Smart Structures and Systems, *Vol. 25, No. 6 (2020) 733-749*
DOI: https://doi.org/10.12989/sss.2020.25.6.733

Autonomous vision-based damage chronology for spatiotemporal condition assessment of civil infrastructure using unmanned aerial vehicle

Tarutal Ghosh Mondal [1a] and Mohammad R. Jahanshahi [*1,2]

[1] *Lyles School of Civil Engineering, Purdue University, West Lafayette, IN, USA*
[2] *School of Electrical and Computer Engineering, Purdue University, West Lafayette, IN, USA*

(*Received October 7, 2019, Revised December 28, 2019, Accepted March 22, 2020*)

Abstract. This study presents a computer vision-based approach for representing time evolution of structural damages leveraging a database of inspection images. Spatially incoherent but temporally sorted archival images captured by robotic cameras are exploited to represent the damage evolution over a long period of time. An access to a sequence of time-stamped inspection data recording the damage growth dynamics is premised to this end. Identification of a structural defect in the most recent inspection data set triggers an exhaustive search into the images collected during the previous inspections looking for correspondences based on spatial proximity. This is followed by a view synthesis from multiple candidate images resulting in a single reconstruction for each inspection round. Cracks on concrete surface are used as a case study to demonstrate the feasibility of this approach. Once the chronology is established, the damage severity is quantified at various levels of time scale documenting its progression through time. The proposed scheme enables the prediction of damage severity at a future point in time providing a scope for preemptive measures against imminent structural failure. On the whole, it is believed that the present study will immensely benefit the structural inspectors by introducing the time dimension into the autonomous condition assessment pipeline.

Keywords: damage chronology; vision-based inspection; autonomous condition assessment; UAV; civil infrastructures

1. Introduction

1.1 Motivation

Civil infrastructures undergo deterioration over time owing to overloading or unfavorable environmental conditions. This calls for periodic inspection of structures in order to prevent sudden failure or to avoid any untoward human casualties caused by unserviceable infrastructure conditions. The existing inspection techniques are predominantly manual, and consequently time consuming, expensive, subjective, and risky. Numerous studies in recent years focused on autonomous inspection techniques based on the latest advancements made in the areas of computer vision and deep learning (Jahanshahi and Masri 2014, Adhikari *et al*. 2014, Duan *et al*. 2019). Spencer *et al*. (2019) provides an exhaustive review of available literature on this topic. A number of investigations in the past explored autonomous damage identification from visual data exploiting various image processing (Yamaguchi and Hashimoto 2010), machine learning and convolutional neural networks (CNN) (Cha *et al*. 2017, Kim *et al*. 2019, Chen and Jahanshahi 2018, 2019) based approaches. Damage quantification also gained some attraction from the research community (Jahanshahi and Masri 2013, Shan *et al*. 2016, Jahanshahi *et al*. 2017a). However, most of the previous studies found in literature were invariably agnostic to the time dimension. It is often important to understand how fast a damage is progressing and how long it may take to reach the limit state of collapse. However, disregarding the temporal information in the state-of-the-art damage assessment pipeline makes such information scarce, preventing the inspectors act preemptively to minimize the cost incurred due to the damage. It is therefore necessary to address this knowledge gap existing in this important area of research, which is the focus of this study.

1.2 Related works

A number of studies in the past explored time-based evaluation of structural defects. Digital image correlation (DIC) is exploited by many researchers (Ghorbani *et al*. 2015) to measure full-field displacement and strain. However, this technique relies on static camera, and therefore can only be used in situations where the damage location is known a priori. Moreover, it necessitates painting of speckle patterns on the structure under investigation to produce distinct visual features, which is not feasible in large structures like buildings and bridges and in situations where the surface of the structure is physically inaccessible. The approach presented in the current study is free from all such limitations. A movable (hand-held) SLR camera and a camera mounted on an unmanned aerial vehicle (UAV) are used for data collection which eliminated the need of prior knowledge about damage locations. Besides, this method is contactless in true

*Corresponding author, Ph.D., Assistant Professor,
E-mail: jahansha@purdue.edu
[a] Ph.D. Student, E-mail: tghoshmo@purdue.edu

http://www.techno-press.org/?journal=sss&subpage=7
ISSN: 1738-1584 (Print), 1738-1991 (Online)

sense of the term, as it does not require any speckle pattern to be painted on the surface to be inspected. Kong and Li (2019) used image overlapping technique to detect fatigue cracks in civil infrastructures. The authors relied on differential image features engendered by crack breathing as indicators for crack identification. The proposed approach permits camera movement and environmental changes to a certain extent. However, it requires prior knowledge about the defect location. Detection of fatigue crack in steel bridges was also studied by Kong and Li (2018) exploiting video-based feature tracking. Movement of each feature was tracked through a video stream and the presence of a crack was indicated by differential movement pattern exhibited by the feature points inside a localized circular region. However, optical flow-based feature tracking process cannot effectively deal with viewpoint changes (Tanathong and Lee 2013) requiring fixed camera orientation. These limitations are dispelled in the current study in many ways. The approach presented in this study is scale invariant, and robust against noise intrusion and changes in illumination condition to a great extent. It can detect large motion. Moreover, it affords the flexibility of capturing images from varied camera positions and orientations, which is a major advantage of this approach. Jahanshahi *et al*. (2011) proposed a vision-based approach for estimating damage evolution through multi-image stitching and scene reconstruction. However, the camera was constrained in this study in regard to translation. In the present study, this constraint is relaxed enabling the camera to rotate and translate without any restrain. Besides, the approach presented in this study (Jahanshahi *et al*. 2011) is not fully autonomous in the sense that a human inspector needs to compare the current scene with its previous condition and deduce the damage evolution manually. In other words, the proposed technique relied on inspector's judgment vis-à-vis evolution of the damage. This makes the entire procedure tedious, labor-intensive, subjective and qualitative. The present study addresses this limitation by including an autonomous localization and quantification module in the damage assessment pipeline making the entire process faster and more efficient. Additionally, quantitative and time-based evaluation of damage severity makes it possible to predict residual life of a structure and to take precautionary measures, if necessary.

1.3 Contribution

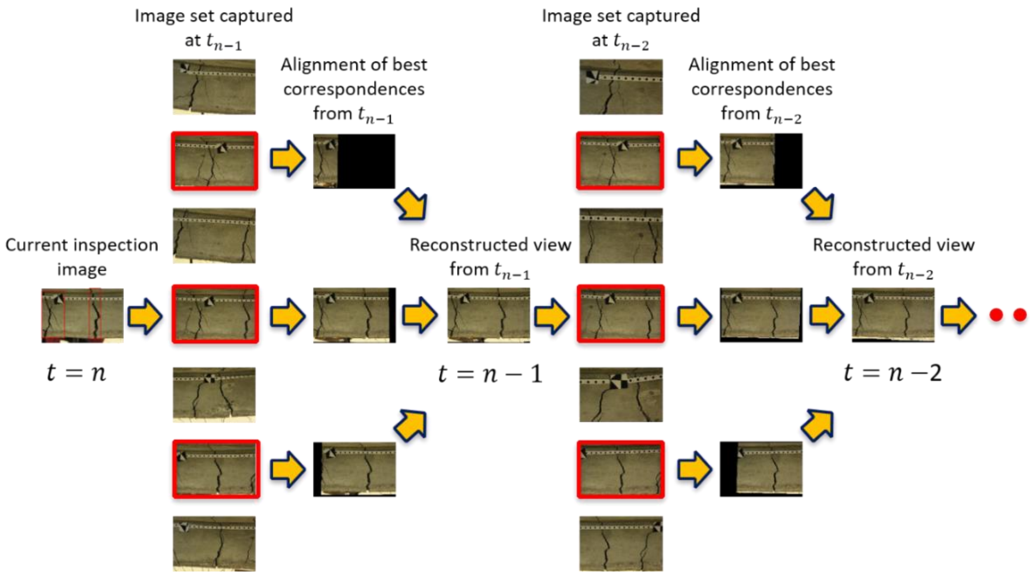
This study presents a novel comprehensive approach to health monitoring of civil infrastructures by introducing a time dimension into the vision-based condition assessment pipeline. It is shown that useful information can be extracted from an archive of inspection images by employing computer vision-based algorithms. Identification of a damage during the course of a recent inspection initiates an exhaustive search into the historical data collected during the previous rounds of inspection. Corresponding images are identified and synthesized to generate a reconstructed view of the scene pertaining to each inspection round. Regions of interest are subsequently extracted from the reconstructed scenes leveraging a CNN-based detection model. This is followed by damage segmentation and quantification exploiting state-of-the-art morphological and image processing techniques paving the way for time-based evaluation of damage severity and cognizant decision making. The methodology presented in this work is robust against noise intrusion and changes in illumination condition. It does not assume any prior knowledge about damage locations and provides a great deal of flexibility with regard to camera poses and orientations. The proposed approach can be applied to data collected by human inspectors using hand-held cameras (e.g., smartphone camera). However, it is most appropriate for autonomous inspection assisted by vision systems mounted on mobile robots including UAVs (Aliakbar *et al*. 2016, Jahanshahi *et al*. 2017b, Aliakbar *et al*. 2019). Cracks on concrete surface is used as a case study to demonstrate the feasibility of this approach. However, it can be extended to other defect categories such as spalling and corrosion, with appropriate modifications. Availability of an inspection database which is complete in terms of coverage of the damaged areas and that affords adequate overlap with adjacent images is a prerequisite for this approach. Besides, the algorithm may not perform well in absence of adequate visual features in the inspection images. In such situations, IMU and GPS information can be exploited for accurate scene reconstruction which is a scope for future research.

1.4 Scope

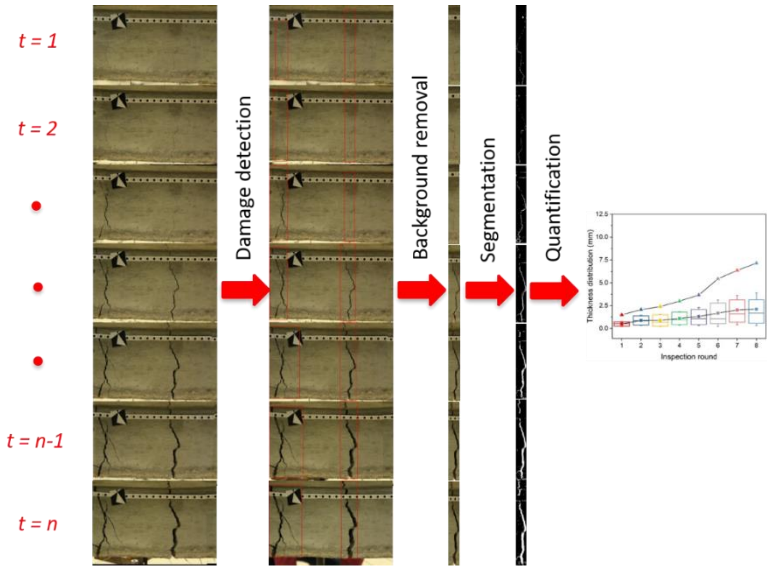
The remaining of the manuscript is arranged in the following order. Section 2.1 presents an overview of the test protocol and data collection procedure. Section 2.2 deals with various components of the correspondence identification technique adopted in this study. The details of damage detection approach are presented in Section 2.3. The necessary theoretical background for damage quantification is presented in Section 2.4. The results are presented and discussed in Section 3. Conclusions are summarized in Section 4. Finally, scope for future research is outlined in Section 5.

