Original Article

Steel bridge corrosion inspection with combined vision and thermographic images

Hyung Jin Lim[1], Soonkyu Hwang[2], Hyeonjin Kim[2] and Hoon Sohn[2]

Abstract

In this study, a faster region-based convolutional neural network is constructed and applied to the combined vision and thermographic images for automated detection and classification of surface and subsurface corrosion in steel bridges. First, a hybrid imaging system is developed for the seamless integration of vision and infrared images. Herein, a three-dimensional red/green/blue vision image is obtained with a vision camera, and a one-dimensional active infrared (IR) amplitude image is obtained from the infrared camera for temperature measurements with halogen lamps as the heat source. Subsequently, the three-dimensional red/green/blue vision image is converted to a two-dimensional chroma blue- and red-difference (CbCr) image because the CbCr image is known to be more sensitive to surface corrosion than the red/green/blue image. A combined three-dimensional (CbCr-IR) image is then constructed by fusing the two-dimensional CbCr image and the one-dimensional infrared image. For the automated corrosion detection and classification, a faster region-based convolutional neural network is constructed and trained using the combined three-dimensional CbCr-IR images of surface and subsurface corrosion on steel bridge structures. Finally, the performance of the trained, faster region-based convolutional neural network is evaluated using the images acquired from real bridges and compared with faster region-based convolutional neural networks trained by other vision and IR-based images. The uniqueness of this study is attributed to the (1) corrosion detection reliability improvements based on the fusion of vision and infrared images, (2) automated corrosion detection and classification with a faster region-based convolutional neural network, (3) detection of subsurface corrosion that is not detectable using vision images only, and (4) application to field bridge inspection.

Keywords

Corrosion detection, active infrared image, chroma blue chroma-red difference vision image, faster region-based convolution neural network, steel bridge inspection

Introduction

Corrosion is one of the most frequently found defects in steel structure members. Corrosion on the surface or below the coating (subsurface) of a steel structural component can reduce the load capacity of the component.[1] Therefore, reliable and early detection of corrosion is crucial, and visual inspection is mandatory for steel bridge maintenance, as stipulated by bridge authorities, such as the Ministry of Land, Infrastructure, and Transportation in the Republic of Korea and the Federal Highway Administration in the United States.[2,3] However, some inspection areas are hard to reach, and the inspection of these inaccessible areas can be dangerous to inspectors. Furthermore, the use of special equipment, such as scaffold and aerial work platforms, often causes traffic delays.

In response to these issues, there are ongoing efforts that use the vision cameras mounted on unmanned aerial vehicle (UAV) and unmanned ground vehicle (UGV) for bridge inspection.[4] Corrosion detection and quantification techniques using an edge detector[5] and wavelet transform[6] have been proposed, and several steel bridges are actually inspected using UAVs and UGVs.[7,8] For corrosion detection using a vision

[1]School for Engineering of Matter, Transport, and Energy, Arizona State University, Tempe, AZ, USA
[2]Department of Civil and Environmental Engineering, KAIST, Daejeon, South Korea

Corresponding author:
Hoon Sohn, Department of Civil and Environmental Engineering, KAIST, Daejeon 34141, South Korea.
Email: hoonsohn@kaist.ac.kr

camera with visible wavelength of 400 to 700 nm, an optimal color space is proposed.[9] The results indicate that the chroma blue- and red-difference (CbCr) color space converted from the conventional red/green/blue (RGB) color space yields a higher sensitivity to surface corrosion than the RGB color space. However, the performance of the vision-based technique is highly affected by the operational conditions, such as the incident angle, illuminance, and undesired contaminants in the air or backgrounds. Moreover, invisible subsurface corrosion cannot be detected with the vision-based technique.

Active thermographic infrared (IR) imaging techniques are also proposed for corrosion identification. External heat is applied to the structure, the corresponding thermal responses are measured using an IR camera, and corroded regions are visualized.[10,11] These IR techniques are applied to maritime and offshore structures.[12] In civil applications, the IR techniques are used for corrosion detection in reinforced concrete and steel member connections.[13,14] In particular, the active IR techniques are attractive for the detection of invisible subsurface and surface corrosion. However, the IR techniques are affected by the non-flat, surface structural geometries, such as blot and welded joints.

To tackle the limitations of vision and IR images, the two images are fused. IR images are combined with vision images for human face recognition to minimize the adverse effects caused by the illumination variations of the vision images.[15] The results indicate that the combined images improve robustness in face recognition in a wide range of illumination conditions. It is also reported that fusing vision and IR images can enhance the accuracy of fault diagnosis in welding processes.[16] A concrete crack detection technique is proposed that combines vision and IR images.[17] The results show that macro- and microcracks on concrete surfaces can be detected with a reduced number of false alarms.

Recently, machine learning algorithms, such as convolutional neural networks (CNNs), have been developed and applied to image-based classification and regression problems. The performance of vision-based corrosion detection techniques was evaluated, and the results indicated that the CNN-based techniques outperformed the wavelet analysis-based techniques.[18] A CNN-based technique has also been proposed for corrosion detection and severity estimation of pipelines.[19] For the identification and localization of multiple objects in a single image, region-based CNN (R-CNN) was developed,[20] and the performance was improved with the use of faster R-CNN.[21] A vision-based, faster R-CNN was designed to detect surface corrosion on plane surfaces and bolts, and concrete cracking.[22] The

results indicated that the damages were accurately detected and classified.

In this study, vision and IR images are simultaneously recorded, and a faster R-CNN model is applied to the combined vision-IR images for detection and classification of surface and subsurface corrosion in steel bridges. First, a hybrid imaging system (HIS) is developed by integrating vision and IR cameras. Herein, a vision image with a three-dimensional (3D) RGB color space is obtained from a vision camera, and a one-dimensional (1D) active lock-in amplitude IR image is obtained from an IR camera with halogen lamps for the heating. Subsequently, the color space of the 3D RGB vision images are converted to two-dimensional (2D) CbCr color space images because 2D CbCr is known to be more sensitive to surface corrosion.[9] The 2D CbCr and the 1D amplitude IR images are fused, and a combined 3D CbCr-IR image is established. Subsequently, a faster R-CNN is constructed and trained with the 3D CbCr-IR images. The performance of the trained faster R-CNN is examined using the CbCr-IR images obtained from real steel bridges and compared with the performances of the faster R-CNNs trained with the use of other vision and IR-based images. The uniqueness of this study is attributed to the (1) improvement of the corrosion detection reliability based on the fusion of vision and IR images, (2) automated corrosion detection and classification with a faster R-CNN, (3) detection of subsurface corrosion—which is not detectable with vision images—and (4) application to actual steel bridge inspections.

