



Assessment and monitoring of bridges using various camera placements and structural analysis

Yongsheng Bai^{1,2} · Aydin Demir^{1,3} · Alper Yilmaz¹ · Halil Sezen¹

Received: 7 April 2023 / Accepted: 3 October 2023
© Springer-Verlag GmbH Germany, part of Springer Nature 2023

Abstract

This study presents a new vision-based deep learning method to monitor and evaluate the structural health of in-service infrastructure. For this purpose, three different camera placements, including remote, structure-mounted, and drone-mounted cameras, are proposed to capture the vibrations or displacements of bridges. The vision-based deep learning method is verified by an optical flow approach. Various techniques, such as visual data denoising and camera motion removal, are utilized to process the test data for displacement measurements and extract the structural frequencies. Structural models of bridges are analyzed to validate the measurements and assess the structural health of several pedestrian, traffic, and railway bridges without interfering with traffic. Measurements in the field experiments and results from the structural analysis on tested bridges show that the proposed framework works successfully and can be potentially engineered to monitor the structural health of existing bridges.

Keywords Camera placements · Deep learning · HR Mask R-CNN · Optical flow · UAVs · Structural analysis · Structural health monitoring

1 Introduction

Bridges are essential to commute, travel and perform other activities resulting in external loads. These structures are designed to be strong enough to support these applied loads

Yongsheng Bai, Aydin Demir, Alper Yilmaz and Halil Sezen have contributed equally to this work.

✉ Halil Sezen
sezen.1@osu.edu

Yongsheng Bai
bai.426@osu.edu

Aydin Demir
demir.19@osu.edu; aydindemir@sakarya.edu.tr

Alper Yilmaz
yilmaz.15@osu.edu

¹ Civil, Environmental and Geodetic Engineering, The Ohio State University, 2070 Neil Avenue, Columbus, OH 43210, USA

² Neural Image Corporation, 1 Lake Bellevue, Bellevue, WA 98005, USA

³ Department of Civil Engineering, Sakarya University, Universite Cd., 54050 Serdivan, Sakarya, Turkey

and must have sufficient stiffness to avoid excessive deformations or vibrations. Their dynamic performance can be evaluated using their strength, stiffness, mass, damping, and other properties with machine learning tools [1, 2] and various sensors [3–5]. Structural engineers can monitor the potential damage and structural response of important bridges by considering environmental factors [6, 7] and using structural analysis [8, 9]. In missions of Structural Health Monitoring (SHM), vibration-based technologies, such as accelerometers permanently or temporarily attached to structural members, have been used as a conventional approach. However, these technologies have several disadvantages, including higher cost, fewer desirable instrumentation locations on in-service tall structures, and limitations as contact sensors [10–17].

In contrast, emerging vision-based technologies can be alternative and effective in achieving this goal because more advanced computer vision techniques and structural analysis tools are available (e.g., Unmanned Aerial Vehicles (UAVs) and less expensive optical devices). In this research, aiming to engineer a pipeline using computer vision methods at low cost and high efficiency in SHM missions, different camera placements are validated by monitoring the vibrations or displacements of the same bridges. Both vibration- and

vision-based technologies are used to collect time history data. Then, signal processing techniques are developed to analyze and interpret the visual data using deep learning and computer vision methods. Structural models, as the last step, are analyzed to verify the measurements and evaluate the conditions of these bridges.

Figure 1 shows that the motion of a target on a bridge can be captured by a camera in spatial and temporal spaces. The target's moving distance AB in the real world can be projected onto the image plane, where the origin is on the left corner, while the x and y axes are denoted from left to right and from top to bottom using pixels as its unit. Since the target's size in the real world and image plane are known, the scale factor is defined to convert coverage of one pixel into millimeters or inches on the target in the scene plane. Thus, the motion of the target in pixel unit can be effectively scaled to the length unit by multiplying it with the scalar. The higher the camera resolution, the more accurate this conversion is. ab is the trajectory of the target on the image plane, where dx and dy are the projected distance of ab in x and y directions, respectively. The target's motion can be recorded with a digital video of N frames or images. These consecutive images are defined as $I^0, \dots, I^i, \dots, I^N$, where I is the intensity value, and i and t represent the i th frame and time. To gain a high-level understanding of information with a video, machine learning and deep learning algorithms can be applied to quantify this motion for a structure or its

component in the laboratory and field experiments [18, 19]. The key to measure the target's motion is to use a motionless infrastructure or a camera itself in the scene as a reference. For example, a stationary camera itself can be the reference to record the movements of targets, otherwise, the surrounding buildings or other standstill objects that are not affected by the target's motion can be the reference to remove the relative camera motion from the measurements.

In this study, three camera placements are considered to measure the displacement or vibration of a bridge, including (i) structure-mounted cameras (when the cameras are placed on the bridge and nearby structures are used as a reference), (ii) remote cameras (when the cameras are fixed remotely to the observed bridges and part of their appearance is captured in the videos), and (iii) drone-mounted cameras (i.e., UAVs). Figure 2 illustrates these three different camera placements to measure the vibrations of a pedestrian bridge. These methods demonstrate how vision-based technologies can be used to monitor and measure the displacements and vibrations of structures or structural elements in a laboratory experiment or a field test. New techniques are developed to use deep learning as a better feature extraction tool and suppress the low- or high-frequency motion in the visual data or the camera's own vibrations. To the best of our knowledge, structure-mounted camera placement and frequency subtraction are first proposed and implemented in this research to monitor the response of bridges and remove

Fig. 1 Motion of a target projected in spatial space (a) and recorded in temporal space (b) with a camera

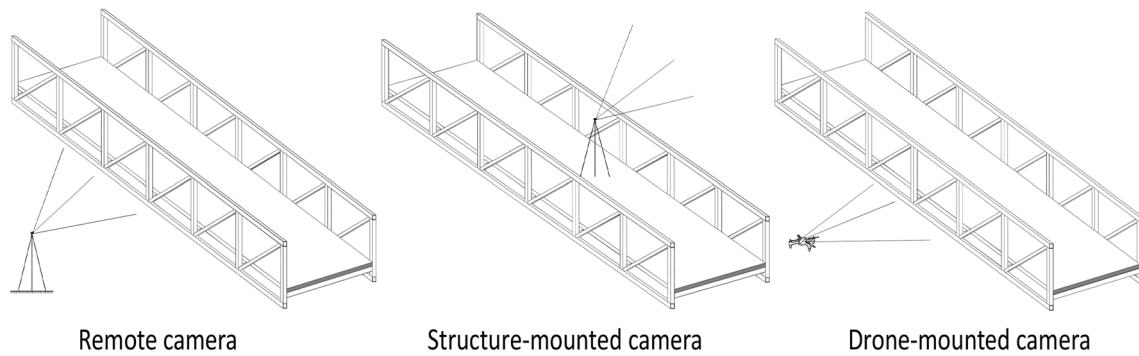
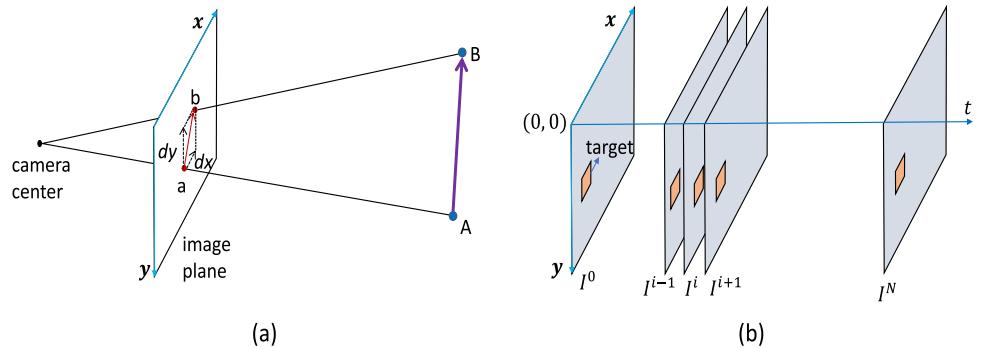


Fig. 2 Three camera placements for vibration measurement of a pedestrian bridge

camera motion [20]. The main objective of demonstrating the use of a structure-mounted camera system and camera motion removal in this research is to provide more options for vibration and displacement measurements with cameras. Field experiments on existing bridges with cameras and accelerometers and structural analysis are utilized to verify the proposed methods.

2 Related work

Wireless accelerometers have been commonly used as contact sensors to detect the dynamic response of various bridge structures these days. Gheitasi et al. demonstrated how to assess the serviceability of an in-service footbridge and gain ideal results when wireless accelerometers were used for vibration data acquisition [21]. Gibbs et al. showed how to process data from wireless accelerometers as the ground truth for a framework to extract dynamic characteristics for rural footbridges [22]. Also, several wireless accelerometers were deployed for measuring vibration responses of traffic bridges in [23, 24]. We used the data from wireless accelerometers as the reference to verify the camera measurements in our field experiments.

There are many computer vision techniques to measure the displacement of structures or their components with traditional computer vision methods. The experimental studies [18, 25–28] help us find a solution to accurately measure displacements in the field or laboratory. We were inspired by the applications of vision sensors for cost-effective structural health monitoring and verified with the field tests of the Manhattan Bridge in New York, USA [29] and a short-span railway bridge [30]. Also, the main idea of this study is based on the template matching or phase-based methods to measure structural displacements [8, 19, 31–33]. Furthermore, the Lucas-Kanade template tracker algorithm used in [34, 35] is employed to measure dynamic displacement for large-scale structures. For displacement measurements with cameras in these laboratory and field experiments, an accuracy of 0.016–0.25 mm and 0.64–3.5 mm is reported, respectively. All these methods did not include a structure-mounted camera placement in experiments, which we first proposed and tested in this study.