2. Methodology

As a case study, cracks on concrete surface are used to illustrate the nuts and bolts of this approach. A reinforced concrete beam was tested in the laboratory subjecting it to a gradual load increment in order to simulate a progressive damage. Cracks appearing on the beam surface were photographed after every stage of load increment and the images were time-stamped and saved in specific folders. This was accomplished by a hand-held SLR camera and a camera mounted on a UAV to emulate actual robot-based data collection where camera positions and orientations are not fully controllable. After the final round of load increment, the entire data set representing time evolution of concrete cracks were available for further analysis. Fig. 1 presents an overview of the proposed algorithm. Identification of damage in the most current data set engenders an exhaustive search in the immediately



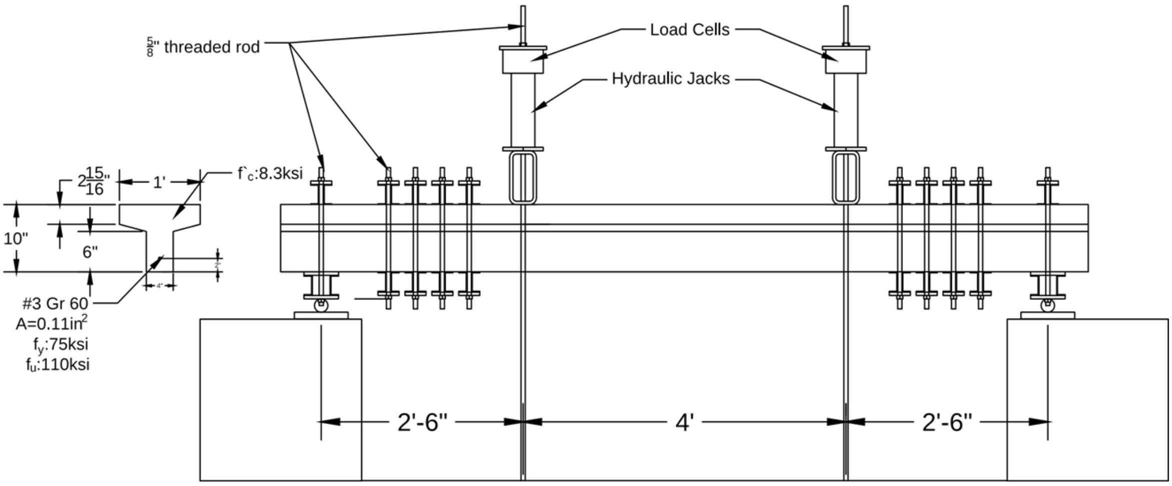
(a)



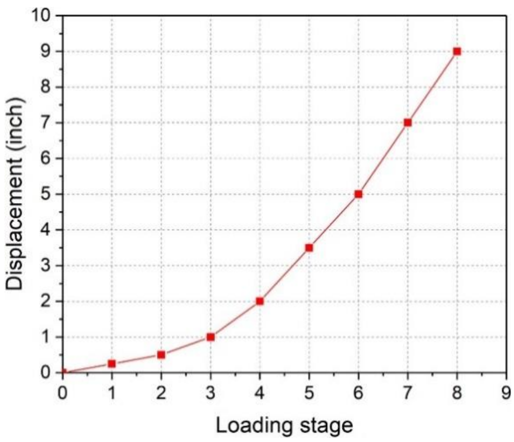
(b)

Fig. 1 The layout of the proposed approach – (a) Correspondence identification from the preceding data set based on spatial proximity, registration of the best correspondences onto the plane of the current reference image, and repetition of the same procedure over all previous data sets to generate a temporally ordered set of 2D reconstructions of the concerned damaged area. (b) Detection of damage on the reconstructed views from the past, extraction of interest area to remove nonessential background, damage segmentation, followed by quantification and time-based visualization

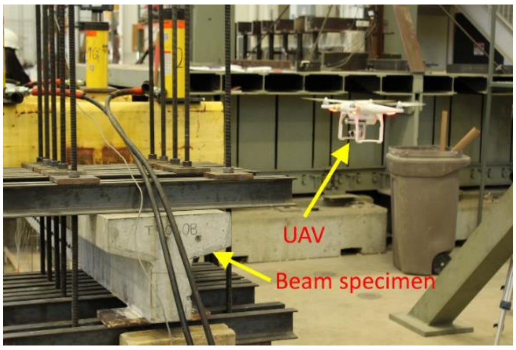
preceding image set looking for correspondences. Speeded up robust features (SURF) (Bay *et al*. 2006) algorithm is used to identify interest points in the current inspection image and also in every single image in the previous data set. Feature matching is carried out based on Euclidean distance between two descriptor vectors and the candidates with large number of matched features are designated as potential correspondences. Homography transformation is computed for each selected correspondence through linear least square method and subsequent nonlinear refinement using Levenberg-Marquardt algorithm (Levenberg 1944, Marquardt 1963), which is followed by registration of the corresponding images onto the plane of the current reference image. The warped images are then stitched to form a complete 2D reconstructed view of the concerned damage region from the immediately preceding data set. This procedure is repeated for all the previous data sets captured at different points in time considering the reconstructed view from the immediately succeeding data set as the reference. Temporally ordered set of 2D reconstructions thus produced chronicles the evolution of a damage in a manner conducive to time-based reasoning and lucid visual interpretation, and forms the basis for the next stage of the proposed algorithm, namely, damage identification and quantification. A notable detection algorithm called Faster RCNN (Ren *et al*. 2015) is leveraged to this end to localize the cracked area in the reconstructed images. The relevant portion of the images



(a) Experimental setup



(b) Loading protocol



(c) Data collection

Fig. 2 Experimental setup for data collection

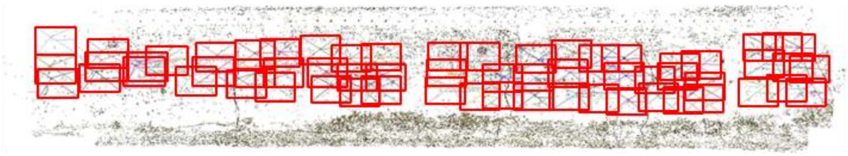
containing the cracks are then cropped out to get rid of the remaining nonessential background (undamaged), inclusion of which may have debilitating effect on the performance of the subsequent segmentation and quantification processes due to noise infusion. The cropped pixels are then segmented using a morphological approach, forming the basis for crack thickness quantification using distance transform method (Lee *et al*. 2013, Zhu *et al*. 2011). The approach presented in this study can be extended to other defect categories such as spalling and corrosion, with appropriate modifications.

2.1 Experimental setup and data collection

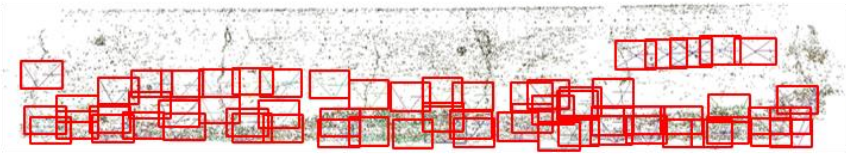
The database required for validation of the proposed approach was generated by testing a reinforced concrete T-beam in the laboratory under gradually increasing load in four-point bending configuration as shown in Fig. 2(a). The beam was tested in displacement control mode, and the applied displacement is shown in Fig. 2(b) as a function of loading step. After every step of displacement increment, an intermission was appropriated during which the entire span of the beam was photographed using a hand-held SLR camera as well as a camera mounted on a UAV to capture the cracks that appeared on the surface (Fig. 2(c)). The SLR camera was displaced laterally to photograph different segments along the span and depth of the beam ensuring adequate overlap between successive images (Fig. 3). The camera movement was not controlled, and the data collection path varied over inspection round as evident from Fig. 3. An archive of time-stamped images representing various levels of degradation was thus produced mimicking time-evolution of damage in concrete structures. This data set formed the basis for subsequent analyses which are described in the following sections.

2.2 Correspondence detection and alignment

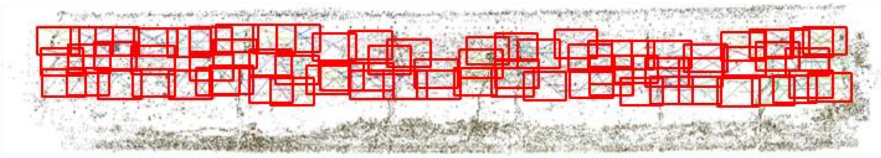
Theoretical formulation of this algorithm presupposes the availability of a comprehensive visual data set built perennially through collection of images over several rounds of routine inspection by a human inspector or by an inspection robot (Esser and Huston 2005, Boller *et al*. 2015, Myung *et al*. 2014). If a defect is detected during the inspection of a structure, it becomes necessary to know the history of evolution of the defect. That necessitates probing into the data collected during previous rounds of inspection. The first challenge that is confronted to this end is identifying the relevant images corresponding to the defective region from a large database of archival images



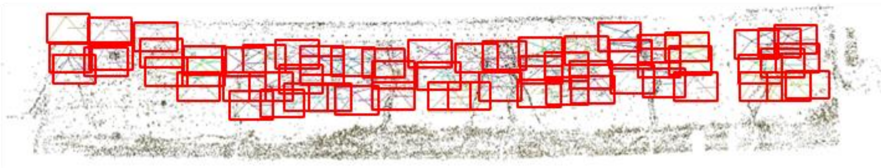
(a) Inspection round - 1



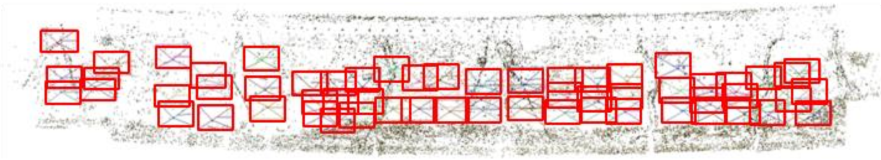
(b) Inspection round - 2



(c) Inspection round - 3



(d) Inspection round - 4



(e) Inspection round - 5

Fig. 3 Data collection path of a hand-held SLR camera for different inspection rounds. The rectangular boxes denote the camera poses and orientations, and the point clouds denote the 3D scene reconstructions of the beam for each inspection round. It should be noted that the data collection path was not constant and it varied over inspection round

(Yeum *et al*. 2017, 2019). This can be achieved through a sequence of widely used computer vision algorithms such as feature detection, feature matching and image registration, as explained in the following sections.

