ABSTRACT The orthotropic steel-box girder (OSG) is widely used in the construction of a large-scale bridges. Since cumulative damages caused by the heavy vehicles and initial flaws of welding, the bridges with OSGs frequently suffer from fatigue cracks, which are commonly distributed around U-ribs. Hence, the management of fatigue cracks is mandatory in practical engineering. Although some techniques have been adopted for the detection of cracks, the workflow is often labor-intensive, time-consuming, and of low-temporal resolution. Considering the optical visibility of a crack and the limitation of the shape of an over-welding-hole around the U-rib, a machine vision-based monitoring methodology for the fatigue cracks in U-rib-to-deck weld seams is proposed in this paper. To be specific, a specific Internet of Things (IoT) based image acquisition device is first developed and introduced to obtain precisely part-view images of a fatigue crack. As followed, a novel image rectification and stitching method based on a specified coded calibration board is innovated and described for generating a measurable panoramic fatigue crack image. Furthermore, a deep learning-based crack detection-segmentation integrated algorithm is developed to detect and segment the crack areas. Afterwards, a feature extraction procedure based on image processing is explored to obtain the morphological features of a crack, involving its area, length and width. Finally, a field experiment was carried out on a real steel suspension bridge. By comparing the measurements both from manual measuring and vision-based monitoring, the results indicate that the proposed methodology is very promising to monitor the fatigue cracks in U-rib-to-deck weld seams, and the root-mean-square errors in length and width measuring could be 3.0195 mm and 0.003 mm, respectively. This work is not only of practical value to the management and maintenance of the OSG bridges in engineering, but also critical for the researches on fatigue cracks propagation.

INDEX TERMS Orthotropic steel deck, machine vision, IoT based image acquisition system, fatigue crack measuring, image stitching, deep learning.

I. INTRODUCTION

In past decades, the orthotropic steel-box girder (OSG) has been widely applied in long-span cable-supported bridges, benefiting from the advantages of low weight, high torsional stiffness, and convenience of manufacturing and construction. However, the main disadvantage of this structure is the fatigue cracks will inevitably be developed in weld seams from welding residual stress under the actions of repeated vehicle loads. Critically, the cracks grow continuously year by year, which has become a common issue, leading to lower security and lower service performance [1], [2].

To make a reasonable maintenance decision for the fatigue cracks, people have to detect and monitor the status of each crack periodically. Currently, there are two main categories techniques for the detection of steel cracks, involving destructive testing (DT) and non-destructive testing (NDT) techniques. In traditional NDT techniques, the eddy current testing (ECT) [3], [4], ultrasonic testing (UT) [5], [6] and acoustic emission testing (AET) [7], [8] are the most common approaches for researchers and workers to detect steel cracks. However, inspections using these
techniques are time-consuming and labor-intensive, especially for large-scale inspection. Besides, skilled inspectors and professional instruments are required for a routine inspection, which is expensive and time-delayed as well.

To improve the feasibility and efficiency of crack detection on OSG, some advanced testing techniques have been studied in recent years. On one hand, immersion ultrasonic testing (IUT) technique [9] was proposed for the examination of weld seams. Cracks in the weld could be identified and localized by comparing typical flaw patterns. Phased-array ultrasonic testing (PAUT) [10] technique has been used to detect the length and depth of the fatigue cracks in U-rib-to-deck weld seams. The soft elastomeric capacitor (SEC) [11] was applied and verified to sense distortion-induced fatigue cracks specifically. On the other hand, inspired by the developments in computer vision field over the last two decades, some new detection and evaluation techniques have been proposed [12] by using optical color cameras [13], infrared cameras [14] or laser scanners [15] as well.

With the image processing and computer vision based algorithms, cracks can be identified from the overall image scene. Compared with other methods, the vision-based methods commonly are inexpensive to implement, feasible for the complex scenes and adaptive for surface cracks, thus lots of researches have been done to detect cracks automatically using computer vision techniques, including road pavement cracks, concrete bridge cracks, and steel board cracks. However, there is less research on the monitoring of fatigue cracks in U-rib-to-deck weld seams. In this study, a machine vision-based methodology is proposed for the monitoring of fatigue cracks growth, with a specific system developed and utilized. Here, the major contributions of this paper are as follows:

- For the acquisition of crack part-view images precisely, we developed and introduced a new IoT based image acquisition device, which can be set to cross the narrow over-welding-hole around U-ribs and observe the fatigue cracks precisely in the U-rib-to-deck weld seams.
- For the generation of a measurable panoramic fatigue crack image, we proposed a novel image rectification and stitching method based on coded calibration board, which is validated by the experiments and the results indicate it is stable in the real scene and can be used for the steel surface image stitching in a low feature points case.
- For quantitative description precisely and long-term monitoring of the growth of cracks, based on the obtained panoramic fatigue crack image, a deep learning-based crack recognition method, involving detection-segmentation integrated crack identification algorithm and post-processing of morphological operations, is explored to calculate the width and length of a crack.

The remainder of this paper is organized as follows. In Section 2 we provide a literature review of vision-based crack detection equipment and methods. Then, Section 3 presents an overview of our methodology. As followed, Section 4 and Section 5 introduce the panoramic crack image acquisition system and the crack recognition methods for monitoring of fatigue cracks in U-rib-to-deck weld seams, respectively. Meanwhile, the effectiveness of the proposed methodology is validated by an experiment conducted in a real OSG bridge in Section 6. In Section 7, conclusions are introduced and some suggestions are provided for future work.

II. RELATED WORKS

A. THE DEVELOPMENT OF CRACK IMAGE ACQUISITION EQUIPMENT

Cracks in civil and infrastructure engineering mainly consist of pavement cracks, concrete cracks and steel cracks, which share similar characteristics. There are some image acquisition systems for crack detection. For example, the robot scanning system [16], [17] was developed and applied for inspection of road conditions. The robot was mounted with cameras for image acquisition and was set to autonomous navigation. In addition, unmanned aerial vehicles (UAVs) [18], [19] equipped with color or infrared cameras are promising to carry out a quick inspection of concrete deck or piers. Terrestrial laser scanning (TLS) system [20], [21] was also introduced for the acquisition and analysis of cracks in pavements or concrete structures at high precision. Consumer-grade camera [13] was used to get crack images inside the steel box girder, whose space was too limited for large-scale instruments. Recently, intelligent climbing robot [22] was invented to move around and inspect the steel plates inside the steel box girder, but it was still in the test stage.

Currently, to the best knowledge of the authors, there is few image acquisition system available for crack monitoring with high precision, especially, for the scene of U-rib-to-deck welded joint area. Thus the monitoring of crack growth in steel box girders is limited for years.

