. SURF-100, SURF-500, and SURF-1000) and those of the CNN-based method with three different minibatch sizes (i.e. CNN-50, CNN-100,

Table 1. Specifications of used cameras.

	EOS-ID X	Coolpix 900S
Manufacture	Canon	Nikon
Image resolution	17.9 MP	15.9 MP
Focal length	100 mm	4.3–357 mm

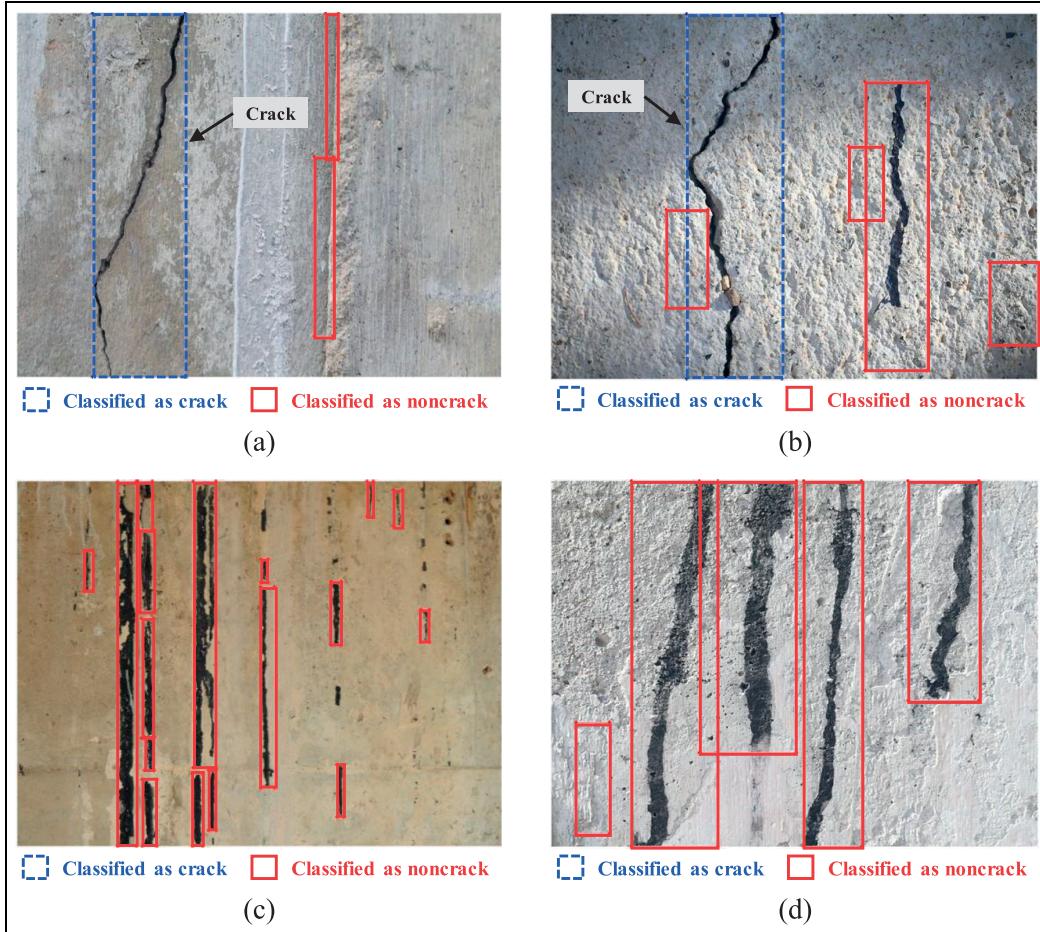


Figure 6. Typical classification results of cracks and noncracks from the CCRs (both the SURF-based and CNN-based methods correctly classify the CCRs): (a) sample 1, (b) sample 2, (c) sample 3, and (d) sample 4.

and CNN-200). Here, the following five performance metrics are selected to compare the models:

- *Precision*: $TP/(TP + FP)$;
- *Recall*: $TP/(TP + FN)$;
- *F1 score*: $2 \times (precision \times recall)/(precision + recall)$;
- *Accuracy*: $(TP + TN)/(TP + FP + FN + TN)$;
- *Computational time* in the training stage.