2.2.1 Feature detection

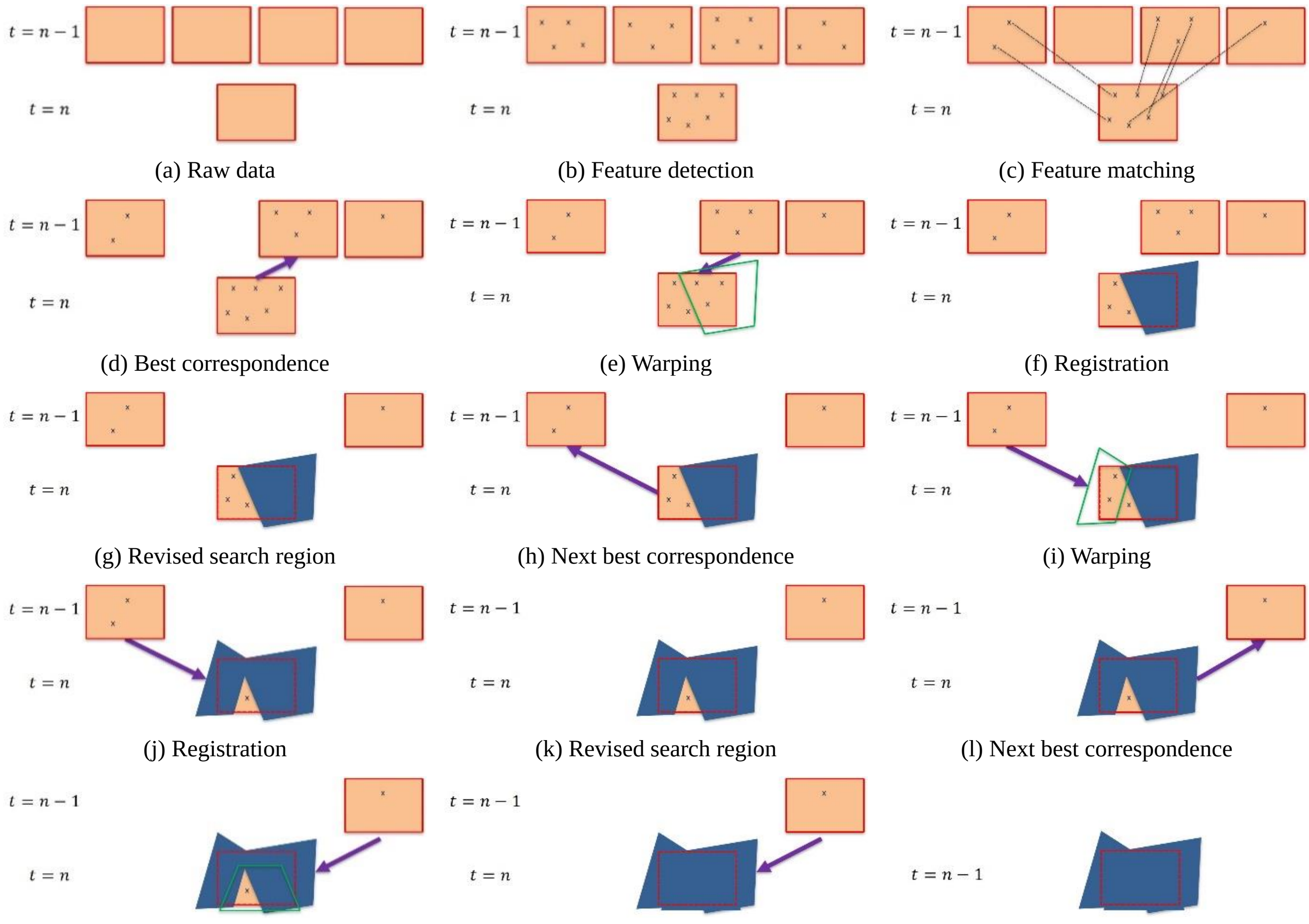
The first step in the correspondence identification pipeline is the detection of features or interest points (Fig. 4(b)). Features are unique patterns which can be easily tracked and compared across several images. There are a number of techniques available in literature for detecting interest points in images. SURF algorithm is one such technique which is leveraged in this study. This algorithm locates high-variance interest points in an image which are invariant to scale, viewpoint and illumination changes. A local dominant direction is associated with each interest point and a 64 element normalized descriptor vector representing the local gray level variations with respect to the dominant direction is computed at each such point. The reader may refer to the original paper by Bay *et al*. (2006) for more detailed discussion about this algorithm.

2.2.2 Feature matching

Feature detection is followed by feature matching (Fig. 4(c)), the objective of which is to identify the best match for a feature in one image from all the features in another image. Number of matched features is an indication of degree of resemblance between two images. Brute-Force matcher is used in this study, where the Euclidean distance between two descriptor vectors is used for similarity comparison. Two best matches are drawn for each feature in the first image. On occasion, the second best match is found to be very close to the best match owing to noise or other reasons. Such anomalies are tackled by computing the ratio of the closest distance to the second closest distance, and discarding all matches where this ratio is greater than 0.75 as suggested by Lowe (2004). This eliminates 95% of the false matches as shown in Fig. 5. However, a small number of outliers are retained at this stage.

2.2.3 Image registration

The image in the immediately previous data set having the largest number of matched features vis-à-vis the



(a) Raw data (b) Feature detection (c) Feature matching

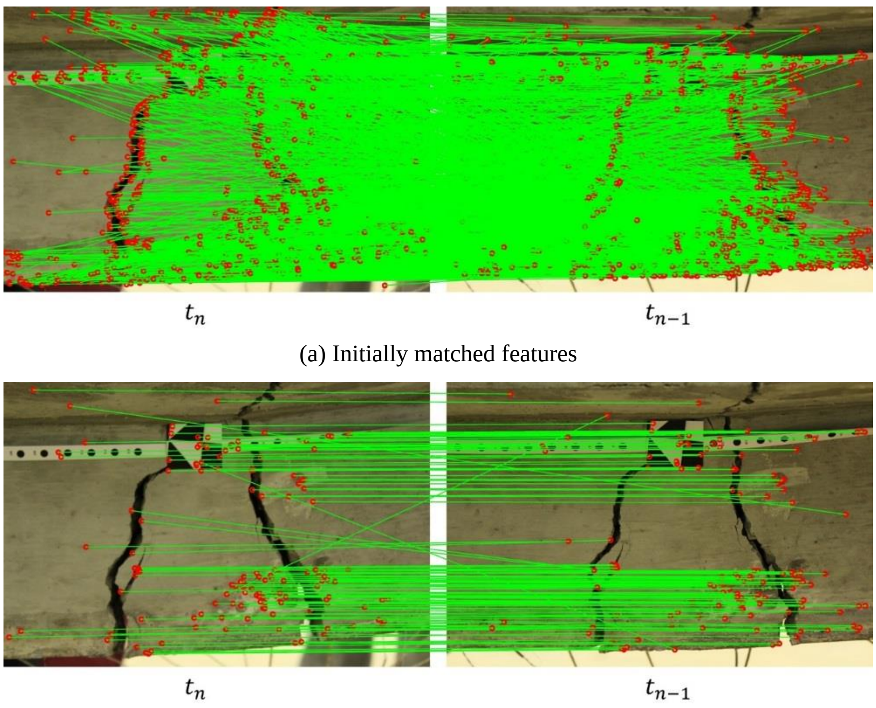
(d) Best correspondence (e) Warping (f) Registration

(g) Revised search region (h) Next best correspondence (i) Warping

(j) Registration (k) Revised search region (l) Next best correspondence

(m) Warping (n) Registration (o) Final reconstruction

Fig. 4 Illustrative diagrams outlining the steps for generating reconstructed view from previous data set



(a) Initially matched features

(b) Matched features after applying Lowe's ratio test

Fig. 5 Feature detection and matching

reference image in the current data set is designated as the best correspondence (Fig. 4(d)). Damage chronology can only be established when the corresponding images from previous data sets are aligned to the plane of the reference image. This requires estimation of the homography matrix between the reference image and the correspondences in the immediately preceding data set. It may be noted here that the homography is a 3×3 transformation matrix which maps the points in one image to the corresponding points in another image. Linear least square method is exploited in combination with an outlier rejection algorithm called RANdom SAmple Consensus (RANSAC) (Fischler and Bolles 1981) to obtain an initial estimate for the homography matrix. This is followed by a nonlinear refinement of the estimated homography matrix using Levenberg-Marquardt algorithm based on the inlier points alone. The estimated homography matrix is then used to warp (Fig. 4(e)) and register (Fig. 4(f)) the best correspondence on the plane of the reference image. Following this, the matched features corresponding to the best correspondence are eliminated from the list of available features for the reference image (Fig. 4(g)), and the next best correspondence is determined based on the revised list of matched features (Fig. 4(h)). This next best correspondence is then registered on the plane of the reference image (Figs. 4(i) and (j)) in a similar fashion following the same procedure mentioned previously in this section. This process of correspondence identification and alignment is continued until the number of residual matched points corresponding to the reference image drops below a predefined threshold (100 in this study) or the number of identified correspondences reaches a preset value (which is set to 10 in this study). Upon completion of this process, all the warped correspondences are stitched together producing a complete 2D reconstruction (Fig. 4(o)) depicting the prior condition of the scene in the reference image (Fig. 6). The reconstructed view acts as a reference image for the next round of iteration, where the correspondences are identified from the immediately preceding data set. Eventually, an ordered set of reconstructed views are obtained portraying the evolution of a scene through time.

