Hindawi
Structural Control and Health Monitoring
Volume 2024, Article ID 1866665, 19 pages
https://doi.org/10.1155/2024/1866665

Research Article

Computer Vision-Based Pier Settlement Displacement Measurement of a Multispan Continuous Concrete Highway Bridge under Complex Construction Environments

Yulin Zhan,[1] **Yuanyuan Huang**,[1] **Zihao Fan**,[1] **Binghui Li**,[1] **Jiawei An**,[1] **Junhu Shao**,[2] **and Yongding Tian**[1]

[1]*School of Civil Engineering, Southwest Jiaotong University, Chengdu 610031, Sichuan, China*
[2]*School of Architecture and Civil Engineering, Chengdu University, Chengdu 610031, Sichuan, China*

Correspondence should be addressed to Yongding Tian; civil_tyd@swjtu.edu.cn

Received 31 July 2023; Revised 28 October 2023; Accepted 8 December 2023; Published 9 January 2024

Academic Editor: Young-Jin Cha

Various concrete bridges have been built across oceans, valleys, and mountains; however, the settlement displacement of bridge piers caused by environmental changes or self-weight during construction phases often leads to uneven stresses, cracking, and eventual collapse. To address the labor-intensive and high-cost issues of pier displacement monitoring using contact-type sensors, this paper proposes an automatic vision-based method for measuring pier settlement displacement under complex construction environments, such as complex image backgrounds, varying ambient light, and camera movement. In the proposed method, a deep learning network was first employed to eliminate the adverse effect of complex image backgrounds and varying ambient light on the accuracy of target detection; then, an adaptive displacement extraction algorithm without a human-computer interaction process was developed to automatically extract the center coordinates of targets attaching to the bridge piers and reference platform; finally, the pier settlement displacement was calculated by using the relative displacements obtained by a dual camera system to eliminate the measurement error caused by camera translation and rotation movements. Laboratory tests of a cantilever beam and field tests of a continuous multispan concrete girder highway bridge under construction have successfully validated the effectiveness and robustness of the developed methodology. The results obtained in this paper can provide some insights for engineers in applying computer vision technology for the real-time monitoring of bridge displacements.

1. Introduction

Various concrete, steel, and composite bridges have been built all around the world to span oceans, mountains, and valleys. Concrete bridges are the most common type of bridge in engineering. Bridge piers are necessary components for transferring traffic loads from the bridge deck to the foundations [1, 2]. The settlement displacement of bridge piers caused by environmental changes or self-weight during construction phases is ubiquitous, which can lead to uneven stresses, cracking, and eventual collapse. Therefore, pier settlement displacement is one of the most important metrics for assessing the construction quality and safety of bridges, and it is essential to be monitored by advanced sensing technologies [3–7]. Current displacement measurement technologies mainly depend on contact sensors, such as displacement transducers and fibre Bragg grating (FBG) sensors [8]. However, sensor installation and data transmission are difficult when the bridge pier is difficult to access. To overcome the challenging problems of contact-type sensing technologies, researchers have developed various noncontact equipment for bridge displacement monitoring, such as total station [9], global position system (GPS) [10], and microwave radar [10–13]. Although these noncontact technologies have been applied to measure structural displacements and pier settlements, they still have some limitations in practical applications.

With the maturity of optical cameras and artificial intelligence, computer vision technology has been recognized as an inexpensive displacement measurement method with the capability of multipoint displacement measurement, high precision, and remote sensing [14–19]. The current computer vision-based displacement extraction algorithms can be roughly divided into image intensity-based methods and phase-based methods. Most of the computer vision-based displacement extraction methods use the image intensity-based method, such as target detection algorithms, edge detection algorithms, feature tracking algorithms, and template matching algorithms. For example, Ye et al. developed a continuous edge detection algorithm for structural deformation measurement by close-range digital photogrammetry system [20]; Feng et al. used a single camera to measure the multipoint dynamic displacement of a simply supported beam by template matching algorithm [21]; Tian et al. used the gradient-based Hough transform (GHT) method to measure the multipoint dynamic displacement of a cantilever beam under static load and impact load [22]; Havaran et al. used the random Hough transform algorithm to track the movement of elliptical markers attached to the structural surface to measure the multipoint displacement of the structure [23]; Tian et al. developed a line segment detection (LSD) and matching algorithm to calculate the dynamic displacements of bridge cables [24]; Shao et al. developed a novel monocular vision system for 3D vibration displacement measurement by using deep neural networks to learn the depth of scenes from captured images [25]. In addition, researchers have developed various phase-based algorithms for structural displacement measurement. Chen et al. developed a phase-based motion magnification algorithm for structural displacement measurement and modal identification of simple structures [26]; Cha et al. developed a phase-based optical flow method for structural displacement measurement and bolt loosening detection using the unscented Kalman filters [27]; Valente et al. quantified the amount of physical motion with the degree of magnification in phase-based displacement extraction method [28]; Shao et al. developed a target-free three-dimensional (3D) tiny displacement measurement method by deep learning and motion magnification technique [29, 30]; Luo et al. proposed a broadband phase-based motion magnification and line tracking algorithm for cable displacement and cable tension force estimation [31]. Although the computer vision-based method has been widely investigated for 2D and 3D displacement measurement of civil infrastructure, there are two challenging problems that need to be addressed when applying this method to complex construction sites.

The first problem is how to remove the negative effect of complex image backgrounds on the accuracy of target detection for subsequent displacement calculation. Due to the complex environment on the bridge construction site, the camera-captured images contain many invalid backgrounds, such as trees, buildings, construction machinery, workers, and so on. These complex backgrounds usually lead to false matching in the target detection and feature tracking processes for displacement calculation. In recent years, several deep learning algorithms have been developed to remove complex backgrounds from captured images to enhance the robustness of the vision-based displacement measurement methods. For instance, Zhang et al. used a pretrained fully convolutional network (FCN) model to remove the complex background information (i.e., pedestrian movement, buildings, trees) contained in the drone-captured video of an urban footbridge for dynamic displacement extraction of the bridge cables with an improved line segment detection algorithm [32]; Cheng et al. used the Yolov4 target detection network to remove the invalid background contained in the captured image of bridge piers; and the displacement trajectory of a bridge pier in the lifting process was extracted by tracking elliptical targets [33]. However, no studies have comprehensively investigated the effect of various complex image backgrounds and varying ambient light on the accuracy of target detection from captured images of bridge piers in complex construction sites with deep learning methods.

The second problem is how to compensate for the measurement error caused by camera movement and avoid the human-computer interaction process for real-time monitoring applications. On the bridge construction site, many uncontrollable factors interfere with the camera, such as the motion caused by the construction machinery, accidental hand touch, and wind-induced vibration, resulting in large errors in the subsequent displacement extraction process. To overcome this problem, researchers have attempted to use stationary objects in the background, such as nearby mountains or buildings, to compensate for the measurement errors caused by camera movement [34–38]. However, it is difficult to find fixed reference points on the bridge construction site in the captured image due to the limited lens and resolution of the camera. On the other hand, computer vision algorithms, such as target detection, template matching, and feature detection algorithms, are usually employed to extract static or dynamic displacements from camera-captured images. In those algorithms, the GHT target detection algorithm has been widely investigated in civil engineering for circular target detection and displacement measurement because of its unique advantage of high accuracy [39]. To cope with the effect of the perspective view on target detection results, a randomized Hough detection algorithm was used to track the movement of ellipse markers attached to the structure for displacement measurement [23]. However, the computational efficiency is seriously affected by the requirement to predefine the radius value when extracting the coordinates of the target center, which makes it difficult to automatically extract the displacement of bridge piers for real-time applications.

Aiming to address the above problems faced by computer vision-based displacement monitoring in complex construction sites, this paper proposes a computer vision-based automatic and highly robust method for the pier displacement measurement of a high bridge pier under construction. The structure of this paper is organized as follows. Section 2 describes the theoretical framework of the proposed method, including the deep learning-based image background removal for accurate target detection, the automatic pier settlement displacement extraction algorithm,

and the camera motion compensation method. Then, laboratory validation of the developed method is presented in Section 3. Subsequently, the robustness and effectiveness of the developed method are further verified by full-scale tests of a long-span bridge under construction in Section 4. Finally, the conclusions are presented in Section 5.

2. Proposed Methodology

2.1. Framework. The framework of the proposed method for displacement measurement of bridge piers under construction from captured video by optical cameras is shown in Figure 1. In the proposed method, a dual-camera system was designed to capture the video of both the bridge pier and the reference target. Then, a deep learning network was trained to remove the unnecessary complex background and varying ambient light contained in the captured video, from which circular targets attached to the bridge pier and reference points were accurately detected. Subsequently, the region of interest (ROI) was determined and mapped to the original image for displacement calculation, in which an adaptive threshold-based GHT algorithm was developed to extract multipoint displacements of the bridge pier and reference points. Finally, the pier settlement displacement is obtained by combining the raw displacements extracted from captured images of the main camera with camera motion calculated from the collected images of the reference target.

2.2. Deep Learning-Based Image Segmentation for Accurate Target Detection. In construction sites, the video recorded by optical cameras usually contains the physical targets, trees, pedestrian movement, and construction machinery, which leads to the difficulty of accurate target detection and displacement extraction. In addition, the varying ambient light affects the intensity distribution of camera-captured images in long-term measurement. Therefore, the U^2-net deep learning architecture was employed to remove the invalid background contained in the recorded video (Figure 2). U^2-Net is a two-level nested U-structure that is designed for object detection without the use of any pretrained backbones from image classification [40]. The U^2-net network consists of a six-stage encoder, a five-stage decoder, a saliency map fusion module that is attached to the decoder stages, and the final encoder stage [40]. The encoder and decoder structures include five residual network structures, namely, RSU-7, RSU-6, RSU-5, RSU-4, and RSU-4F, where the RSU-4F uses dilated convolutions to replace the upsampling and the downsampling. The fusion module then fuses the saliency maps produced by each layer to produce the final predicted probability map.

The loss function in the training process is defined as follows:

$$L = \sum_{m=1}^{M} w_{\text{side}}^{(m)} l_{\text{side}}^{(m)} + w_{\text{fuse}} l_{\text{fuse}}\, L = \sum_{m=1}^{M} w_{\text{side}}^{(m)} l_{\text{side}}^{(m)} + w_{\text{fuse}} l_{\text{fuse}}, \tag{1}$$

where $l_{\text{side}}^{(m)}$ and l_{fuse} are the loss of the side output salience map and the loss of the final fusion output salience map, respectively, and $w_{\text{side}}^{(m)}$ and w_{fuse} are the weights of loss of $l_{\text{side}}^{(m)}$ and l_{fuse}, respectively;

The standard binary cross-entropy is used to calculate each loss term l:

$$l = -\sum_{(r,c)}^{(H,W)} \left[P_{G(r,c)} \log P_{S(r,c)} + \left(1 - P_{G(r,c)}\right) \log\left(1 - P_{S(r,c)}\right)\right], \tag{2}$$

where (H, W) and (r, c) are the pixel coordinates and the height and width size of the captured image, and $P_{G(r,c)}$ and $P_{S(r,c)}$ are the pixel values of the ground truth and the predicted saliency probability map, respectively.

The evaluation metrics max F_β and MAE [41, 42] are used in the U^2-net network training process to assess whether the model has converged. F_β is used to evaluate both the precision and recall.

$$F_\beta = \frac{\left(1 + \beta^2\right) \times \text{precision} \times \text{recall}}{\beta^2 \times \text{precision} + \text{recall}}, \tag{3}$$

where β is a number from 0 to 1, and it is set to 0.3 in this training process. The max F_β is chosen as the evaluation index, and the higher the value obtained, the better the training accuracy obtained.