B. THE DEVELOPMENT OF CRACK DETECTION TECHNIQUES

Based on the crack images, image processing techniques (IPTs) are often used for crack detection, especially in earlier studies. Basically, due to their lower intensity in the gray-scale image than the background, cracks can be directly extracted by thresholds [23], [24] or by using morphological bottom-hat transform to reduce the effect of the nonuniform illumination [25]. Moreover, as edge-like features that cracks take, edge detection and texture feature extraction algorithms are introduced to obtain the profiles of cracks. Histogram of Oriented Gradients (HOG) features [26] were used to detect the location of the cracks near bolt, and the edge of cracks were determined by Canny edge detector. A Sobel edge detector [27] was used in the segmentation technique to identify cracks and other damages. Comparison among Haar transform, Fourier transform, Sobel edge detector and
Canny edge detector [28] was provided to identify cracks in bridges. In addition, since a fatigue crack is prone to open and close when subjected to repeated loading, dynamic features can be extracted and analyzed alongside the crack. Thus the crack is sensed and quantified through video feature tracking [29]. Besides, some literature shows that the location and depth of the cracks can also be determined and estimated by the frequency-domain methods by processing the images as two-dimensional discrete signals. Gabor filter [30] was used to identify longitudinal and transverse cracks in road surfaces and further promoted by the Adaboost classifier. After calculating the continuous wavelet transform of the slope of the mode shape, locations and depth of the cracks in a concrete beam were detected and estimated [31], [32].

However, those cracks detected from IPTs may contain real cracks and image noise. Thus, more advanced methods are needed to distinguish between cracks and non-crack noise. After detecting major cracks in the image, iterative applications of genetic programming algorithm [33] were used to eliminate the residual noise. More generally, machine learning-based classifiers are more universal in this task. Support Vector Machine (SVM), Adaboost and Random Forest (RF) [16] were studied and evaluated on gradient-based features and scale-space features, serving as classifiers to classify the crack and non-crack region. Artificial Neural Network (ANN) [34] was designed to classify whether an image object is crack or non-crack after the proposed feature extraction method. ANN, SVM and K-Nearest Neighbour (KNN) [35] were researched and evaluated by accuracy, precision, sensitivity, and specificity, in order to classify cracks from the non-crack pattern.

Though the robustness and accuracy are improved with the machine learning classifiers, these IPTs-based methods suffer from two aspects: (1) The in-situ scenes of structures are more complex than the laboratory, tricky spots or non-uniform light may cause the algorithms invalid. (2) The methods depend heavily on the human crafted features, which is very difficult to describe high-level features, such as the semantic information of the cracks. Recently, the state-of-the-art deep learning technique has been widely applied in the domain of computer vision [36]. Convolutional Neural Network (CNN) [37] is one of the most basic blocks of deep learning. CNN is able to extract features in the image from low-level (i.e., edges, colors) to high-level (i.e., locations, width), and predicting a class label to the image. The layer-by-layer features learned by CNN show more accuracy and robustness than the features extracted using traditional IPTs-based methods [38].

In the last five years, deep learning methods have been applied in civil engineering, such as the identification of bridge defects and road damage. Based on patch-wise and pixel-wise classification respectively, there are two main deep learning approaches to identify cracks of structures.

The patch-wise approaches detect the location of the cracks and select the crack region by using a bunch of image bounding boxes. NB-CNN [39] was proposed to detect the crack region by scanning the frame with 120 × 120 patches. And the Naive Bayes (NB) data fusion scheme helped discard non-crack patches more effectively than other methods. DDLNet [40] was trained to estimate the location of the suspicious defect region with the bounding box and predict the label of the defect. Thus the scheme can detect and classify multiple types of defects simultaneously. A crack detection workflow was proposed by combining a CNN and then the sliding window strategy [41]. The crack maps consisting of the positive patches in an image were obtained consequently. Faster Region-based CNN (Faster R-CNN) [42] architecture visual inspection was proposed to detect five damage types of concrete and steel infrastructure. The defect regions were selected with bounding box regression, improving the performance of the sliding window method. The trained model achieved high accuracy and fast speed in the detection of [43].

Though the location and range of cracks can be detected by the patch-wise approach in a short time, the quantification properties remain unknown since the approach merely indicates the boundary of cracks with rectangle boxes, while the pixel-wise approaches identify each pixel in an image to determine whether the pixel belongs to a crack or not. Fully Convolutional Network (FCN) [44] was applied to the detection of concrete cracks [45], [46]. Each pixel in the image was classified so that the cracks were segmented precisely from the background. Finally, the morphological features of the cracks were measured by the skeleton algorithm [45]. U-Net [47] was adopted to build a deep learning model for crack segmentation [48]. The U-Net structure was found to be more effective, more accurate than the patch-wise CNN approach in [41] and still reached high accuracy without a large training set in [45], [49]. A crack segmentation scheme with dual-scale deep CNN assisted by image enhancement and thresholding was proposed [50]. Meanwhile, the measurement of thin cracks was calculated accurately based on Zernike Moment Operator (ZMO), showing lower error than the traditional methods.

To conclude, however, the calibration are not conducted and the poses of cameras are not controlled in most studies. Thus the ratio between real world coordinates and image coordinates is unknown, which makes it impossible for the precise measurement of the cracks. On the other hand, most studies pay great attention on the algorithms, but neglecting the standard strategies for inspection in engineering, which is very significant for quantitative evaluation and long-term monitoring of fatigue cracks.

### III. OVERVIEW OF OUR METHODOLOGY

In this paper, we propose a machine vision-based methodology, which can be regarded as an overall solution, containing imaging strategy, data processing algorithms and usable descriptions for the evaluation and monitoring of fatigue cracks in U-rib-to-deck weld seams, which prefer to suffer from fatigue cracks much more frequently than other details.

As indicated in Fig. 1, this framework consists of two major parts. In panoramic crack image acquisition, we developed an image acquisition device to capture the raw crack...
part-view images. Based on IoT technology, the device consists of three major modules, including image acquisition module, mechanical driving module and IoT based controller module. Then by utilizing a specifically designed coded calibration board, the image rectification, image registration, and blending algorithms are integrated to generate a measurable panoramic crack image. In crack recognition and morphological features extraction, a deep-learning based cascade crack recognition workflow is introduced including crack region detection, crack segmentation and crack fragments linking, so that the segmentation mask of the crack is obtained. Through image morphological operations, mainly the skeleton algorithm, the crack length, width and average width are extracted.