2.3 Damage detection

This section describes the process for autonomous segmentation of damages in the chronologically ordered reconstructed views of the scene under consideration. Previous studies primarily focused on two approaches for detecting cracks in images, namely, edge-based techniques and morphological techniques. Jahanshahi *et al*. (2009) compared the pros and cons of the two approaches and concluded that morphological techniques outperform edge-based techniques in presence of non-crack edges. Therefore, morphological approach is adopted in this study for extracting cracks from the images. However, presence of surface irregularities may produce false positives leading to inaccurate segmentation (Jahanshahi and Masri 2013). This can be averted by secluding the damaged region of interest from the remaining image. Deep learning-based approaches have been used by several researchers in the past (Cha *et*

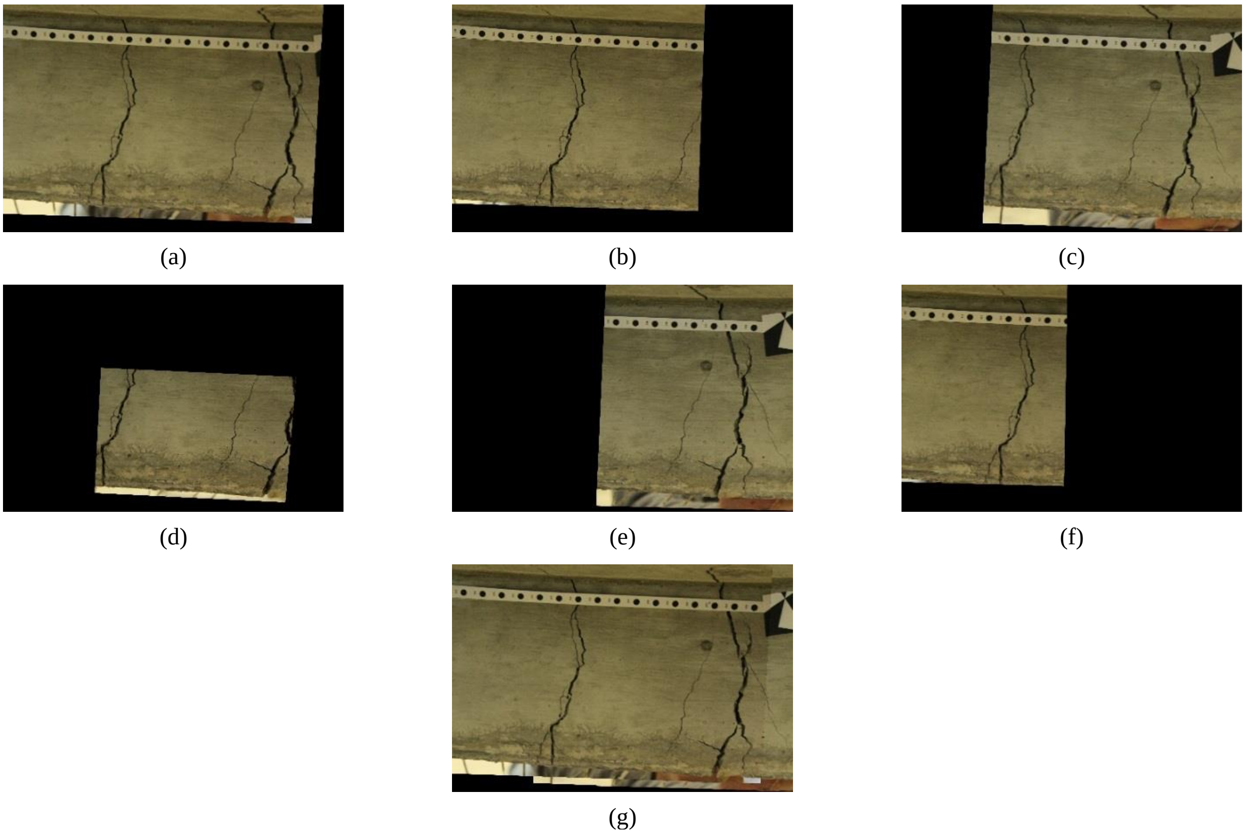


Fig. 6 (a)-(f) Warping and registration of correspondences, (g) View synthesis producing complete 2D reconstruction *al*. 2017, 2018, Chen and Jahanshahi 2018, 2019) to localize defects in images. This study leverages Faster RCNN algorithm to this end. This eliminates a large part of the nonessential background significantly diminishing the scope of noise infusion in the morphology-based segmentation process.

2.3.1 Damage localization using Faster RCNN

In Faster RCNN, a CNN is first used to generate a feature map from the input image. Inception-ResNet-v2 network (Szegedy *et al*. 2001a), which incorporates Residual connections (He *et al*. 2001) and Inception module (Szegedy *et al*. 2001b), is used to this end in this study. Thereafter, Region Proposal Network (RPN) (Ren *et al*. 2015) is used to generate region proposals. RPN is a fully convolutional network trained to predict object bounds and objectness scores. Following this, Fast RCNN (Girshick 2015) module is utilized to classify the region proposals and to refine the bounding box coordinates. The RPN and the Fast RCNN modules are unified into a single network enabling sharing of convolutional layers (Fig. 7). The details of Faster RCNN algorithm can be found in Ren *et al*. (2015).

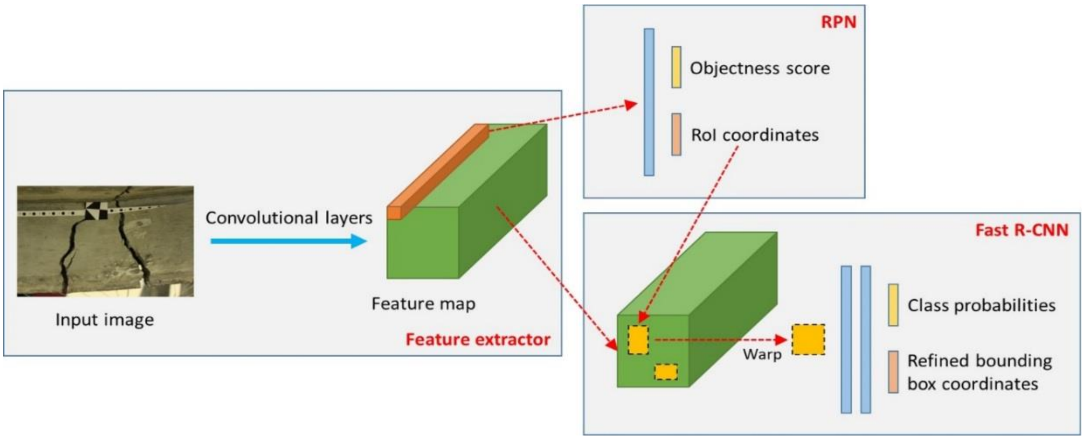


Fig. 7 Faster RCNN architecture

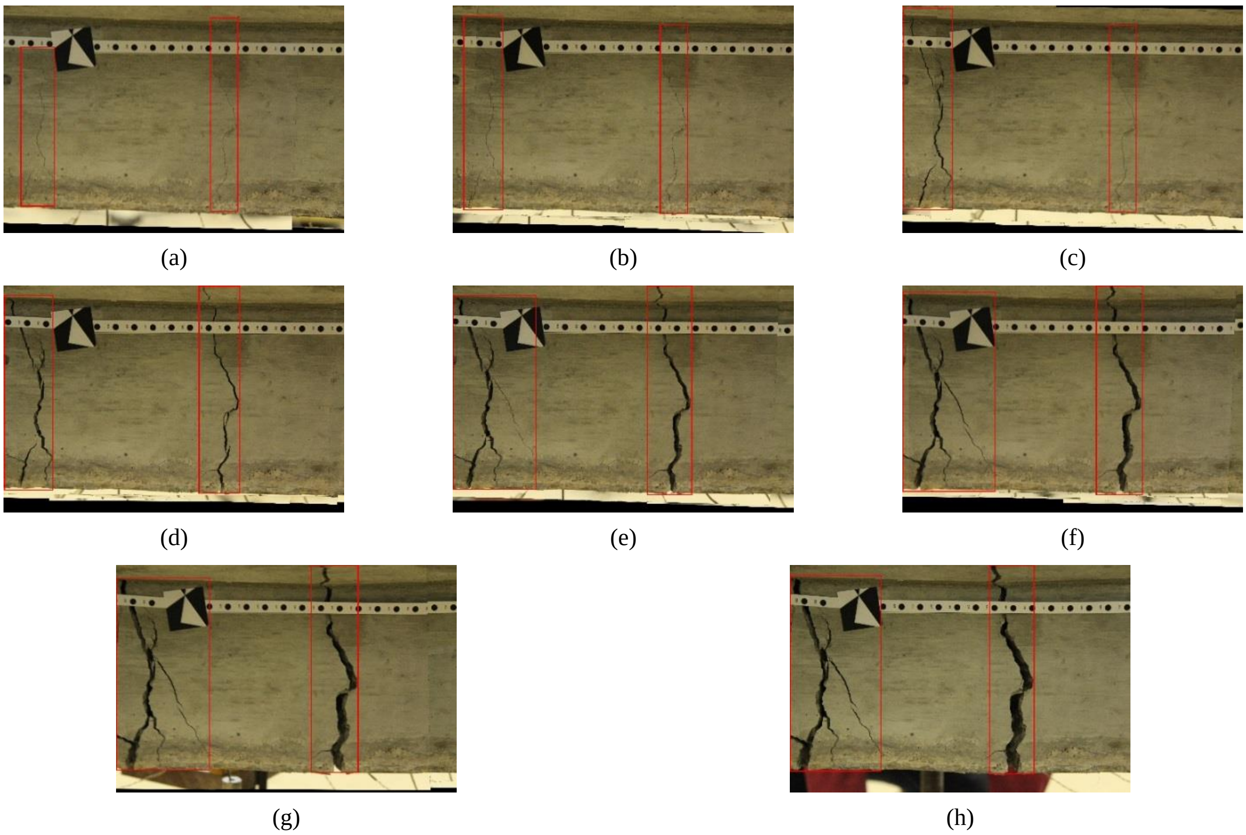


Fig. 8 Damage chronology produced by successive view synthesis and alignment of correspondences from previous inspection data sets. Cracks detected by Faster RCNN algorithm are highlighted by rectangular bounding boxes