This study is organized as follows. In section "Development of a corrosion detection and classification technique based on combined vision and IR images," the hardware configuration of HIS, the combination of 3D CbCr-IR image, and the architecture construction and training of a faster R-CNN are outlined. The application of the developed corrosion detection and classification technique to actual steel bridges is presented in section "In-situ Bridge Testing." In section "Comparison of Corrosion Detection Performance using different input images," the performance of the 3D CbCr-IR image is compared with those of other vision and IR-based images. Finally, the concluding remarks are presented in section "Conclusion."

Development of a corrosion detection and classification technique based on combined vision and IR images

Development of HIS

Figure 1 shows the HIS that combines the vision and active IR imaging techniques. The HIS consists of three
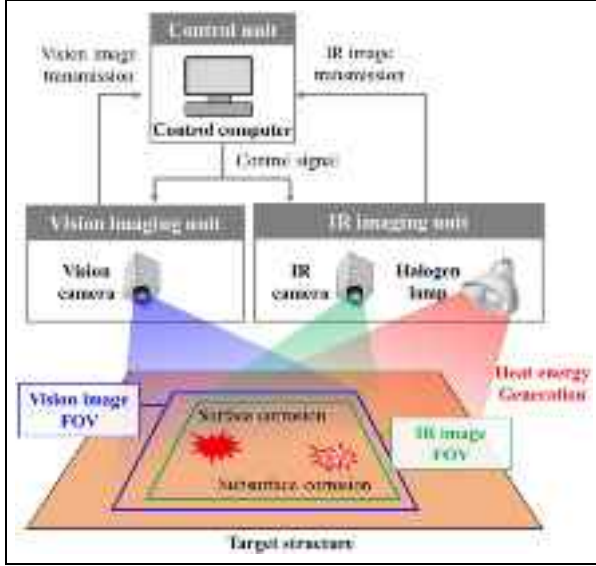
Figure 1. Schematic of the proposed hybrid imaging system (HIS) combining vision and thermographic (IR) images for corrosion detection and classification.

units: the vision and the IR imaging units and the control unit. The vision imaging unit captures vision images of a target inspection surface. The IR imaging unit applies heating to the target structure using a halogen lamp and measures the resulting temporal and spatial variation of temperature with an IR camera. The vision and IR imaging units are controlled and synchronized by the control unit. The acquired vision and IR images are stored and processed in the control unit as well. The detailed image fusion and corrosion detection processes are described in sections "Fusion of vision and IR images" and "Automated Corrosion Detection and Classification using Faster R-CNN," respectively, and the hardware configuration and specifications are provided in section "Test configuration."

Fusion of vision and IR images

The color space of conventional vision images is 3D RGB that constitutes the combination of the red (R), green (G), and blue (B) components. When it comes to corrosion detection in steel members, it has been reported that the RGB color space may not be the best one. The optimal color space for corrosion detection was examined by comparing 3D RGB, 3D YCbCr, 2D CbCr, and the 1D grayscale color spaces, and CbCr yielded the best performance.[8] Herein, the YCbCr color space consists of a luma (Y) and two Cb and Cr differences.

In this study, an RGB vision image (Figure 2(a)) is converted to a CbCr image (Figure 2(b)) using equation (1) and combined with the corresponding IR image

$$\begin{pmatrix} Y \\ Cb \\ Cr \end{pmatrix} = \begin{pmatrix} 0.299 & 0.587 & 0.114 \\ -0.169 & -0.331 & 0.500 \\ 0.500 & -0.419 & -0.081 \end{pmatrix} \begin{pmatrix} R \\ G \\ B \end{pmatrix} + \begin{pmatrix} 0 \\ 128 \\ 128 \end{pmatrix}$$

$$(1)$$

where R, G, B, and Y, Cb, Cr represent the intensities of three components in the RGB and YCbCr color spaces, respectively, and these values range between 0 and 255. For example, the pure red color in RGB can be represented as (255, 0, 0).

When the halogen lamp applies heat energy to the surface of the target structure, thermal waves are generated on the surface based on the photothermal effect and propagate along the thickness direction of the structure. The presence of corrosion alters the thermal conductivity, reflectivity, and emissivity of the material and produces spatial and temporal variations of the thermal wave propagation around the corrosion. Furthermore, subsurface corrosion that often occurs at the interface between the coating layer and the substrate steel member creates discontinuity in the thermal wave propagation. A lock-in amplitude IR image, referred to as the 1D lock-in amplitude IR image in this study, is one way of visualizing temporal and spatial disturbances. When the heat energy is applied to the structure with a duration of $\tau/2$, and the thermal images are acquired over a time period τ, the 1D lock-in amplitude IR image can be obtained as follows[23,24]

$$A(x,y) = \sqrt{\left\{I_0(x,y) - I_{\tau/2}(x,y)\right\}^2 + \left\{I_{\tau/2}(x,y) - I_\tau(x,y)\right\}^2}$$

$$(2)$$

where $I_t(x,y)$ is the temperature of the raw IR image at x and y coordinates recorded by the IR camera at time t, and $A(x,y)$ is the 1D lock-in amplitude IR image at the corresponding coordinates.