Deep learning methods are widely used to gain a high-level understanding of visual data and extract desired information. Like aforementioned traditional methods, the deep learning methods perform displacement or vibration measurements using data processing techniques to automatically convert visual information to datapoints in experiments. Precise motion measurements can be achieved from images and videos with these methods. Xiao et al. proved that it is applicable to a proposed SHM system using deep learning algorithms to evaluate the structural response, in which

visual data and measurements from conventional sensors are fused for a more reliable diagnosis of bridges' health [36]. Dong et al. implemented an optical flow algorithm named FlowNet2 to measure the displacements and vibrations of bridge structures [37]. Dong and Catbas applied Visual Graph Visual Geometry Group to extract features on the target of a two-span bridge during the traffic time [38]. The measured accuracy of displacements with deep learning varied from 0.0087 to 0.08 mm in the laboratory tests. These methods indicate how cameras can be used efficiently to monitor and measure the displacements or vibrations of structural members or a structure in the laboratory or field. Li and Sun applied a deep learning algorithm to detect structural damages by monitoring the continuous bridge deflection based on fiber optic sensing technology [39]. Inspired by these applications, a deep learning framework with High-resolution Mask Regional Convolutional Neural Network (HR Mask R-CNN) was proposed by us to track and accurately measure the deflection of three concrete beams, and the vibrations of three masses on a shaking table in the laboratory tests. This HR Mask R-CNN was trained following standard data annotation, loss regulation and parameter settings of such a deep neural network. A measurement-smoothing technique referred to as the Scale-Invariant Feature Transformation (SIFT) [40] was also introduced for high-accuracy measurements. Thus, the average error of deflection measurements from HR Mask R-CNN + SIFT for the three test beams is 0.13 mm, and the difference between the extracted and input frequencies is less than 9% by identifying all the intended frequencies. An optical flow method referred to as Lucas-Kanade (LK) tracker [41] was also utilized in these studies as a validation method [20, 42]. It is critical to test both methods on real structures, so we apply this framework and LK tracking algorithm to monitor and evaluate the response of in-service bridges in this paper.

In addition, drones have become an important platform to mount high-definition cameras and investigate the response of simulated structures or existing infrastructure. Yoon et al. displayed how to measure the vibrations of structures in laboratory and field tests with UAVs. They recovered the motion of the UAVs and got correct measurements during the flight [43]. Tomiczek et al. developed a small unmanned aircraft system that can be used to perform bridge inspections using a laser range finder and an optical flow sensor [44]. Chen et al. employed homography transformation to estimate the motion of the UAV and then corrected the measurement of a model bridge with Digital Image Correlation (DIC) [19]. Hoskere et al. used a UAV to measure the dynamic structural displacements from videos and obtained the model of a full-scaled suspension pedestrian bridge [9]. Ribeiro et al. used the visual and Inertial Measuring Unit (IMU) data from UAVs to measure the in-plane displacement, in which IMU data represented the motion of the UAV.

The visual data described the combined motion of the target, and then the subtraction of both actions was the absolute displacement of the target [45]. Perry and Guo offered a framework to use stereo and infrared cameras for 3D displacement of structures, which was verified by an indoor shaking table test [46]. Khuc et al. employed UAVs to measure the swaying displacement of small-scale structures [47]. Different from Khuc et al.'s study using homography transformation to align the adjacent frames in the video, we think there is another way to remove the camera motions during the flights. In our study, each frame is aligned to the first frame with Affine Transformation for removing drone motion.

3 Methodologies

There are two steps in our framework, including visual data acquisition and processing, to perform an accurate measurement and assess the vibrations or deflections of a bridge under normal loading excitation with vision-based technologies [20, 48]. In the data acquisition step, three different ways to place cameras in field experiments were studied to investigate the accuracy of vision-based vibration methods and their limitations for real applications. If a camera is seated on an intended location of a bridge and moves with the structure, or it is fixed remotely, or a drone flies and keeps at a certain height to focus on the moving part of the bridge and nearby stationary objects, these camera placements are called structure-, remote and drone-mounted cameras, respectively (see Fig. 2). In the visual data processing step, the HR Mask R-CNN method [49–51] is employed to track targets or regions of interest (ROIs) on the bridges in these videos, where spatial information (i.e., coordinates in the image plane) is included in each frame and the temporal information can be obtained by the duration and frame rate used to record the videos (see Fig. 1). In order to achieve sub-pixel accuracy of displacement measurement, SIFT is also used to refine and smooth the measurements. Figure 3 is an example that a target (e.g., a joint region) on a bridge is recognized and tracked with the HR Mask

R-CNN. In Fig. 4, the keypoints on the detected target with our deep learning method are matched with the SIFT. Then, the average coordinate change (dx^i and dy^i) of these keypoints in the image plane, where (x_j^0, y_j^0) and (x_j^i, y_j^i) are the coordinates of j th matching keypoints on the first and i th frames in the video, respectively, will be used to represent the target's movement between these two frames. In our experimental studies, the measured movement was converted from pixel unit to metric unit based on the target's size in the image plane and real world. The LK tracker [41] is employed to verify the proposed HR Mask R-CNN + SIFT in this study. Both methods were validated using laboratory tests in [20, 42, 48].

Some signal processing techniques are applied in our framework of displacement or vibration measurements. Butterworth Band Filter [52] is used to get rid of noises caused by very low or high frequencies and to correct the wrong trend of the visual data. Then, Faster Fourier Transform (FFT) [53] is utilized to extract the modal frequencies of structures or structural components with these filtered vibration signals. Since a camera can be affected by the wind or movement of the ground where it stands or may shake if it is attached

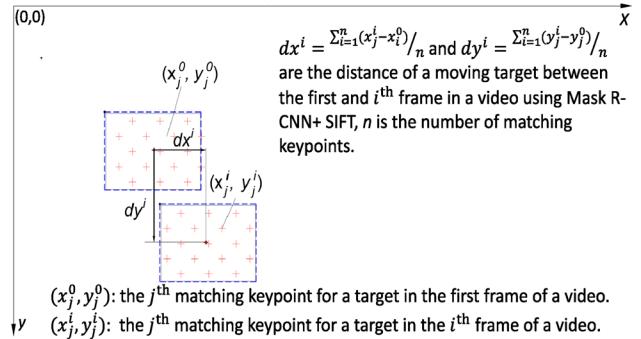
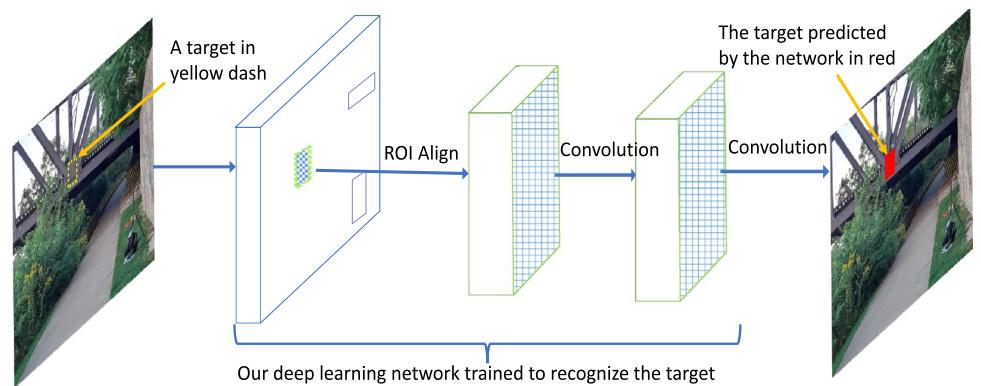


Fig. 4 An illustration to get the average coordinate change of multiple matching keypoints on a target in Fig. 3 between two frames in a video with Mask R-CNN + SIFT. The target's movement is represented with this change and can be converted from pixel unit to metric unit if the scale factor is known

Fig. 3 An example of a target on Columbus Downtown North Bridge detected and tracked by the proposed HR Mask R-CNN



to the drone, the camera motion can be isolated and eliminated when appropriate techniques are applied. For example, displacement and frequency subtractions are used to remove the drone motion in field experiments [20, 48]. To be more specific, the former is a subtraction of movement between the drone's motion and the targets on the bridge. The latter is the subtraction between the frequencies of the drone and bridge motions in the camera. The flowchart of the proposed HR Mask R-CNN + SIFT to track and measure the vibration of a bridge with cameras is illustrated in Fig. 5. As a validation method, the LK tracker method also has a similar pipeline.

The proposed framework with various camera placements can save time and effort for monitoring the vibrations of existing bridges. This research provides structural engineers with more options to select a place to fix a camera or fly a drone to an ideal observation point. Then, higher-quality data can be obtained in addition to eliminating the need for markers on the bridges and reference objects. The data processing is automatic since the acquired visual data are the only input and the frequencies of bridges are the output. Currently, it may take several minutes to do this if the training to detect and track the targets are needed. In the proposed framework, this can be real-time if the vibration measurements are repeated using the same setup that was used in previous tests on the same bridge.

4 Experimental studies

The framework proposed in the previous section to capture and measure the vibrations of structures or their components was validated with two shaking table tests in the laboratory [20, 42, 48]. The HR Mask R-CNN + SIFT and the LK tracker

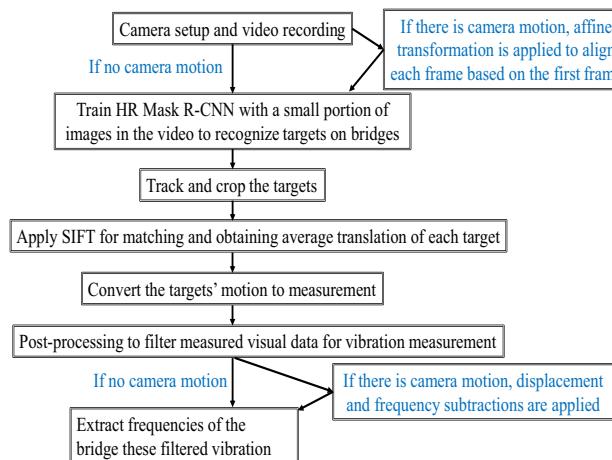


Fig. 5 A flowchart of HR Mask R-CNN + SIFT used for vibration measurement of a bridge

achieved sub-pixel accuracy in these tests with remote cameras as non-contact sensors. In this section, all three camera placements and techniques to remove camera motion for in-service bridges are tested. Some measurements were compared with the results from the structural analysis performed using SAP2000 [54]. These bridges are shown in Fig. 6.