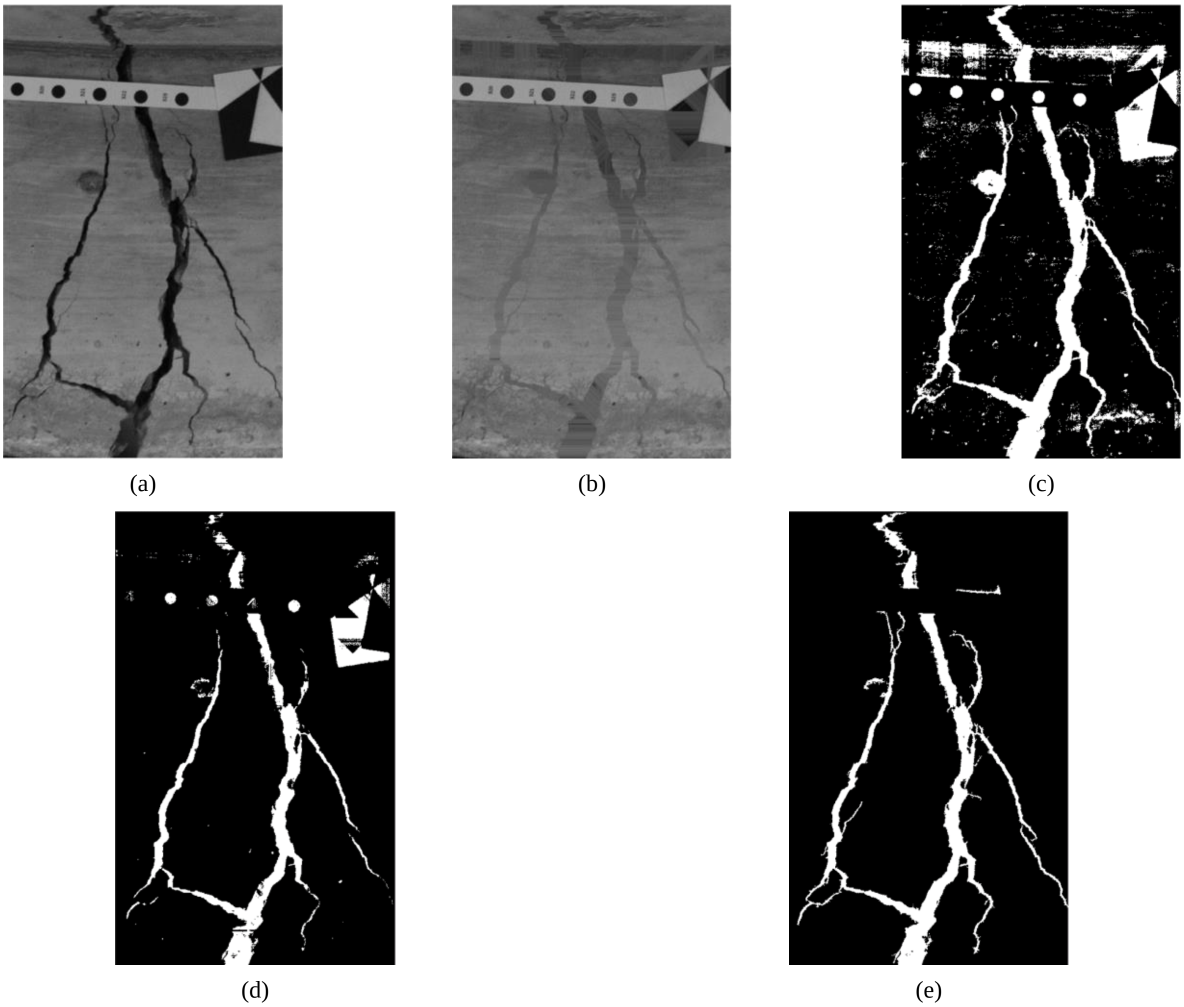


Fig. 9 Steps involved in the segmentation process – (a) Grayscale image (I); (b) Result of $max[\,(I \circ S_{\{0°,45°,90°,135°\}}) \bullet S_{\{0°,45°,90°,135°\}}, I]$; (c) Crack map ($T$) generated by Eq. (1); (d) Binary image obtained by applying Otsu's threshold to T, and; (e) Final segmentation mask obtained after post-processing a noise removal

The Faster RCNN algorithm is implemented using TensorFlow open-source library. The input images are horizontally flipped randomly with a probability of 0.5 to execute data augmentation. Subsequently, features are extracted from the input image using a sequence of convolutional layers which are a part of the Inception-ResNet-v2 network. A 3×3 sliding window is applied to the feature map generated by the last shared convolutional layer mapping it down to a lower dimension. At each sliding window location, a set of 9 anchor boxes having different scales and aspect ratios are considered as region proposals. The anchor boxes are strided by 8 pixels along the height and the width. A large number of region proposals are generated for each image leading to multiple detections. Duplicate boxes are eliminated using a greedy technique called non-maximum suppression (NMS) (Hosang *et al.* 2017). The weights of the Inception-ResNet-v2 network are initialized by a model pretrained on MSCOCO data set (Lin *et al.* 2001) and fine-tuned thereon using Stochastic Gradient Descent (SGD) algorithm (Bottou 2010) with a momentum value of 0.9. Gradient clipping is employed to avert the problem of exploding gradient. The initial learning rate is set to 0.003 and is gradually reduced thereafter with training steps.

The network is trained on 686 images containing 1023 crack instances. The training data is generated by loading a T-beam as shown in Fig. 2 and taking pictures of the resulting cracks by means of a movable (hand-held) SLR camera and a camera mounted on a UAV. The performance of the trained network is evaluated on the test data comprising 100 images and 255 crack instances. The test data is produced by loading another T-beam with slightly different cross-section and reinforcement distribution, and photographing the evolving cracks in a similar manner. The predicted bounding boxes (Fig. 8) are compared with ground truth boxes and the results are reported in terms of precision and recall. The proposed algorithm produces a precision of 95.5%, which means that 95.5% of all predicted boxes classified as crack can be designated as correct detections. On the other hand, the recall value is evaluated as 98.6%, indicating that 98.6% of all annotated cracks are correctly detected. It takes roughly about 0.95 seconds at this stage to process a single image of 5184×3456 resolution using a NVIDIA Titan X (Pascal) GPU. It is important to ensure that the predicted bounding boxes enclose respective damage regions completely. This calls for rigorous training of the detection algorithm with stringent requirement imposed on the predicted boxes vis-à-vis overlap with ground truth boxes.

2.3.2 Damage segmentation using morphological techniques

Morphological approach for image segmentation is motivated by the developments in the fields of s*et al*gebra (Minkowski 1989) and topology (Matheron 1975). Dilation and erosion are two rudimentary operations which all morphological methods are based upon. Dilation expands the bright portions of an image, while erosion shrinks the same. These operations applied sequentially form important building blocks for morphological noise removal. Erosion followed by dilation is called morphological opening, which removes bright sharp details from the image. On the other hand, the same operations when applied in the reverse order constitute morphological closing which seeks to remove dark details from an image. Salembier (1990) integrated these morphological concepts with bottom-hat transform to propose an algorithm for identification of dark defects in images. The following equation shows a slightly modified version of the algorithm (Jahanshahi *et al*. 2013) which is used in this study.

$$T = max[\,(I \circ S_{\{0°,45°,90°,135°\}}) \bullet S_{\{0°,45°,90°,135°\}}, I] - I \quad (1)$$

where, ‘ ∘ ‘ and ‘ • ‘ denote morphological opening and closing operations, respectively. I is the gray-scale image and S is a structuring element. A structuring element is a matrix that decides which neighborhood pixels are included in the morphological operations. The structuring element should be suitably chosen as it determines the shape and size of the cracks that can be extracted from an image. Jahanshahi *et al*. (2013) proposed an adaptive approach for estimating an appropriate structural element size based on crack size, camera parameters and camera-to-object distance. A linear structuring element (a structuring element which is line-shaped) with four different orientations $(0°, 45°, 90°, 135°)$ is used in this study to make the filter invariant to crack orientation. The crack map so generated (T) was subjected to Otsu’s thresholding (Otsu 1979), followed by a series of post-processing and noise removal strategies (Fig. 9) to obtain the final binary segmentation mask as shown in Fig. 10. The post-processing scheme involves removing small areas, filling small holes, bridging unconnected pixels, and removing spur pixels and isolated pixels. The entire crack region is segmented in this approach, unlike edge detection-based techniques where only the crack boundaries are extracted. Apart from that, morphological approaches are divested of the time-consuming and tedious data annotation and training processes which are required by typical deep learning-based semantic segmentation algorithms.

2.4 Damage quantification

Thickness is an effective indicator for severity of cracks. This section presents the crack thickness quantification algorithm that is used in this study. Some researchers in the past (Yu *et al*. 2007) resorted to boundary-to-boundary approach to visually measure the crack thickness. In this approach, the crack thickness at a boundary point is evaluated as the distance to the nearest point on the other boundary. However, the limitation of this approach is that the crack thickness line is usually not normal to the centerline. Moreover, the thickness measured at a boundary point may not be identical to the same measured at the corresponding thickness point located on the other boundary of the crack. These limitations can be redressed by employing centerline-based techniques such as orthogonal line method (Jahanshahi *et al*. 2013, Jahanshahi and Masri 2013) and distance transform method (Lee *et al*. 2013, Zhu *et al*. 2011). The latter approach is adopted in this study. This method begins by finding the centerline of the crack. Researchers in the past exploited various methods for locating the crack centerline in an image. Jahanshahi *et al*. (2017a) used fast marching algorithm which was originally proposed by (Van Uitert and Bitter 2007). A number of studies (Jahanshahi *et al*. 2013, Jahanshahi and Masri 2013), on the other hand, employed morphological thinning operation on binary crack maps, which is followed in this study. The thickness at a given centerline point is given by twice the shortest distance to any of the boundaries. Quasi-Euclidean distance transform is used in this study to find the closest pixel on the boundaries. The effectiveness of the segmentation and quantification approach adopted in this study was previously established by Jahanshahi *et al*. (2013) and the same is not repeated here. Each of the black circles on the tape as observed in Fig. 8 had a diameter of 5 mm and was represented by 92 pixels in the image. This information is exploited in this study to convert the unit of crack thickness from pixels to mm.

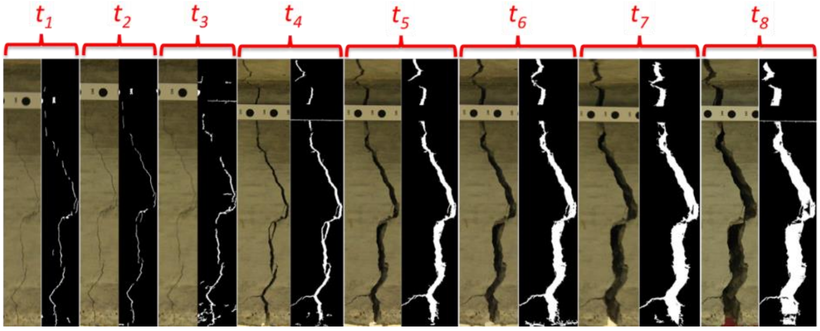


Fig. 10 Illustrative examples of original image and generated segmentation mask for a crack at different points in time