Once vision and IR images are synchronized and collected simultaneously, additional image registration and fusion processes are required to construct the proposed combined 3D CbCr-IR images. First, the CbCr and the 1D amplitude IR images are adjusted so that the same objects in the two images have identical sizes. The misalignment between the two images is fine-tuned using 2D cross-correlation.[25] The numbers of vertical and horizontal pixels to be shifted to achieve the best alignment between the two images are calculated by searching for the pixel shifts that maximize the cross-correlation value. Finally, the two images are overlaid on each other, and a region-of-interest (RoI) of 640×480 pixels is cropped from the fused CbCr-IR image (Figure 2(d)). Note that the size and misalignment adjustments need to be performed only once and the parameters for these adjustments can be used for all subsequent images.
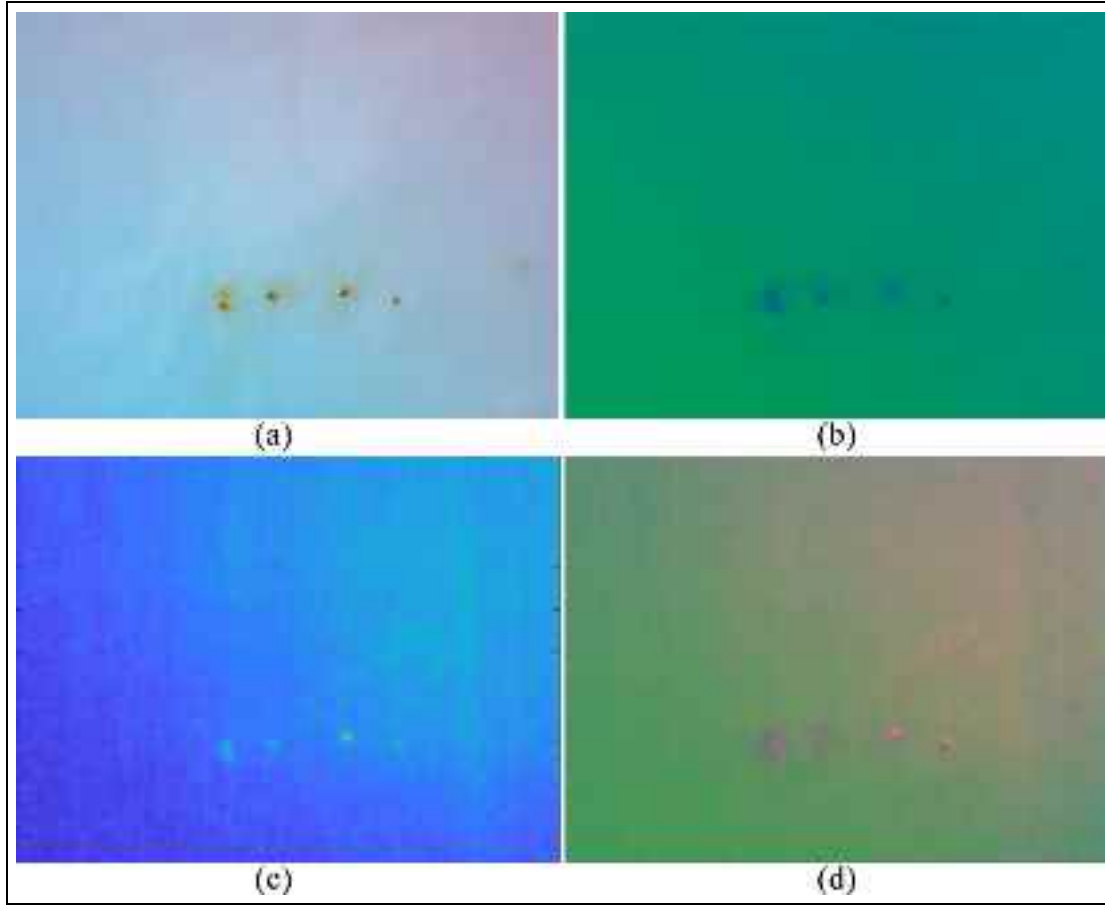
Figure 2. Fusion of vision and IR images: (a) three-dimensional (3D) red/green/blue (RGB) color space vision image, (b) two-dimensional (2D) chroma blue and red-difference (CbCr) color space vision image, (c) one-dimensional (1D) lock-in amplitude IR image, and (d) 3D combined CbCr-IR image.

## Automated corrosion detection and classification using faster R-CNN

Several machine learning techniques, such as the faster R-CNN, you only look once (YOLO),[26] single shot detector (SSD),[27] and mask R-CNN[28] have been applied to the object detection and segmentation of vision images. In faster R-CNN, YOLO, and SSD, target objects are identified by placing rectangular bounding boxes around the objects. Meanwhile, mask R-CNN partitions an image into multiple sets of pixels and determines if each pixel set overlaps with the target object or not. Because object detection is performed at each pixel level, mask R-CNN can quantify the size of the target object. However, the establishment of the training dataset is labor-intensive and time-consuming. YOLO and SSD are suitable for real-time object detection from video streaming because of their faster computation at the expense of lower object detection accuracy.[29,30] In this study, faster R-CNN is chosen because this technique offers the best object detection

performance, and the computation time is not a significant concern for the target application.

The schematic of the faster R-CNN used in this study is presented in Figure 3. Detailed explanations for each component of faster R-CNN and their functions are provided in the literature.[21,22] The overall concept is briefly reviewed here for the completeness of the paper.

First, the input image is fed into the CNN which consists of convolutional (Conv) layers,[31] rectifier linear unit (ReLu) activation functions,[32] and pooling layers, so that a unique feature map can be extracted from the input image. Subsequently, the region proposal network (RPN) takes the feature map as the input and outputs a set of rectangular regions called region proposals (RPs) that may contain target objects. Finally, the fast R-CNN identifies and classifies the target objects from the RPs and localizes them on the input image.[33]

A four-step process is used for the training of faster R-CNN.[21] First, the CNN and the RPN (① and ② in Figure 3) are trained with the initial weights, and RPs
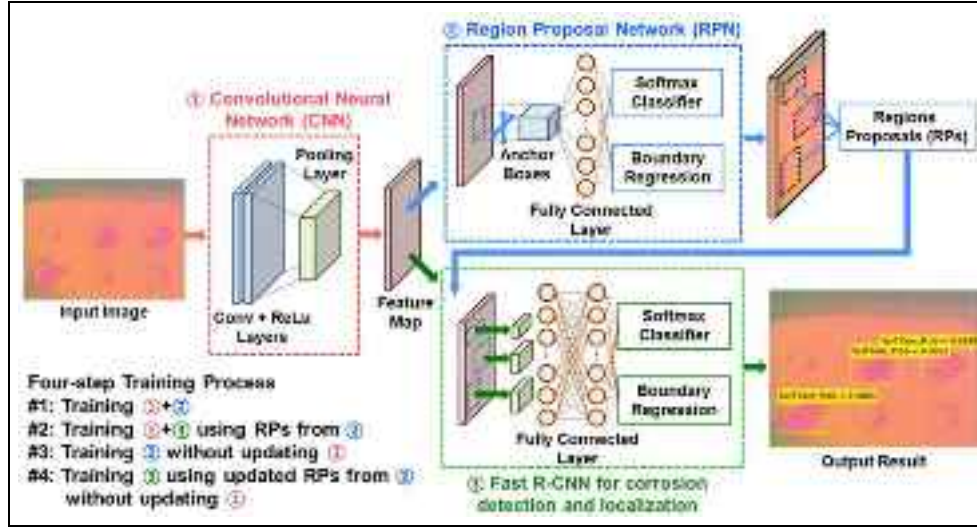
Figure 3. Schematic of a faster R-CNN used for the corrosion detection.

Table 1. Structure and specification of proposed faster R-CNN for corrosion detection.

| Layer	Type	Specifications
1	Conv + ReLu	No. of filters: 128, Size: 6 × 6, Stride: 1, Shared
2	Conv + ReLu	No. of filters: 128, Size: 6 × 6, Stride: 1, Shared
3	Conv + ReLu	No. of filters: 128, Size: 6 × 6, Stride: 1, Shared
4	Max pooling	Size: 3 × 3, Stride: 1, Shared
5	Sliding Conv + ReLu	No. of filters: 128, Size: 3 × 3, Stride: 1, Region proposal network (RPN)
6	Fully connected	No. of nodes: 128, RPN
7	Fully connected	No. of nodes: 128, RPN
8	Softmax and Regression	Region proposal, RPN
9	Fully connected + ReLu	No. of nodes: 256, Fast R-CNN
10	Fully connected + ReLu	No. of nodes: 256, Fast R-CNN
11	Fully connected + ReLu	No. of nodes: 128, Fast R-CNN
12	Fully connected	No. of nodes: 2, Fast R-CNN
13	Softmax and Regression	Corrosion detection and classification, fast R-CNN

are obtained. Next, the CNN and the fast R-CNN (① and ③ in Figure 3) are trained using the RPs from the previous training. The RPN is then trained again without updating the weights in the CNN. Finally, the faster R-CNN is trained again with the updated RPs from the RPN without updating the CNN. Because the RPN and the fast R-CNN share the CNN in the faster R-CNN, the computational efficiency and automaticity are improved compared with the fast R-CNN.