The mean absolute error (MAE) is calculated to evaluate the average per-pixel difference between a predicted saliency map and its ground truth.

$$\text{MAE} = \frac{1}{H \times W} \sum_{r=1}^{H} \sum_{c=1}^{W} |P(r,c) - G(r,c)|, \tag{4}$$

where $P(r, c)$ is the predicted probability map and $G(r, c)$ is the corresponding ground truth. The lower the value of MAE, the better the training performance.

2.3. Automatic Displacement Extraction and Camera Motion Compensation. After removing the complex image background and varying ambient light effect in the captured video, target-tracking algorithms can be further employed to extract structural displacements. An adaptive threshold-based GHT algorithm was developed to automatically detect the radius of circular targets contained in the captured video and calculate structural displacements, as shown in Figure 3. The basic idea of the developed algorithm is described as follows.

First, the radius of circular targets contained in the captured images is estimated for subsequent automatic displacement calculation, as shown in Figure 3(a). In this step, the Gaussian filter was employed to eliminate various noises contained in the captured images, which is expressed as follows:

$$H[i, j] = \frac{1}{2\pi\sigma^2} e^{-\left((i-k-1)^2 + (j-k-1)^2 / 2\sigma^2\right)}, \tag{5}$$

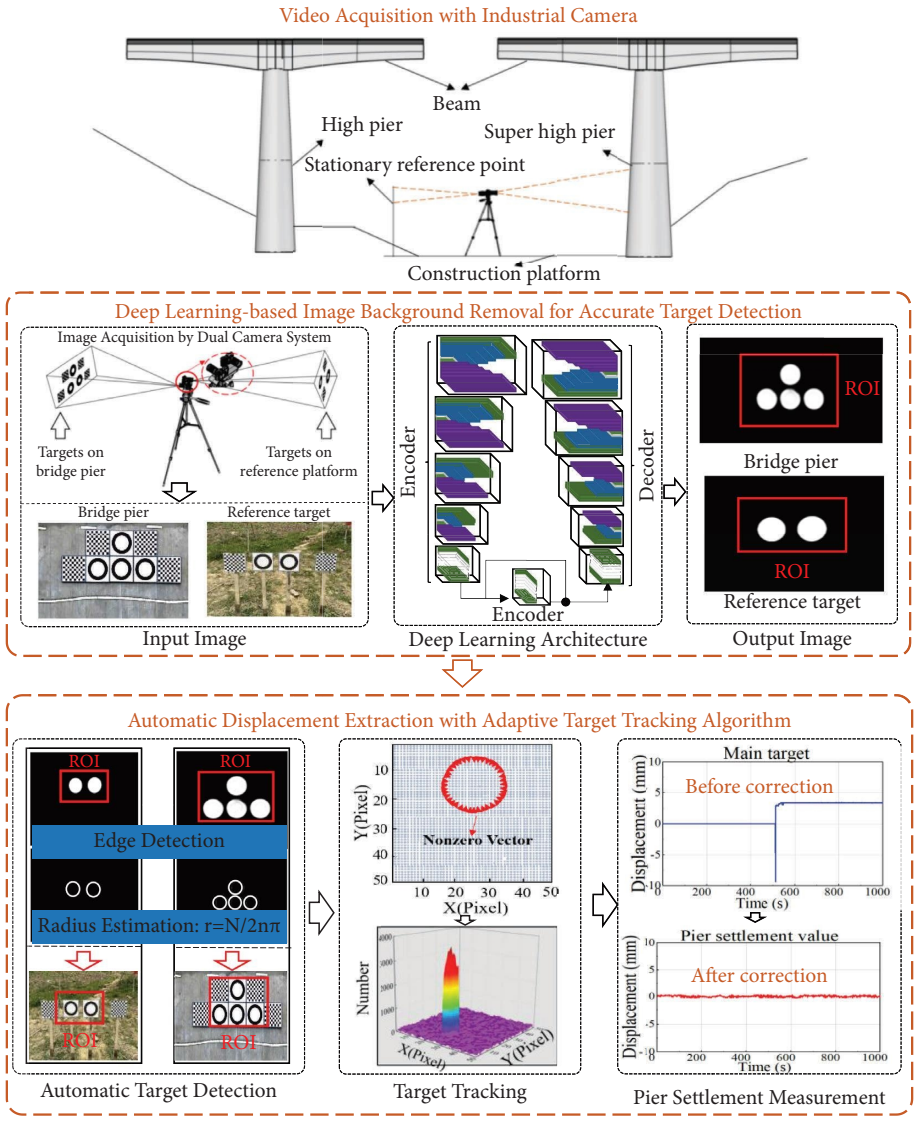


FIGURE 1: Framework of the proposed method.

where k and σ are the dimension size and standard deviation of the Gaussian filter, respectively, and i and j are the coordinates in the X and Y direction.

After applying the Gaussian filter to the raw images, the image gradient in the X and Y direction can be calculated as follows:

$$\mathbf{G}_x = \mathbf{S}_x * \mathbf{I}, \mathbf{G}_y = \mathbf{S}_y * \mathbf{I}, \tag{6}$$

where $\mathbf{G}_x$ and $\mathbf{G}_y$ are the image gradient in the X and Y directions, $\mathbf{S}_x$ and $\mathbf{S}_y$ stands for the Sobel operator, and $\mathbf{I}$ is the intensity matrix of the recorded image.

By combining the image gradients expressed in equation (6), the image gradient at pixel (i, j) is calculated as $G_m(i, j) = \sqrt{G_x(i, j)^2 + G_y(i, j)^2}$. Then, the non-maximum suppression (NMS) algorithm was adopted to eliminate the errors caused by edge detection. The linear interpolation of two adjacent gradient values was calculated as follows:

$$G_{\text{up}}(i, j) = (1 - t)G_m(i, j + 1) + tG_m(i - 1, j + 1), \tag{7a}$$

$$G_{\text{down}}(i, j) = (1 - t)G_m(i, j - 1) + tG_m(i + 1, j - 1), \tag{7b}$$

where $t = |G_y(i, j)|/G_x(i, j)$ is the proportional ratio for gradient value calculation.

If the current gradient value is larger than the computed gradient value in the positive and negative directions, the current pixel point is considered the edge point. Then, the binary images of attached circular targets can be obtained by a defined threshold. After that, the number of pixels with grey values of 1 in the binary image was counted, and the pixel radius value of the circle in each image can be estimated as follows:

$$r_i = \frac{N_i}{2n\pi} (i = 1, 2, 3, \cdots k), \tag{8}$$

where r_i is the radius value (pixel) of the i-th frame of the recorded images, n is the number of circular targets to be detected, N is the number of pixels with a grey value of 1 in the binary image, and k is the number of images.

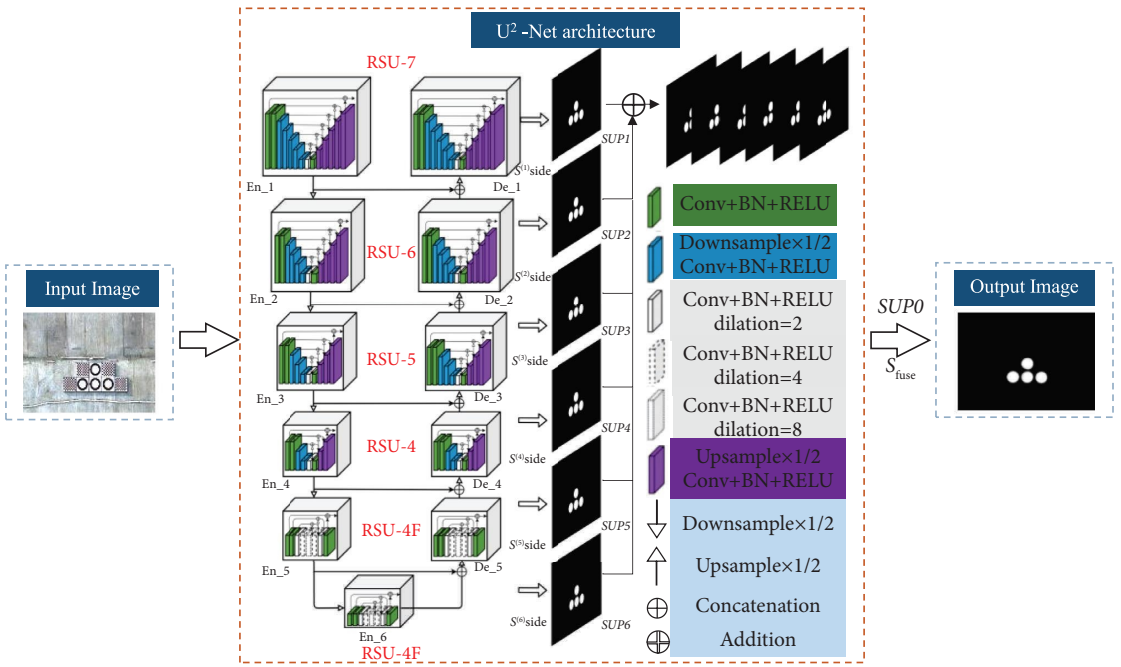


FIGURE 2: Framework of deep learning-based image segmentation for target detection.

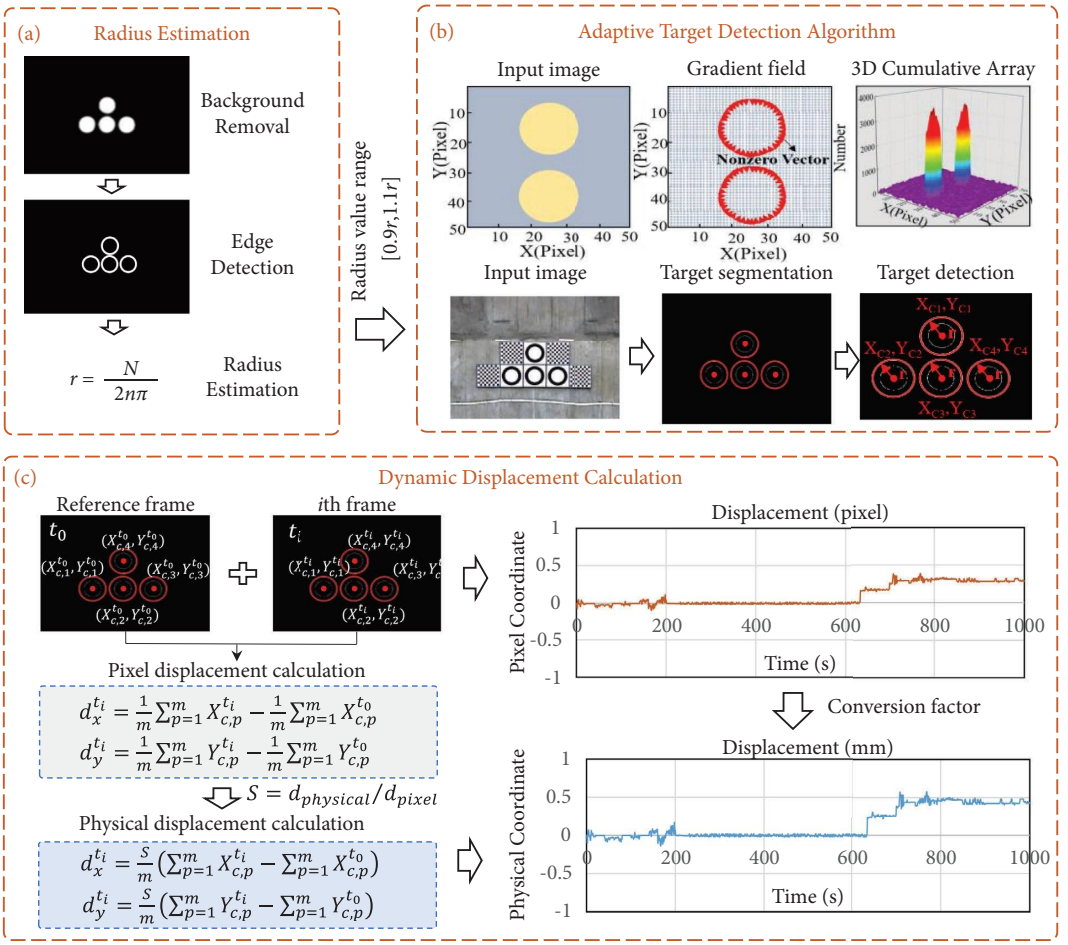


FIGURE 3: Basic principle of the developed algorithm for automatic displacement measurement: (a) radius estimation; (b) circle center detection; and (c) displacement calculation.