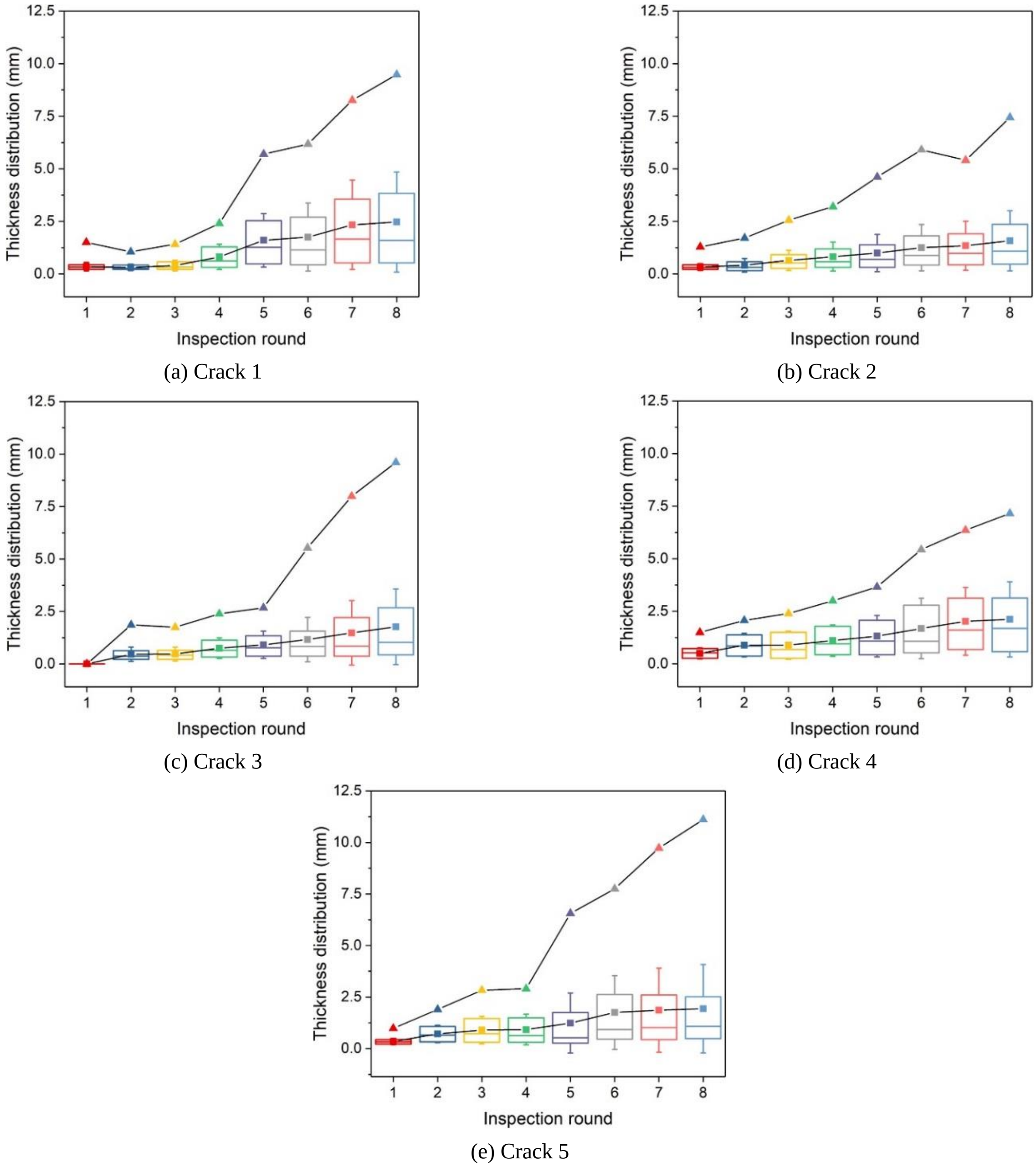


Fig. 11 Time evolution of crack thickness distribution for five different cracks. The rectangular boxes denote the range between the first and the third quartiles. The horizontal lines inside the boxes represent the second quartile, also known as the median. The small solid squares inside the boxes symbolize the mean values whereas the whiskers protruding out from the boxes signify one standard deviation on either side of the mean value. The small triangles outside the rectangular boxes represent the maximum values

3. Results and discussions

The damage evolution dynamics for five different cracks scattered along the span of the tested beam was studied and is shown in Fig. 11. The distribution of crack thickness evaluated at several points along the centerline of the cracks is plotted against time which is characterized by inspection round. The rectangular boxes denote the range between the first and the third quartiles. The horizontal lines inside the boxes represent the second quartile, also known as the median. The small solid squares inside the boxes symbolize the mean values whereas the whiskers protruding out from the boxes signify one standard deviation on either side of the mean value. All the parameters discussed above are indicators of damage severity and its evolution with time. However, the one parameter which is of highest interest to the inspectors is the maximum thickness. It is represented by small triangles, which are connected by straight lines for better depiction of its evolution with time. It was observed that the maximum as well as the mean crack thickness increases almost monotonically with increase in load. The segmentation algorithm used in this study presumes that the cracks are darker compared to the background. However, this hypothesis is violated at times when light penetrates inside thick cracks making a portion of the crack interior appear bright. This leads to inaccurate segmentation and therefore underestimation of crack thickness as indicated by abrupt dip in the maximum thickness value (Fig. 11(b)). However, similar dip observed at lower stages of loading (Fig. 11(a)) can be attributed to the debilitating effect of image noise on segmentation of very thin cracks. Increase in load also resulted in higher dispersal in the thickness values due to increase in crack thickness as well as appearance of new branches. Besides, increase in loading intensity increased the difference between the maximum and mean thicknesses.

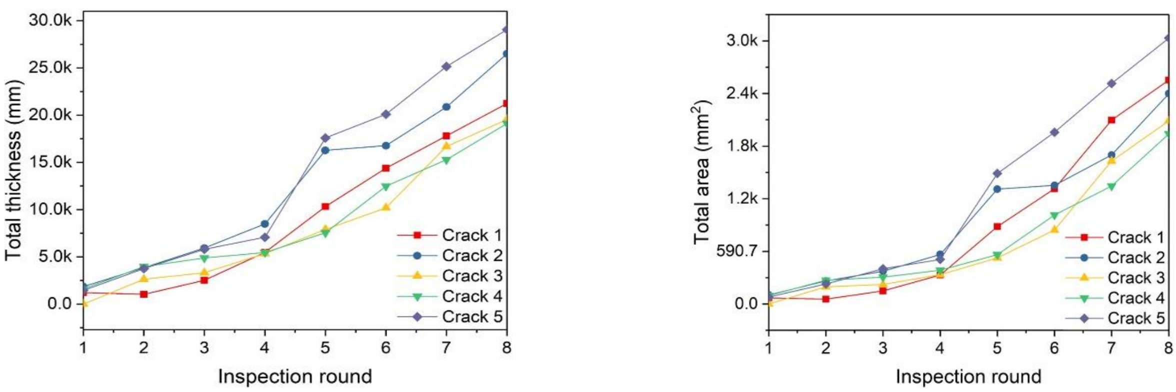
Many a time, evolution of old crack is accompanied by appearance of new branches, which are not accounted for by the maximum thickness. Therefore, total crack thickness together with total area of the cracks, which take into account the main crack as well as its branches, are plotted against time (characterized by inspection round) in Fig. 12. Total crack thickness is estimated as the summation of crack thicknesses at different locations. On the other hand, the area of a crack is measured by evaluating the number of pixels in a region enveloped by an 8-connected component in the binary crack map. This figure (Fig. 12) presents an overall estimate of how fast the crack is growing as a whole.

The deterioration rate (rate of change in crack thickness) is plotted against time (inspection round) in Fig. 13. It is nothing but the first derivative of crack thickness with respect to time. It is evident from the figure that rate of change in thickness is not monotonic, in contrast with thickness itself. This indicates that the growth rate is not proportional to the applied load. An illustration of this sort will make it possible to single out the two inspection rounds in between which a crack has grown at the fastest pace. For instance, it can be inferred from Fig. 13 that the crack-5 suffered the worst degradation in between the fourth and the fifth rounds of inspection. Similar conclusions can likewise be drawn for other cracks as well. Such information may prove to be crucial for chronologically connecting the extent of degradation with extreme events from the past such as seismic vibration, fire, mechanical overload, etc. This will facilitate zeroing in on the most probable reason for damage among several possibilities which are otherwise equally likely.

There are occasions when the inspectors are privy to the data recorded by accelerometers or displacement sensors installed in different floors of a building or at different places along the span of a bridge, in addition to images captured by visual sensors. This provides a scope for correlating component level damage severity with peak acceleration or displacement experienced by the structure. Fig. 14, which shows the variation of crack thickness with the displacement induced by the actuators at loading points, illustrates this concept. The abscissa in this figure should be suitably chosen so as to serve the specific need of the problem at hand. Peak seismic ground motion, mid-span deflection of a bridge or top story deflection of a building are some of the possible alternatives, to name a few. An analysis as such will enable the structural engineers to anticipate the possible damage in a structure that may be induced by a future earthquake of any given intensity.

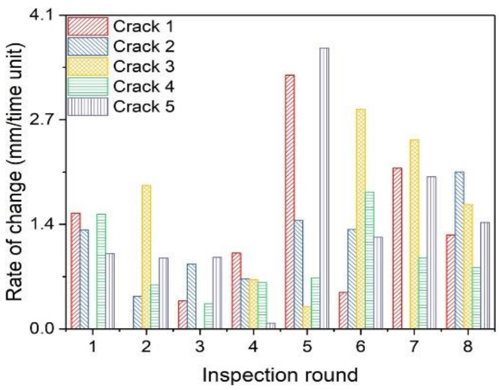
All the figures presented in this section provide a clear picture of how fast the crack is growing and thereby facilitate an informed decision making with regard to any immediate follow-up action where necessary. The state-of-the-art approaches for autonomous condition assessment of civil infrastructures are deprived of this crucial time dimension which prohibits any rationale prognostication about an imminent structural failure. However, inclusion of precious chronological intelligence, as suggested in this study, into the condition assessment and monitoring framework, will significantly narrow down this limitation, making it possible to estimate the residual life of a structural component and take preemptive measures as needed. It will also be instrumental in recommending any requisite adjustment in the frequency of future routine inspections.

Success of the proposed approach depends greatly on the quality of inspection data. Therefore, certain standards should be followed during data collection, omission of which may lead to inaccurate scene reconstruction. First of all, each point on the structure under investigation should be visible at least in two images. Successive images should have more than 60% overlap to secure best possible results. This can be achieved by taking pictures at a regular time interval in case of a UAV flying at a constant speed (Choi *et al*. 2018). The flight path of the UAV should be designed so as to ensure complete coverage of the area under inspection.

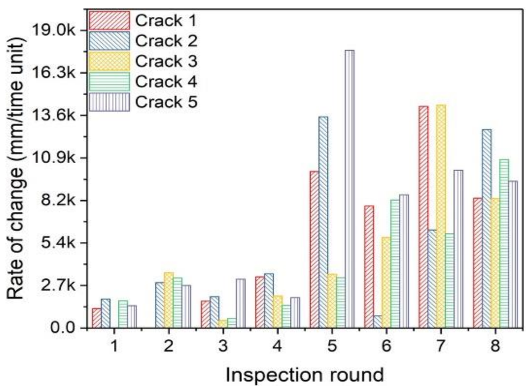


(a) Total thickness (b) Total area

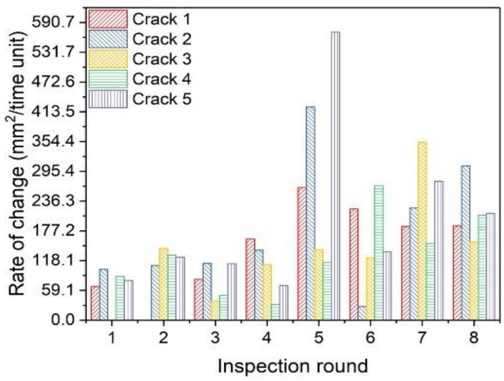
Fig. 12 Time evolution of total crack thickness and area