**IV. PANORAMIC FATIGUE CRACK IMAGE ACQUISITION**

**A. IoT BASED IMAGE ACQUISITION SYSTEM**

The idea of the IoT has been proposed and thousands of practical applications came true in the last decade [51], [52].
The IoT based controller module. It includes a micro single-board computer and various interfaces. The computer is the controller of the mechanical driving module, the processor of the obtained images, data transfer of the panorama and extracted features. The interfaces serve as the interaction among the modules, the electricity, and the WIFI. Besides, the computer has been combined with IoT components to create an internet-connected monitor, which can collect and process the images in real-time and send the processed data to the cloud server as well.

Each module can be connected by cables easily and installed on the ribs, and then, a series of raw crack part-view images can be captured, as shown in Fig. 2a.

B. PANORAMIC CRACK IMAGE GENERATION BASED ON CODED CALIBRATION BOARD

Unlike the images taken in normal scenes, the images obtained from the proposed IoT device are restricted, mesoscopic, and of small view field. In order to obtain the real size of a crack from images, the series of crack part-view images in a scanning should be rectified and stitched as a measurable orthographic panorama. To do that, it is essential to find enough pairs of points with the same features from the adjacent raw images and know their relative coordinates for perspective transformation. However, the surface of steel materials usually is less rough and it is almost impossible to measure the coordinates manually in such a crowded millimeter-level region. Hence, it is hard to find enough reference points, which may lead to most automated feature-based image stitching methods invalid and unstable in practice.

In this study, we proposed a novel framework for image rectification and stitching by using a specific coded calibration board, which is specially set for the images photographed in this special case. As the workflow illustrated in Fig. 3, once the fatigue crack image acquisition device is installed on the ribs, the images of every scanning will be captured at the fixed locations, since the shooting positions and poses of the camera are precisely and strictly limited by the stepping motor in the mechanical driving module. Thus, initially, we utilize a specific coded calibration board to obtain the image transform and registration parameters for each position. And then, the images in every scanning can be rectified and stitched by reusing these parameters.
1) THE CODED CALIBRATION BOARD
As shown in Fig. 2b, the proposed coded calibration board is patterned by the chessboard (frequently used in camera calibration to resolve the intrinsic parameters of a camera [55]) and concentric circle markers (used in machine vision technology to match a series of images or serve as a control point [56]). Concretely, the pattern of the chessboard consists of 5 x n squared lattices, whose corners can be detected accurately and very fast. Due to the size of each lattice is previously known, the pairs of control points can be determined directly. In addition, several series of concentric circles with a different number of circle edges are designed to overlay the central white lattices of the chessboard cyclically, which will help the algorithm to know the orders of raw images in scanning and obtain the spatial relationship between them.

2) PERSPECTIVE RECTIFICATION BASED ON CODED CALIBRATION BOARD
Since it is difficult to keep the optical axle of a camera perpendicular to the surface of crack during the observation, perspective distortion is inevitable. Accordingly, the length in the image counted by the pixels will be unpredictable and cannot represent the real size of a crack. As usual, the perspective transformation is used to rectify a distort image by using the homography matrix $H$. Define $x$ and $x'$ as the origin and transformed coordinates of points, respectively. So the perspective transformation formula is:

$$x' = Hx \quad (1)$$

and then:

$$\begin{pmatrix} x' \\ 1 \end{pmatrix} = \begin{bmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{bmatrix} \begin{pmatrix} x \\ y \\ 1 \end{pmatrix} \quad (2)$$

So, the transformed coordinates of points are:

$$\begin{align*}
x &= a_1x' + a_2y' + a_3 \\
y &= b_1x' + b_2y' + b_3 \\
x &= c_1x' + c_2y' + 1 \\
y &= c_3x' + c_3y' + 1 \quad (3)
\end{align*}$$

For solving the eight unknown parameters in Eq. 3, four pairs of points are required at least. In practice, based on the coded calibration board, we first detect the corners and define the 4 most outside corners as the control points. Once the pairs of control points are obtained for each shooting position, the $H$ matrix can be solved.

Once the parameters of $H$ are obtained, by applying perspective transform point-by-point, the image will be rectified. Finally, by recording and using the four point-pairs, the raw images in each shooting position are rectified.

3) IMAGE REGISTRATION
As the key step of image stitching, image registration is to align the overlapping images by solving the spatial transformation among them, such as translation, rotation, and scaling. To avoid wrong and unstable registration in crack image stitching, by comparing with the area-based registration [57], feature-based registration [58] and phase-based registration [59] methods, a code-based method is proposed in this study, which will be color and texture-independent.

As seen from Fig. 4, we define the concentric circle markers with different edge numbers from 2 to 5 as the code from ‘A’ to ‘D’. For the images in different shooting position, we can extract the ordered point dictionary by using the concentric circles to represent the partial zone of chessboard and the order of corners to determine the order of feature points, i.e. $\text{List}_{\text{Name,Coordinates}} = [\text{List}_{2}, \text{List}_{3}, \ldots, \text{List}_{n}]$. While the ordered feature points are figured out in two adjacent images, we can use the coded name to make the pairs of corresponding points directly, then the registration parameters can be solved by a least square method and use to align the crack part-view images. Note that, the transformation here in this study is between two adjacent rectified images, so the registration parameter is only the parameters of translation displacement in the plane.

4) IMAGE BLENDING
Due to the change of illumination and inconsistent exposure time in each shooting, the stitching seam of two registered images is existed obviously as usual, without image blending processing, which is used to eliminate the stitching seam and adjust the color space of the overlapping region.
By comparing the weighted method [60] and multi-band method [61], the multi-band method is applied in this study, due to the weighted method often suffers from local blur and pixel-level error.

In the multi-band method, the overlapping area from image $I_1$ and image $I_2$ are denoted as $A$ and $B$, respectively. Then the Laplacian pyramids of $A$ and $B$ are built, denoted as $LA$ and $LB$. A new Laplacian pyramid $LS$ is built from $LA$ and $LB$, by using Eq. 4.

$$LS_s(i, j) = \begin{cases} 
LA_s(i, j) & \text{when } i < w/2 \\
LA_s(i, j) + LB_s(i, j)/2 & \text{when } i = w/2 \\
LB_s(i, j) & \text{when } i > w/2 
\end{cases}
$$

where $s$ is the level of the Laplacian pyramid, $(i, j)$ is the pixel coordinate of the image, $w$ is the width of the overlapping area.

Finally, the images in the $LS$ are upsampled from the bottom to the top. In each level, the resulting image will be the average of the upsampled image and the origin image.

An example of blending two registered images is shown in Fig. 5. The stitching seam is completely eliminated, and there is a smooth transition of color in the overlapping area.