After obtaining the radius value of each image, the averaged radius value of all images $r = \sum_1^k r_i/k$ was calculated as the final results for the subsequent calculation. An optimized range with the calculated radius value was used to automatically detect possible circular targets in the gradient Hough transform algorithm, as shown in Figure 3(b). In this step, a discrete characteristic curve is defined on the image gradient field to determine the circle center coordinates in the image.

Assume that the image sequences corresponding to the time $t_0, t_1, \cdots t_n$ are collected, and the GHT algorithm with an optimized radius threshold is applied to all the captured images to automatically detect the circle center. Then, the two-dimensional displacements of the bridge piers and reference targets in pixel coordinates can be calculated by subtracting the circle coordinates of subsequent images from those of the reference image, as shown in Figure 3(c).

$$d_x^{t_i} = \frac{1}{m}\sum_{p=1}^{m} X_{c,p}^{t_i} - \frac{1}{m}\sum_{p=1}^{m} X_{c,p}^{t_0}, \tag{9a}$$

$$d_y^{t_i} = \frac{1}{m}\sum_{p=1}^{m} Y_{c,p}^{t_i} - \frac{1}{m}\sum_{p=1}^{m} Y_{c,p}^{t_0}, \tag{9b}$$

where $d_x^{t_i}$ and $d_y^{t_i}$ are the vertical and horizontal displacements of the target at the time t_i; $X_{c,p}^{t_i}$ and $X_{c,p}^{t_0}$ are the circle center of the target p at the time t_i and t_0 in the horizontal direction respectively; $Y_{c,p}^{t_i}$ and $Y_{c,p}^{t_0}$ are the circle center of the target p at the time t_i and t_0 in vertical direction, respectively; p is the number of artificial targets attached to the bridge surface; and m is the total number of the attached targets.

Finally, the pixel displacements need to be converted into physical displacements using the pixel-to-displacement conversion factor. In this study, the known target dimensions in the captured video are used to calculate the conversion factor, $S = d_{\text{physical}}/d_{\text{pixel}}$.

The above equation can extract the raw displacements of the bridge piers and the reference target from the captured video, but the extracted displacement contains the measurement error caused by the camera movement, such as the vertical translation motion and rotation motion in the vertical and horizontal direction, respectively. To address this issue, a dual-camera system was used to compensate for the effect of camera motion on extracted displacements, as shown in Figure 4. If the support platform has a vertical translation movement (shown in Figure 4(a)), the true displacement of the pier settlement is calculated by subtracting the raw vertical displacement of the main target attaching to the bridge pier from that of the reference target. The calculation formula is expressed as follows:

$$d_y^{t_i} = d_{y,C_1}^{t_i} S_{C_1} + \Delta d = d_{y,C_1}^{t_i} S_{C_1} - d_{y,C_2}^{t_i} S_{C_2}, \tag{10}$$

where $d_y^{t_i}$ is the relative displacement obtained by the dual camera system; $d_{y,C_1}^{t_i}$ and $d_{y,C_2}^{t_i}$ denote the vertical raw displacement of the bridge pier and reference target, respectively; S_{C_1} and S_{C_2} are the conversion factors of camera 1 and camera 2, respectively; and C_1 and C_2 stand for the main camera and secondary camera for video recording of bridge piers and reference targets, respectively.

If the support platform has a rotation movement in the vertical direction (as shown in Figure 4(b)), the true displacement of the pier settlement can be calculated by the following equation:

$$d_y^{t_i} = d_{y,C_1}^{t_i} S_{C_1} + \left(d_{y,C_2}^{t_i} S_{C_2}\right)\frac{L_1}{L_2} = d_{y,C_1}^{t_i} S_{C_1} + \left(d_{y,C_2}^{t_i} S_{C_2}\right)\frac{S_{C_1}}{S_{C_2}} = d_{y,C_1}^{t_i} S_{C_1} + d_{y,C_2}^{t_i} S_{C_1}. \tag{11}$$

If the support platform has a rotation movement in the horizontal direction (shown in Figure 4(c)), the false displacements induced by the two cameras are the same; therefore, the true pier settlement is calculated as follows:

$$d_y^{t_i} = d_{y,C_1}^{t_i} S_{C_1} + d_{y,C_2}^{t_i} S_{C_2}. \tag{12}$$

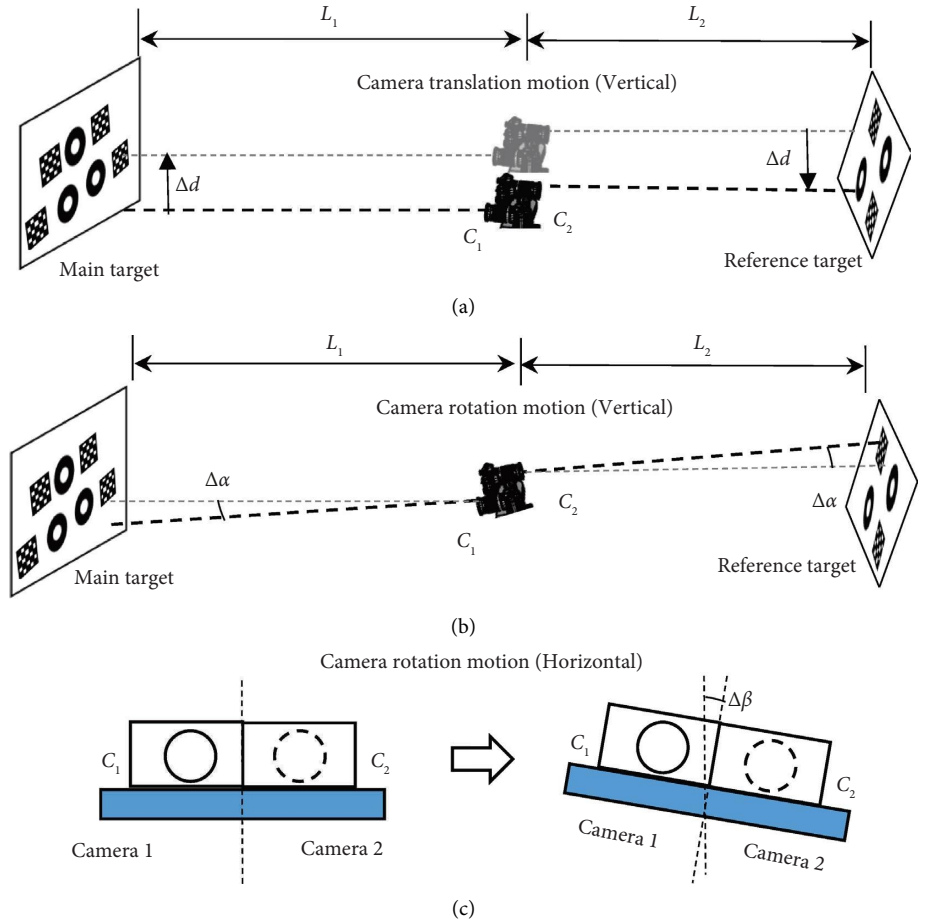


FIGURE 4: Schematic diagram of different camera motion: (a) translation motion; (b) rotation motion in vertical direction; (c) rotation motion in horizontal direction.

Substituting the displacement calculation formula in equation (9b) into (10)–(12) results in the final expression for pier settlement calculation by considering three types of camera movements, namely,

$$d_y^{t_i} = \frac{S_{C_1}}{m}\left(\sum_{p=1}^{m} Y_{c,p}^{t_i} - \sum_{p=1}^{m} Y_{c,p}^{t_0}\right) - \frac{S_{C_2}}{n}\left(\sum_{p=1}^{n} Y_{c,p}^{t_i} - \sum_{p=1}^{n} Y_{c,p}^{t_0}\right), \tag{13a}$$

$$d_y^{t_i} = \frac{S_{C_1}}{m}\left(\sum_{p=1}^{m} Y_{c,p}^{t_i} - \sum_{p=1}^{m} Y_{c,p}^{t_0}\right) + \frac{S_{C_1}}{n}\left(\sum_{p=1}^{n} Y_{c,p}^{t_i} - \sum_{p=1}^{n} Y_{c,p}^{t_0}\right), \tag{13b}$$

$$d_y^{t_i} = \frac{S_{C_1}}{m}\left(\sum_{p=1}^{m} Y_{c,p}^{t_i} - \sum_{p=1}^{m} Y_{c,p}^{t_0}\right) + \frac{S_{C_2}}{n}\left(\sum_{p=1}^{n} Y_{c,p}^{t_i} - \sum_{p=1}^{n} Y_{c,p}^{t_0}\right), \tag{13c}$$

where m and n denote the total number of attached targets on the bridge pier and reference platform, respectively. It should be noted that equations (13a)–(13c) are used to eliminate the vertical translation movement and rotation movement in the vertical and horizontal direction, respectively.

3. Laboratory Testing of a Cantilever Beam

3.1. Experimental Setup. The effectiveness of the proposed methodology is first verified by laboratory testing of a cantilever beam. The experimental setup of the cantilever beam under investigation is shown in Figure 5. The cantilever beam has a length of 2 m, and the height of the cross-section is 0.4 m. A total of six circular targets were attached to the cantilever beam for displacement extraction, and six displacement gauges were also installed at the corresponding position to verify the accuracy of the proposed method. In this experiment, a high-speed camera with a resolution of 1024×1024 pixels and a frame rate of 1000 Hz was used to capture image sequences of the circular targets when the cantilever beam was excited by an impact hammer.

3.2. Experimental Results. After acquiring the images of the investigated beam under impact loading, the U^2-net was used for image background removal. Then, the developed displacement extraction algorithm was used to detect circular targets from the preprocessed images. Figure 6(a) shows the circular target detection results of the cantilever beam without complex background segmentation with a radius range of $[0.8r, 1.2r]$ in the conventional gradient

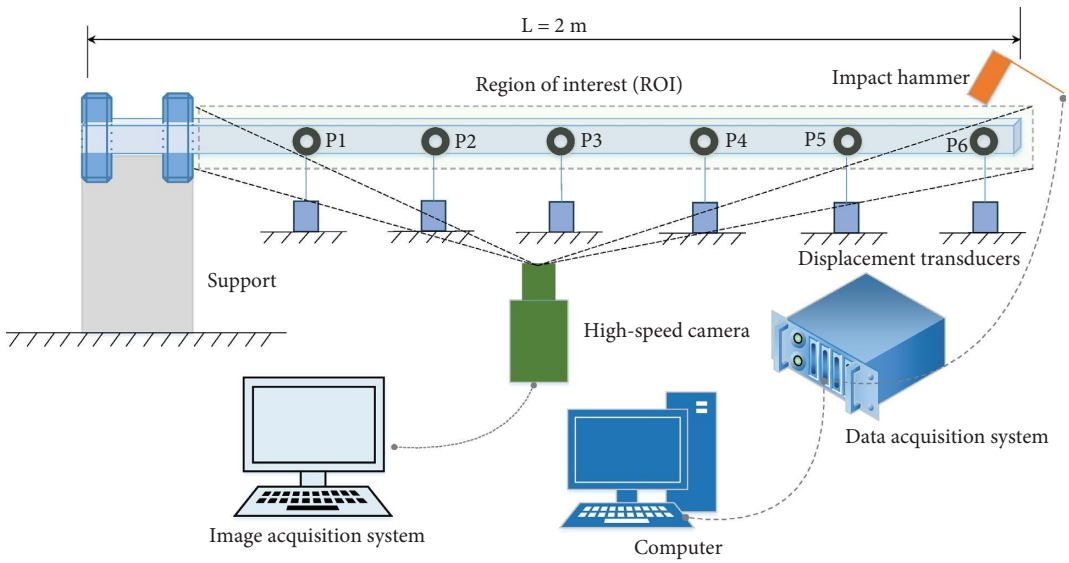


Figure 5: Experimental setup.