Where TP, FP, FN, and TN denote true positive, false positive, false negative, and true negative, respectively. As shown in Figure 7(b), the *recall* values corresponding to the SURF-based and CNN-based methods exhibit increasing trends with respect to the number of CCRs. However, the *recall* value of the SURF-based method decreases when the largest size of the CCRs is employed (i.e. 3000) because of over-fitting. In terms of the *precision*, as shown in Figure 7(a), the *precision* of the CNN-based method is higher than that of the SURF-based method and is reflected in the high *F1 score* (Figure 7(c)) and *accuracy* (Figure 7(d)). In

particular, the *F1 score* and the *accuracy* of CNN-50 significantly increase higher than those of the SURF-based method when 3000 CCRs are used in the training. Thus, when a sufficient minibatch size is used, CNN is observed to exhibit consistently high-performance metrics. In addition, the *computational time* for generating each classification model exhibits increasing trends in accordance with the number of CCRs, as shown in Figure 7(e). Although the CNN-based method is slightly better than the SURF-based method, it is difficult to directly compare them because the SURF-based and CNN-based methods are implemented on different processing units of CPU and GPU, respectively. Overall, the CNN-based method outperforms the SURF-based method in most cases in the crack and noncrack classifications.

The classification models of the SURF-based and CNN-based methods can be compared for specific CCR cases to qualitatively understand their identification characteristics. In particular, SURF-1000 and CNN-200 are used to categorize the CCRs in concrete