The structure and specifications of the proposed faster R-CNN for corrosion detection and classification are listed in Table 1. The CNNs (layers 1–4 in Table 1 and ① in Figure 3) are designed based on VGG16 and contain several Conv + ReLu layers followed by a pooling layer,[34] and the detailed structure and parameters of the faster R-CNN are modified and optimized by the authors.

The training dataset was constructed with 36 3D CbCr-IR images (640 × 480 pixels). Herein, a data augmentation technique was adopted to enlarge the number of the training dataset and to enhance the corrosion detection performance.[35] Thus, 72 images were used in total for the training by adding 36 rotated versions (by 180°) of the original images. During the annotation process, 202 objects (134 surfaces corrosions and 68 subsurface corrosions) were manually labeled with their coordinates and sizes.

For the training, the stochastic gradient descent with a momentum (SGDM) optimizer[36] with a learning rate of 1e-5 and a momentum of 0.8 was used, and the training was stopped after 1,250 epochs. During the training, 58 images were randomly selected out of 72 images for training and 14 images for validating the training process. In addition, the training and
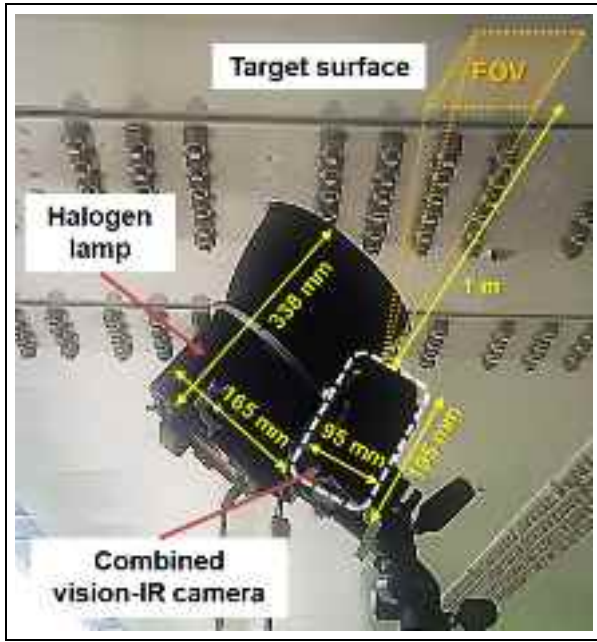
Figure 4. Hardware configuration of HIS for in-situ bridge inspection.

validation datasets were shuffled for every epoch. Mean squared error (MSE) was calculated at each iteration, and the MSE value for the validation dataset at the end of the training was 2e-3. To avoid overfitting, L2 regularization[37] with a regularization parameter of 5e-5 was applied. For the region proposal, 2,048 regions were arbitrarily selected for each image. The faster R-CNN construction, training, and tests were performed with MATLAB® 2019a and CUDA 10.0 environment on a computer with an Intel Core i7-4790 at 3.60 central processing unit (CPU), 16 GB double data rate (DDR) memory, and an 8 GB memory GeForce RTX 2080 super graphics processing unit (GPU).

In-situ bridge testing

Test configuration

Figure 4 shows the HIS used for in-situ bridge inspection consisting of a combined vision-IR camera (T650SC, FLIR, $143 \times 195 \times 95$ mm^3, 1.3 kg), a halogen lamp (H25 S, Hedler, $165 \times 165 \times 338$ mm^3, 2.65 kg), and a laptop (XPS159570, Dell, $357 \times 235 \times 11$ mm^3, 2.0 kg). The combined vision-IR camera acquires vision and IR images with 95×71 cm^2 and 33×44 cm^2 field-of-views (FOVs), respectively, and the resolution is 2592×1944 and 640×480 pixels, respectively. The thermal resolution and the spectral range of the IR image are 0.02 K and long-wave IR

(7.5–13 μm), respectively. For the inspection, HIS is placed 1 m apart from the target surface compromising the inspection speed and the spatial resolution of acquitted images. The incident angle of the halogen lamp to the target surface was set to 45° to minimize the effect of the reflected light on image acquisitions. Once the halogen lamp (maximum power consumption of 2.5 kW) introduces the heat energy onto the target surface, the combined IR-vision camera initiated the acquisition of vision and IR images. Herein, the halogen lamp applied heating for 10 s, and the IR images were recorded for 20 s (10 s heating and additional 10 s cooling) with a 30 Hz frame rate. These parameter values were selected considering the halogen lamp power, data acquisition time, and thermal properties of the coating material. The vision image (snapshot) of the target surface was then acquired. Finally, the 1D amplitude IR image was obtained with equation (2), and the combined 3D CbCr-IR image was then obtained. The detailed information regarding the images acquired from these bridges is not provided in this article because of a confidentiality agreement with the bridge authority.

Target bridges

The performances of HIS and the faster R-CNN-based corrosion detection technique were examined through the in-situ bridge tests at the First Jindo Grand Bridge and the Deung-Sun Bridge in South Korea. The First Jindo Grand Bridge is a three-span steel-box girder cable-stayed bridge composed of one long main span (344 m) and two long side spans (70 m) with widths equal to 11.7 m. Herein, the steel-box girder is streamlined and supported by the stay cables connected to two A-type pylons on concrete piers. Because the First Jindo Grand Bridge is located near the ocean and is exposed to a strong sea breeze, corrosion is one of the main concerns for bridge maintenance, and corrosion is often spotted around bolts and welded joints. In the case of the First Jindo Grand Bridge, the previous inspections reported that corrosion was most frequently observed in the first and third spans (side spans). Therefore, these spans were inspected by the HIS placed on the bridge inspection carriage (Figure 5(a)).