---

### V. CRACK RECOGNITION AND MORPHOLOGICAL FEATURE EXTRACTION

In recent years, Deep Convolutional Neural Networks (DCNNs) [37], [62]–[67] have shown significant superiority than traditional methods in image classification, object detection and semantic segmentation. For achieving an accurate measuring of cracks, the pixel-level classification is required. However, the semantic segmentation for a large image usually is time-consuming and not efficient. Since the area of crack is less than the background part, it will be effective to detect the suspected crack regions before accurate crack segmentation to speed up the computation in the whole processing of crack recognition.

In this study, a cascade crack recognition method is developed. As shown in Fig. 6, there are three stages involved. Firstly, the sliding windows are utilized to re-sample a series of sub-images from a panoramic crack image without overlaps. Secondly, a DCNN based crack classifier is employed to distinguish whether these sub-images are crack image or not. Finally, with the crack sub-images, an image segmentation model is followed to segment the crack pixels from background.

#### A. CRACK REGION DETECTION

Cracks not only differ the shape but the color and texture from the background, so the DCNN used to classify cracks should be good at learning their high-level features. The VGGNet [62], ResNet [64], MobileNet [67], Inception [63] and DenseNet [66] are five classic DCNNs for image classification, which have shown great performance in extracting image features. The VGG block in VGGNet, firstly proposed by [62], is a typical and effective architecture in DCNN. As given in Table 1, the VGGNet-13 is beginning with 5 vgg blocks, each containing two convolution layers and one ReLU activation layer and one max pooling layer. Among them, the convolution layer and ReLU layer extract the features and generate feature maps at each level. While the max pooling layer reduce the size of the feature map by down-sampling. Finally, after three fully connected layers, the input image will be translated into a $1 \times n$ feature vector, where $n$ indicates the number of classification labels, for instance, crack or non-crack, etc.

#### B. CRACK SEGMENTATION

Based on the crack sub-images, the crack segmentation is utilized to predict the semantic label of each pixel. Fully Convolutional Network (FCN) [44], DeepLab [68] and PSPNet [69] are three of the most commonly used DCNN for image segmentation. By introducing the Laplacian pyramid idea into the CNN, the FCN can be regarded as a down-sampling to the up-sampling scheme. In the down-sampling stage, the backbone network (such as VGGNet, etc.) is used to extract the features of the image. In the up-sampling stage,
the de-convolution is applied to raise the size of the feature maps. In this study, the architecture of the FCN is grafted on the VGG blocks as listed in Table 2. Its output will be a mask matrix with the same size as the input image, whose cell will reveal the score for each classification labels and present the pixel position.

### C. CRACK FRAGMENTS LINKING

Considering the performance of DCNNs and the inevitable error in crack segmentation, the mask of the crack regions may be fragmented, which leads to wrong measurements in crack measuring, i.e. smaller area and length than the actual values. Thus, it is necessary to link all crack fragments together to obtain a complete crack mask. Commonly, the close operation is used to de-fragment the contiguous regions, which is defined by Eq. 5. Particularly, in close operation, the dilation operation will be taken firstly to enlarge the foreground object (i.e. crack) area and then followed by an erosion operation with the same structuring element template to shrink those enlarged areas by dilation and tiny isolated object (e.g. noise). Both operations here can be regarded as contrary operation. So in the end, the close operation can link all fragments and remove those tiny noise points.

\[
    A \cdot B = (A \oplus B) \ominus B
\]  

(5)

where, \( A \oplus B \) and \( A \ominus B \) means dilation and erosion operation, respectively.

**Dilation:**

\[
    A \oplus B = \{z | ((B)_z \cap A) \subseteq A\}
\]  

(6)

**Erosion:**

\[
    A \ominus B = \{z | (B)_z \subseteq A\}
\]  

(7)

where, A is the binary mask of crack and B is a template-like structuring element, also called kernel.

### D. SKELETON EXTRACTION

Considering the actual shape of cracks, the length of the crack could be defined as the length of its center line. Generally, the skeleton extraction algorithm is often used to do this, by extracting the skeleton from the segmentation mask.

The mathematical skeleton algorithm is defined by Eq. 8. The width of the crack mask is thinned by executing several times of erosion operation from the boundary, only remaining the center line of the crack mask. When the pixels in the mask no longer change after the erosion, the remainders will be the skeleton of the crack.

\[
    S_k(A) = (A \ominus kB) - (A \ominus kB) \circ B
\]  

(8)

Here, \( S_k(A) \) is the skeleton of A, \( (A \ominus kB) \) means k times consecutive erosion to A. the \( \circ \) means open operation, for
example $A \circ B$ is defined as Eq. 9

$$A \circ B = (A \ominus B) \oplus B \quad (9)$$

**E. MORPHOLOGICAL FEATURES OF A CRACK**

To evaluate a crack, the quantification of crack measuring is necessary. In this study, we defined three typical morphological features of a crack to describe it, including area, length, width. Based on the panoramic crack mask and its corresponding skeleton mask, all features can be easily calculated by following:

1) Area extraction. The area of the crack is simply defined as the number of pixels of the segmentation mask. Thus the area feature of the crack is defined in Eq. 10.

$$A = N_{\text{pixels}}(M) \quad (10)$$

Here, $M$ is the segmentation mask of the crack.

2) Length extraction. Since the skeleton of the crack is one-pixel width and represents the shape of the crack, the length of the crack can be regarded as the number of pixels of the skeleton, as in Eq. 11 and Eq. 8.

$$L = N_{\text{pixels}}(S_k(M)) \quad (11)$$

3) Width extraction. The width of the crack varies along the path of the crack. However, extracting widths at multiple positions along the crack is meaningful. The crack width is measured from the horizontal width of the segmentation mask. Besides, the average width is also significant to describe the crack. Thus the width feature of the crack is denoted as in Eq. 12.

$$\begin{cases} W_p = N_{\text{pixels}}(\Delta X_p(M)) \\ W_{\text{avg}} = \frac{A}{L} \end{cases} \quad (12)$$

where $\Delta X_p(M)$ is the horizontal width at point $p$, $A$ and $L$ are the area and length of the crack, respectively.

**VI. APPLICATION**

In this section, the proposed methodology was applied and evaluated in the steel box girder of a 20-year suspension bridge. Firstly, the image acquisition device was developed and installed above the U-rib-to-deck weld seams. Fourteen series of clear crack images were captured, proving the feasibility of the device. Then fourteen rectified panoramic crack images were generated by the proposed board-based image stitching method. Afterwards, several image classification and image segmentation models were trained and evaluated on a dataset with 6000 samples, which were randomly clipped from nine panoramic images. Finally, the precision of crack recognition was tested and verified on the other three panoramic images, by comparing the value of lengths and widths calculated by our method with the values of manual measurements.