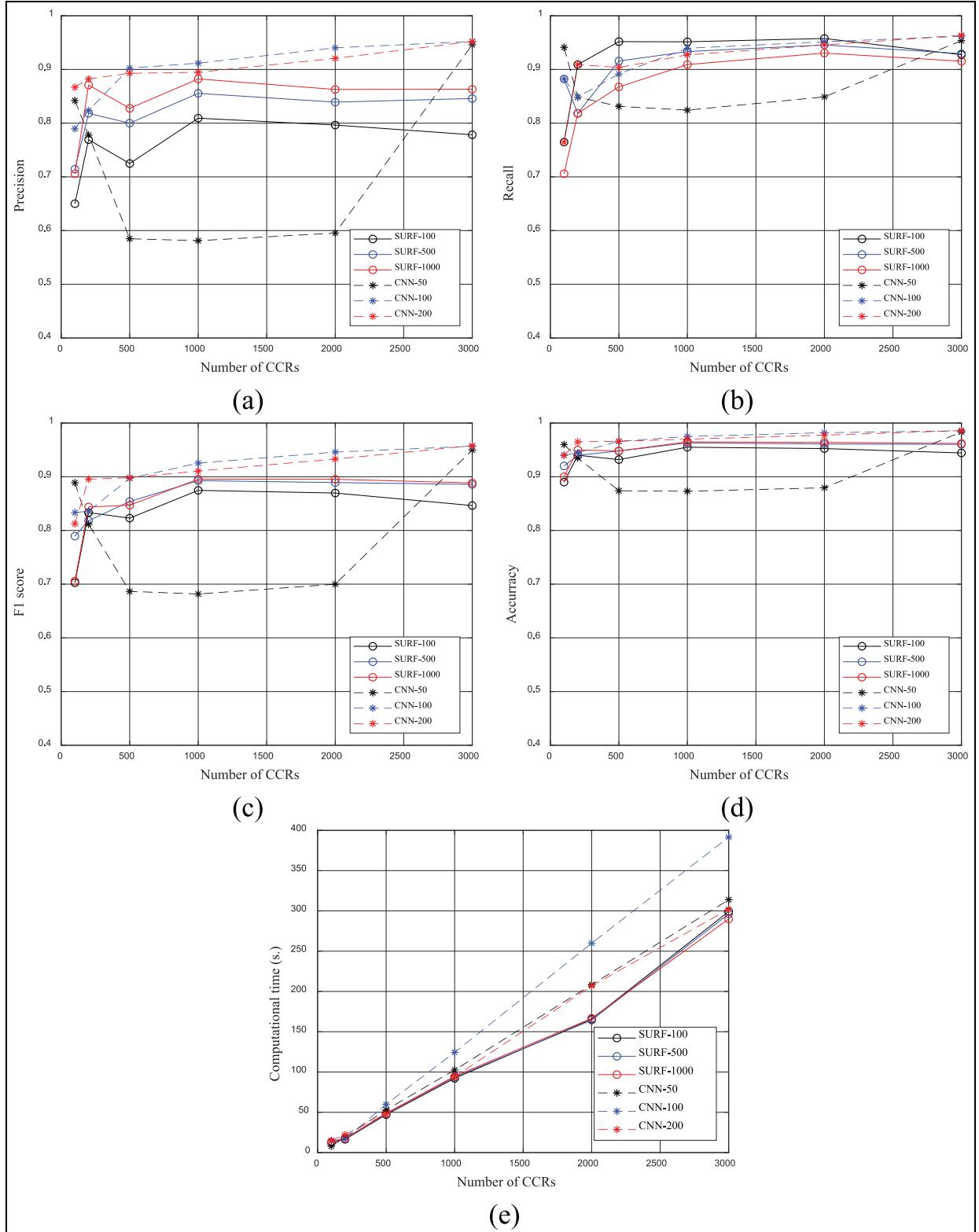


Figure 7. Comparison of the SURF-based and CNN-based methods in terms of (a) precision, (b) recall, (c) F1 score, (d) accuracy, and (e) computational time.