The Deung-Sun Bridge is 2 km long and 11 m wide. The Deung-Sun Bridge has a hybrid girder system composed of 20 steel box (1st to 10th and 31th to 40th) and 20 prestressed concrete (PSC) box (11th to 30th) girders. Herein, the 2nd to 10th and the 36th to 38th steel girders are vulnerable to corrosion owing to the strong vortices induced by topographical influences. For the Deung-Sun Bridge, the 3rd to 6th spans were inspected. Unlike the First Jindo Grand Bridge, the Deung-Sun Bridge does not have bridge inspection carriages. Thus,
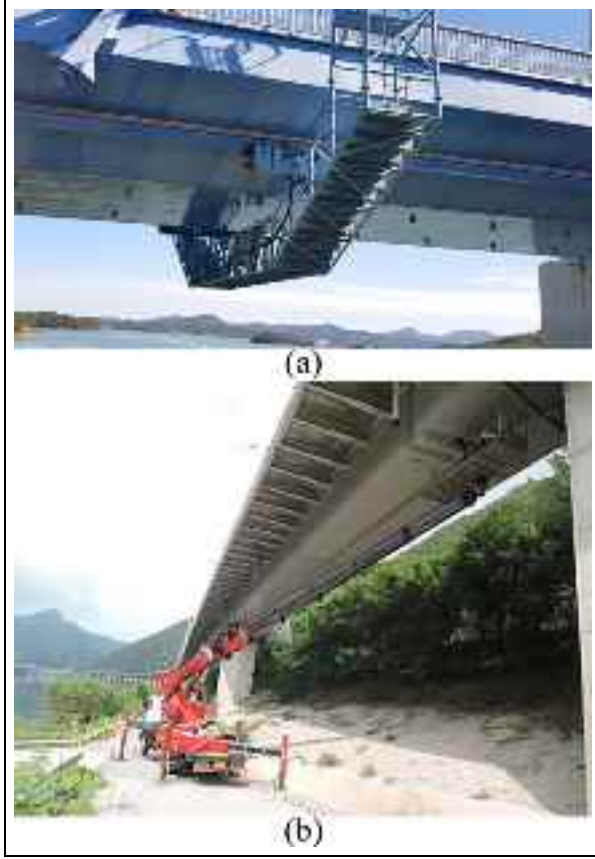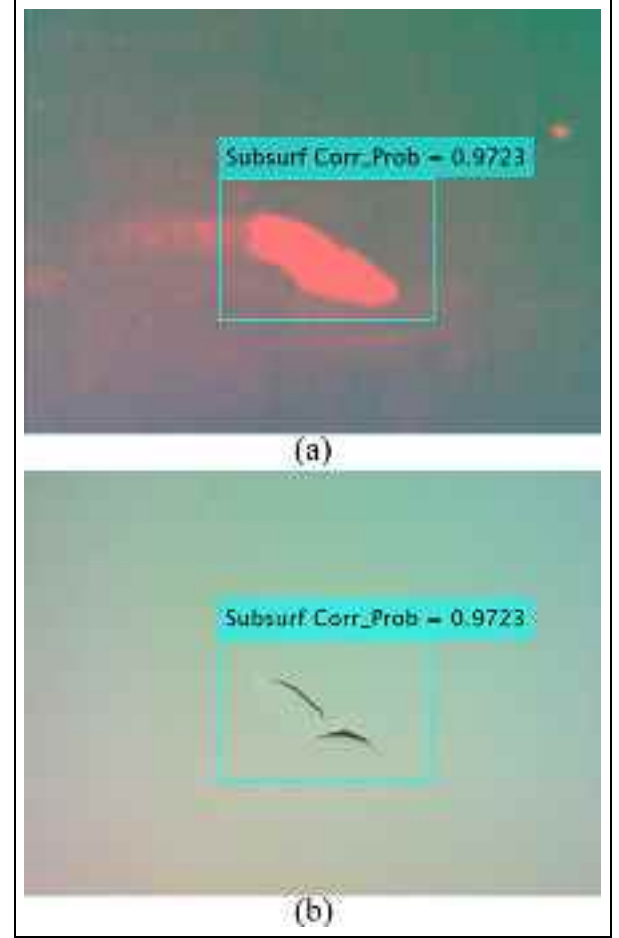
Figure 5. In-situ bridge inspection with an inspection carriage and aerial work platform. (a) First Jindo Grand Bridge inspection with the bridge inspection carriage and (b) Deung-Sun Bridge inspection with the aerial work platform.
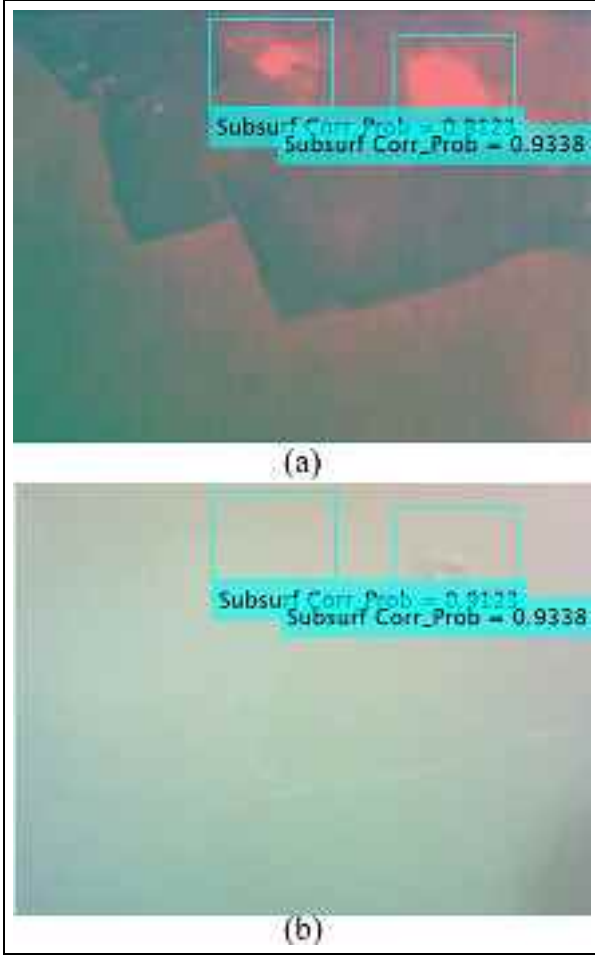

Figure 6. Detection of subsurface corrosion on plane surface: (a) detection result using the CbCr-IR image and (b) detection result overlapped on RGB vision image.

the aerial work platform shown in Figure 5(b) was used to access the areas beneath of the bridge girders, and the HIS was installed on the tip of the platform.

Test results

To evaluate the performance of the trained faster R-CNN for corrosion detection, additional images that were not used for the training were acquired, and representative results are shown in Figures 6–11. In all test results, only the objects identified with a confidence of 0.8 or higher were marked on the images. Figure 6(a) shows that subsurface corrosion on a planar surface is correctly detected. In Figure 6(b), the location of the identified subsurface corrosion is overlaid on the original RGB vision image for display. In particular, Figure 7(a) demonstrates that the proposed corrosion detection technique can even detect invisible subsurface corrosion, which is impossible to identify using conventional vision-based techniques (Figure 7(b)). In Figure 8, the surface corrosion on a plane surface is

successfully detected. Figures 9 and 10 show the surface corrosion detected around bolts and a welded joint, respectively. The results indicate that the trained faster R-CNN can detect surface corrosion on steel structures with complex geometries as well as simple plates. The effect of a complex geometry on corrosion detection (false positive) was also investigated in Figure 11. Although the welded joint was visible in the CbCr-IR image, no false alarm was triggered.