In the experiments, the original unit for the vibration measurement of these bridges is in pixels. One pixel is the smallest 2D square to divide the image plane evenly in two directions and can be used as the length unit (Fig. 1). For example, the convention to describe the pixel resolution of an image is to use integer numbers in both width and height directions, where width is the number of pixel columns in x direction and height is the number of pixel rows in y direction. The measurements were converted into millimeters (mm) with the measured scale factors, which are in a range between 1.10 and 3.80 mm/pixel in our field experiments. We rely on the SIFT to obtain sub-pixel accuracy, thus, the error of displacement measurements in a laboratory test can be 0.13 mm [42]. The same subpixel precision can be achieved in field experiments but may be less accurate because of the longer distance with our current cameras. Also, we did not model the Lane Ave Bridge because the documents at hand are insufficient to model and analyze this prestressed and cable-stayed concrete bridge in Fig. 6.

4.1 Vibrations of pedestrian bridges measured with three camera placements

Six pedestrian bridges on the main campus of the Ohio State University (OSU) were tested. All of them are truss structures with steel tube members and concrete decks. These bridges are less rigid and tend to vibrate more when people pass. In this study, their dynamic responses under the excitation of jumps were measured with the camera placements illustrated in Fig. 2. The tracking targets are all on the midspans of these bridges.

Cannon Bridge, a three-span bridge, was used as a detailed example to address the application of the proposed deep learning-based method to quantify its dynamic characteristics. The drawing and views of this bridge are shown in Fig. 7. The midspan of the second span was selected as the location for tests with remote, structure-mounted, and drone-mounted cameras. The frame rates of remote and structure-mounted cameras are set as 45 and 60, respectively, with an image size of 1920×1200. The drone-mounted camera has an image size of 1920×1080 and records at a speed of 48 frames per second. Original measured visual data are filtered or aligned in case of camera motion. FFT is applied to extract the frequencies of the bridges [20, 48].

The measured vibrations and extracted fundamental frequency of this bridge with the proposed framework are shown in Fig. 8. The fundamental frequency obtained from

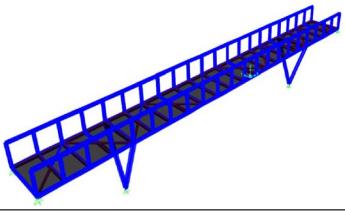
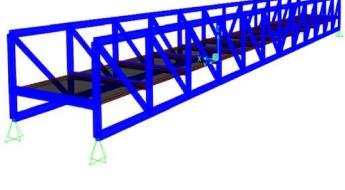
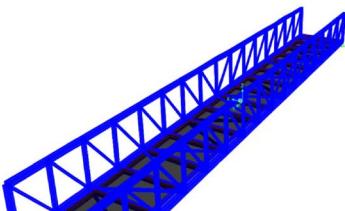
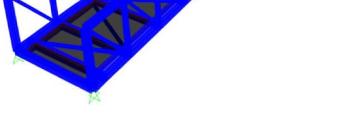
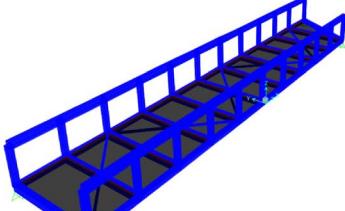
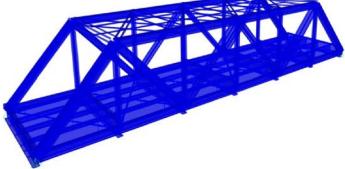
Name	Views	SAP2000	Information
Cannon Bridge			Three-span bridge with steel tubes
Schottenhamstein Bridge			Simply supported steel truss bridge
Carmack Bridge 1			Simply supported steel truss bridge
Carmack Bridge 2			Simply supported steel truss bridge
Morrill Bridge			Simply supported steel truss bridge
Lincoln Bridge			Simply supported steel truss bridge
Lane Ave Bridge			Two-span fixed-ended cable-stayed bridge with precasted concrete
Columbus Downtown North Bridge			Five simply-supported steel truss bridge

Fig. 6 Eight bridges tested in this study

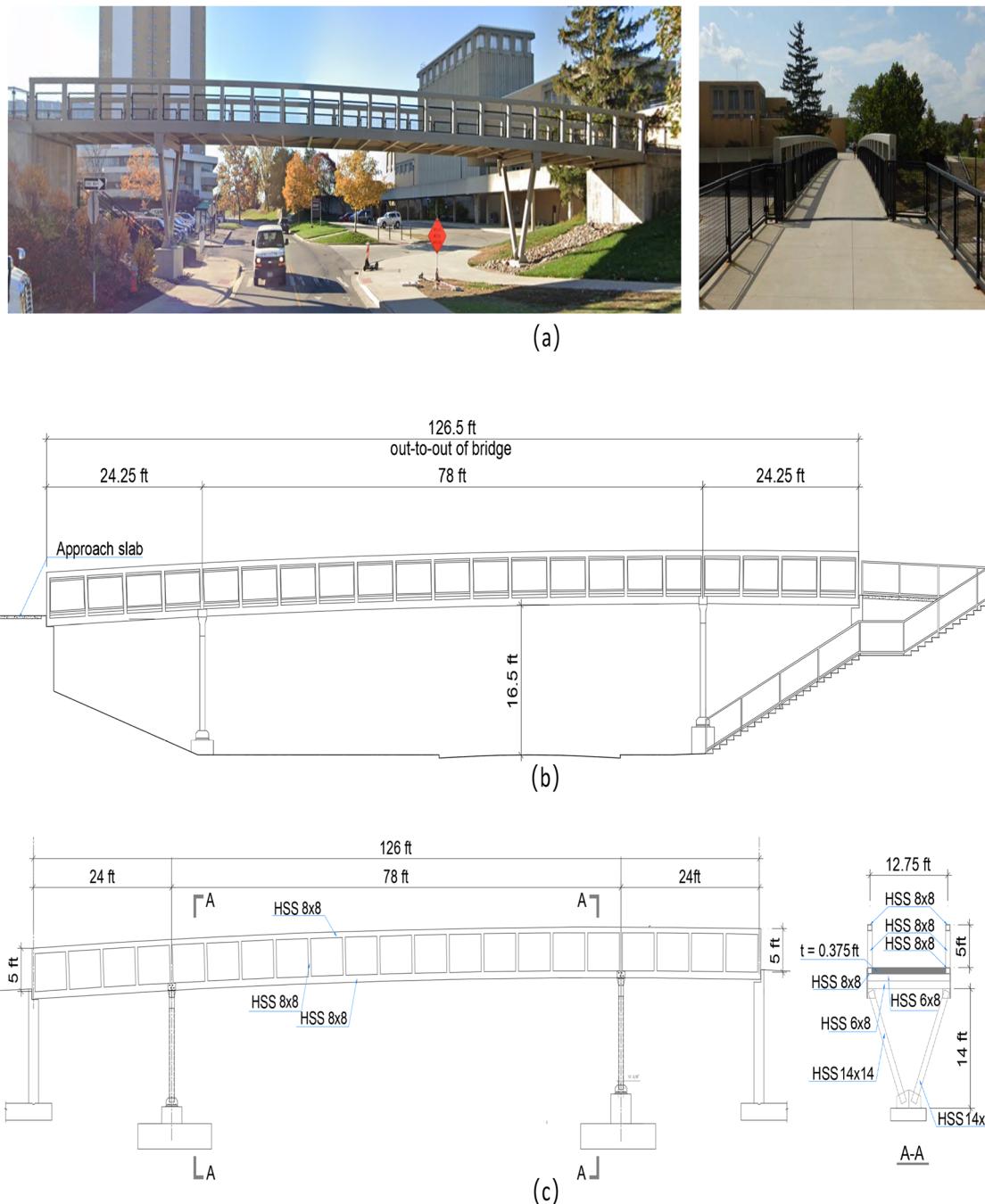


Fig. 7 Views and drawings for Cannon Bridge. **a, b** and **c** are two views, architectural and structural drawings for the bridge, respectively (1 ft = 30.48 cm)

the accelerometers is approximately 4.0 Hz. Similarly, the fundamental frequency calculated by analyzing the structure with SAP2000 is around 4.0 Hz (see Table 1). Using three different camera placements in Fig. 2, its vibrations can be captured successfully and the fundamental frequency can be determined accurately. In addition, it can be observed that the structure-mounted camera (Fig. 8c) can acquire more accurate shapes of a free-damped vibration system as

one jump excited the bridge to vibrate from the peak to the standstill state due to the damping, but the remote and drone-mounted cameras (Fig. 8a and b) cannot clearly capture this shape. For this reason, we focused more on the frequency extraction than vibration magnitude in this paper.

Table 1 summarizes the measured and calculated fundamental frequencies of these six pedestrian bridges, although a study in [20] shows that the structure-mounted camera