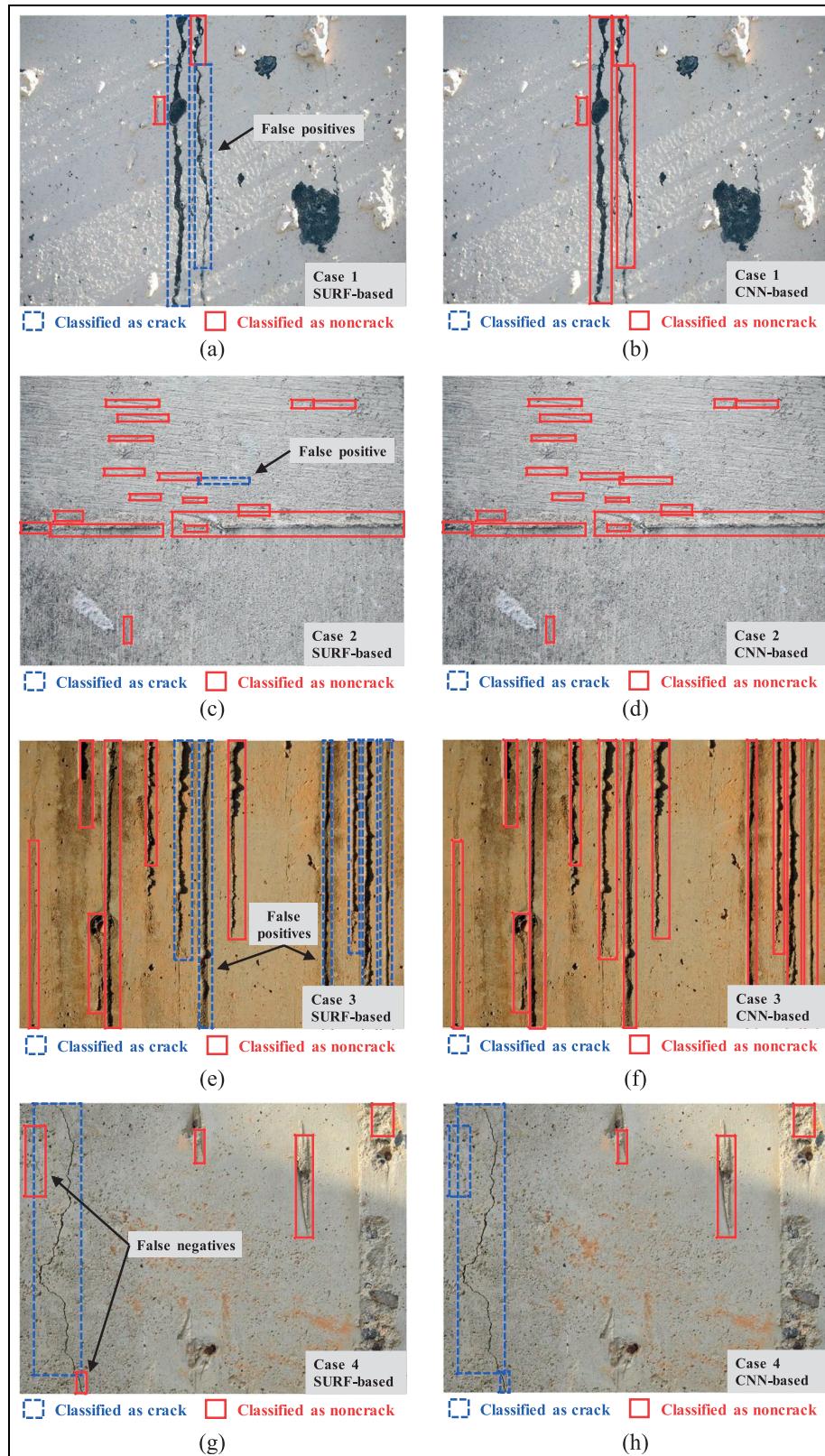


Figure 8. Classification of cracks and noncracks from the CCRs: (a) case 1 with the SURF-based method, (b) case 1 with the CNN-based method, (c) case 2 with the SURF-based method, (d) case 2 with the CNN-based method, (e) case 3 with the SURF-based method, (f) case 3 with the CNN-based method, (g) case 4 with the SURF-based method, and (h) case 4 with the CNN-based method.

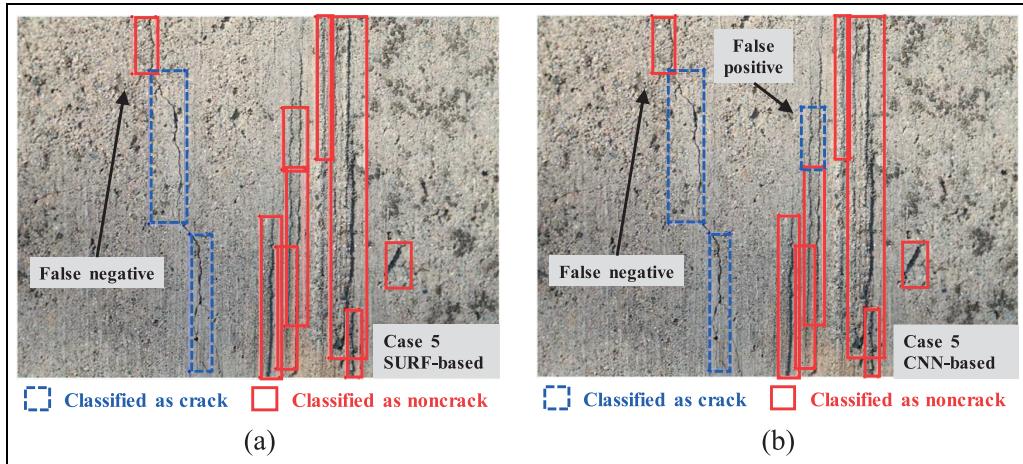


Figure 9. Classification of cracks and noncracks from the CCRs: (a) case 5 with the SURF-based method and (b) case 5 with the CNN-based method.

surface images that are not used in the training stage. Figure 8 shows the classification results for the four cases. Note that cases 1, 2, 3, and 4 represent dark stains flowing down from the top, protruding lumps generated between the casts, cement leaking from the cast, and surface cracks, respectively. As shown in Figure 8(b), (d), (f), and (h), CNN-200 correctly classifies all the CCRs in the four cases as either a crack or as a noncrack, as indicated in the blue and red boxes, respectively. In particular, the crack-like noncracks in cases 1, 2, and 3 that share similar geometry and colors with those of cracks are successfully identified as noncracks. Furthermore, the cracks with small widths are accurately recognized in case 4. In contrast to the CNN-based method, false positives and negatives are found in case of the SURF-based method (see Figure 8(a), (c), (e), and (g)). These examples show that the overall performance of the CNN-based method is better than that of the SURF-based method. Nevertheless, for the images used in this study, both the SURF-based and CNN-based methods correctly classify cracks and noncracks in most cases.