Average precision (AP) is extensively used for the evaluation of the performance of machine learning-based object detection algorithms.[38] The AP value is calculated using precision and recall (sensitivity) values obtained as follows

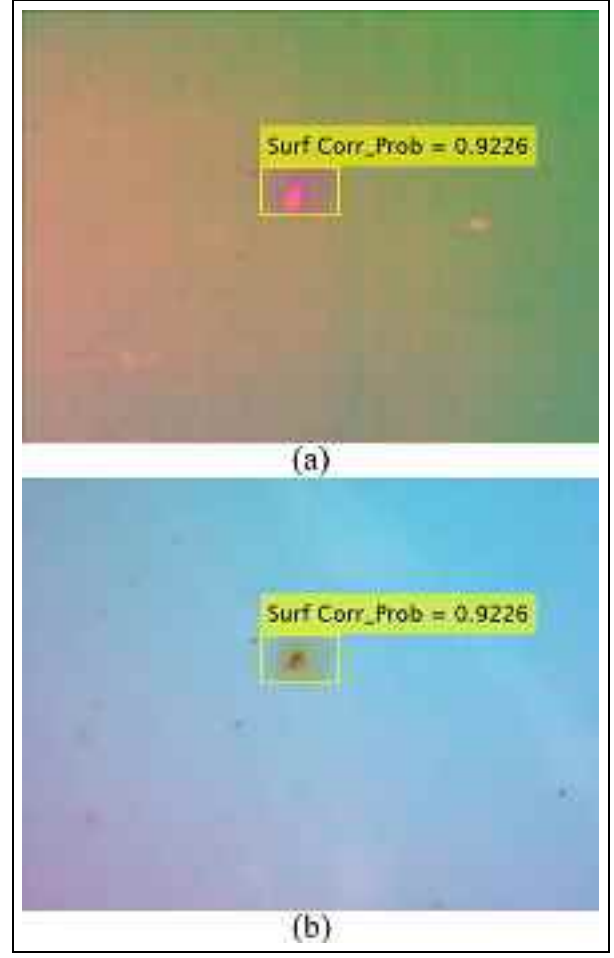$$Precision = \frac{TP}{TP + FP} \tag{3}$$

and

Figure 7. Detection of an invisible subsurface corrosion: (a) detection result using the CbCr-IR image and (b) detection result overlapped on RGB vision image.

Figure 8. Detection of surface corrosion on a plane surface: (a) detection result using CbCr-IR image and (b) detection overlapped on RGB vision image.

$$Recall = \frac{TP}{TP + FN} \qquad (4)$$

where *TP*, *FP*, and *FN* represent the number of true positive (correctly detected corrosion), false positive (falsely detected non-corrosion), and false negative (missed corrosion) detections, respectively. The precision value measures the accuracy of the corrosion detection algorithm, while the recall value represents the fraction of the successfully detected corrosion among all target corrosion. First, the precision and recall values are computed using the detected regions and the ground truths. Subsequently, the precision-recall curve is constructed, and the AP value is obtained by computing the area under the precision-recall curve.

In Table 2, the corrosion detection performance was evaluated with the AP values for surface (AP_1) and subsurface (AP_2) corrosion separately. The test dataset that consisted of 20 images was constructed with the ten originally acquired CbCr-IR images that were not used

for the training, and the additional ten images obtained by rotating the original ten images by 180°. The AP_1 and AP_2 values of 88.84% and 88.59% were obtained for surface and subsurface corrosions, respectively. It is expected that the corrosion detection performance can be improved when more images become available for training.

Comparison of corrosion detection performance using different input images

In this section, four additional faster R-CNNs were constructed using 3D RGB, 2D CbCr, 1D lock-in amplitude IR, and 3D IR images as inputs, respectively, to compare their performances with the proposed faster R-CNN with 3D CbCr-IR images. Herein, the 3D IR images were constructed by assembling the three raw 1D IR images acquired at 0, $\tau/2$, and τ for the 1D amplitude IR image in equation (2) in a single
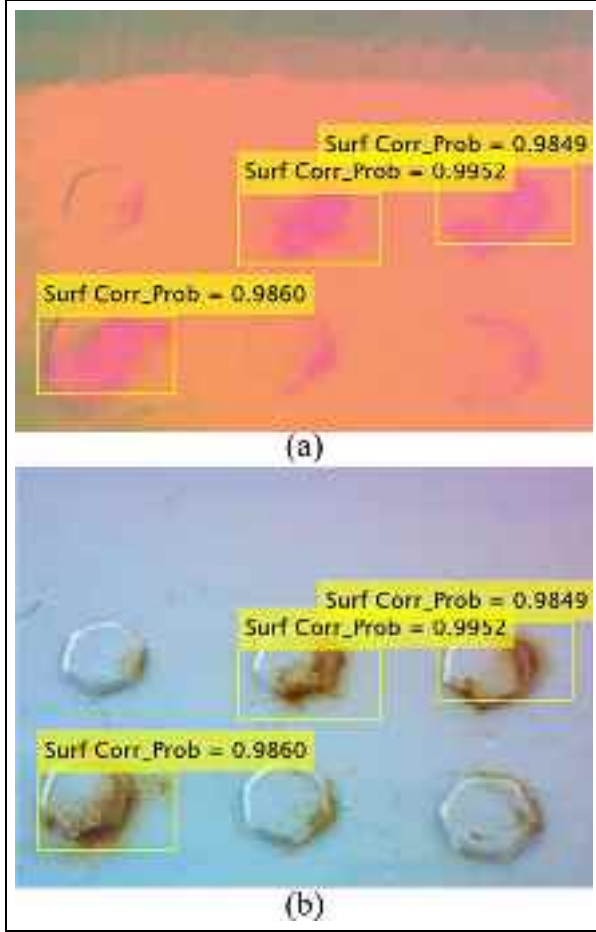
Figure 9. Detection of surface corrosion around bolts: (a) detection result using the CbCr-IR image and (b) detection result overlapped on RGB vision image.
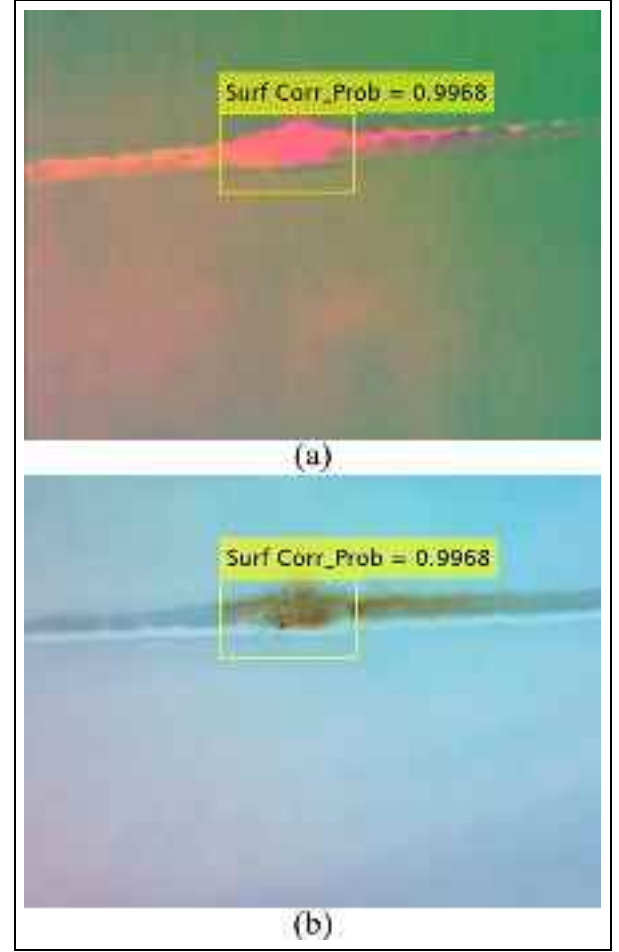
Figure 10. Detection of surface corrosion on a welded joint: (a) detection result using the CbCr-IR image and (b) detection result overlapped on RGB vision image.