**A. GENERAL INFORMATION**

The experiment was carried out in the steel box girder of a 20-year cable suspension bridge. The main span of the bridge is 1385 meters. The width and height of the girder are 36.9 and 3.0 meters, respectively. The bridge contains a total of six lanes, each with a 3.75-meters width. The section drawing of the steel box girder is displayed in Fig. 7.

The IoT based image acquisition device was installed as Fig. 8 shown. The modules were connected by cables and attached to the steel surface by strong magnets. The imaging parameters were listed in Table. 3. It is worth mentioning that after rectification, the scale factor in the crack image is 0.02 mm/px.

The MXNet framework [70] was used to perform all the experiments and studies in this article. To implement the
B. IMAGE ACQUISITION

Considering the verification of the device and the variety of the dataset, the device was installed on fourteen different locations in the girder at each time. The captured board and crack images were all vivid and clear. Three series of board and crack images (partly) were shown in Fig. 9.

Furthermore, to obtain the panoramic and measurable image for each crack, the proposed rectification and stitching method were applied to process the part-view images. Clear, measurable and seamless panoramic crack images were obtained eventually. Three images were shown in Fig. 10.

C. DCNNs MODEL TRAINING AND EVALUATION

1) DATASET

In this study, we made a dataset manually by ourselves, including sub-images and the corresponding segmentation mask. Concretely, all 14 panoramic crack images were applied and divided into three groups: 9 images for training, 2 images for validation and the rest 3 images for test. In image classification, to fit the input size of DCNN models, we used a 128 px × 128 px window to re-sample the image samples randomly from each panoramic image. While in image segmentation, considering the balance of area between crack and non-crack, we used a 64 px × 64 px window. To enlarge the number of samples, we also applied the image augmentation techniques, including flipping, random cropping and resize, and color jitter, etc. Consequently, there were 6,000 positive samples (i.e. crack regions) and 6,000 negative samples (i.e. background regions) obtained for training, and 200 positive samples and 200 negative samples for validation, which will be used to monitor the generalization ability of the models in the training process. Finally, 60 crack and 534 non-crack sub-images, clipped from the test panoramic images, were used to evaluate the performance of the model for inference.

2) TRAINING AND EVALUATION

For crack image classification and segmentation, comparative study of different architectures was applied. The main indices of model training are listed in Table 4. The loss function and optimizer are 2D Cross-Entropy and Adam [71], respectively. The learning rate (LR) will be reduced by 0.5 every 5 epochs. To monitor the training process, the loss curve of training set of these models are illustrated in Fig. 11a and Fig. 11b. Obviously, the loss is decreasing with iteration epochs, which means the errors are decreasing and the parameters are optimized gradually. Besides, to evaluate the saved models, the Precision-Recall curve (P-R curve) and the average precision (AP) are employed. As Table 5 and Table 6 shown, the highest average precision for image classification

![Fatigue crack examples](image)

**FIGURE 9.** Three series of board and crack images (partly).

**TABLE 3.** Parameters of the IoT imaging system.

| Parameters     | Units     | Values     |
|----------------|-----------|------------|
| Acquisition    |           |            |
| Total Length   | mm        | 200        |
| Part-view images|          | 50         |
| Camera         |           |            |
| Image size     | px × px   | 640 × 480  |
| Image resolution|         | 300,000    |
| Focal length   | mm        | 4          |
| Aperture size  | mm        | 1          |
| Angle of field-of-view | ° | 40 |
| Measurement    |           |            |
| Size of the coded board | mm × mm | 500 × 10 |
| Size of the lattice | mm × mm | 1.60 × 1.60 |
| Scale factor   | mm/px     | 0.02       |
D. CRACK RECOGNITION AND FEATURE EXTRACTION

Based on the trained classification and segmentation models, the cracks are recognized. The threshold is defined as the minimum score an image has to achieve to be recognized as crack after fed to the neural network. Firstly, the VGGNet-13 model with threshold 0.5 is used to detect the crack regions by a 128 px × 128 px sliding window with 128 px stride, shown as the blue boxes in Fig. 12b. Then the detected crack regions are fed into the FCN model with threshold 0.4 to obtain the pixel-level segmentation mask, shown in the red area in Fig. 12c.

However, the segmentation mask is fragmented in some regions due to the inevitable error of the model. Thus the close

is 0.978, achieved by the VGGNet-13 model. While the highest average precision for image segmentation is 0.932, achieved by the FCN (backbone: VGGNet-13) model. The P-R curve of some models are illustrated in Fig. 11c and Fig. 11d. In addition, the inference time for one sub-image is also obtained. Among the studied models, the VGGNet and FCN based architectures are most time-efficient. To conclude, considering the precision and efficiency, the trained VGGNet-13 model and FCN (backbone: VGGNet-13) are applied for the crack recognition procedure.

TABLE 5. Evaluations of the image classification architectures.

| Architectures | AP of test set | Inference time /ms (128 × 128 px) |
|---------------|----------------|----------------------------------|
| VGG-11        | 0.962          | 1.748                            |
| VGG-13        | 0.978          | 2.022                            |
| VGG-16        | 0.954          | 2.431                            |
| VGG-19        | 0.967          | 3.024                            |
| ResNet_V2-18  | 0.948          | 3.368                            |
| ResNet_V2-34  | 0.930          | 5.777                            |
| ResNet_V2-50  | 0.950          | 7.492                            |
| ResNet_V2-101 | 0.952          | 14.237                           |
| MobileNet_V1  | 0.944          | 3.825                            |
| MobileNet_V2  | 0.966          | 7.113                            |
| Inception_V3  | 0.939          | 14.607                           |
| DenseNet-121  | 0.959          | 18.646                           |

VOLUME 8, 2020
operation is applied to link all crack-fragments together to obtain a complete mask of the crack, shown in the white area in Fig. 12d.

Since the skeleton of the test crack image was extracted by the skeleton algorithm, the length of the crack is obtained by counting the continuous pixels of the crack mask in Fig. 12e. Finally, Width of the crack from four positions with the same adjacent interval are obtained, denoted as \( w_1, w_2, w_3 \) and \( w_4 \), shown in Fig. 12d. While the average width of the crack is obtained by Eq. 12.

The same workflow was applied to the other two test images. The actual lengths were obtained from manual inspection reports in the year 2019. The actual widths were obtained from the ground truth mask from manual annotations. As Table 7 shown, the average error was 14.57% for the lengths while 7.91% for the widths. The root-mean-square errors of length and width are 3.0195 mm and 0.003 mm, respectively. It is worthy noted that, the length measurements from the inspection reports is obtained by the DT devices, so the invisible part of a crack could be detected. Thus, it should be greater than the value obtained by our method.