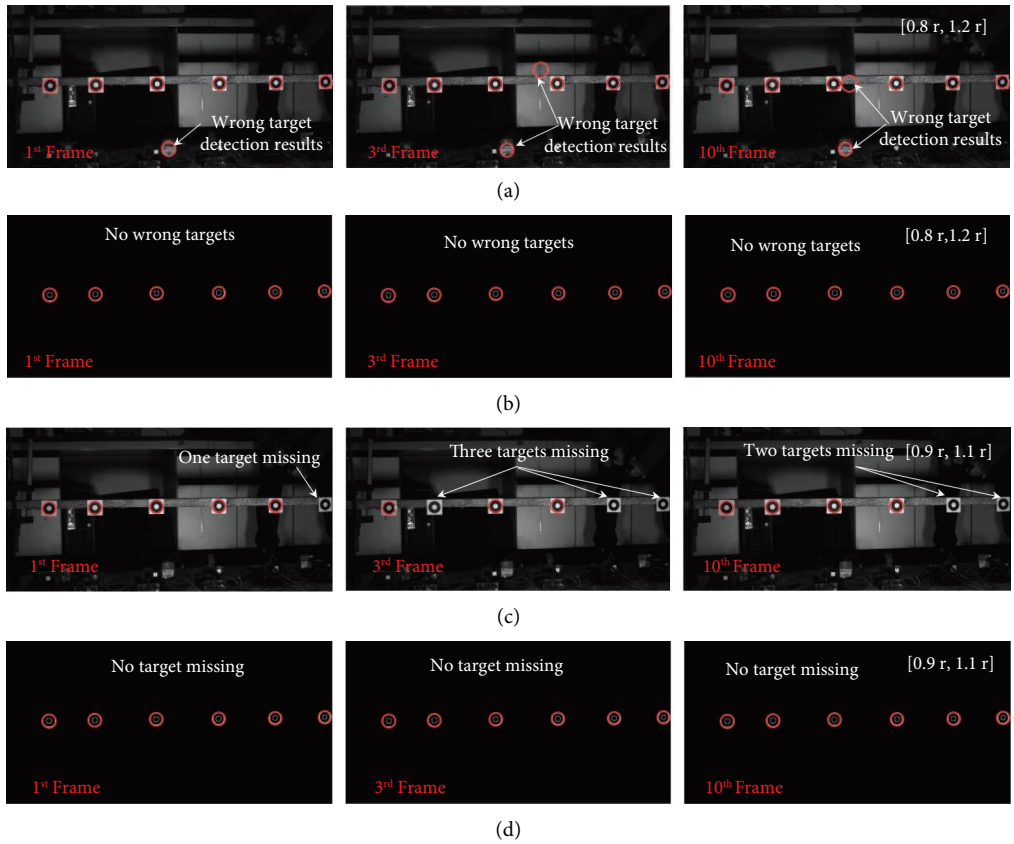


Figure 6: Target detection results of the studied cantilever beam. (a) Target detection without background removal with a radius of $[0.8r, 1.2r]$; (b) target detection with background removal with a radius of $[0.8r, 1.2r]$; (c) target detection without background removal with a radius of $[0.9r, 1.1r]$; (d) target detection with background removal with a radius of $[0.9r, 1.1r]$.

Hough transform algorithm. The conventional method requires a human-computer interface by selecting the region of interest (ROI) and artificially determining the radius for circular center detection. It can be seen that incorrect target detection results occur in the process of circular target detection with a larger radius range of $[0.8r, 1.2r]$.

In comparison, there are no false circular target detection results when the raw image is processed by the proposed method (Figure 6(b)). However, the detected

circular targets are missing when the circle radius is set to a smaller range $[0.9r, 1.1r]$, as shown in Figure 6(c). Specifically, one, three, and two circular targets are lost in the first, third, and tenth frames of the recorded images of the studied cantilever beam, making it impossible to calculate displacements at these targets. However, in the preprocessed images using the proposed method, there are no missing targets (Figure 5(d)). It is concluded that incorrect target detection results and target-missing phenomena occur in the process of circular target detection by the conventional method due to the complex image backgrounds. In addition, the change in the radius range affects the circular target detection results due to the invalid image background.

After the image background segmentation, the developed automatic displacement extraction algorithm is used to extract the dynamic displacements of the six circular targets attached to the cantilever beam. The extracted displacements are compared with the directly measured displacements collected by the displacement meters to verify the accuracy of the developed algorithm (Figure 7). It can be seen that the extracted displacements by the developed algorithm are in good agreement with the directly measured displacements by the displacement meter, thus verifying the correctness of the developed algorithm.

Furthermore, three metrics—the correlation coefficient (ρ), the root mean square error (RMSE), and the coefficient of determination (R^2)—are used to quantitatively evaluate the measurement accuracy of the developed algorithm.

$$\begin{aligned} \mathrm{RMSE} &= \sqrt{\frac{\sum_i \left(x_v(i) - x_c(i)\right)^2 / n}{n}}, \\ \rho &= \frac{\left|\sum_i \left(x_v(i) - \mu_v\right) \times \left(x_c(i) - \mu_c\right)\right|}{\sqrt{\sum_i \left(x_v(i) - \mu_v\right)^2} \times \sqrt{\sum_i \left(x_c(i) - \mu_c\right)^2}}, \\ R^2 &= 1 - \frac{\sum_i \left(x_v(i) - x_c(i)\right)^2}{\sum_i \left(x_c(i) - \mu_c\right)^2}, \end{aligned} \tag{14}$$

where x_v and x_c are displacements obtained by the developed method and the displacement meters, respectively, n represents the total number of captured images, μ_v and μ_c are the average values of calculated displacement time histories, ρ stands for the correlation degree between two kinds of displacements, and R^2 represents the matching degree of two curves.

The specific values of the three metrics between the displacements calculated by the developed algorithm and the results measured by the displacement meters are given in Table 1. It can be seen that the displacements calculated by the developed method are in good agreement with the displacements measured by the displacement meters, and the maximum RMSE of the two displacement curves is 0.0613, the minimum ρ and R^2 are 0.9198 and 0.8954, respectively, validating the accuracy of the developed algorithm for displacement extraction. In addition, the target missing rate by using the developed method and the traditional method was also compared. It is seen that measurement points 1 to 6 have different proportions of the target missing rate in the traditional method, whereas there is no target missing in the proposed method, verifying the correctness of the developed method.

4. Field Testing of a Long-Span Highway Bridge

To verify the robustness and stability of the proposed method in complex environments, field tests were carried out on a continuous multispan long-span highway bridge under complex construction environments. This section provides details of the field tests and the results of the pier settlement displacement of the investigated bridge.

4.1. Bridge Description. The bridge under investigation, called the Lugou River Bridge, is a continuous rigid frame bridge, as shown in Figure 8. The superstructure is a prestressed continuous rigid-frame with a cast-in-place cantilever construction method; the substructure adopts double-legged solid piers and hollow thin-walled piers; and the foundation is a group pile foundation. Furthermore, the bridge is a separate two-line bridge with a main bridge span arrangement of $(96 + 5 \times 180 + 96)$ m, a deck width of 16.25 m, and a maximum bridge height of 209.57 m. The cross-section of the main girder is a single box girder, and the height of the box girder is 4.0 m at the center of the span and 11.5 m at the top of the pier.

4.2. Overview of Field Tests. To minimize the measurement error caused by the camera movement during the pier settlement monitoring, a dual camera system was designed for the measurement, using two optical cameras (MV-CA050-10GC) mounted on a tripod with a resolution of 2448×2048 pixels, as shown in Figure 9(a). Image sequences of the bridge pier (main target) and the reference point (subtarget) were acquired simultaneously, with the reference point being a fixed platform chosen to suit the site conditions. The main target at the bridge pier and the subtarget at the reference point are shown in Figures 9(b) and 9(c), respectively. The physical radius of the outer circle of the circular target used in the field test is 100 mm, and the pixel size is approximately 60. The camera is roughly located in the middle of the main target and subtarget, about 30 meters away from both sides.

In the field test, three test conditions were designed, as shown in Table 2. Specifically, case 1 is set up to validate the effectiveness of the camera motion compensation, and cases 2 and 3 are used to validate the feasibility of the proposed method in monitoring pier settlement displacement.

4.3. Monitoring Results

4.3.1. Complex Background Removal for Accurate Target Detection. Before extracting the pier settlement displacement, the deep learning-based image background removal method is used to remove the invalid background contained in the image. The training database contains images with circular targets and common backgrounds (i.e., pedestrians, buildings, trees, etc.). In the training phase, the initial

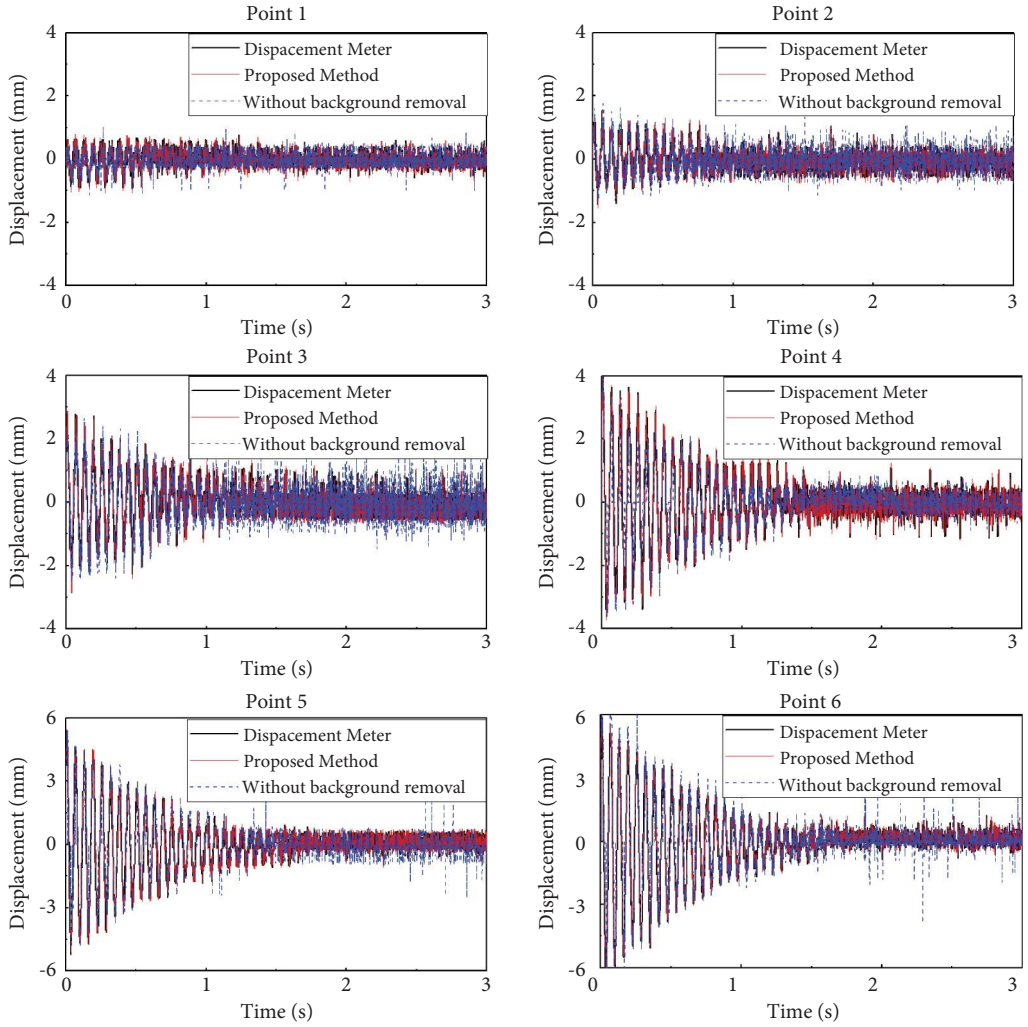


Figure 7: Displacement calculation of the cantilever beam with the developed automatic algorithm.