(a) Maximum crack thickness

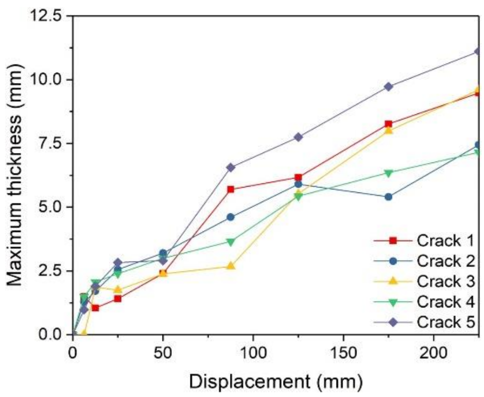


(b) Total crack thickness

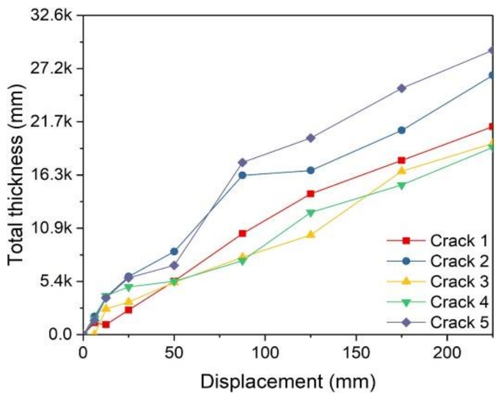


(c) Total crack area

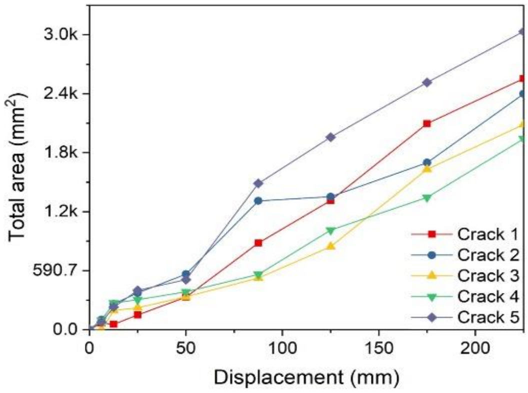
Fig. 13 Deterioration rate during the intervening time between successive inspections



(a) Maximum crack thickness



(b) Total crack thickness



(c) Total crack area

Fig. 14 Variation of crack thickness and area with respect to induced displacement

However, capturing multiple images from the same position should be avoided as it creates panoramic effects giving rise to complication in image alignment. High resolution images contain more visual details. Therefore, pictures should be taken with the highest resolution possible, and saved as raw images without any compression. Good lighting condition is paramount for efficient scene reconstruction, particularly in indoor environment. Static light produces better results as light movement creates a false impression of moving features. Still photos are preferred to videos as slow shutter may introduce blurring effect in sequential frames of videos. If rolling shutter cameras are used, care should be taken that the camera is not shaky while taking pictures. This is particularly challenging in UAV-based data collection where vibration dampers are normally used to alleviate the effect of platform instability. Depth of field of the camera should be suitably chosen as high depth of field causes diffraction resulting in sharpness reduction. Lens with low distortion is better than fish-eye lens having considerable barrel distortions. Exposure of the camera should be carefully set, as under or over exposure causes loss of useful details, and on the other hand, inconsistent exposure produces bright and dark patches.

Civil infrastructures undergo visual changes with time due to accumulation of dirt, rust, stains, etc. Although SURF algorithm is used for feature detection which is invariant to illumination changes, the performance of the correspondence identification process can be affected by contamination of visual features leading to reduction in the number of matched points. In extreme cases, this may result in the failure of feature matching and image alignment exercises if there are not sufficient features to estimate the homography matrix accurately. However, in presence of adequate interest points, as in the case of present study, the proposed algorithm will perform reasonably well without any appreciable loss of accuracy. The detection algorithm can be made robust against such surface irregularities by diversifying the training data with regard to all possible noise intrusion and illumination conditions. A lot of noises will be disposed of at this stage by rejection of nonessential background. Furthermore, proper pre- and post-processing techniques can reduce the image noise significantly.

The beam specimens considered in this study were subjected to flexural failure. Therefore, most of the cracks that appeared on the surface were predominantly vertical. However, inspectors often run into situations where structural elements fail in shear giving rise to cracks that are primarily diagonal. In such cases, the Faster RCNN algorithm will predict a larger bounding box enhancing the scope for noise ingression. However, an appropriate post-processing in the segmentation stage will ensure that the noises are duly identified and eliminated. Moreover, linear structuring element with different orientations (Eq. (1)) renders the segmentation technique invariant to crack orientation. In the same way, the quantification approach used in this study is efficient for vertical and diagonal cracks alike. It is not uncommon to encounter situations where two initially unconnected cracks intersect and become inseparable with increase in load. In such situations, it is recommended that the pair of cracks should be treated as a single entity and evaluated jointly. Cracking in concrete is used as a case study to validate the efficacy of the proposed approach. However, the same techniques can be extended to any other defect category or to multiple defect categories with appropriate modifications. The data sets used for training and validation of the detection algorithm should be suitably updated to include instances from all the defects being investigated. Defect-specific segmentation and quantification algorithms should be invoked to put in place a comprehensive condition assessment pipeline.

4. Conclusions

This study was motivated by the observation that most of the published works in the area of vision-based autonomous structural inspection and health monitoring are agnostic to the time dimension. Ignoring vital historical information, which can otherwise be a key to time-based analysis of damage growth, makes it impossible for inspectors to act preemptively to avert any imminent structural failure and consequent human and financial losses. This study aimed at filling this research gap by proposing a novel computer vision-based approach to leverage from the crucial chronological intelligence embedded in archival images captured by mobile inspection robots or UAVs. Strategies are proposed for autonomous exploration into the erstwhile inspection data looking for correspondences, view synthesis from multiple correspondences and alignment to the current scene under consideration, localizing damage in the reconstructed scenes from the past, segmenting damage, and finally quantifying the damage to extract necessary information and derive meaningful conclusions, after a damage is detected in the current data set. Time history of damage is graphically presented facilitating easier interpretation in addition to predictive and quantitative evaluations. Cracks on concrete surface are used as a case study to demonstrate the feasibility of this approach, which can be potentially extended to any type of structural defects, namely, spalling and corrosion. However, effective implementation of the proposed algorithm makes it necessary to have complete coverage of the damaged areas and adequate overlap between successive images at each batch of inspection data. Besides, this algorithm will fail if there are insufficient visual features or if the structure under investigation is not planer. Incorporation of IMU and GPS information may lead to more robust scene reconstruction in such situations, which is a scope for further study. Future research should also investigate the feasibility of the proposed approach through field experiments with varieties of inspection robots (Jahanshahi *et al*. 2017c) endowed with online data processing capability (Wu *et al*. 2019).

5. Future work

Estimating the service life of a structure is important for the sake of scheduling future maintenance. Mechanics-based models are widely used and usually most reliable in this matter (Cruse and Besuner 1975, Dhawan *et al*. 2019, Besuner 1976). However, in absence of proper analytical model, statistical data driven approaches are adopted which rely on observed data from the past (Wang *et al*. 2014, Agrawal *et al*. 2010, Mohanty *et al*. 2010). Probabilistic evaluation of historical damage evolution data helps predict the expected timeline for a specified serviceability limit state. This calls for establishing the chronology of a damage by exploring an archive of visual inspection data, which was not thoroughly studied by researchers in the past. The present study will potentially fill that knowledge gap and will make it possible to anticipate the remaining life of a civil infrastructure system by exploiting a statistics-based prognostic model, the detailed investigation of which is beyond the scope of the present study.

Besides, estimation of loss due to possible seismic events is an important interest area for planners, government organizations and insurance agencies. It helps them in disaster planning, formulating risk reduction policies, decision making on retrofit and mitigation strategies, and in calculating insurance rating. Evaluating the probability of reaching or exceeding a damage state given a specific value of intensity measure is a prerequisite for seismic loss estimation and risk assessment of infrastructure systems. This probability represented graphically is known as the fragility curve. Professional judgment provided by a panel of experts is one of the commonly used approaches for generating fragility curves, even though it lacks credibility on account of being subjective and dependent on expertise of individual experts. The damage prognostication approach alluded in this paper can open up a new avenue of research in the direction of image-based fragility curve generation exploiting the chronological information embedded in archival data.

Acknowledgments

This study was supported in part by a support from Bentley Systems, Inc. The authors would like to thank Dr. Santiago Pujol, Dr. Aishwarya Puranam and Mr. Rih-Teng Wu for their help to conduct the experiment and data collection.

References

Adhikari, R.S., Bagchi, A. and Moselhi, O. (2014), "Automated condition assessment of concrete bridges with digital imaging", *Smart Struct. Syst., Int. J.*, **13**(6), 901-925. https://doi.org/10.12989/sss.2014.13.6.901

Agrawal, A.K., Kawaguchi, A. and Chen, Z. (2010), "Deterioration rates of typical bridge elements in New York", *J. Bridge Eng.*, **15**(4), 419-429. https://doi.org/10.1061/(ASCE)BE.1943-5592.0000123

Aliakbar, M., Qidwai, U., Jahanshahi, M.R., Masri, S. and Shen, W.M. (2016), "Progressive image stitching algorithm for vision based automated inspection", *Proceedings of 2016 International Conference on Machine Learning and Cybernetics (ICMLC)*, **1**, 337-343.

Aliakbar, M., Qidwai, U. and Jahanshahi, M.R. (2019), "An evaluation of image-based structural health monitoring using integrated unmanned aerial vehicle platform", *Struct. Control Health Monitor.*, **26**(1), e2276. https://doi.org/10.1002/stc.2276

Bay, H., Tuytelaars, T. and Van Gool, L. (2006), "Surf: Speeded Up Robust Features", *Proceedings of the European Conference on Computer Vision*, Springer, Berlin, Heidelberg.

Besuner, P.M. (1976), "Residual life estimates for structures with partial thickness cracks", *Proceedings of Mechanics of Crack Growth*, ASTM International.

Boller, C., Starke, P., Dobmann, G., Kuo, C.M. and Kuo, C.H. (2015), "Approaching the assessment of ageing bridge infrastructure", *Smart Struct. Syst., Int. J.*, **15**(3), 593-608. https://doi.org/10.12989/sss.2015.15.3.593

Bottou, L. (2010), "Large-scale machine learning with stochastic gradient descent", *Proceedings of COMPSTAT'2010*, Physica-Verlag HD.