Although the classification performance of the CNN-based method is better in classifying actual cracks and crack-like noncrack objects, some of the CCRs could be successfully categorized only using the SURF-based method. As shown in Figure 9, both the SURF-based and CNN-based methods yield false negatives; however, the CNN-based method has an additional false detection from the lump on the concrete surface. Thus, the local features extracted using the SURF can in some instances correctly classify the CCRs that were incorrectly categorized using the CNN-based method. Hence, the combined use of deep neural networks and SVM classifiers with local/global features is found to have a potential to improve the classification performance.

To clearly show the advantage of the proposed crack identification, a comparative analysis has been conducted for three different classification models of previous studies and the proposed approach. Model A represents a classical classification constructed with k -means clustering and SVM. Widely used features for training in the literature^{23–26} are selected, including geometric patterns and statistical properties of crack and crack-like noncrack objects on concrete surface images. Based on the work by Cha et al.,³¹ model B is constructed using CNN with cracks and intact surfaces, while crack-like noncracks are not used. Model C built with CNN represents the proposed approach. All the number of CCRs in the training set are constant in each model (i.e. 527 cracks and 2659 intact surfaces or crack-like noncracks), and the parameters corresponding to the highest performance shown in Figure 7 are selected here. In the validation stage, a 10-fold cross-validation is conducted, in which all the classification models are applied to the CCRs containing largely cracks and crack-like noncracks. The training configuration for the three models is summarized in Table 2.

The validation results in Table 2 clearly show the efficacy of the proposed approach. The low-performance metrics of model A reveals that the geometric patterns and statistical properties are inadequate features to distinguish cracks and crack-like noncracks. In addition, without using crack-like noncracks results in poor classification results in model B. As such, the CNN features trained with cracks and crack-like noncracks are the critical enablers for successful crack identification.

Conclusion

This article proposes a machine learning approach to determine the existence and location of cracks in

Table 2. Comparison of classification models with CCRs containing largely cracks and crack-like noncracks.

		Classification model A	Classification model B	Classification model C ^a
Training configuration	Features	Geometric patterns and statistical properties	CNN features	CNN features
	Classification model	SVM	CNN	CNN
	Training data	Cracks and crack-like noncracks	Cracks and intact surfaces	Cracks and crack-like noncracks
Validation results	Precision	0.51	0.24	0.94
	Recall	0.49	1.00	0.96
	F1 score	0.50	0.38	0.95
	Accuracy	0.84	0.47	0.98

CCR: crack candidate region; CNN: convolutional neural network; SVM: support vector machine.

^aProposed approach.

concrete surface images containing possible crack-like noncrack objects. The main contribution of this article was to propose a classification framework based on the CCRs for identifying cracks in the presence of non-crack objects that share similar image characteristics (i.e. shape and color). In the training stage, concrete surface images with cracks and noncracks were prepared, from which CCRs were automatically extracted using image binarization. After the CCRs were generated, the SURF-based and CNN-based methods were applied to the CCRs to extract the important features of the cracks and noncracks, which were subsequently used to construct classification models. The obtained crack identification models were validated using concrete surface images that were not part of the training set. The experimental results confirmed that the proposed framework could successfully identify both cracks and crack-like noncracks using CCRs. Furthermore, the CNN-based method was found to be more accurate and efficient than the SURF-based method for crack identification. The experimental results can be summarized as follows:

1. Cracks and noncrack objects were effectively extracted and categorized from concrete surface images using the proposed crack identification framework based on the extracted CCRs.
2. The overall performance of the CNN-based method was better than that of the SURF-based method in most cases. The *precision* and *F1 score* were higher for the CNN-based method provided that sufficiently large minibatch sizes and CCR set sizes were used. The *recall* and *accuracy* of the CNN-based and SURF-based methods were largely the same.
3. In some cases, the SURF-based method was able to classify CCRs that were incorrectly classified using the CNN-based method. Combining deep neural networks and SVM classifiers with

local/global features could enable improved classification performance compared to using each method separately.

4. By introducing various crack-like noncracks in the form of CCRs in the training, the proposed framework enables accurate identification of cracks from concrete surface images in the presence of noncrack objects.

The proposed machine-learning-based crack identification approach has a strong potential for automated crack assessment of concrete structures.

Declaration of conflicting interests

The author(s) declare no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This research was supported by a grant (18SCIP-B103706-04) from the Construction Technology Research Program funded by Ministry of Land, Infrastructure and Transport of Korean government.