Table 1: Performance metrics of the ATGHT algorithm.

Measurement points	Proposed method				Traditional method			
	R^2	RMSE	ρ	Target missing rate (%)	R^2	RMSE	ρ	Target missing rate (%)
1	0.8954	0.0604	0.9198	0	0.7650	0.4229	0.7004	0
2	0.9243	0.0609	0.9210	0	0.6151	0.4508	0.7293	0.1
3	0.9432	0.0611	0.9326	0	0.7687	0.4289	0.8156	0
4	0.9648	0.0613	0.9372	0	0.5184	0.6348	0.7876	2.1
5	0.9724	0.0604	0.9433	0	0.7467	0.6303	0.9117	1.2
6	0.9841	0.0603	0.9501	0	0.8220	0.6396	0.9327	11

learning rate, the number of training epochs, and the batch size are set to 0.001, 50, and 2, respectively. The segmentation results of circular targets after model training are shown in Figure 10. It can be seen that the U^2-net model can effectively detect circular targets in various complex backgrounds.

The loss function curve during model training is shown in Figure 11(a), which shows that the training loss and tar loss decrease rapidly with training epochs and then fluctuate around 0.08 and 0.01, respectively. In addition, the evaluation metrics are also shown in Figure 11(b), where max F_β gradually increases and stabilizes at 0.96, and MAE gradually decreases and stabilizes at 0.005.

The circular target detection results of the bridge pier without and with complex image background segmentation are shown in Figure 12. It can be seen that false circular target detection results occur in the captured images, as shown in Figures 12(a) and 12(c). Several false circular targets are detected in the original image when the radius

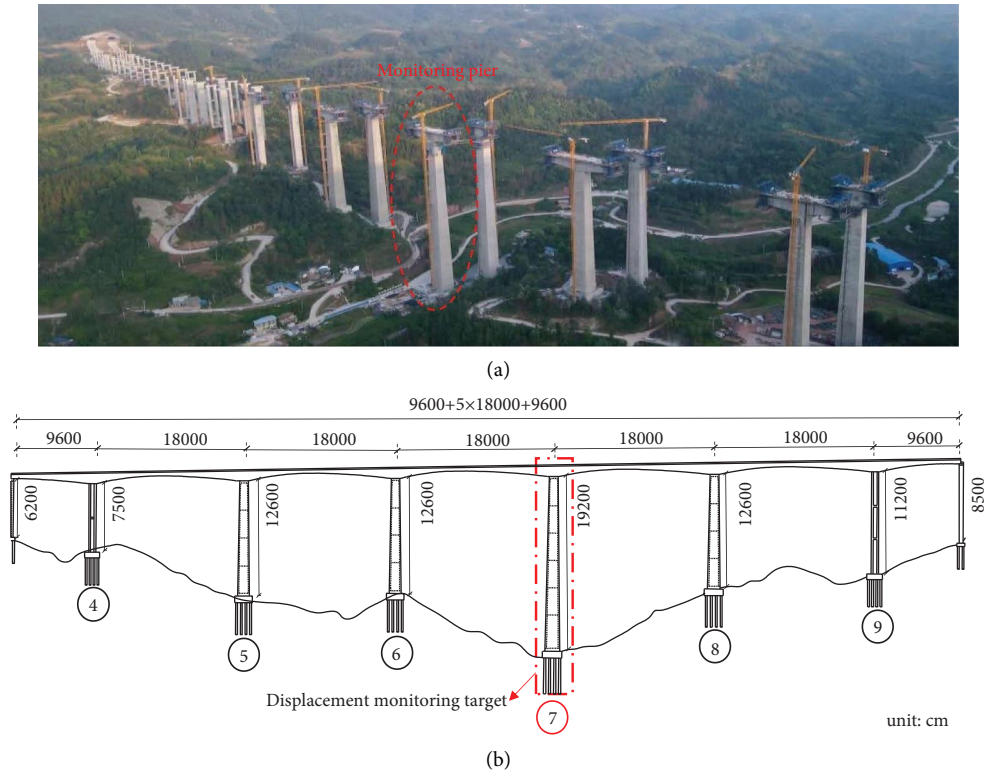


Figure 8: The studied Lugou river bridge: (a) panoramic view and (b) elevation view.

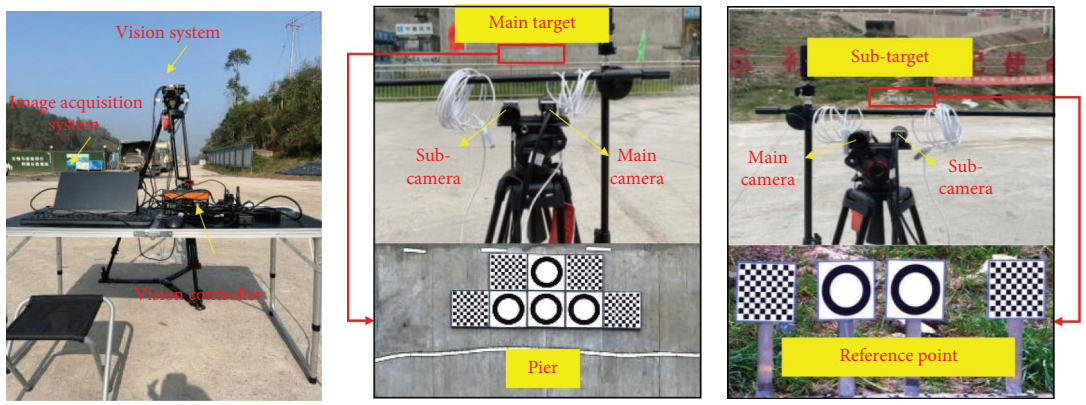


Figure 9: Field testing setup: (a) image acquisition system; (b) main target at the pier; (c) subtarget at the reference point.

Table 2: Field test conditions.

Test conditions	Testing contents
Case 1	Pier settlement monitoring under artificial disturbance
Case 2	Pier settlement monitoring during lifting of hanging baskets
Case 3	Pier settlement monitoring during concrete pouring

value is set to a larger range (Figure 12(a)), while the circular targets are lost when the radius value is set to a smaller range (Figure 12(c)). This is due to the inaccurate results of circle radius estimation by manual selection with randomness and subjectivity. On the contrary, after the image background segmentation, the radius of the circular targets to be detected can be estimated automatically, and the change in the radius range does not affect the circular target detection results (Figures 12(b) and 12(d)). This example shows that the background segmentation to remove the invalid background

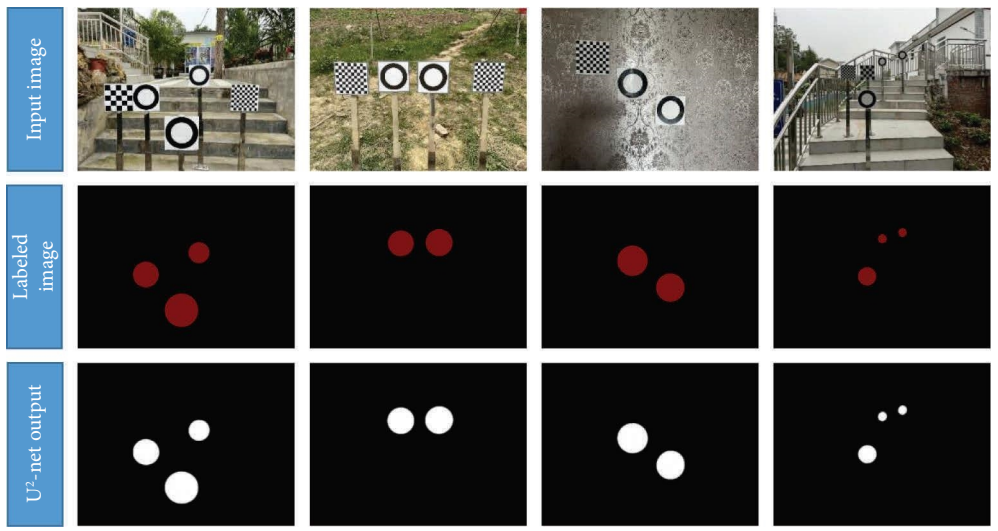


Figure 10: Validation of U^2-net training.

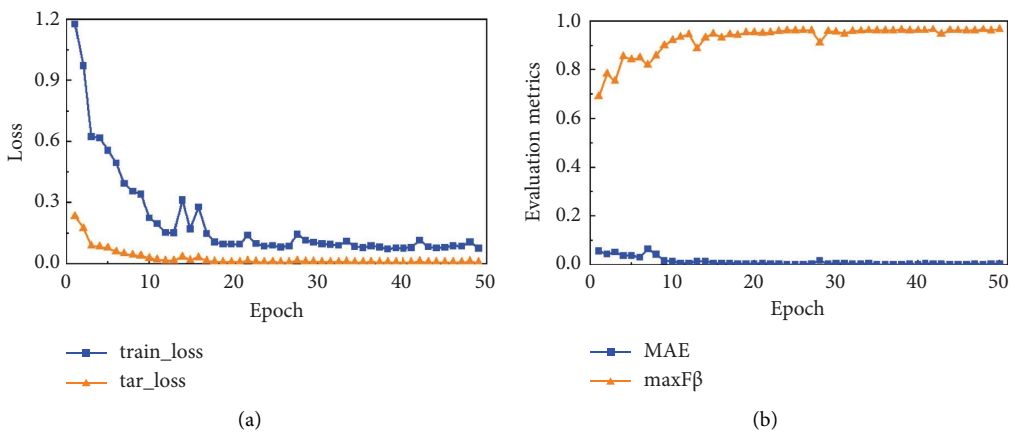


Figure 11: Train process for complex background removal by U^2-net: (a) training loss curve; (b) evaluation metrics.