E. DISCUSSION OF DEVIATIONS
Even though the extraction results indicate low deviations in width measuring, certain errors are still obvious in length measuring. In the second test image, the detected length is less than the true value by 15.40%, which reveals that the segmentation masks of the cracks are not perfectly recognized. The errors of the image acquisition and the image recognition models are responsible for the measurement deviations. To be concrete, there are three major error resources should be noted.

- **Errors From the Dynamic Status of a Crack:** Due to the running vehicles above the OSG, the fatigue cracks are not always opened during the crack image acquisition process. When it is slightly closed, the tip of a crack can be invisible on the images. Thus, the panoramic crack image may not display the true ending of a crack, which would lead to deviated results.
- **Errors From the Cracking Identification:** Since the image dataset employed in this study only contains a small number of crack samples, the trained DCNNs can not fully master the generalized features of the crack
VII. CONCLUSION

A novel machine vision-based methodology for monitoring of fatigue cracks in U-rib-to-deck weld seams is proposed in this study. The IoT based image acquisition device is firstly designed and developed to observe the fatigue cracks in the mesoscopic scale. The system consists of three modules and can be easily installed above the seams, but only part-view crack images are obtained. In order to acquire a panoramic crack image, we also propose a novel framework for image rectification and stitching by using a specific coded calibration board. The panoramic crack image can be obtained after specific perspective rectification, image registration and image blending algorithm. Afterward, a cascade crack recognition method is developed, containing crack region detection, crack semantic segmentation, the morphological close operation and the skeleton extraction algorithm. In the end, the length and width of a crack are measured from the segmented mask and skeleton mask. By applying the whole system and algorithms to monitor 14 fatigue cracks in real OSGs bridges, its feasibility and precision region, though the image augmentation methods have been applied to increase the number of trained images.

- **Errors From the Morphological Operations**: Based on the wrong segmented pixel, the linking processing may be invalid, while the distance of the fragments is greater than the kernel size of closing operation. Therefore, the area of a crack may be smaller than the truths.

TABLE 7. The extracted lengths and width of the test images.

| Number | Items | Result value (mm) | True value (mm) | Deviations (mm) | Error (%) |
|--------|-------|-------------------|-----------------|----------------|-----------|
| 1      | l     | 43.10             | 50              | -6.90          | 13.80%    |
| 1      | w₁    | 0.30              | 0.28            | 0.02           | 7.14%     |
| 1      | w₂    | 0.58              | 0.68            | -0.10          | 14.71%    |
| 1      | w₃    | 0.78              | 0.78            | 0.00           | 0.00%     |
| 1      | w₄    | 0.62              | 0.68            | -0.06          | 8.82%     |
| 1      | w_{avg} | 0.47           | -               | -              | -         |
| 2      | l     | 33.84             | 40              | -6.16          | 15.40%    |
| 2      | w₁    | 0.54              | 0.52            | 0.02           | 3.85%     |
| 2      | w₂    | 0.64              | 0.54            | 0.10           | 18.52%    |
| 2      | w₃    | 0.34              | 0.36            | -0.02          | 5.56%     |
| 2      | w₄    | 0.54              | 0.54            | 0.00           | 0.00%     |
| 2      | w_{avg} | 0.43           | -               | -              | -         |
| 3      | l     | 17.10             | 20              | -2.90          | 14.50%    |
| 3      | w₁    | 0.74              | 0.70            | 0.04           | 5.71%     |
| 3      | w₂    | 0.30              | 0.32            | -0.02          | 6.25%     |
| 3      | w₃    | 0.44              | 0.54            | -0.10          | 18.52%    |
| 3      | w₄    | 0.36              | 0.34            | 0.02           | 5.88%     |
| 3      | w_{avg} | 0.41           | -               | -              | -         |

Average error (for Length)  | 14.57% |
Average error (for Width)   | 7.91%  |
Root mean square error (for Length) | 3.0195 |
Root mean square error (for Width) | 0.0030 |
are validated with satisfied results, which indicates that the machine vision-based methodology proposed in this study is useful for fatigue crack detection and monitoring with mesoscopic precision. By analyzing the deviation reasons in measuring, the errors from the superficial status of a crack, the size of training dataset and the morphological operations are discussed as well. In summary, it can be seen that the proposed methodology is effective in the monitoring of the fatigue cracks in U-ribs-to-deck weld seams and the practical management of crackings can benefits a lot from the conclusions above directly, while it is more helpful for the further mechanism analysis of crack propagation.

In the future, the long-term monitoring of cracks in real OSGs bridge will be conducted to achieve the dynamic feature descriptions of fatigue cracks for fatigue studies. The images in the same camera position will be multi-sampled to reduce the error from dynamic status of the cracks. Besides, the effect of DCNN models used in crack recognition will be further explored to improve the measuring precision and their generalization abilities in real application.

ACKNOWLEDGMENT

The authors greatly appreciate the suggestions from the anonymous reviewers.

REFERENCES

[1] J. W. Fisher and C. C. Menzemet. (1990). Fatigue Cracking in Welded Steel Bridges. [Online]. Available: http://onlinepubs.trb.org/Onlinepubs/trr/1990/1282/1282-016.pdf

[2] S. Ya, K. Yamada, and T. Ishikawa, “Fatigue evaluation of Rib-to-Deck welded joints of orthotropic steel bridge deck,” J. Bridge Eng., vol. 16, no. 4, pp. 492–499, Jul. 2011.

[3] B. Helfa, A. Oulhajdi, A. Benbelghit, I. K. LeRafia, F. Boubeneder, and D. Boufassouma, “Detection and measurement of surface cracks in ferromagnetic materials using eddy current testing,” NDT E Int., vol. 39, no. 5, pp. 384–390, Jul. 2006.

[4] R. J. Ditchburn, S. K. Burke, and M. Posada, “Eddy-current nondestructive inspection with thin spiral coils: Long cracks in steel,” J. NDEstruct. Eval., vol. 22, no. 2, pp. 63–77, 2003.

[5] F.-X. Xie, B.-H. Ji, Z.-Y. YuanZhou, Z.-Q. Fu, and H.-B. Ge, “Ultrasound detecting method and repair technology based on fatigue crack features in steel box girder,” J. Perform. Constructed Facilities, vol. 30, no. 2, Apr. 2016, Art. no. 04015006.

[6] T. Watanabe, H. T. H. Trang, K. Harada, and C. Hashimoto, “Evaluation of corrosion-induced crack and rebar corrosion by ultrasonic testing,” Construct. Building Mater., vol. 67, pp. 197–201, Sep. 2014.