ORCID iD

Sung-Han Sim  <https://orcid.org/0000-0002-7737-1892>

References

1. Haynes C, Todd MD, Flynn E, et al. Statistically-based damage detection in geometrically-complex structures using ultrasonic interrogation. *Struct Health Monit* 2013; 12(2): 141–152.
2. Larrosa C, Lonkar K and Chang FK. In situ damage classification for composite laminates using Gaussian discriminant analysis. *Struct Health Monit* 2014; 13(2): 190–204.
3. Qiu L, Yuan S and Boller C. An adaptive guided wave-Gaussian mixture model for damage monitoring under time-varying conditions: validation in a full-scale aircraft fatigue test. *Struct Health Monit* 2017; 16(5): 501–517.

4. Liu P, Lim HJ, Yang S, et al. Development of a “stick-and-detect” wireless sensor node for fatigue crack detection. *Struct Health Monit* 2017; 16(2): 153–163.
5. Karthick SP, Muralidharan S, Saraswathy V, et al. Effect of different alkali salt additions on concrete durability property. *J Struct Integ Maint* 2016; 1(1): 35–42.
6. Domaneschi M, Sigurdardottir D and Glišić B. Damage detection based on output-only monitoring of dynamic curvature in concrete-steel composite bridge decks. *Struct Monit Maint* 2017; 4(1): 1–15.
7. Xu J, Fu Z, Han Q, et al. Micro-cracking monitoring and fracture evaluation for crumb rubber concrete based on acoustic emission techniques. *Struct Health Monit*. Epub ahead of print 15 Spetember 2017. DOI: 10.1177/1475921717730538.
8. Reagan D, Sabato A and Niezrecki C. Feasibility of using digital image correlation for unmanned aerial vehicle structural health monitoring of bridges. *Struct Health Monit*. Epub ahead of print 10 October 2017. DOI: 10.1177/1475921717735326.
9. Hu WH, Said S, Rohrmann RG, et al. Continuous dynamic monitoring of a prestressed concrete bridge based on strain, inclination and crack measurements over a 14-year span. *Struct Health Monit*. Epub ahead of print 30 October 2017. DOI: 10.1177/1475921717735505.
10. Liu Y, Cho S, Spencer BF Jr, et al. Automated assessment of cracks on concrete surfaces using adaptive digital image processing. *Smart Struct Syst* 2014; 14(4): 719–741.
11. Kim H, Ahn E, Cho S, et al. Comparative analysis of image binarization methods for crack identification in concrete structures. *Cement Concrete Res* 2017; 99: 53–61.
12. Kim H, Lee J, Ahn E, et al. Concrete crack identification using a UAV incorporating hybrid image processing. *Sensors* 2017; 17(9): E2052.
13. Abdel-Qader I, Abudayyeh O and Kelly ME. Analysis of edge-detection techniques for crack identification in bridges. *J Comput Civil Eng* 2003; 17(4): 255–263.
14. Hutchinson TC and Chen Z. Improved image analysis for evaluating concrete damage. *J Comput Civil Eng* 2006; 20(3): 210–216.
15. Jahanshahi MR, Masri SF, Padgett CW, et al. An innovative methodology for detection and quantification of cracks through incorporation of depth perception. *Mach Vision Appl* 2013; 24(2): 227–241.
16. Lee BY, Kim YY, Yi S-T, et al. Automated image processing technique for detecting and analysing concrete surface cracks. *Struct Infrastruct Eng* 2013; 9(6): 567–577.
17. Jahanshahi MR., Kelly JS, Masri SF, et al. A survey and evaluation of promising approaches for automatic image-based defect detection of bridge structures. *Struct Infrastruct Eng* 2009; 5(6): 455–486.
18. Koch C, Georgieva K, Kasireddy V, et al. A review on computer vision based defect detection and condition assessment of concrete and asphalt civil infrastructure. *Adv Eng Inform* 2015; 29(2): 196–210.
19. Yamaguchi T and Hashimoto S. Fast crack detection method for large-size concrete surface images using percolation-based image processing. *Mach Vision Appl* 2010; 21(5): 797–809.