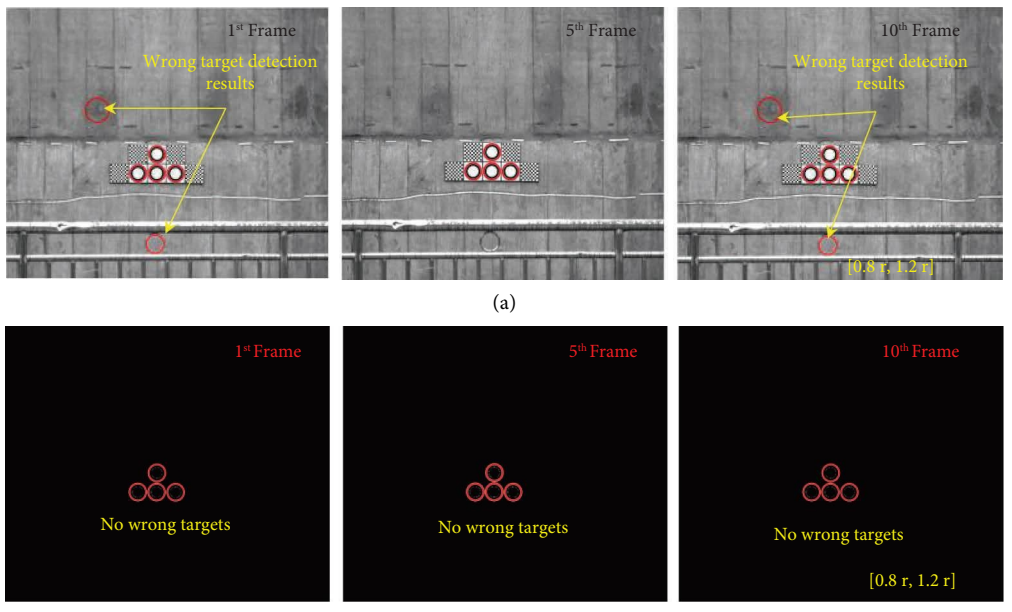


Figure 12: Continued.

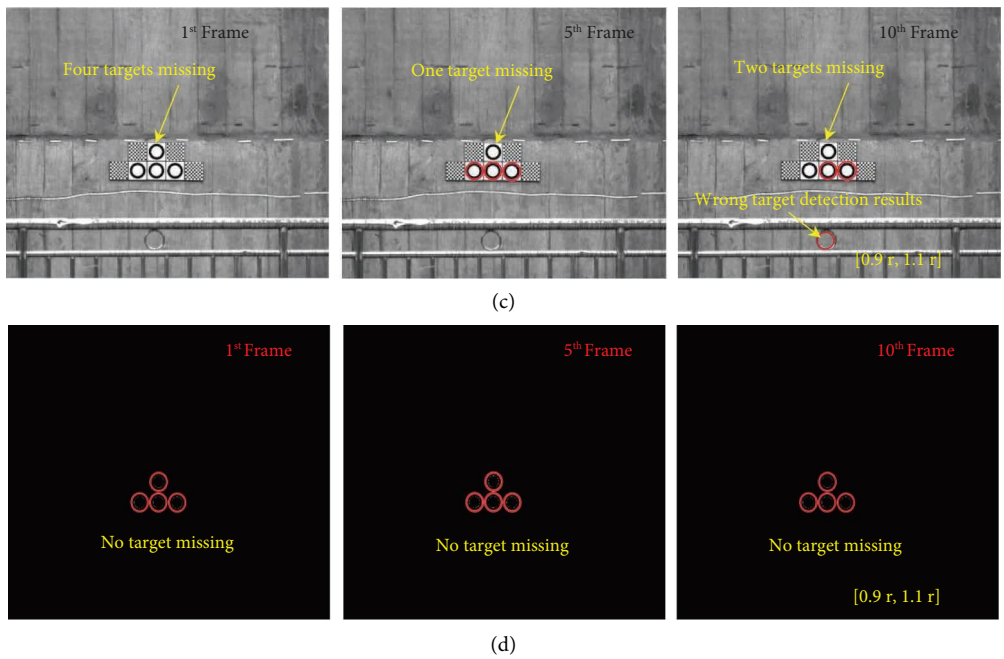


Figure 12: Target detection results of the studied bridge pier: (a) target detection without background removal with a radius of [0.8r, 1.2r]; (b) target detection with background removal with a radius of [0.8r, 1.2r]; (c) target detection without background removal with a radius of [0.9r, 1.1r]; (d) target detection with background removal with a radius of [0.9r, 1.1r].

not only improves the robustness of the displacement extraction but also avoids the human-computer interaction process, which greatly improves the computational efficiency.

4.3.2. Pier Settlement Displacement Monitoring under Complex Environments

(1) Case 1: Pier Settlement Monitoring under Artificial Disturbance. To verify the effectiveness of the camera motion compensation method, three groups of experiments were designed in Case 1, where the camera was artificially disturbed 1, 2, and 3 times, respectively, during the measurement process. The experimental setup and the recorded images are shown in Figure 13. It can be seen that the brightness of the captured images varies with the measurement time except for the artificial disturbances.

Then, the developed automatic displacement extraction algorithm is applied to the segmented circular targets. The extracted displacements of the bridge pier and the true pier displacement are shown in Figure 14. There is an error of about −10 mm after a single artificial disturbance, and then the error returns to 4 mm. The proposed camera motion compensation algorithm was applied to process the extracted raw displacements, and the maximum measurement error was reduced to 0.09 mm.

The results of pier settlement monitoring under two artificial disturbances are shown in Figure 14(b). It can be seen that the camera system produces measurement errors of approximately 3 mm and 9 mm under the first and second disturbances. After applying the camera motion compensation algorithm, the maximum measurement error is reduced to 0.15 mm. Figure 14(c) shows the measurement error of the camera under three times of disturbances, the maximum measurement error is up to 11 mm but is reduced to 0.13 mm by the proposed method. The averaged absolute displacement values under one, two, and three times of disturbance by the proposed method are 0.015 mm, 0.2750 mm, and 0.088 mm, respectively. In contrast, the averaged absolute displacement values in the three cases without background removal are 0.595 mm, 0.791 mm, and 0.559 mm, respectively, which is much larger than the results obtained by the proposed method. It demonstrates the high accuracy of the developed method for displacement measurement. It can be seen that the measured displacements have obvious deviations under the camera disturbance, but the proposed method can reduce the measurement error. It should be noted that the accuracy of the measured displacements is affected by the pixel/physical size scale and camera noise. The systematic investigation of the displacement measurement errors will be conducted in future work.

(2) Case 2: Pier Settlement Monitoring during Lifting of Hanging Baskets. The pier settlement was monitored using the developed dual camera system when the main beam basket was lifted, and the captured images are shown in Figure 15. It is seen that the brightness of the captured image varies with the natural lighting conditions and even contains shadows in the captured image of the reference target at the time of 6400 s. Therefore, the complex image background needs to be removed by the deep learning method. The circular targets mounted on the bridge pier and the reference target are accurately detected regardless of various complex image backgrounds, as shown in Figures 15(a) and 15(b).

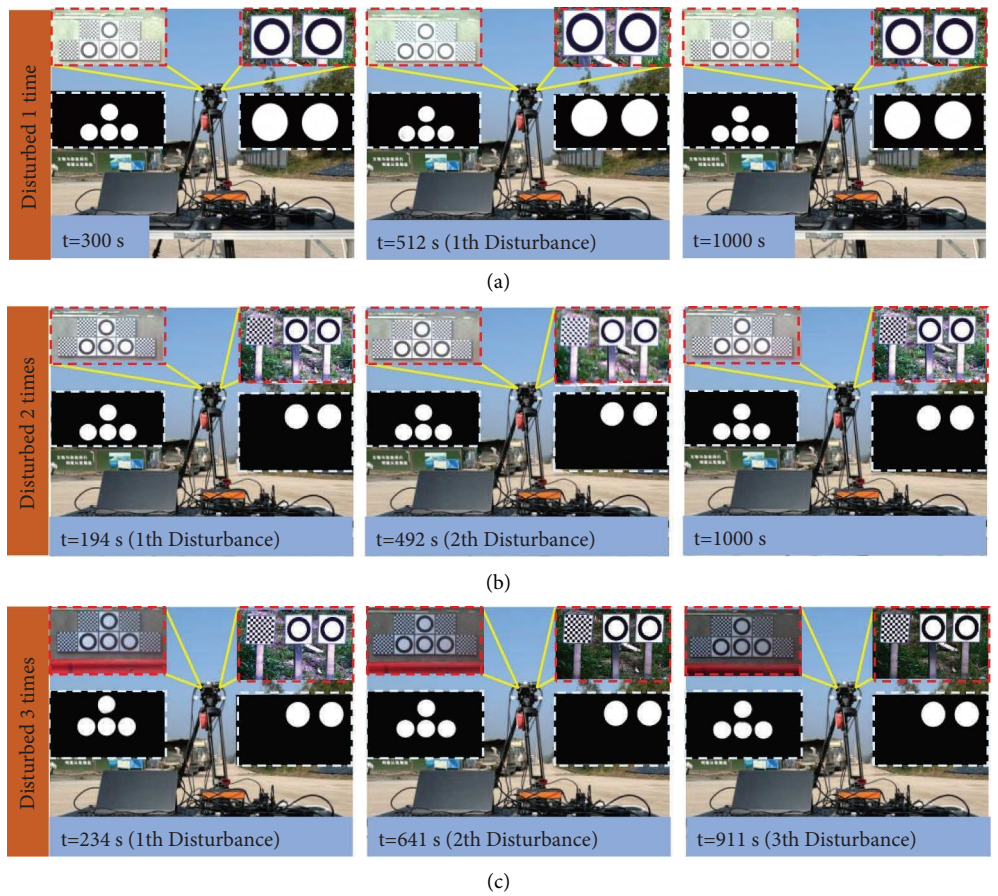


FIGURE 13: Captured images of the bridge pier and reference target in case 1 (a) disturbed 1 time; (b) disturbed 2 times; (c) disturbed 3 times.