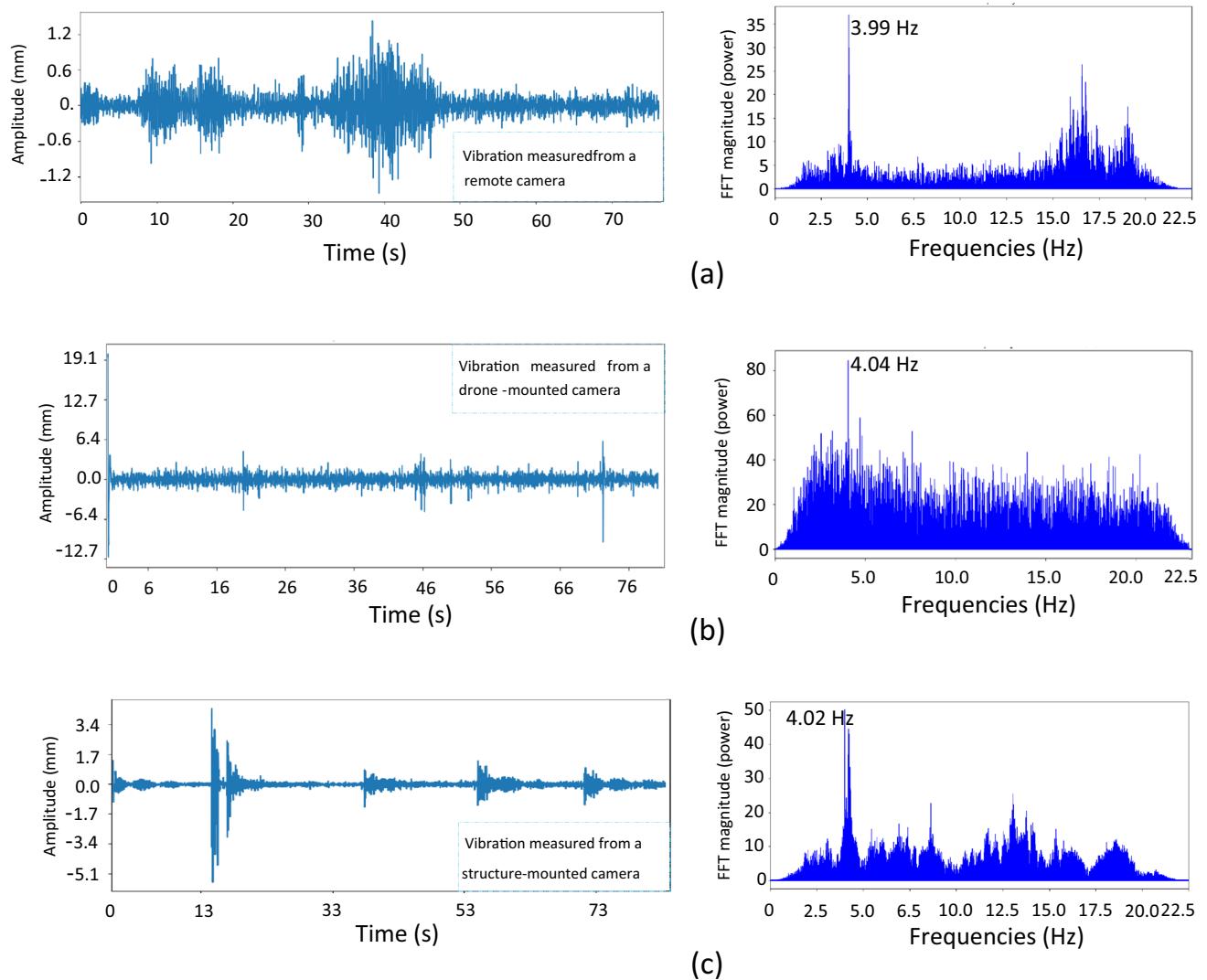


Fig. 8 Filtered vibrations (left) with three different camera placements and extracted frequency (right) with FFT for Cannon Bridge. **a, b** and **c** are the results from a remote, drone-mounted, and structure-mounted camera, respectively

Table 1 Fundamental frequency (Hz) of six pedestrian bridges measured and extracted by three different camera placements and accelerometers

Bridge	Remote			Structure-mounted			Drone-mounted			SAP2000
	HR	LK	Acc	HR	LK	Acc	Disp	Freq	Acc	
Carmack1	4.68	4.68	4.66	4.69	4.62	4.64	4.70	4.68	4.59	4.75
Carmack2	4.61	4.61	4.62	4.61	4.62	4.61	4.68	4.68	4.60	4.75
Lincoln	5.44	5.44	5.44	5.49	5.47	5.47	5.45	4.94	5.45	5.44
Morrill	5.43	5.43	5.43	5.47	5.45	5.45				5.44
Cannon	3.99	3.99	4.03	4.02	4.02	3.99	4.04	4.02	3.99	4.04
Schottenstein	3.84	3.84	3.85	3.84	3.84	3.85	3.91	3.87	3.84	3.88

*HR and LK refer to HR Mask R-CNN + SIFT and LK tracker. Acc., Disp., and Freq. stand for accelerometers, displacement subtraction, and frequency subtraction

placement can identify the top four frequencies of Schottenstein Bridge. The drone-mounted camera placement is not tested for Morrill Bridge because this bridge is almost

identical to Lincoln Bridge. These results indicate that our proposed methods can accurately capture the fundamental frequencies of the bridges.

4.2 A test on a traffic bridge

Lane Avenue Bridge is a cable-stayed bridge over the Olentangy River near the OSU campus (see Fig. 6). The total length of this two-span bridge is 113.1 m (371 ft). There are six lanes for vehicle traffic with additional two pedestrian lanes on the bridge. It is a stiff bridge made of prestressed concrete girders, a reinforced concrete deck, and steel cables connected to the reinforced concrete towers. The motion of the bridge induced by the regular traffic was measured and evaluated.

In the experiment, a camera was fixed on the bridge's sidewalk about 6.1 m (20 ft) far from the reference target, the bridge's tower, to record the deck's motion during normal traffic. The frame rate was set as 45 frames per second and the image size is 1920×1200. Both HR Mask R-CNN and the LK tracker were utilized to capture the vibration with sub-pixel precision, although the magnitude of the vibration is very small (e.g., the magnitude of vibrations is less than 0.2 mm). The measured and calculated results are shown in Fig. 9. Compared to the measurement from accelerometers, the first two modal frequencies, such as 7.86 and 9.63 Hz, were successfully detected by our visual data processing method. But the bridge's vibration caused by running vehicles is not significant. Thus, the remote and drone-mounted

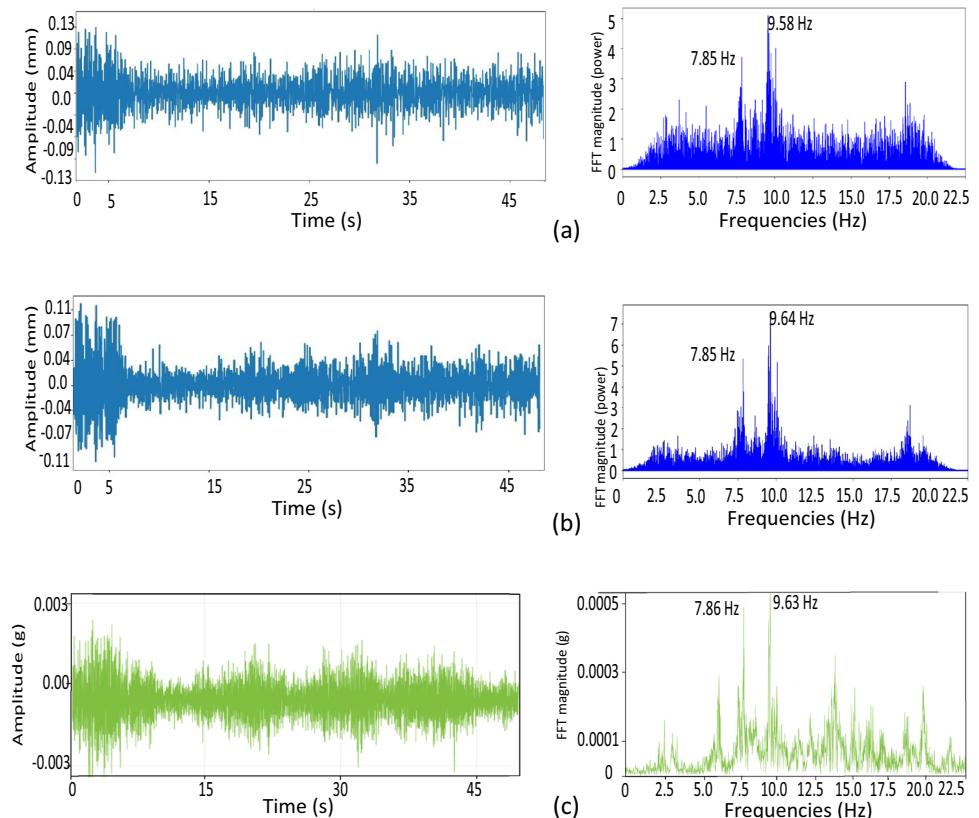
camera placements did not work on this bridge with our current cameras.

4.3 Vibrations and deflection measured for Columbus Downtown North Bridge

This railway bridge is a steel truss structure built in the 1930s in Downtown Columbus, Ohio. It is currently open for railway transportation. It has five spans, but pin-hinged supports independently restrain each span and make each of them a simply supported structure. The bridge has composite steel sections and an open deck with square timber members placed laterally on the floor beams. The views and information about the structure and steel members are shown in Fig. 10. The largest span length is 38.1 m (125 ft). Tests were conducted on both end spans of this bridge as described below.

(1) A structure-mounted camera with the jump excitation in the north end span: This was a test to excite the bridge by jumps. Obviously, this kind of excitation is far smaller than the vibration that the bridge experiences when a train passes over it. A camera was placed right on the top of a steel beam of the bridge, and its frame rate was set as 60 frames per second. The image size is 1920×1200. The camera was focused on the nearby retaining wall as a reference of stationary objects.