[7] C. K. Mukhopadhyay, T. Jayakumar, T. K. Haneef, S. Suresh Kumar, B. P. C. Rao, S. Goyal, S. K. Gupta, V. Bhasin, S. Vishnuvadhan, R. Ragha, and P. Gandhi, “Use of acoustic emission and ultrasonic techniques for monitoring crack initiation/growth during ratcheting studies on 304LN stainless steel straight pipe,” Int. J. Pressure Vessels Piping, vol. 116, pp. 27–36, Apr. 2014.

[8] J. Yu, P. Zhiel, B. Zárate, and J. Caicedo, “Prediction of fatigue crack growth in steel bridge components using acoustic emission,” J. Comput. Civil Eng., vol. 33, no. 9, pp. 1254–1260, Aug. 2011.

[9] M. Tabataeaeipour, J. Hettler, S. Delrue, and K. Van Den Abeele, “Non-destructive ultrasonic examination of root defects in friction stir welded butt-joints,” NDT E Int., vol. 80, pp. 23–34, Jun. 2016.

[10] Y.-L. Zhou, X. Qian, A. Birnie, and X.-L. Zhao, “A reference free ultrasonic phased array to identify surface cracks in welded steel pipes based on transmissibility,” Int. J. Pressure Vessels Piping, vol. 168, pp. 66–78, Dec. 2018.

[11] X. Kong, J. Li, W. Collins, C. Bennett, S. Laflamme, and H. Jo, “Sensing-distortion-induced fatigue cracks in steel bridges with capacitive skin sensor arrays,” Smart Mater. Struct., vol. 27, no. 11, Nov. 2018, Art. no. 115008.

[12] C. Koch, K. Georgiev, V. Kasidreddy, B. Akinci, and P. Fieguth, “A review on computer vision based defect detection and condition assessment of concrete and asphalt civil infrastructure,” Adv. Eng. Informat., vol. 29, no. 2, pp. 196–210, Apr. 2015.

[13] Y. Xu, Y. Bao, J. Chen, W. Zuo, and H. Li, “Surface fatigue crack identification in steel box girders of bridges by a deep fusion convolutional neural network based on consumer-grade camera images,” Struct. Health Monitor., vol. 18, no. 3, pp. 653–674, May 2019.

[14] T. Sakagami, “Remote nondestructive evaluation technique using infrared thermography for fatigue cracks in steel bridges,” Fatigue Fract. Eng. Mater. Struct., vol. 38, no. 7, pp. 755–779, Jul. 2015.

[15] A. Landstrom and M. J. Thurley, “Morphology-based crack detection for steel slabs,” IEEEJ. Sel. Topics Signal Process., vol. 6, no. 7, pp. 866–875, Nov. 2012.

[16] P. Prasanna, K. J. Dana, N. Gucumski, B. B. Basily, H. M. La, R. S. Lim, and H. Parvardeh, “Automated crack detection on concrete bridges,” IEEE Trans. Autom. Sci. Eng., vol. 13, no. 2, pp. 591–599, Apr. 2016.

[17] R. S. Lim, H. M. La, and W. Sheng, “A robotic crack inspection and mapping system for bridge deck maintenance,” IEEE Trans. Autom. Sci. Eng., vol. 11, no. 2, pp. 367–378, Apr. 2014.

[18] A. Ellenberg, A. Kontsos, F. Moon, and I. Bartoli, “Bridge deck delamination identification from unmanned aerial vehicle infrared imagery,” Autom. Construct., vol. 72, pp. 155–165, Dec. 2016.

[19] C. Eschmann and T. Wundsam, “Web-based georeferenced 3D inspection and monitoring of bridges with unmanned aircraft systems,” J. Surv. Eng., vol. 143, no. 3, Aug. 2017, Art. no. 04017003.

[20] W. Ouyang and B. Xu, “Pavement cracking measurements using 3D laser-scan images,” Meas. Sci. Technol., vol. 24, no. 10, Oct. 2013, Art. no. 105204.

[21] D. W. Law, D. Silcock, and L. Holdenh, “Terrestrial laser scanner assessment of deteriorating concrete structures,” Struct. Control Health Monitor., vol. 25, no. 5, e2156, May 2018.

[22] R. Wang and Y. Kawamura, “A magnetic climbing robot for steel bridge inspection,” in Proc. 11th World Congr. Intell. Control Autom., Jun. 2014, pp. 3303–3308.

[23] H. D. Cheng and M. Miyojin, “Automatic pavement distress detection system,” Inf. Sci., vol. 108, nos. 1–4, pp. 219–240, Jul. 1998.

[24] H. Oliveira and P. L. Correia, “Automatic road crack segmentation using entropy and image dynamic thresholding,” in Proc. 17th Eur. Signal Process. Conf., Aug. 2009, pp. 622–626.

[25] H.-B. Yun, S. Mokhtari, and L. Wu, “Crack recognition and segmentation using morphological image-processing techniques for flexible pavements,” Transp. Res. Rec., J. Transp. Res. Board, vol. 2523, no. 1, pp. 115–124, Jan. 2015.

[26] C. M. Yeum and S. J. Dyke, “Vision-based automated crack detection for bridge inspection,” Comput.-Aided Civil Infrastruct. Eng., vol. 30, no. 10, pp. 759–770, Oct. 2015.

[27] M. O’Byrne, B. Ghosh, F. Schoefs, and V. Pakrashi, “Regionally enhanced multiphase segmentation technique for damaged surfaces,” Comput.-Aided Civil Infrastruct. Eng., vol. 29, no. 9, pp. 644–658, Oct. 2014.

[28] I. Abdel-Qader, O. Abdurahim, and M. E. Kelly, “Analysis of edge-detection techniques for crack identification in bridges,” J. Comput. Civ. Eng., vol. 17, no. 4, pp. 255–263, Oct. 2003.

[29] X. Kong and J. Li, “Vision-based fatigue crack detection of steel structures using video feature tracking,” Comput.-Aided Civil Infrastruct. Eng., vol. 33, no. 9, pp. 783–799, Sep. 2018.

[30] E. Zalama, J. Gómez-García-Bermejo, R. Medina, and J. Llamas, “Road crack detection using visual features extracted by Gabor filters,” Comput.-Aided Civil Infrastruct. Eng., vol. 29, no. 5, pp. 342–358, May 2014.

[31] X. Jiang, Z. J. Ma, and W.-X. Ren, “Crack detection from the slope of the mode shape using complex continuous wavelet transform,” Comput.-Aided Civil Infrastruct. Eng., vol. 27, no. 3, pp. 187–201, Mar. 2012.