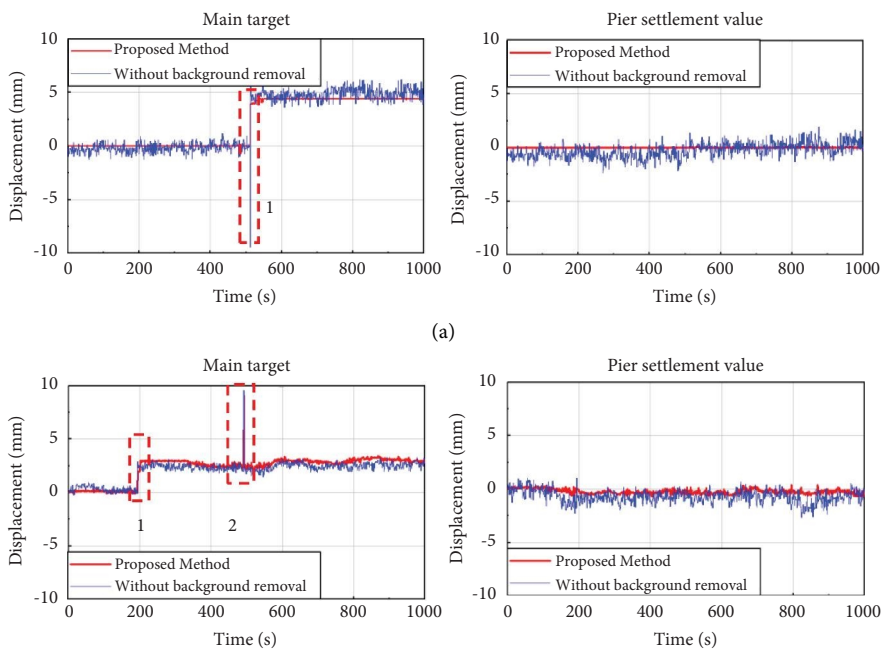
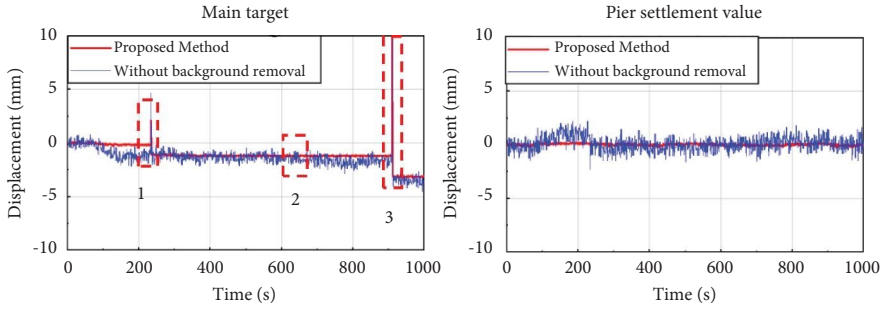
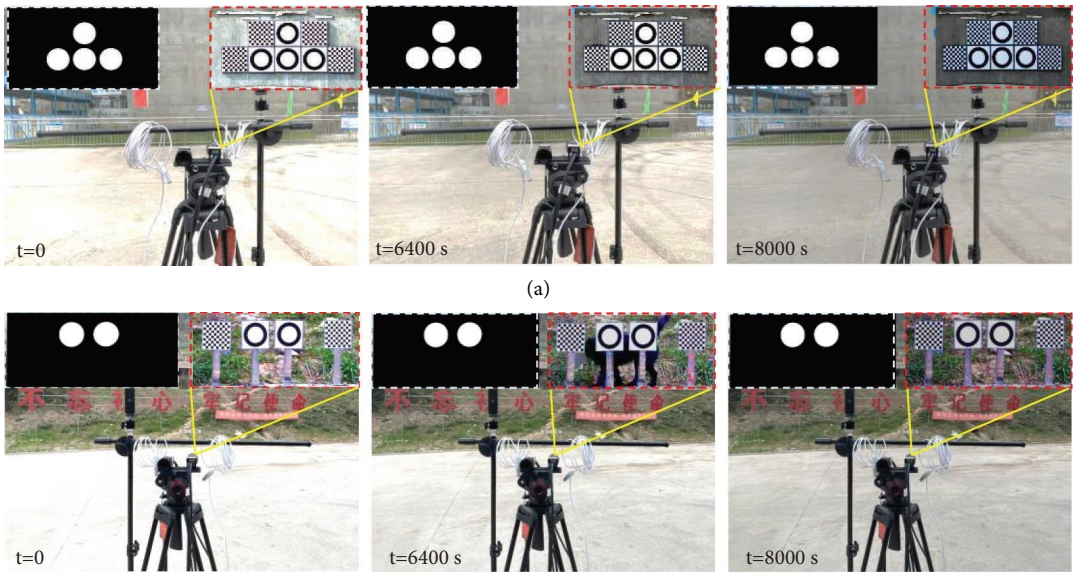


FIGURE 14: Continued.



(c)

Figure 14: Pier settlement monitoring results of case 1 (a) disturbed 1 time; (b) disturbed 2 times; (c) disturbed 3 times.



(a)

(b)

Figure 15: Captured images of the studied bridge in case 2 with complex image background and varying ambient light: (a) bridge pier; (b) reference target.

The extracted displacements of the bridge pier in Case 2 are shown in Figure 16. The results show that the camera was slightly disturbed during the lifting process of the hanging basket (Figures 16(a) and 16(b)), resulting in measurement errors. The true pier settlement displacement by the camera motion compensation method is shown in Figure 16(c). It can be seen that the pier settlement displacement fluctuates around 0 mm, and there is no obvious settlement tendency. As the bridge foundation is a deep pile foundation with a length of 40 m, there will be no settlement of the piers when the hanging basket is lifted.

(3) Case 3: Pier Settlement Monitoring during Concrete Pouring. The extracted displacement of the bridge pier during the concrete concreting process of segment 1 of the main beam is shown in Figure 17. It is seen that the extracted raw displacement contains large spikes without the deep learning-based image segmentation method (Figures 17(a) and 17(b)). The pier settlement displacement during the test period is divided into two stages, which gradually increase from 0 mm and stabilize at 0.15 mm, as shown in Figure 17(c). During the pouring process of segment 1, the pier settlement increases by about 0.15 mm due to the gradual increase of the concrete wet load. This is because the ratio of the weight of the hanging basket to the weight of the poured concrete segment is generally 0.3 to 0.5 in the actual project. Therefore, it is reasonable that the settlement value of the pier in Case 3 is relatively large. The test results of both Cases 2 and 3 have verified the robustness of the proposed method for pier settlement monitoring of long-span bridges.

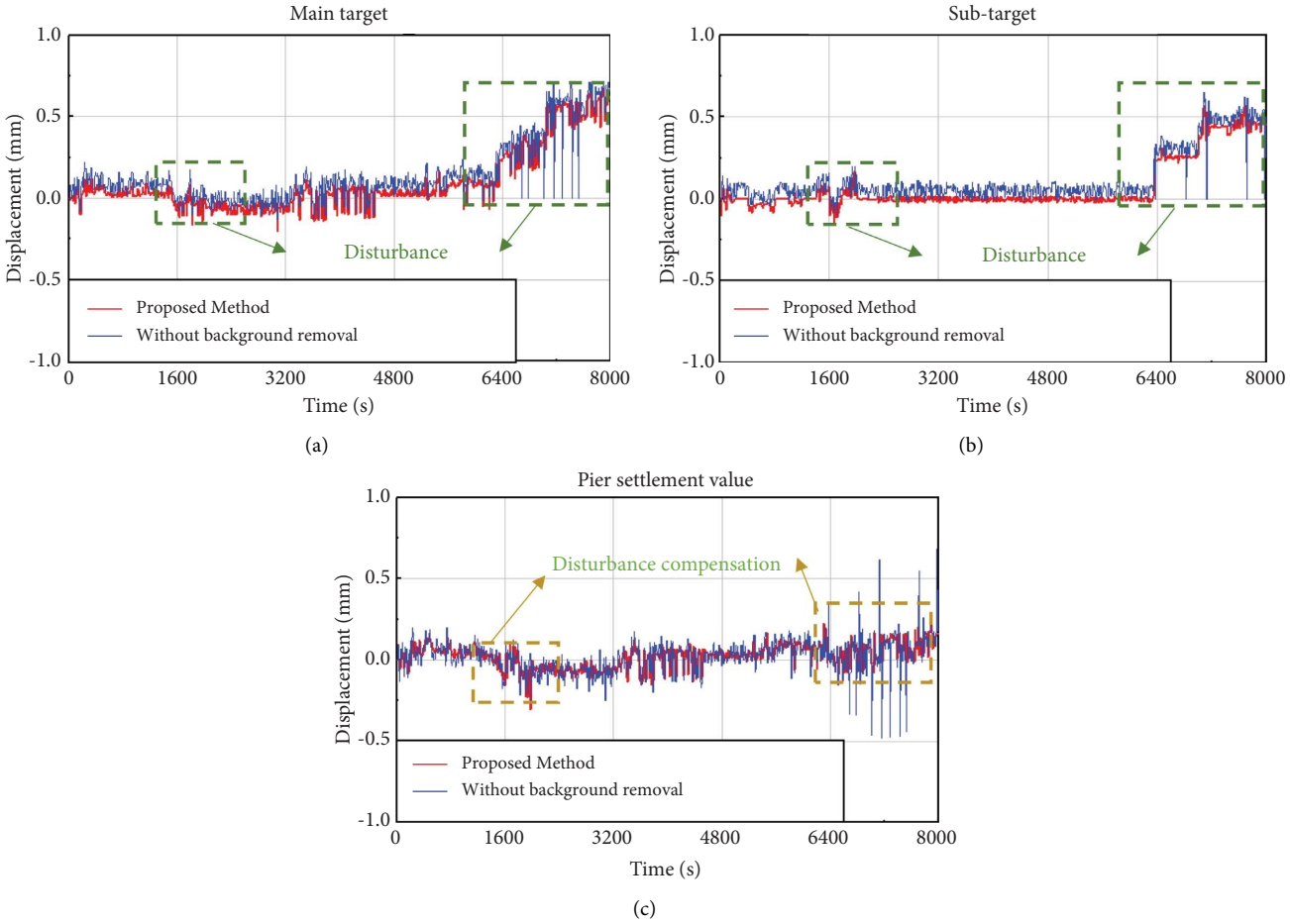


Figure 16: Displacement results of case 2. (a) Raw displacement of the pier; (b) displacement at the reference point; (c) true pier settlement displacement.

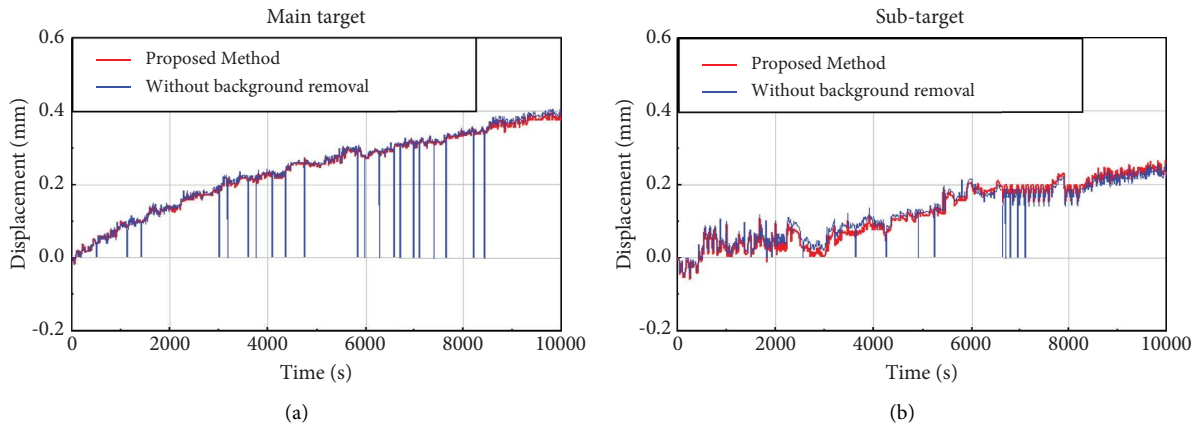
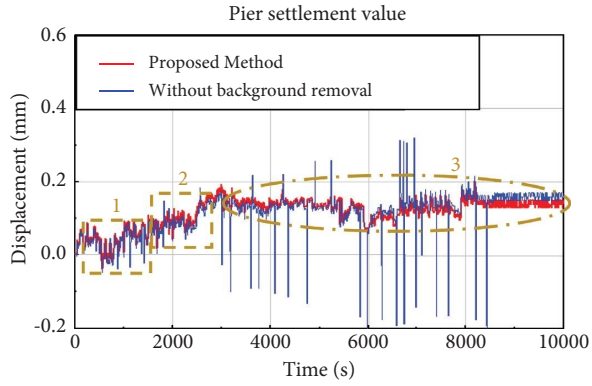


Figure 17: Continued.



(c)

Figure 17: Displacement results of case 3. (a) Raw displacement of the bridge pier; (b) raw displacement of the reference point; (c) true pier settlement value.

5. Conclusions and Future Works

The complex image background, varying ambient light, and camera movement affect the robustness of the computer vision-based method for displacement monitoring, and the human-computer interaction process in the conventional method prevents automatic and real-time displacement extraction. To overcome the above problems, a computer vision-based automatic and highly robust pier settlement measurement method was developed using deep learning technologies. The detailed results are as follows:

(1) A deep learning-based complex background removal method was used to eliminate the effect of invalid backgrounds (i.e., pedestrian movement, trees, construction machinery, etc.) and the varying ambient light on the accuracy of target detection. Then, an adaptive threshold-based GHT algorithm was developed to accurately and automatically calculate pier settlement displacement. The relative displacement of the bridge pier concerning the reference target was considered as the true pier displacement to compensate for the camera movement induced by construction machinery on the construction site and accidental hand contact by workers.