image. All these faster R-CNNs share identical architectures, training parameters, and training and test datasets with those used by the previous version of the faster R-CNN with 3D CbCr-IR images.

The performances of these five input images are compared in Table 2. Vision images (3D RGB and 2D CbCr) yielded high-detection accuracy for surface corrosion. In addition, 2D CbCr images showed superior surface corrosion detection performance ($AP_1 = 83.13\%$) than 3D RGB images ($AP_1 = 78.97\%$), as reported in the literature.[8] However, these vision images were not successful in identifying subsurface corrosion. On the contrary, IR images (1D amplitude IR and 3D IR) successfully detected subsurface corrosion but produced false alarms for surface corrosion because of structural features, such as bolts and welded joints. If only subsurface corrosion was concerned, the performance of 1D amplitude IR images ($AP_2 = 87.02\%$) was comparable to that of the proposed 3D CbCr-IR images

($AP_2 = 88.59\%$). For the overall performance evaluation, the mean average precision (mAP) values were obtained by averaging the AP_1 and AP_2 values. The proposed 3D CbCr-IR images yielded the best performance (mAP = 88.72\%) regarding the simultaneous detection of surface and subsurface corrosion. The results indicate that the vision and IR images are complementary to each other, and the proposed HIS can detect invisible subsurface corrosion as well as surface corrosion with higher accuracy than vision or IR system alone. However, HIS takes a longer time for inspection, bulkier and heavier than the conventional vision-based inspection system.

Conclusion

In this study, a hybrid inspection system (HIS) was developed for automated detection and classification of corrosion in bridge structures by (1) combining vision
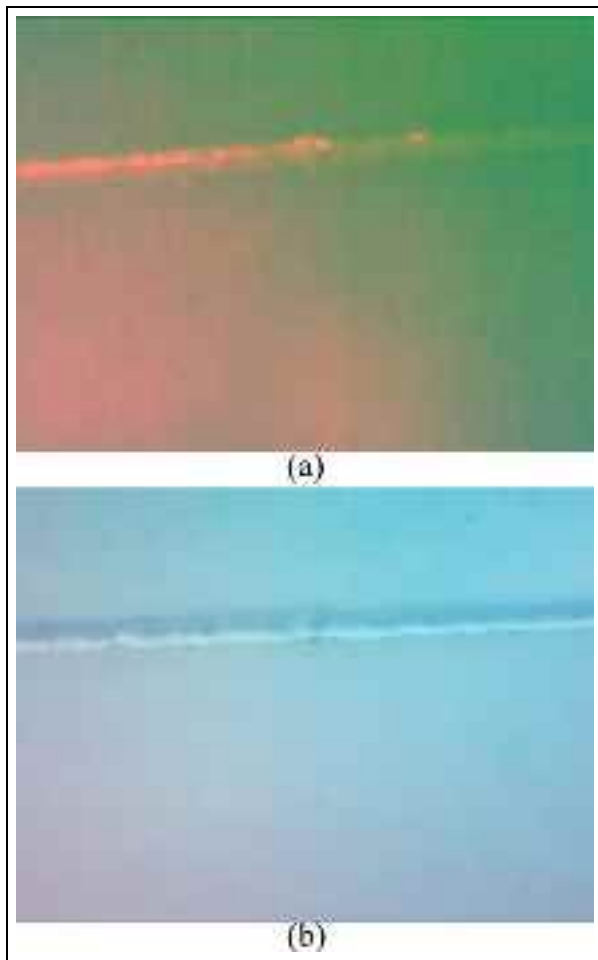
Table 2. Comparison of Corrosion Detection Performance.

Input image	Surface corrosion AP_1^a (%)	Subsurface corrosion AP_2 (%)	mAP^b (%)
3D CbCr-IR	**88.84**	**88.59**	**88.72**
3D RGB	78.97	11.54	45.26
2D CbCr	83.13	3.45	43.29
1D amplitude IR	55.52	87.08	71.30
3D IR	51.75	79.23	65.49

Note: The boldface values represent the highest values.
[a]Average precision.
[b]Mean average precision ($mAP = (AP_1 + AP_2)/2$).

Figure 11. Result of false-positive alarm test on a plane surface with a welded joint: (a) detection result using the CbCr-IR image and (b) detection result overlapped on RGB vision image.

and active thermographic (IR) images and (2) training a faster R-CNN model with these images. First, a 3D RGB and a 1D active lock-in amplitude IR image were simultaneously captured with the developed HIS. The 3D RGB vision image was converted to a 2D chroma blue- and red-difference (CbCr) image and fused with the 1D amplitude IR image to construct a combined 3D CbCr-IR image. Subsequently, a faster R-CNN was constructed and trained with the 3D CbCr-IR images. The uniqueness of this study comprises (1) improvement of corrosion detection performance by fusing vision and IR images, (2) automated corrosion detection and classification with a faster R-CNN, (3) detection of subsurface corrosion, which is not feasible with vision images, and (4) application to actual bridge inspections.

Based on vision and IR images obtained from the First Jindo Grand Bridge and the Deung-Sun Bridge in South Korea, the proposed 3D CbCR-IR images successfully detected most of surface and subsurface corrosions with the average precision (AP) values of 88.84% and 88.59%, respectively. Furthermore, when the corrosion detection performances of 3D RGB, 2D CrCb, 1D amplitude IR, and 3D IR images were compared with those of the proposed 3D CbCR-IR images, the proposed 3D CbCR-IR images yielded the best overall performance for the detection of both surface and subsurface corrosions. As more training images become available, the accuracy and robustness of the proposed technique are expected to be improved.

The delamination (such as blisters and bubbles) is also one of the coating defects that should be inspected. It is observed that the thermal characteristics of corrosion and delamination are quite different. Based on this observation, a machine learning-based coating defect classification technique is being developed using HIS. New findings will be reported in a follow-up study. In addition, the size, the weight, and the power consumption of HIS are being minimized to allow its integration with UAV and/or UGV for bridge inspections. With such integration with UAV and UGV, the data acquisition and signal processing can be further automated and the time and labor for bridge inspection will be reduced.