Fig. 9 Processed visual data measured by a structure-mounted camera (top recordings) and accelerometers (bottom recording) on Lane Avenue Bridge. **a**, **b** and **c** are measured vibrations (left) and extracted frequencies (right) from HR Mask R-CNN, LK tracker, and accelerometers, respectively



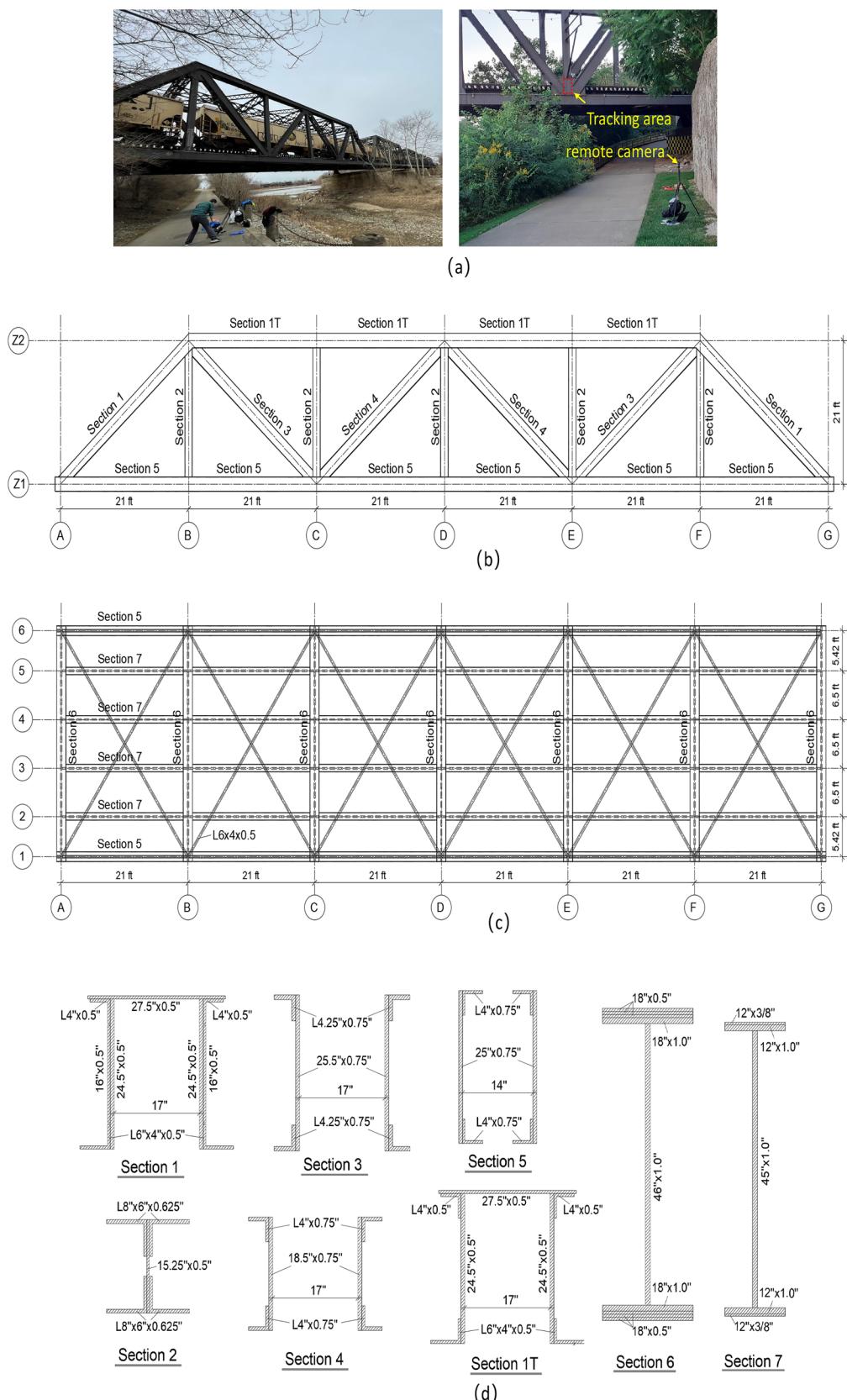


Fig. 10 Views and drawings for Columbus Downtown North Bridge. (a) is its views, (b) to (d) are the elevation and plan views, and sections for major members of the bridge, respectively (1 ft = 12 in. = 30.48 cm)

Figure 11 shows the testing results and Table 2 indicates the difference of extracted frequencies between the camera and accelerometers. It can be seen that only the third frequency was detected by both HR Mask R-CNN and the LK tracker, but the frequencies of 5.07(5.10) and 15.63 Hz are not well captured by both proposed methods. We think it is because the jumps cannot excite the bridge to a significant overall movement.

(2) Remote cameras focusing on two locations of the south end span of the railway bridge: Two cameras were placed 12.2 m (40 ft) far away from the south end span of this bridge. They were focused on the midspan and one-quarter span of the bottom chord of this steel truss bridge. The synchronization of data acquisition by two cameras was tested in this study. The frame rate and image size of each camera were set as 45 frames per second and 1920×1200, respectively. Visual data during the train going over that end span was processed by our proposed framework, whereas only the LK tracker was used to capture the bridge's vibrations at two locations. Details of the experiment data processing and applications of the proposed frameworks can be found in [20].

Figures 12 and 13 show that the filtered vibrations and extracted frequencies for two observed locations on the bridge can be captured using the method described in the previous section. Three major frequencies, such as 6.32, 10.90 (11.11), and 15.26 (15.91) Hz, at both locations were

Table 2 Three frequencies (Hz) measured and extracted by the frameworks of HR Mask R-CNN and LK tracker and with accelerometers for Columbus Downtown North Bridge

Accelerometers	HR Mask R-CNN		LK tracker	
	Measure- ment	Difference %	Measure- ment	Difference %
7.39	5.07	−31.4	5.07	−31.4
11.03	11.08	0.4	11.0	0
18.02	15.63	−13.3	15.63	−13.3

detected by the camera and accelerometer data, but the frequency of 4.89 Hz was not identified from the accelerometer data. Their differences in extracted frequencies for the one-quarter and midspan are shown in Table 3. Based on the measurements from accelerometers, the camera placement under the train-running excitation has less difference in extracted frequencies than the one under the jump excitation. Compared with Table 2, these extracted frequencies from different camera placements are also not close. This indicates that the former excitation can cause a structural response for such a heavy and stiff bridge but the latter cannot achieve this. Also, the mass of this bridge was changing when the train went over the span. Thus, the dynamic response of this train-structure system would be less steady. It was hard to achieve an accurate vibration measurement

Fig. 11 Two different methods used for processing data from a structure-mounted camera for Columbus Downtown North Bridge with jumps. **a**, **b** and **c** are measured vibrations (left) and extracted frequencies (right) from HR Mask R-CNN, LK tracker, and accelerometers, respectively

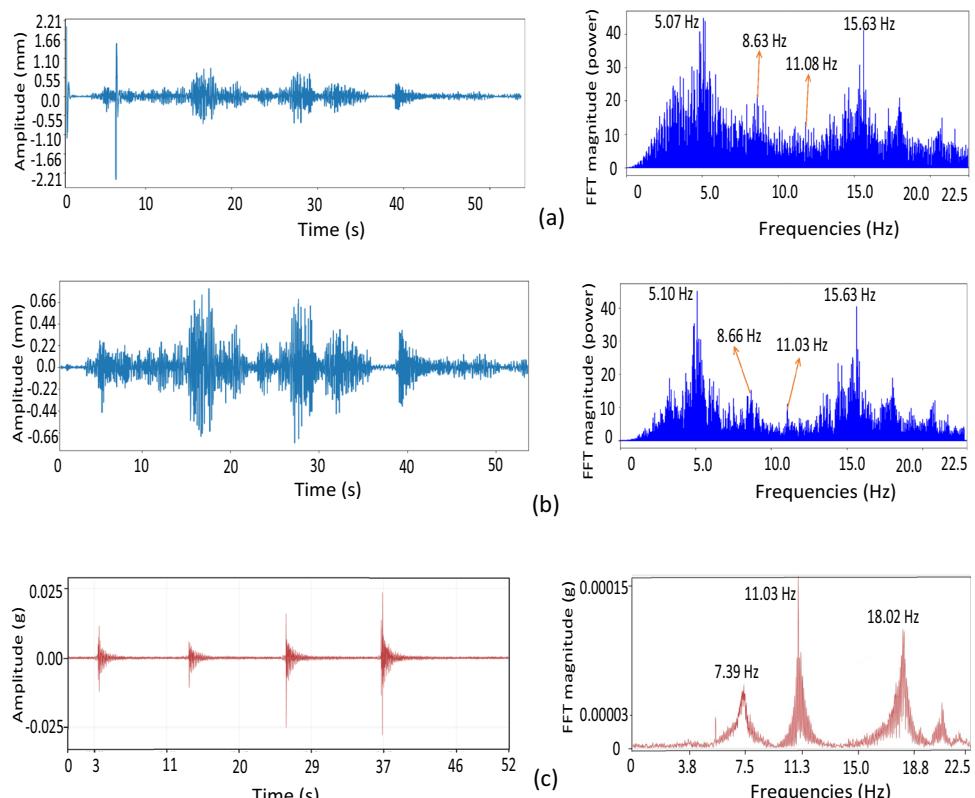


Fig. 12 Remote-camera data for the one-quarter span of the south end span of Columbus Downtown North Bridge **a** and **b** are measured vibrations (left) and extracted frequencies (right) from the LK tracker and accelerometers

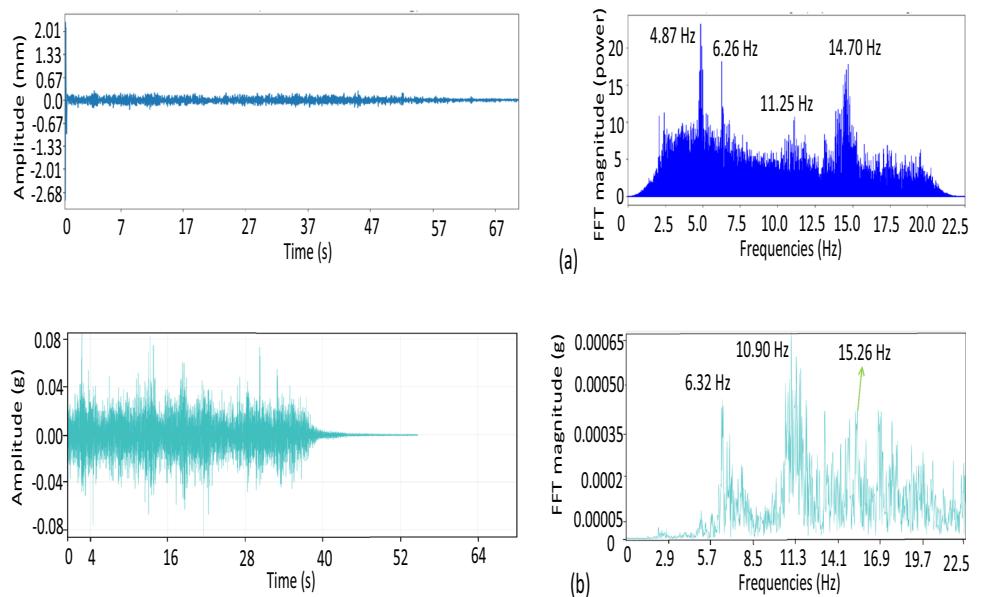


Fig. 13 Remote-camera data for the midspan of the south end span of Columbus Downtown North Bridge. **a** and **b** are measured vibrations (left) and extracted frequencies (right) from the LK tracker and accelerometers