20. Lattanzi D and Miller GR. Robust automated concrete damage detection algorithms for field applications. *J Comput Civil Eng* 2012; 28(2): 253–262.
21. Cortes C and Vapnik V. Support-vector networks. *Mach Learn* 1995; 20(3): 273–297.
22. Breiman L. Random forests. *Mach Learn* 2001; 45(1): 5–32.
23. Zhang W, Zhang Z, Qi D, et al. Automatic crack detection and classification method for subway tunnel safety monitoring. *Sensors* 2014; 14(10): 19307–19328.
24. Prasanna P, Dana KJ, Gucinski N, et al. Automated crack detection on concrete bridges. *IEEE T Autom Sci Eng* 2016; 13(2): 591–599.
25. Shi Y, Cui L, Qi Z, et al. Automatic road crack detection using random structured forests. *IEEE T Intell Transp* 2016; 17(12): 3434–3445.
26. Li G, Zhao X, Du K, et al. Recognition and evaluation of bridge cracks with modified active contour model and greedy search-based support vector machine. *Automat Constr* 2017; 78: 51–61.
27. Lindeberg T. Feature detection with automatic scale selection. *Int J Comput Vision* 1998; 30(2): 79–116.
28. Lowe DG. Distinctive image features from scale-invariant keypoints. *Int J Comput Vision* 2004; 60(2): 91–110.
29. Bay H, Ess A, Tuytelaars T, et al. Speeded-up robust features (SURF). *Comput Vis Image Und* 2008; 110(3): 346–359.
30. Juan L and Gwun O. A comparison of SIFT, PCA-SIFT and SURF. *Int J Image Process* 2009; 3(4): 143–152.
31. Cha Y-J, Choi W and Büyüköztürk O. Deep learning-based crack damage detection using convolutional neural networks. *Comput-Aided Civ Inf* 2017; 32(5): 361–378.
32. Gopalakrishnan K, Khaitan SK, Choudhary A, et al. Deep convolutional neural networks with transfer learning for computer vision-based data-driven pavement distress detection. *Constr Build Mater* 2017; 157: 322–330.
33. Tong Z, Gao J, Han Z, et al. Recognition of asphalt pavement crack length using deep convolutional neural networks. *Road Mater Pavement* 2017; 13: 1–16.
34. Zhang A, Wang KC, Li B, et al. Automated pixel-level pavement crack detection on 3D asphalt surfaces using a deep-learning network. *Comput-Aided Civ Inf* 2017; 32(10): 805–819.
35. LeCun Y, Boser B, Denker JS, et al. Backpropagation applied to handwritten zip code recognition. *Neural Comput* 1989; 1(4): 541–551.
36. Csurka G, Dance CR, Fan L, et al. Visual categorization with bags of keypoints. In: *Proceedings of the ECCV*, Prague, 11–14 May 2004.
37. Duda O, Hart PE and Stork DG. *Pattern classification*. Hoboken, NJ: John Wiley & Sons, 2000.
38. Krizhevsky A, Sutskever I and Hinton GE. Imagenet classification with deep convolutional neural networks. In: *Proceedings of the advances in neural information processing systems*, Lake Tahoe, NV, 3–8 December 2012.
39. Zheng L, Yang Y and Tian Q. SIFT meets CNN: a decade survey of instance retrieval. *IEEE T Pattern Anal*.

- Epub ahead of print 30 May 2017. DOI: 10.1109/TPAMI.2017.2709749
40. Ren S, He K, Girshick R, et al. Faster R-CNN: towards real-time object detection with region proposal networks. *IEEE T Pattern Anal* 2017; 39(6): 1137–1149.
41. Cha Y-J, Choi W, Suh G, et al. Autonomous structural visual inspection using region-based deep learning for detecting multiple damage types. *Comput-Aided Civ Inf*. Epub ahead of print 28 November 2017. DOI: 10.1111/mice.12334.
42. Niblack W. *An introduction to digital image processing*. Upper Saddle River, NJ: Prentice Hall, 1985.
43. Sauvola J and Pietikäinen M. Adaptive document image binarization. *Pattern Recognit* 2000; 33(2): 225–236.
44. Wolf C and Jolion JM. Extraction and recognition of artificial text in multimedia documents. *Pattern Anal Appl* 2004; 6(4): 309–326.
45. MATLAB. *Neural network toolbox release*. Natick, MA: The MathWorks, 2017.