Declaration of conflicting interests

Funding

ORCID iD

Hoon Sohn https://orcid.org/0000-0001-9337-6653

References

1. Kim I, Dao D, Jeong Y, et al. Effect of corrosion on the tension behavior of painted structural steel members. *J Constr Steel Res* 2017; 133: 256–268.
2. Federal Highway Administration. *Steel bridge design handbook corrosion protection of steel bridges*. Washington, DC: FHWA, 2015.
3. Ministry of Land, Infrastructure and Transport. Guidelines for the implementation of safety maintenance of infrastructure. 2019. https://www.mlit.go.jp/road/road_e/s3_maintenance.html
4. Spencer BF, Hoskere V and Narazaki Y. Advances in computer vision-based civil infrastructure inspection and monitoring. *Engineering* 2019; 5(2): 199–222.
5. Pakrashi V, Schoefs F, Memet JB, et al. ROC dependent event isolation method for image processing based assessment of corroded harbour structures. *Struct Infrastruct Eng* 2010; 6(3): 365–378.
6. Ghanta S, Karp T and Lee S. Wavelet domain detection of rust in steel bridge images. In: *Proceedings of the IEEE international conference on acoustics, speech, and signal*, Prague, Czech, 22-27 May 2011, pp. 1033-1036. New York: IEEE.
7. Li Y, Kontsos A and Bartoli I. Automated rust-defect detection of a steel bridge using aerial multispectral imagery. *J Infrastruct Syst* 2019; 25(2): 04019014.
8. Lee JY, Sim C, Detweiler C, et al. Computer-vision based UAV inspection for steel bridge connections. In: *Proceedings of the 12th international workshop on structural health monitoring*, Stanford, CA, 10-12 September 2019.
9. Jahanshahi MR and Masri SF. Parametric performance evaluation of wavelet-based corrosion detection algorithms for condition assessment of civil infrastructure systems. *J Comput Civ Eng* 2013; 27(4): 345–357.
10. Solla M, Lagüela S, Fernández N, et al. Assessing rebar corrosion through the combination of nondestructive GPR and IRT methodologies. *Remote Sens* 2019; 11(14): 1705.
11. HeY Tian G, Cheng L, Zhang H, et al. Parameters influence in steel corrosion evaluation using PEC thermography. In: *Proceedings of the 17th international conference on automation and computing*, Huddersfield, 10 September 2011, pp. 255-260. New York: IEEE.
12. Bhandari J and Khan F. Modelling of pitting corrosion in marine and offshore steel structures—A technical review. *J Loss Prevent Proc* 2015; 37: 39–62.
13. Milovanović B and Pečur IB. Review of active IR thermography for detection and characterization of defects in reinforced concrete. *J Imaging* 2016; 2(2): 11.
14. Earls CJ. Bayesian inference of hidden corrosion in steel bridge connections: non-contact and sparse contact approaches. *Mech Syst Signal Pr* 2013; 41(1-2): 420–432.
15. Kong SG, Heo J, Boughorbel F, et al. Multiscale fusion of visible and thermal IR images for illumination-invariant face recognition. *Int J Comput Vis* 2007; 71: 215–233.
16. Fidali M and Jamrozik W. Diagnostic method welding process based on fused infrared and vision images. *Infrared Phys Technol* 2013; 61: 241–253.
17. An YK, Jang K, Kim B, et al. Deep learning-based concrete crack detection using hybrid images. *In: Proceedings of the SPIE smart structures and materials + nondestructive evaluation and health monitoring, Denver, CO, 4-8 March* 2018.
18. Atha DJ and Jahanshahi MR. Evaluation of deep learning approaches based on convolutional neural networks for corrosion detection. *Struct Health Monit* 2018; 17: 1110–1128.
19. Bastiana BT, Jaspreeth N, Ranjith SK, et al. Visual inspection and characterization of external corrosion in pipelines using deep neural network. *NDT & E Int* 2019; 107: 102134.
20. Girshick R, Donahue J, Darrell T, et al. Rich feature hierarchies for accurate object detection and semantic segmentation. *arXiv*: 13112524v5; 2014.
21. Ren S, He K, Girshick R, et al. Faster RCNN: towards real-time object detection with region proposal networks. *arXiv*: 150601497v3; 2016.
22. Cha YJ, Choi W, Suh G, et al. Autonomous structural visual inspection using region-based deep learning for detecting multiple damage types. *Comput-aided Civ Inf* 2018; 33(9): 731–747.
23. Liu J, Wang Y and Dai J. Research on thermal wave processing of lock-in thermography based on analyzing image sequences for NDT. *Infrared Phys Techn* 2010; 53: 348–357.
24. Song H, Lim HJ, Lee S, et al. Automated detection and quantification of hidden void in Triplex adhesive layers using active lock-in thermography. *NDT&E Int* 2015; 74: 94–105.
25. Haralick RM and Shapiro LG. *Computer and robot vision, Vol. 2*. Boston, MA: Addison-Wesley Publishing Company, 1993.
26. Redmon J, Divvala S, Girshick R, et al. You only look once: unified, real-time object detection. *arXiv*: 150602640v5; 2016.
27. Liu W, Anguelov D, Erhan D, et al. SSD: single shot multibox detector. *arXiv*: 151202325v5; 2016.
28. He K, Gkioxari G, Dollár P, et al. Mask R-CNN. *arXiv*: 170306870v3; 2018.
29. Huang J, Rathod V, Sun C, et al. Speed/accuracy trade-offs for modern convolutional object detectors. *arXiv*: 161110012v3; 2017.
30. Zhao ZQ, Zheng P, Xu S, et al. Object detection with deep learning: a review. *arXiv*: 180705511v2; 2019.
31. LeCun Y, Boser B, Denker JS, et al. Backpropagation applied to handwritten zip code recognition. *Neural Comput* 1989; 1(4): 541–551.
32. Nair V and Hinton GE. Rectified linear units improve restricted Boltzmann machines. *In: Proceedings of the 27th international conference on machine learning, Haifa, Israel, 21-24 June* 2010, pp. 807-814. Madison, WI: Omnipress.

33. Girshick R. Fast R-CNN. *arXiv*: 150408083v2; 2015.
34. Simonyan K and Zisserman A. Very deep convolutional networks for large-scale image recognition. *In: Proceedings of the international conference on learning representations (ICLR)*, San Diego, CA, 7-9 May 2015.
35. Shorten C and Khoshgoftaar TM. A survey on image data augmentation for deep learning. *J Big Data* 2019; 6: 60.
36. Qian N. On the momentum term in gradient descent learning algorithms. *Neural Netw* 1999; 12(1): 145–151.
37. Bishop CM. *Pattern recognition and machine learning*. New York: Springer, 2006.
38. Manning CD, Raghavan P and Schütze H. *Introduction to information retrieval*. New York: Cambridge University Press, 2008.