Cha, Y.J., Choi, W. and Büyüköztürk, O. (2017), "Deep learning-based crack damage detection using convolutional neural networks", *Comput.-Aided Civil Infrastruct. Eng.*, **32**(5), 361-378. https://doi.org/10.1111/mice.12263

Cha, Y.J., Choi, W., Suh, G., Mahmoudkhani, S. and Büyüköztürk, O. (2018), "Autonomous structural visual inspection using region-based deep learning for detecting multiple damage types", *Comput.-Aided Civil Infrastruct. Eng.*, **33**(9), 731-747. https://doi.org/10.1111/mice.12334

Chen, F.C. and Jahanshahi, M.R. (2018), "NB-CNN: deep learning-based crack detection using convolutional neural network and Naïve Bayes data fusion", *IEEE Transact. Indust. Electron.*, **65**(5), 4392-4400. https://doi.org/10.1109/TIE.2017.2764844

Choi, J., Yeum, C.M., Dyke, S.J. and Jahanshahi, M.R (2018), "Computer-aided approach for rapid post-event visual evaluation of a building façade", *Multidisciplinary Digital Publishing Institute*, **18**(9), 3017. https://doi.org/10.3390/s18093017

Cruse, T.A. and Besuner, P.M. (1975), "Residual life prediction for surface cracks in complex structural details", *J. Aircraft*, **12**(4), 369-375. https://doi.org/10.2514/3.44458

Dhawan, S.K., Bindal, A., Bhalla, S. and Bhattacharjee, B. (2019), "Expected residual service life of reinforced concrete structures from current strength considerations", *Adv. Struct. Eng.*, **22**(7), 1631-1643. https://doi.org/10.1177/1369433218818001

Duan, Y., Chen, Q., Zhang, H., Yun, C.B., Wu, S. and Zhu, Q. (2019), "CNN-based damage identification method of tied-arch bridge using spatial-spectral information", *Smart Struct. Syst., Int. J.*, **23**(5), 507-520. https://doi.org/10.12989/sss.2019.23.5.507

Esser, B. and Huston, D.R. (2005), "Versatile robotic platform for structural health monitoring and surveillance", *Smart Struct. Syst., Int. J.*, **1**(4), 325-338. https://doi.org/10.12989/sss.2005.1.4.325

Fischler, M.A. and Bolles, R.C. (1981), "Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography", *Communicat. ACM*, **24**(6), 381-395. https://doi.org/10.1145/358669.358692

Ghorbani, R., Matta, F. and Sutton, M.A. (2015), "Full-field deformation measurement and crack mapping on confined masonry walls using digital image correlation", *Experim. Mech.*, **55**(1), 227-243. https://doi.org/10.1007/s11340-014-9906-y

Girshick, R. (2015), "Fast R-CNN", *Proceedings of the IEEE International Conference on Computer Vision*.

Girshick, R., Donahue, J., Darrell, T. and Malik, J. (2014), "Rich feature hierarchies for accurate object detection and semantic segmentation", *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*.

Glorot, X. and Bengio, Y. (2010), "Understanding the difficulty of training deep feedforward neural networks", *Proceedings of the*

13th International Conference on Artificial Intelligence and Statistics.

He, K., Zhang, X., Ren, S. and Sun, J. (2001), "Deep residual learning for image recognition", *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*.

Hosang, J.H., Benenson, R. and Schiele, B. (2017), "Learning non-maximum suppression", *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*.

Jahanshahi, M.R. and Masri, S.F. (2013), "A new methodology for non-contact accurate crack width measurement through photogrammetry for automated structural safety evaluation", *Smart Mater. Struct.*, **22**(3), 035019. https://doi.org/10.1088/0964-1726/22/3/035019

Jahanshahi, M.R. and Masri, S.F. (2014), "Image-based crack detection", Google Patents, **22**(3), 8,873,837.

Jahanshahi, M.R., Kelly, J.S., Masri, S.F. and Sukhatme, G.S. (2009), "A survey and evaluation of promising approaches for automatic image-based defect detection of bridge structures", *Struct. Infrastruct. Eng.*, **5**(6), 455-486. https://doi.org/10.1080/15732470801945930

Jahanshahi, M.R., Masri, S.F. and Sukhatme, G.S. (2011), "Multi-image stitching and scene reconstruction for evaluating defect evolution in structures", *Struct. Health Monitor.*, **10**(6), 643-657. https://doi.org/10.1177/1475921710395809

Jahanshahi, M.R., Masri, S.F., Padgett, C.W. and Sukhatme, G.S. (2013), "An innovative methodology for detection and quantification of cracks through incorporation of depth perception", *Mach. Vision Applicat.*, **24**(2), 227-241. https://doi.org/10.1007/s00138-011-0394-0

Jahanshahi, M.R., Chen, F.C., Joffe, C. and Masri, S.F. (2017a), "Vision-based quantitative assessment of microcracks on reactor internal components of nuclear power plants", *Struct. Infrastruct. Eng.*, **13**(8), 1013-1026. https://doi.org/10.1080/15732479.2016.1231207

Jahanshahi, M.R., Chen, F.C., Ansar, A., Padgett, C.W., Clouse, D. and Bayard, D.S. (2017b), "Accurate and robust scene reconstruction in the presence of misassociated features for aerial sensing", *J. Comput. Civil Eng.*, **31**(6), 04017056. https://doi.org/10.1061/(ASCE)CP.1943-5487.0000702

Jahanshahi, M.R., Shen, W.M., Mondal, T.G., Abdelbarr, M., Masri, S.F. and Qidwai, U.A. (2017c), "Reconfigurable swarm robots for structural health monitoring: a brief review", *Int. J. Intel. Robot. Applicat.*, **1**(3), 287-305. https://doi.org/10.1007/s41315-017-0024-8

Kim, H., Ahn, E., Shin, M. and Sim, S.H. (2019), "Crack and noncrack classification from concrete surface images using machine learning", *Struct. Health Monitor.*, **18**(3), 725-738. https://doi.org/10.1177/1475921718768747

Kong, X. and Li, J. (2018), "Vision-based fatigue crack detection of steel structures using video feature tracking", *Comput.-Aided Civil Infrastruct. Eng.*, **33**(9), 783-799. https://doi.org/10.1111/mice.12353

Kong, X. and Li, J. (2019), "Non-contact fatigue crack detection in civil infrastructure through image overlapping and crack breathing sensing", *Automat. Constr.*, **99**, 125-139. https://doi.org/10.1016/j.autcon.2018.12.011

Lee, B.Y., Kim, Y.Y., Yi, S.T. and Kim, J.K. (2013), "Automated image processing technique for detecting and analysing concrete surface cracks", *Struct. Infrastruct. Eng.*, **9**(6), 567-577. https://doi.org/10.1080/15732479.2011.593891

Levenberg, K. (1944), "A method for the solution of certain non-linear problems in least squares", *Quarterly Appl. Mathe.*, **2**(2), 164-168.

Lin, T.Y., Maire, M., Belongie, S., Hays, J., Perona, P., Ramanan, D., Dollár, P. and Zitnick, C.L. (2001), "Microsoft COCO: Common Objects in Context", *Proceedings of the European Conference on Computer Vision*, Springer, Cham.

Lowe, D.G. (2004), "Distinctive image features from scale-invariant keypoints", *Int. J. Comput. Vision*, **60**(2), 91-110. https://doi.org/10.1023/B:VISI.0000029664.99615.94

Marquardt, D.W. (1963), "An algorithm for least-squares estimation of nonlinear parameters", *J. Soc. Indust. Appl. Mathe.*, **11**(2), 431-441. https://doi.org/10.1137/0111030

Matheron, G. (1975), "Random Sets and Integral Geometry", **91**(4), 1301-1328.

Minkowski, H. (1989), "Volumen und oberfläche", *Proceedings of Ausgewählte Arbeiten zur Zahlentheorie und zur Geometrie*, Springer, Vienna, Austria.

Mohanty, S., Chattopadhyay. A. and Peralta, P. (2010), "Adaptive residual useful life estimation of a structural hotspot", *J. Intel. Mater. Syst. Struct.*, **21**(3), 321-335. https://doi.org/10.1177/1045389X09357972

Myung, H., Wang, Y., Kang, S.C. and Chen, X. (2014), "Survey on robotics and automation technologies for civil infrastructure", University of Canterbury, Mechanical Engineering.

Otsu, N. (1979), "A threshold selection method from gray-level histograms", *IEEE Transactions on Systems, Man, and Cybernetics*, **9**(1), 62-66.

Ren, S., He, K., Girshick, R. and Sun, J. (2015), "Faster R-CNN: Towards real-time object detection with region proposal networks", *Proceedings of Advances in Neural Information Processing Systems*.

Salembier, P. (1990), "Comparison of some morphological segmentation algorithms based on contrast enhancement", *Proceedings of 5. European Signal Processing Conference*.

Spencer Jr., B.F., Hoskere, V. and Narazaki, Y. (2019), "Advances in computer vision-based civil infrastructure inspection and monitoring", *Eng.*, **5**(2), 199-222. https://doi.org/10.1016/j.eng.2018.11.030

Shan, B., Zheng, S. and Ou, J. (2016), "A stereovision-based crack width detection approach for concrete surface assessment", *KSCE J. Civil Eng.*, **20**(2), 803-812. https://doi.org/10.1007/s12205-015-0461-6

Szegedy, C., Ioffe, S., Vanhoucke, V. and Alemi, A.A. (2001a), "Inception-v4, Inception-Resnet and the impact of residual connections on learning", *Proceedings of the 31st AAAI Conference on Artificial Intelligence*.

Szegedy, C., Vanhoucke, V., Ioffe, S., Shlens, J. and Wojna, Z. (2001b), "Rethinking the inception architecture for computer vision", *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*.

Tanathong, S. and Lee, I. (2013), "Translation-based KLT tracker under severe camera rotation using GPS/INS data", *IEEE Geosci. Remote Sensing Lett.*, **11**(1), 64-68. https://doi.org/10.1109/LGRS.2013.2245858

Van Uitert, R. and Bitter, I. (2007), "Subvoxel precise skeletons of volumetric data based on fast marching methods", *Medical Physics*, **34**(2), 627-638. https://doi.org/10.1118/1.2409238

Wang, X., Balakrishnan, N. and Guo, B. (2014), "Residual life estimation based on a generalized Wiener degradation process", *Reliabil. Eng. Syst. Safety*, **124**, 13-23. https://doi.org/10.1016/j.ress.2013.11.011

Wu, R.T., Singla, A., Jahanshahi, M.R., Bertino, E., Ko, B.J., Verma, D. (2019), "Pruning deep convolutional neural networks for efficient edge computing in condition assessment of infrastructures", *Comput.-Aided Civil Infrastruct. Eng.*, **34**(9), 774-789. https://doi.org/10.1111/mice.12449

Yamaguchi, T. and Hashimoto, S. (2010), "Fast crack detection method for large-size concrete surface images using percolation-based image processing", *Mach. Vision Applicat.*, **21**(5), 797-809. https://doi.org/10.1007/s00138-009-0189-8

Yeum, C.M., Choi, J. and Dyke, S.J. (2017), "Autonomous image localization for visual inspection of civil infrastructure", *Smart Mater. Struct.*, **26**(3), 035051.

https://doi.org/10.1088/1361-665X/aa510e

Yeum, C.M., Choi, J. and Dyke, S.J. (2019), "Automated region-of-interest localization and classification for vision-based visual assessment of civil infrastructure", *Struct. Health Monitor.*, **18**(3), 675-689. https://doi.org/10.1177/1475921718765419

Yu, S.N., Jang, J.H. and Han, C.S. (2007), "Auto inspection system using a mobile robot for detecting concrete cracks in a tunnel", *Automat. Constr.*, **16**(3), 255-261. https://doi.org/10.1016/j.autcon.2006.05.003

Zhu, Z., German, S. and Brilakis, I. (2011), "Visual retrieval of concrete crack properties for automated post-earthquake structural safety evaluation", *Automat. Constr.*, **20**(7), 874-883. https://doi.org/10.1016/j.autcon.2011.03.004

BS