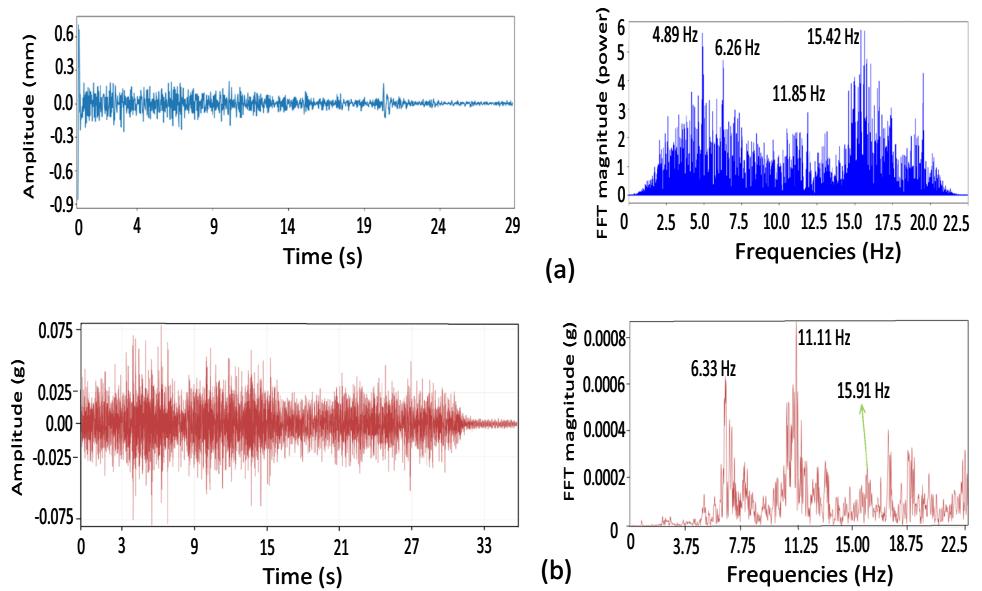


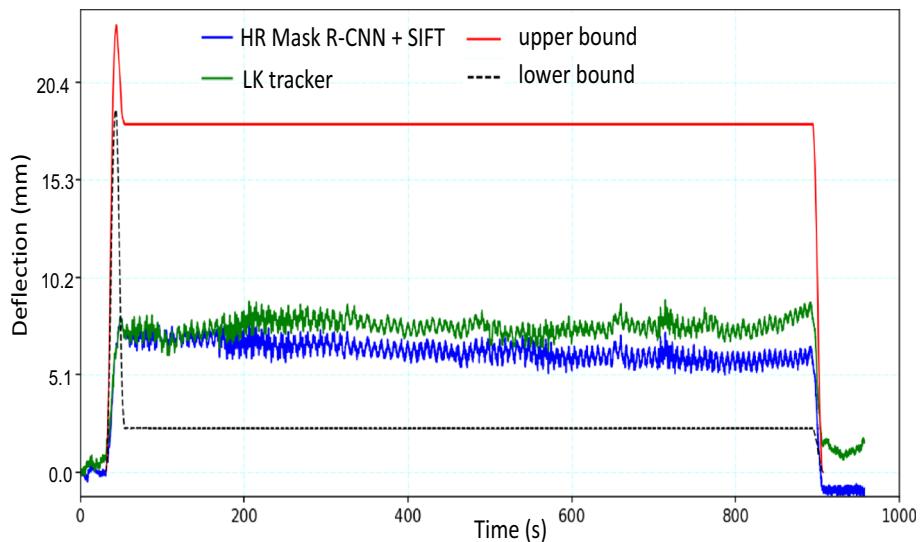
Table 3 Three frequencies measured and extracted by the framework of LK tracker and with accelerometers for the one-quarter span and midspan of the south end span of Columbus Downtown North Bridge

Accelerometers	One-quarter		Midspan	
	Measure- ment	Difference %	Measure- ment	Difference %
6.33 (6.32)	6.26	− 0.1	6.26	− 0.1
10.90 (11.11)	11.25	3.2	11.85	6.7
15.26 (15.91)	14.70	− 3.7	15.42	− 3.1

with the remote camera placement in this case. We may figure out a solution for this in a future study.

(3) Deflection of the north end span of the railway bridge under train loading: A high-resolution camera with an image size of 3260×2440 was used to measure the deflection of the midspan of the north end span of the bridge. The camera was placed approximately 18.3 m (60 ft) far away from it and focused on the joint of the bottom chord in the middle of the bridge. The frame rate was 15 frames per second. There were 30 cars on the train. It recorded the whole process of loading and unloading cycles on the bridge. As shown in Fig. 14, two proposed methods, HR Mask R-CNN + SIFT and the LK

Fig. 14 Two different methods to measure the deflection of midspan of Columbus Downtown North Bridge. Positive deflection means downward movement



tracker, were applied for deflection measurement. There is a sharp increase and decrease for both when the train was on and off the span.

5 Validation and structural health assessment with structural analysis

In the above experiments, the performance of the tested bridges can be captured by various camera placements. Some of their dynamic characteristics can be interpreted with techniques like FFT. But it is insufficient to evaluate the health of bridge structures only with these measurements, especially for in-service bridges. Structural analysis can be introduced to model these bridges and analyze the calculated response with the measurements considering the degradation of materials. The results of experiments can also be used to validate and update finite element models. Thus, a more accurate numerical analysis can be performed during the assessment of the bridge structures.

SAP2000 software (2022) is utilized to model and analyze the selected six pedestrian bridges (e.g., Cannon Bridge) and a railway bridge (e.g., Columbus Downtown North Bridge). The dynamic performance of these bridges under ambient loading conditions is the focus of this study.

Table 4 Information for six pedestrian bridges (1 ft = 30.48 cm)

Bridge	Spans	Span length (ft)	Main components	Materials	Support types
Carmack1	1	80.75	HSS 7 × 7 × 1/4	A36	Simply
Carmack2	1	80.75	HSS 7 × 7 × 1/4	A36	Simply
Lincoln	1	53	HSS 6 × 8 × 1/8	A36	Simply
Morrill	1	53	HSS 6 × 8 × 1/8	A36	Simply
Cannon	3	24,78,24	HSS 8 × 8 × 3/8	A36	Simply
Schottenstein	1	90.5	HSS 6 × 6 × 1/4	A36	Simply

The modulus of elasticity and yield strength of steel material were assumed as 200 GPa (29,000 ksi) and 413.7 MPa (60 ksi), respectively. Details of the structural models, geometric information, and loads are provided in [20].

5.1 Structural models for all six pedestrian bridges

The views of six pedestrian bridges and their structural models in SAP2000 are shown in Fig. 6. The information used for structural modeling is shown in Table 4. Six modes of the Cannon Bridge model are shown in Fig. 15 based on the configuration in Fig. 7. The fundamental frequency corresponding to Mode 1 in the vertical direction is 4.04 Hz and it is very close to the measured frequencies by the cameras and accelerometers. The frequencies of the first mode analyzed by SAP2000 for these pedestrian bridges are shown in Table 1.

5.2 Structural modeling and analysis of Columbus Downtown North Bridge

The general plan and elevation views of the bridge are depicted in Fig. 10. The load-bearing members of the railway bridge were modeled using frame elements in SAP2000, including moment releases defined at both ends of the

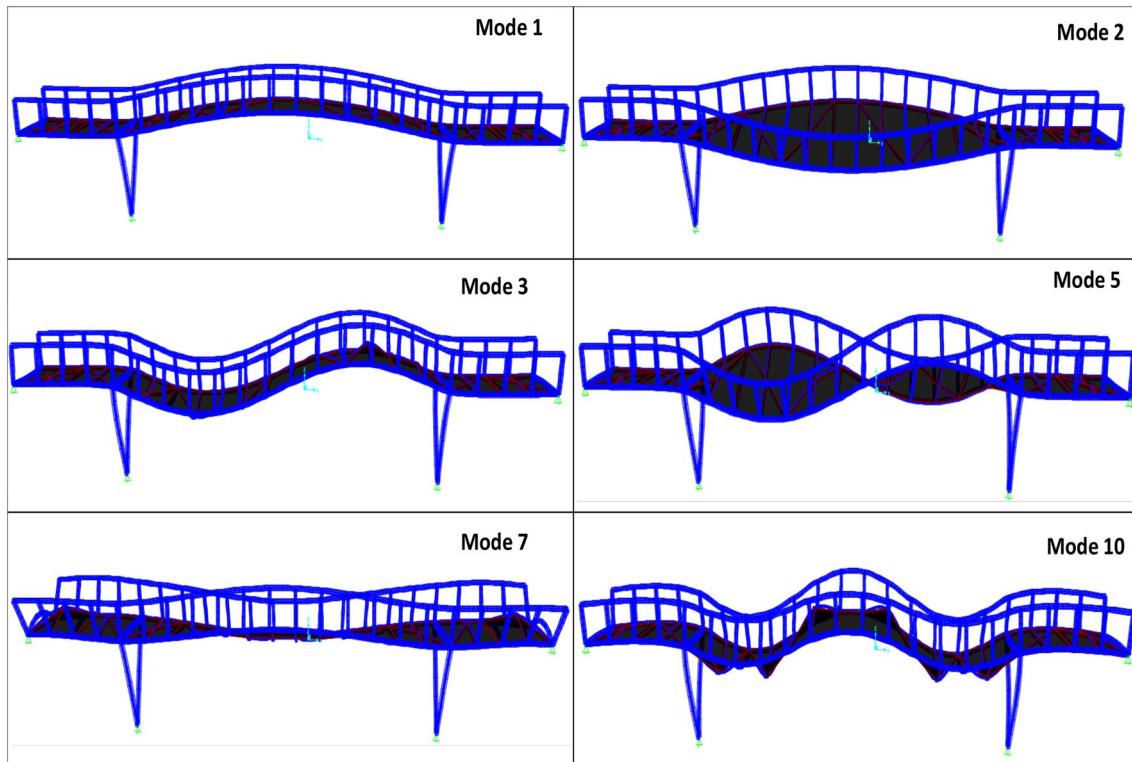


Fig. 15 Six selected modal shapes of the computer model of the Cannon Bridge

members to simulate the hinged end connections. A modal analysis was conducted to calculate the frequencies of the bridge. The first six modes have frequencies of 5.78, 7.15, 9.43, 11.45, 13.43, and 16.52 Hz. Compared to the measurements from cameras in Fig. 11, the difference between the measured and calculated frequencies are 14.0% ($(5.78 - 5.07)/5.07 = 14.0\%$), -17.1% ($(7.15 - 8.63)/8.63 = -17.1\%$), 3.3% ($(11.45 - 11.08)/11.08 = 3.3\%$), and 5.7% ($(16.52 - 15.63)/15.63 = 5.7\%$) for the first, second, fourth, and sixth modes, respectively. The visual measurements can not capture the other modes. It is apparent that the modal analysis results are a good match with the test measurements.