[32] J. Xiang and M. Liang, “Wavelet-based detection of beam cracks using modal shape and frequency measurements,” Comput.-Aided Civil Infrastruct. Eng., vol. 27, no. 6, pp. 439–454, Jul. 2012.
[33] T. Nishikawa, J. Yoshida, T. Sugiyama, and Y. Fujino, “Concrete crack detection by multiple sequential image filtering,” Comput.-Aided Civil Infrastruct. Eng., vol. 27, no. 1, pp. 29–47, Jan. 2012.

[34] L. Wu, S. Mokhtari, A. Nafez, B. Nam, and H.-B. Yun, “Improvement of crack-detection accuracy using a novel crack fragmentation technique in image-based road assessment,” J. Comput. Civil Eng., vol. 30, no. 1, Jan. 2016, Art. no. 04014118.

[35] M. R. Jahnshahsai, S. F. Masri, C. W. Padgett, and G. S. Sukhatme, “An innovative methodology for detection and quantification of cracks through incorporation of depth perception,” Mach. Vis. Appl., vol. 24, no. 2, pp. 227–241, Feb. 2013.

[36] Y. LeCun, Y. Bengio, and G. Hinton, “Deep learning,” Nature, vol. 521, pp. 436–444, May 2015.

[37] A. Krizhevsky, I. Sutskever, and G. E. Hinton, “ImageNet classification with deep convolutional neural networks,” Commun. ACM, vol. 60, no. 6, pp. 84–90, May 2017, doi: 10.1145/3065386.

[38] J. Schmidhuber, “Deep learning in neural networks: An overview,” Neural Netw., vol. 61, pp. 85–117, Jan. 2015.

[39] F. Chen and M. R. Jahnshahsai, “NB-CNN: Deep learning-based crack detection using convolutional neural network and Naïve Bayes data fusion,” IEEE Trans. Ind. Electron., vol. 65, no. 5, pp. 4392–4400, May 2018.

[40] B. Li, Y. Yuan, W. Zhang, and Y. Yuan, “Unified vision-based methodology for simultaneous concrete defect detection and geolocalization,” Comput.-Aided Civil Infrastruct. Eng., vol. 33, no. 7, pp. 527–544, Jul. 2018.

[41] Y.-J. Cha, W. Choi, and O. Büyükköztürk, “Deep learning-based crack damage detection using convolutional neural networks,” Comput.-Aided Civil Infrastruct. Eng., vol. 32, no. 5, pp. 361–378, May 2017.

[42] S. Ren, K. He, R. Girshick, and J. Sun, “Faster R-CNN: Towards real-time object detection with region proposal networks,” IEEE Trans. Pattern Anal. Mach. Intell., vol. 39, no. 6, pp. 1137–1149, Jun. 2017.

[43] Y.-J. Cha, W. Choi, G. Suh, S. Mahmoudkhani, and O. Büyükköztürk, “Autonomous structural visual inspection using region-based deep learning for detecting multiple damage types,” Comput.-Aided Civil Infrastruct. Eng., vol. 33, no. 9, pp. 731–747, Sep. 2018.

[44] E. Shelhammer, J. Long, and T. Darrell, “Fully convolutional networks for semantic segmentation,” IEEE Trans. Pattern Anal. Mach. Intell., vol. 39, no. 4, pp. 640–651, Apr. 2017.

[45] X. Yang, H. Li, Y. Yu, X. Luo, T. Huang, and X. Yang, “Automatic pixel-level crack detection and measurement using fully convolutional network,” Comput.-Aided Civil Infrastruct. Eng., vol. 33, no. 12, pp. 1090–1109, Dec. 2018.

[46] C. V. Dung and L. D. Anh, “Autonomous concrete crack detection using deep fully convolutional neural network,” Autom. Constr., vol. 99, pp. 52–58, Mar. 2019.

[47] O. Ronneberger, P. Fischer, and T. Brox, “U-Net: Convolutional networks for biomedical image segmentation,” in Proc. Int. Conf. Med. Image Comput. Comput.-Assist. Intervent., 2015, pp. 234–241.

[48] Z. Liu, Y. Cao, Y. Wang, and W. Wang, “Computer vision-based concrete crack detection using U-net fully convolutional networks,” Autom. Constr., vol. 104, pp. 129–139, Aug. 2019.

[49] C. V. Dung, H. Sekiya, S. Hirano, T. Okatani, and C. Miki, “A vision-based method for crack detection in gusset plate welded joints of steel bridges using deep convolutional neural networks,” Autom. Constr., vol. 102, pp. 217–229, Jun. 2019.

[50] F. Ni, J. Zhang, and Z. Chen, “Zernike-moment measurement of thin-crack width in images enabled by dual-scale deep learning,” Comput.-Aided Civil Infrastruct. Eng., vol. 34, no. 5, pp. 367–384, 2019, doi: 10.1111/mice.12421.

[51] J. Gubbio, R. Buyya, S. Marusic, and M. Palaniswami, “Internet of Things (IoT): A vision, architectural elements, and future directions,” Future Gener. Comput. Syst., vol. 29, no. 7, pp. 1645–1660, Sep. 2013.

[52] A. Al-Fuqaha, M. Guizani, M. Mohammadi, M. Aledhari, and M. Ayyash, “Internet of Things: A survey on enabling technologies, protocols, and applications,” IEEE Commun. Surveys Tuts., vol. 17, no. 4, pp. 2347–2376, 4th Quart., 2015.

[53] C. Lazo, P. Gallardo, and S. Cespedes, “A bridge structural health monitoring system supported by the Internet of Things,” in Proc. IEEE Colombian Conf. Comput. Commun. Comput. (IEEE COLCOM ), May 2015, pp. 1–6.
YIQING DONG received the M.E. degree in civil engineering from Tongji University, Shanghai, China, in 2019, where he is currently pursuing the Ph.D. degree in bridge engineering. His current research interests include the health monitoring of bridges based on computer vision, including crack recognition, vehicle recognition, and 3D reconstruction.

YUE PAN received the B.Eng. degree in geographic information systems from Tongji University, Shanghai, China, in 2011, where he is currently pursuing the Ph.D. degree in civil engineering. He is currently a Lecturer with the College of Electronics and Information Engineering, Tongji University, and is a Scientific Assistant with the Research Group of Computer Vision and Artificial Intelligence Application in Bridge Engineering. His research interests are in image processing, 3D reconstruction, the Internet of Things, and deep learning.

RUJIN MA received the Ph.D. degree in bridge and tunnel engineering from Tongji University, China, in 2004. He is an Associate Professor with the Department of Bridge Engineering, Tongji University. His research interests include the life cycle design theory of bridge, bridge risk assessment, bridge performance under extreme events, and maintenance and management of bridge structures. He has taken charge of and participated in many national science foundation projects, and projects of national science and technology supporting plan.