(2) Laboratory tests on a cantilever beam were carried out to verify the accuracy and robustness of the developed algorithm. The test results show that the displacements extracted by the proposed method are in good agreement with those directly measured by displacement gauges. The RMSE of the two types of displacement curves is 0.0613, and the minimum correlation coefficient (ρ) and determination coefficient (R^2) are 0.9198 and 0.8954, respectively, verifying the correctness of the developed method.

(3) To validate the effectiveness and robustness of the proposed method, field tests were carried out on a long-span, high-pier highway bridge under construction. The results show that the proposed method can effectively reduce the measurement error caused by the camera movement and eliminate the adverse effects of complex image backgrounds and varying ambient light on the measured displacements. The pier settlement displacements obtained by the proposed method under three types of experimental conditions are consistent with the actual engineering project.

In conclusion, the proposed method has great potential for use in noncontact measurement of pier settlement in harsh environments. Future work will focus on the development of a portable computer vision system for real-time monitoring of bridge displacement. In addition, a stereo vision system and an advanced monocular vision system will be developed to measure the three-dimensional displacements with high accuracy. The effect of the pixel/physical size scale, camera noise, and ambient temperature on the accuracy of the displacement measurement in field tests also needs to be systematically investigated.

Data Availability

Data are available upon reasonable request to the author.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

The authors are grateful for the financial support from the National Natural Science Foundation of China (Grant no. 52308330), the Fundamental Research Funds for the Central Universities (Grant no. 2682022CX077), and the Sichuan Science and Technology Program (Grant No. 2023NSFSC0893).

References

[1] Z. W. Chen and W. M. Zhai, "Theoretical method of determining pier settlement limit value for China's high-speed railway bridges considering complete factors," *Engineering Structures*, vol. 209, Article ID 109998, 2020.

[2] Y. D. Tian, C. Chen, K. Sagoe-Crentsil, J. Zhang, and W. H. Duan, "Intelligent robotic systems for structural health monitoring: applications and future trends," *Automation in Construction*, vol. 139, Article ID 104273, 2022.

[3] B. F. Spencer, V. Hoskere, and Y. Narazaki, "Advances in computer vision-based civil infrastructure inspection and monitoring," *Engineering*, vol. 5, no. 2, pp. 199–222, 2019.

[4] J. H. Deng, A. Singh, Y. Zhou, Y. Lu, and V. C. Lee, "Review on computer vision-based crack detection and quantification methodologies for civil structures," *Construction and Building Materials*, vol. 356, Article ID 129238, 2022.

[5] D. Feng and M. Q. Feng, "Computer vision for SHM of civil infrastructure: from dynamic response measurement to damage detection – a review," *Engineering Structures*, vol. 156, pp. 105–117, 2018.

[6] W. B. Chen, B. F. Yan, J. B. Liao, L. Luo, and Y. Dong, "Cable force determination using phase-based video motion magnification and digital image correlation," *International Journal of Structural Stability and Dynamics*, vol. 22, no. 07, Article ID 2250036, 2022.

[7] Y. Xu and J. Brownjohn, "Review of machine-vision based methodologies for displacement measurement in civil structures," *Journal of Civil Structural Health Monitoring*, vol. 8, no. 1, pp. 91–110, 2018.

[8] Z. Yi, G. Jinshen, and W. Xu, "Intelligent settlement monitoring system of high-speed railway bridge," *Journal of Civil Structural Health Monitoring*, vol. 9, no. 3, pp. 307–323, 2019.

[9] W. J. Merkle and J. J. Myers, "Use of the total station for load testing of retrofitted bridges with limited access," *Smart Structures and Materials 2004: Sensors and Smart Structures Technologies for Civil, Mechanical, and Aerospace Systems*, vol. 5391, pp. 687–694, 2004.

[10] S. B. Im, S. Hurlebaus, and Y. J. Kang, "Summary review of GPS technology for structural health monitoring," *Journal of Structural Engineering*, vol. 139, no. 10, pp. 1653–1664, 2013.

[11] C. Gentile and G. Bernardini, "An interferometric radar for non-contact measurement of deflections on civil engineering structures: laboratory and full-scale tests," *Structure and Infrastructure Engineering*, vol. 6, no. 5, pp. 521–534, 2010.

[12] B. C. Zhang, X. L. Ding, C. Werner et al., "Dynamic displacement monitoring of long-span bridges with a microwave radar interferometer," *International Society for Photogrammetry and Remote Sensing Journal of Photogrammetry and Remote Sensing*, vol. 138, pp. 252–264, 2018.

[13] G. W. Zhang, Y. L. Wu, W. J. Zhao, and J. Zhang, "Radar-based multipoint displacement measurements of a 1200-m-long suspension bridge," *International Society for Photogrammetry and Remote Sensing Journal of Photogrammetry and Remote Sensing*, vol. 167, pp. 71–84, 2020.

[14] S. Bhowmick, S. Nagarajaiah, and Z. Lai, "Measurement of full-field displacement time history of a vibrating continuous edge from video," *Mechanical Systems and Signal Processing*, vol. 144, Article ID 106847, 2020.

[15] Y. J. Cha, W. Choi, and O. Büyüköztürk, "Deep learning-based crack damage detection using convolutional neural networks," *Computer-Aided Civil and Infrastructure Engineering*, vol. 32, no. 5, pp. 361–378, 2017.

[16] D. Kang and Y. J. Cha, "Autonomous UAVs for structural health monitoring using deep learning and an ultrasonic beacon system with geo-tagging," *Computer-Aided Civil and Infrastructure Engineering*, vol. 33, no. 10, pp. 885–902, 2018.

[17] B. Ferrer, P. Acevedo, J. Espinosa, and D. Mas, "Targetless image-based method for measuring displacements and strains on concrete surfaces with a consumer camera," *Construction and Building Materials*, vol. 75, pp. 213–219, 2015.

[18] Y. Zhang, P. Liu, and X. F. Zhao, "Structural displacement monitoring based on mask regions with convolutional neural network," *Construction and Building Materials*, vol. 267, Article ID 120923, 2021.

[19] Y. J. Cha, K. You, and W. Choi, "Vision-based detection of loosened bolts using the Hough transform and support vector machines," *Automation in Construction*, vol. 71, pp. 181–188, 2016.

[20] J. Ye, G. K. Fu, and U. P. Poudel, "Edge-based close-range digital photogrammetry for structural deformation measurement," *Journal of Engineering Mechanics*, vol. 137, no. 7, pp. 475–483, 2011.

[21] D. M. Feng and M. Q. Feng, "Experimental validation of cost-effective vision-based structural health monitoring," *Mechanical Systems and Signal Processing*, vol. 88, pp. 199–211, 2017.

[22] Y. D. Tian, J. Zhang, and S. S. Yu, "Vision-based structural scaling factor and flexibility identification through mobile impact testing," *Mechanical Systems and Signal Processing*, vol. 122, pp. 387–402, 2019.

[23] A. Havaran and M. Mahmoudi, "Markers tracking and extracting structural vibration utilizing Randomized Hough transform," *Automation in Construction*, vol. 116, Article ID 103235, 2020.

[24] Y. D. Tian, C. Zhang, S. Jiang, J. Zhang, and W. H. Duan, "Noncontact cable force estimation with unmanned aerial vehicle and computer vision," *Computer-Aided Civil and Infrastructure Engineering*, vol. 36, no. 1, pp. 73–88, 2021.

[25] Y. Shao, L. Li, J. Li, Q. L. Li, S. An, and H. Hao, "Monocular vision based 3D vibration displacement measurement for civil engineering structures," *Engineering Structures*, vol. 293, Article ID 116661, 2023.

[26] J. G. Chen, N. Wadhwa, Y. J. Cha, F. Durand, W. T. Freeman, and O. Buyukozturk, "Modal identification of simple structures with high-speed video using motion magnification," *Journal of Sound and Vibration*, vol. 345, pp. 58–71, 2015.

[27] Y. J. Cha, J. G. Chen, and O. Büyüköztürk, "Output-only computer vision based damage detection using phase-based optical flow and unscented Kalman filters," *Engineering Structures*, vol. 132, pp. 300–313, 2017.

[28] N. A. Valente, C. T. do Cabo, Z. Mao, and C. Niezrecki, "Quantification of phase-based magnified motion using image enhancement and optical flow techniques," *Measurement*, vol. 189, Article ID 110508, 2022.

[29] Y. Shao, L. Li, J. Li, S. J. An, and H. Hao, "Target-free 3D tiny structural vibration measurement based on deep learning and motion magnification," *Journal of Sound and Vibration*, vol. 538, Article ID 117244, 2022.

[30] Y. Shao, L. Li, J. Li, S. J. An, and H. Hao, "Computer vision based target-free 3D vibration displacement measurement of structures," *Engineering Structures*, vol. 246, Article ID 113040, 2021.

[31] K. Luo, X. Kong, X. Y. Wang, T. J. Jiang, G. T. Frøseth, and A. Rønnquist, "Cable vibration measurement based on broad-band phase-based motion magnification and line tracking algorithm," *Mechanical Systems and Signal Processing*, vol. 200, Article ID 110575, 2023.

[32] C. Zhang, Y. D. Tian, and J. Zhang, "Complex image background segmentation for cable force estimation of urban bridges with drone-captured video and deep learning,"

Structural Control and Health Monitoring, vol. 29, no. 4, Article ID e2910, 2022.

[33] Y. Y. Cheng, F. Z. Lin, W. G. Wang, and J. Zhang, "Vision-based trajectory monitoring for assembly alignment of precast concrete bridge components," *Automation in Construction*, vol. 140, Article ID 104350, 2022.

[34] I. Rodriguez-Padilla, B. Castelle, V. Marieu, and D. Morichon, "A simple and efficient image stabilization method for coastal monitoring video systems," *Remote Sensing*, vol. 12, no. 1, p. 70, 2019.

[35] J. G. Chen, A. Davis, N. Wadhwa, F. Durand, W. T. Freeman, and O. Büyüköztürk, "Video camera–based vibration measurement for civil infrastructure applications," *Journal of Infrastructure Systems*, vol. 23, no. 3, Article ID B4016013, 2017.

[36] H. Yoon, J. Shin, and B. F. Spencer, "Structural displacement measurement using an unmanned aerial system," *Computer-Aided Civil and Infrastructure Engineering*, vol. 33, no. 3, pp. 183–192, 2018.

[37] Z. Zhang, "Flexible camera calibration by viewing a plane from unknown orientations," in *Proceedings of the seventh IEEE international conference on computer vision*, pp. 666–673, IEEE, Kerkyra, Greece, September 1999.

[38] V. Hoskere, J. W. Park, H. Yoon, and B. F. Spencer, "Vision-based modal survey of civil infrastructure using unmanned aerial vehicles," *Journal of Structural Engineering*, vol. 145, no. 7, Article ID 04019062, 2019.

[39] M. M. Wang, W. K. Ao, J. Bownjohn, and F. Y. Xu, "A novel gradient-based matching via voting technique for vision-based structural displacement measurement," *Mechanical Systems and Signal Processing*, vol. 171, Article ID 108951, 2022.

[40] X. Qin, Z. Zhang, C. Huang, M. Dehghan, O. R. Zaiane, and M. Jagersand, "U2-Net: going deeper with nested U-structure for salient object detection," *Pattern Recognition*, vol. 106, Article ID 107404, 2020.

[41] N. Liu, J. Han, and M. H. Yang, "PiCANet: learning pixel-wise contextual attention for saliency detection," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 3089–3098, Melbourne, Australia, June 2018.

[42] L. Zhang, J. Dai, H. Lu, Y. He, and G. Wang, "A bi-directional message passing model for salient object detection," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 1741–1750, Boston, Massachusetts, June 2018.