Additionally, a live load analysis was applied to obtain the deflection time history of this bridge loaded with a train. Since the weight, axle location, and other properties of cars in freight trains can be very different in the U.S., it is very challenging to determine their exact weight and location as observers. Therefore, upper and lower bound limits were defined to include the cases of fully loaded and empty freight trains. The maximum live load was determined for the upper bound analysis by considering design live loads prescribed in the Manual for Railway Engineering of the American Railway Engineering and Maintenance-of-Way Association (AREMA) [55]. The primary live load model of the manual, Cooper's E80 load (see Fig. 16), was implemented in the study. Secondly, the weight of empty freight cars and

corresponding axle load (P), which are 311.4 kN (70 kips) and 77.8 kN (17.5 kips), respectively, were selected for the lower bound analysis. The freight train dimensions and axle configuration were chosen according to Fig. 16 reported by [56]. It should be pointed out that the weight of the locomotives leading to the empty freight cars was assumed as 1174.3 kN (264 kips), and a similar locomotive axle pattern prescribed in Cooper's E80 live load was implemented [56–58].

In the SAP2000 model, the load factor was not considered for dead and live loads since the purpose of this study is to capture the actual behavior as much as possible. A moving load defined on a predetermined path was applied for the live load pattern to represent both fully loaded and empty trains. In addition, a nonlinear static load case simulating a vertical pushover analysis was defined to consider the self-weight of the bridge and the dead loads. Then, two nonlinear multi-step static analysis cases representing an empty and full freight train were employed to apply the moving live loads in sequence [59]. Because the freight trains traverse the bridge at very low speeds, other potential longitudinal and lateral dynamic actions that they may induce were ignored in the simulations. The results of the upper and lower bound live load analyses are shown in Fig. 14 for the fully loaded and empty freight train, respectively. The deflections from camera measurements and structural analyses are consistent

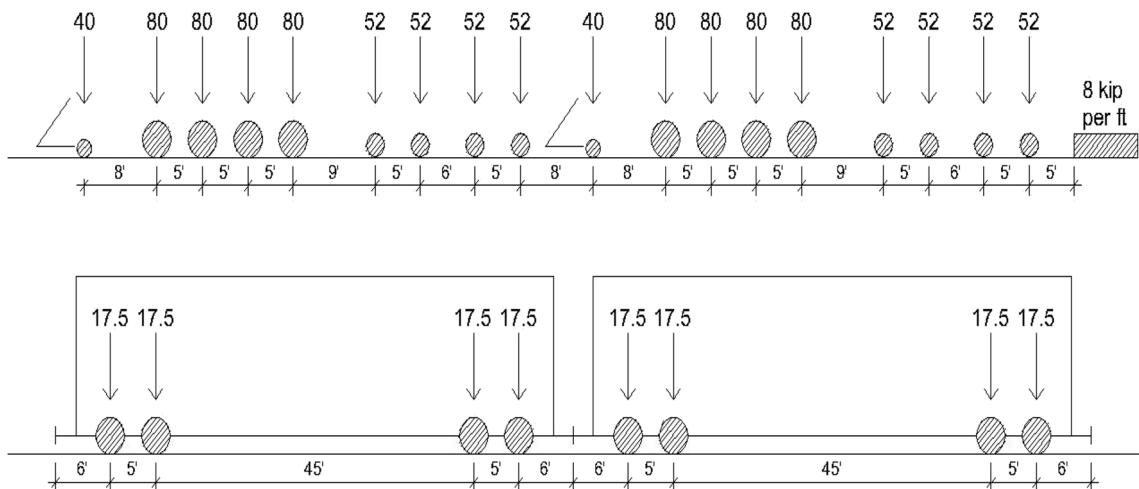


Fig. 16 Implemented live load patterns: Cooper's E80 (top) and idealized empty vehicles (bottom). Loads are in kips unit (1 kip = 4.45 kN)

because the typical freight train loads are unlikely to be as large as the specified conservative design loads. The camera measurements also stay within the lower and upper bound response limits.

6 Discussion

This research investigates the effect of various camera placements on bridge deformation measurements. Accurate results can be achieved with our proposed pipeline and the remote camera placement in laboratory tests, where more structural modes or precise displacements can be captured [42]. However, in field experiments, structure-mounted camera placement is the only one which can capture more accurate vibrations of bridges, as addressed by an ablation study in [20]. That is because this camera placement is used as contact sensors attached to the structures. Remote or drone-mounted cameras are affected harder by the surroundings and weather than structure-mounted cameras. Thus, less accurate shapes or few modes of structural dynamic response can be observed. Although similar experiments were not conducted to quantify the ambient effects or weather on the camera's performance, we believe that high-quality cameras and ideal daytime are essential for more accurate displacement or vibration measurements of bridges. We are also working on performing real-time and automatic measurements based on our proposed framework. For visual data collection and processing, it should be pointed out that high-performance computers (e.g., better Graph Processing Units (GPUs) and large storage space) are required.

As discussed in experimental studies, it is a key factor that the appropriate excitation is applied on these bridges. A simple jump is easier to cause the resonant movement of a pedestrian bridge, but a traffic or railway bridge is so rigid

that it will not vibrate significantly under such an excitation. This is why our framework for some camera placements performs poorly in some experiments on Lane Ave Bridge and Columbus Downtown North Bridge. Also, it is time-consuming and challenging to retrieve design documents and model these existing bridges. However, if a structural model is built and the changes in structural materials (e.g., degradation) and loading conditions can be simulated for a bridge, this analysis can be updated with long-term vibration measurements obtained from the cameras to evaluate the health of this bridge.

On the other hand, the proposed framework with various camera placements was also applied to a progressive collapse study of the field experiment on a multi-story building [20]. Like some tests in this paper, these applications show that this framework can be treated as a remote sensing system using remote and drone-mounted camera placements to focus on more observation locations of a structure. As long as the weather conditions are ideal and the setup of camera placements is right, our proposed framework can work well in field experiments on bridges and buildings.

7 Conclusions

Aiming to explore more advanced vision-based technologies for capturing the dynamic characteristics of in-service bridges, a framework with a deep learning method (i.e., HR Mask R-CNN) and a smoothing technique for displacement measurement (i.e., SIFT) are proposed and tested using field experiment data from eight bridges, including six pedestrian bridges, a traffic bridge, and a railway bridge. Remote, structure- and drone-mounted camera placements (see Fig. 2) were used to acquire the measurement data. Similar to the proposed framework of HR

Mask R-CNN + SIFT, the Lucas-Kanade (LK) tracker is utilized to verify the vibration measurements of these bridges. The accelerometers and analyses of structural models are employed to validate the measurements from three camera placements. There are several conclusions from this research:

- (1) In the data acquisition step, three camera placements were tested with these in-service bridges. These camera placements were found to be practical and effective to capture the structural characteristics of bridges under normal loading excitation. Structure-mounted camera placement can be used as a contact sensor and achieve more accurate measurements in field experiments. But remote and drone-mounted cameras can cover more desirable observation locations and be deployed easily and effectively in these experiments.
- (2) The proposed framework of HR Mask R-CNN + SIFT or the LK tracker can process the visual data well when some filters are employed to suppress the very low and high frequencies. It should be pointed out that other deep learning-based methods for tracking can be used in the proposed framework.
- (3) To eliminate the camera motion for drone-mounted cameras, two approaches (i.e., displacement and frequency subtractions) were proposed and validated. Our experiments on pedestrian bridges show that both methods work successfully to capture the fundamental frequency of each bridge.
- (4) Low-cost cameras can be used to perform accurate measurements of vibrations or deflections of existing bridges if the cameras are set up correctly and pipelines for visual data processing are established effectively. Also, more modal shapes or structural frequencies of monitored bridges can be obtained with sufficient excitation and better camera placements. In this research, visual data can not only be served as documentation but also for quantitative analysis as we did. Structural analysis and our proposed framework of visual data processing can be used together to assess whether the bridges are healthy or not by capturing the changes in their dynamic response over time.

Acknowledgements In this study, Dr. Bing Zha and Jianli Wei helped us do experiments in the field. We appreciate their time and participation. The second author was supported by the 2219-International Postdoctoral Research Fellowship Program of The Scientific and Technological Research Council of Turkey (TUBITAK).

Funding This study was funded by the National Science Foundation (grant number 2036193).

Data availability The data of this study are available from the corresponding author upon reasonable request.

Declarations

Conflict of interest The authors declare that they have no conflict of interest.

References

1. Farrar CR, Worden K (2012) Structural health monitoring: a machine learning perspective. Wiley, New York, pp 4–36
2. Mao J, Yang C, Wang H, Zhang Y, Lu H (2022) Bayesian operational modal analysis with genetic optimization for structural health monitoring of the long-span bridge. *Int J Struct Stab Dyn* 22(05):2250051
3. Zhu Y-F, Ren W-X, Wang Y-F (2022) Structural health monitoring on Yangluo Yangtze river bridge: implementation and demonstration. *Adv Struct Eng* 25(7):1431–1448
4. Garg P, Moreu F, Ozdagli A, Taha MR, Mascareñas D (2019) Non-contact dynamic displacement measurement of structures using a moving laser doppler vibrometer. *J Bridge Eng* 24(9):04019089
5. Lobo-Aguilar S, Zhang Z, Jiang Z, Christenson R (2019) Infra-sound-based noncontact sensing for bridge structural health monitoring. *J Bridge Eng* 24(5):04019033
6. Ni Y, Hua X, Fan K, Ko J (2005) Correlating modal properties with temperature using long-term monitoring data and support vector machine technique. *Eng Struct* 27(12):1762–1773
7. Ye X, Yi T-H, Dong C, Liu T (2016) Vision-based structural displacement measurement: system performance evaluation and influence factor analysis. *Measurement* 88:372–384
8. Chen JG, Davis A, Wadhwa N, Durand F, Freeman WT, Büyüköztürk O (2016) Video camera-based vibration measurement for civil infrastructure applications. *J Infrastruct Syst* 23(3):4016013
9. Hoskere V, Park JW, Yoon H, Spencer BF Jr (2019) Vision-based modal survey of civil infrastructure using unmanned aerial vehicles. *J Struct Eng* 145(7):04019062
10. Casas JR, Moughty JJ (2017) Bridge damage detection based on vibration data: past and new developments. *Front Built Environ* 3:4
11. Wickramasinghe WR, Thambiratnam DP, Chan TH, Nguyen T (2016) Vibration characteristics and damage detection in a suspension bridge. *J Sound Vib* 375:254–274
12. Lin S-W, Yi T-H, Li H-N, Ren L (2017) Damage detection in the cable structures of a bridge using the virtual distortion method. *J Bridge Eng* 22(8):04017039
13. Pan H, Azimi M, Yan F, Lin Z (2018) Time-frequency-based data-driven structural diagnosis and damage detection for cable-stayed bridges. *J Bridge Eng* 23(6):04018033
14. Liu J, Chen S, Bergés M, Bielak J, Garrett JH, Kovačević J, Noh HY (2020) Diagnosis algorithms for indirect structural health monitoring of a bridge model via dimensionality reduction. *Mech Syst Signal Process* 136:106454
15. Alamdar MM, Rakotoarivelo T, Khoa NLD (2017) A spectral-based clustering for structural health monitoring of the Sydney harbor bridge. *Mech Syst Signal Process* 87:384–400
16. Ni F, Zhang J, Noori MN (2020) Deep learning for data anomaly detection and data compression of a long-span suspension bridge. *Comput-Aided Civ Infrastruct Eng* 35(7):685–700
17. Alamdar MM, Dang Khoa NL, Wang Y, Samali B, Zhu X (2019) A multi-way data analysis approach for structural health monitoring of a cable-stayed bridge. *Struct Health Monit* 18(1):35–48
18. Feng D, Feng MQ, Ozer E, Fukuda Y (2015) A vision-based sensor for noncontact structural displacement measurement. *Sensors* 15(7):16557–16575
19. Chen G, Liang Q, Zhong W, Gao X, Cui F (2021) Homography-based measurement of bridge vibration using UAV and DIC method. *Measurement* 170:108683

20. Bai Y (2022) Deep learning with vision-based technologies for structural damage detection and health monitoring. PhD dissertation
21. Gheitasi A, Ozbulut OE, Usmani S, Alipour M, Harris DK (2016) Experimental and analytical vibration serviceability assessment of an in-service footbridge. *Case Stud Nondestruct Test Eval* 6:79–88
22. Gibbs MM, Kwon DK, Kareem A (2019) Data-enabled prediction framework of dynamic characteristics of rural footbridges using novel citizen sensing approach. *Front Built Environ* 5:38
23. Baisthakur S, Chakraborty A (2020) Modified Hamiltonian Monte Carlo-based Bayesian finite element model updating of steel truss bridge. *Struct Control Health Monit* 27(8):2556
24. White R, Alexander N, Macdonald J, Bocian M (2020) Characterisation of crowd lateral dynamic forcing from full-scale measurements on the Clifton suspension bridge. *Structures*, vol 24. Elsevier, Netherlands, pp 415–425
25. Lee JJ, Shinozuka M (2006) Real-time displacement measurement of a flexible bridge using digital image processing techniques. *Exp Mech* 46(1):105–114
26. Ji YF, Chang CC (2008) Nontarget image-based technique for small cable vibration measurement. *J Bridge Eng* 13(1):34–42
27. Lee J, Lee KC, Cho S, Sim S-H (2017) Computer vision-based structural displacement measurement robust to light-induced image degradation for in-service bridges. *Sensors* 17(10):2317
28. Liu B, Zhang D, Guo J (2016) Vision-based displacement measurement sensor using modified Taylor approximation approach. *Opt Eng* 55(11):114103
29. Feng D, Feng MQ (2017) Experimental validation of cost-effective vision-based structural health monitoring. *Mech Syst Signal Process* 88:199–211
30. Xu Y, Brownjohn JMW, Huseynov F (2019) Accurate deformation monitoring on bridge structures using a cost-effective sensing system combined with a camera and accelerometers: case study. *J Bridge Eng* 24(1):05018014
31. Wu ZY, Shenton HW III, Mo D, Hmosze M (2021) Integrated video analysis framework for vision-based comparison study on structural displacement and tilt measurements. *J Struct Eng* 147(9):05021005
32. Chen JG, Adams TM, Sun H, Bell ES, Büyüköztürk O (2018) Camera-based vibration measurement of the World War I memorial bridge in Portsmouth, New Hampshire. *J Struct Eng* 144(11):04018207
33. Su C, Sicheng Z, Zhijun D, Xiaojun L, Yue Z (2019) Visual identity-based earthquake ground displacement testing method. *Shock Vib.* <https://doi.org/10.1155/2019/2585423>
34. Guo J, Zhu C (2016) Dynamic displacement measurement of large-scale structures based on the Lucas-Kanade template tracking algorithm. *Mech Syst Signal Process* 66:425–436
35. Lydon D, Lydon M, Taylor S, Del Rincon JM, Hester D, Brownjohn J (2019) Development and field testing of a vision-based displacement system using a low cost wireless action camera. *Mech Syst Signal Process* 121:343–358
36. Xiao P, Wu Z, Christenson R, Lobo-Aguilar S (2020) Development of video analytics with template matching methods for using camera as sensor and application to highway bridge structural health monitoring. *J Civ Struct Health Monit* 10(3):405–424
37. Dong C, Celik O, Catbas FN, O'Brien EJ, Taylor S (2020) Structural displacement monitoring using deep learning-based full field optical flow methods. *Struct Infrastruct Eng* 16(1):51–71
38. Dong C, Catbas FN (2019) A non-target structural displacement measurement method using advanced feature matching strategy. *Adv Struct Eng* 22(16):3461–3472
39. Li S, Sun L (2020) Detectability of bridge-structural damage based on fiber-optic sensing through deep-convolutional neural networks. *J Bridge Eng* 25(4):04020012
40. Szeliski R (2010) Computer vision: algorithms and applications. Springer, New York, pp 3–20
41. Bouguet JY (2001) Pyramidal implementation of the affine Lucas Kanade feature tracker description of the algorithm. Intel Corp 5(1–10):4
42. Bai Y, Abdullah RM, Sezen H, Yilmaz A (2021) Automatic displacement and vibration measurement in laboratory experiments with a deep learning method. In: 2021 IEEE Sensors, IEEE. pp 1–4
43. Yoon H, Shin J, Spencer BF Jr (2018) Structural displacement measurement using an unmanned aerial system. *Comput-Aided Civ Infrastruct Eng* 33(3):183–192
44. Tomiczek AP, Whitley TJ, Bridge JA, Ifju PG (2019) Bridge inspections with small unmanned aircraft systems: case studies. *J Bridge Eng* 24(4):05019003
45. Ribeiro D, Santos R, Cabral R, Saramago G, Montenegro P, Carvalho H, Correia J, Calçada R (2021) Non-contact structural displacement measurement using unmanned aerial vehicles and video-based systems. *Mech Syst Signal Process* 160:107869
46. Perry BJ, Guo Y (2021) A portable three-component displacement measurement technique using an unmanned aerial vehicle (UAV) and computer vision: a proof of concept. *Measurement* 176:109222
47. Khuc T, Nguyen TA, Dao H, Catbas FN (2020) Swaying displacement measurement for structural monitoring using computer vision and an unmanned aerial vehicle. *Measurement* 159:107769
48. Bai Y, Sezen H, Yilmaz A, Qin R (2023) Bridge vibration measurements using different camera placements and techniques of computer vision and deep learning. *Advances in bridge engineering*
49. He K, Gkioxari G, Dollár P, Girshick R (2017) Mask R-CNN. In: Proceedings of the IEEE International Conference on Computer Vision. pp 2961–2969
50. Sun K, Zhao Y, Jiang B, Cheng T, Xiao B, Liu D, Mu Y, Wang X, Liu W, Wang J (2019) High-resolution representations for labeling pixels and regions. *arXiv preprint arXiv:1904.04514*
51. Bai Y, Zha B, Sezen H, Yilmaz A (2023) Engineering deep learning methods on automatic detection of damage in infrastructure due to extreme events. *Struct Health Monit* 22(1):338–352
52. Press WH, Teukolsky SA (1990) Savitzky–Golay smoothing filters. *Comput Phys* 4(6):669–672
53. Cooley JW, Tukey JW (1965) An algorithm for the machine calculation of complex Fourier series. *Math Comput* 19(90):297–301
54. CSI (2022) SAP2000 advanced structural analysis program, Version 23.0.0. Computers and Structures, Inc., Walnut Creek
55. AREMA (2019) Manual for Railway Engineering. American Railway Engineering and Maintenance-of-Way Association
56. Rakoczy AM (2012) Development of system reliability models for railway bridges. The University of Nebraska-Lincoln, Lincoln, Nebraska, pp 66–80
57. Chen WF, Duan L (2014) Bridge engineering handbook, five volume set, vol 2. CRC Press, Boca Raton, Florida, pp 35–80
58. Rakoczy AM, Nowak AS (2018) Live load spectra for railway bridges in USA. 64 Konferencja Naukowa
59. (2023) CSI Technical Knowledge Base, Computers and Structures, Inc. <https://wiki.csiamerica.com>

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.