



# Use of a roving computer vision system to compare anomaly detection techniques for health monitoring of bridges

Darragh Lydon<sup>1</sup> · Rolands Kromanis<sup>2</sup> · Myra Lydon<sup>3</sup> · Juliana Early<sup>1</sup> · Su Taylor<sup>3</sup>

Received: 31 March 2022 / Revised: 21 July 2022 / Accepted: 1 August 2022 / Published online: 18 August 2022  
© The Author(s) 2022

## Abstract

Displacement measurements can provide valuable insights into structural conditions and in-service behaviour of bridges under operational and environmental loadings. Computer vision systems have been validated as a means of displacement estimation; the research developed here is intended to form the basis of a real-time damage detection system. This paper demonstrates a solution for detecting damage to a bridge from displacement measurements using a roving vision sensor-based approach. Displacements are measured using a synchronised multi-camera vision-based measurement system. The performance of the system is evaluated in a series of controlled laboratory tests. For damage detection, five unsupervised anomaly detection techniques: Autoencoder, K-Nearest Neighbours, Kernel Density, Local Outlier Factor and Isolation Forest, are compared. The results obtained for damage detection and localisation are promising, with an f1-Score of 0.96–0.97 obtained across various analysis scenarios. The approaches proposed in this research provide a means of detecting changes to bridges using low-cost technologies requiring minimal sensor installation and reducing sources of error and allowing for rating of bridge structures.

**Keywords** Computer vision · Structural health monitoring · Anomaly detection · Deep learning

## 1 Introduction

The progressive deterioration of civil infrastructure is now of paramount concern to asset owners and users alike. Structural damage results in a change in the geometric or material properties of bridges which manifests through a change in stiffness or stability of the structure. Traditional bridge inspections are subjective to human error and bias, and can often result in over conservative assumptions on reduced load carrying capacity [1]. Structural health monitoring (SHM) systems provide a means of objectively capturing and quantifying this change under operational conditions. The application of such systems has a significant cost saving potential across the lifespan of bridges and can ensure the

safe operation of our road and rail transport networks. With over one million bridges across Europe, the task of assessing each structure often surpasses the available resources. This shortfall dramatically reduces the resilience of transport networks particularly considering 35% of Europe's rail bridges are over 100 years old [2]. The understanding of the true capacity of this ageing infrastructure is now more critical than ever as in recent years the UK witnessed an increasing number of failure events in clusters of bridges such as those witnessed in Northern Ireland in 2017 and Yorkshire in 2019 [3, 4]. In response to this, during the last two decades, a significant amount of research has been dedicated to the development and enhancement of SHM systems for bridge monitoring. The challenge of accurately detecting and quantifying damage in civil infrastructure still exists globally. Only a few SHM systems have been deployed and verified on real bridges [5].

There are two approaches currently in practice throughout SHM applications. The first approach uses detailed models of the structure in what is referred to as physics-based models. The second approach, which shall be used in this study, is data-driven monitoring which is based on information

✉ Darragh Lydon  
D.Lydon@qub.ac.uk

<sup>1</sup> School of Mechanical and Aeronautical Engineering, Queen's University Belfast, Belfast, United Kingdom

<sup>2</sup> Department of Civil Engineering, University of Twente, Drienerlolaan 5, 7522 NB Enschede, The Netherlands

<sup>3</sup> School of Natural and Built Environment, Queen's University Belfast, Belfast, United Kingdom

collected from the structure under varying conditions and analysed through various methods [6, 7].

The most commonly utilised technique in recent decades is modal-based damage detection [8–13]. Modal analysis is a widely accepted method of bridge analysis and has had very successful commercial applications worldwide. Traditionally changes in natural frequencies, mode shapes, mode curvatures or damping ratios have been used as a damage indicator. A comparative study of the modal-based damage detection techniques has found that the method is very sensitive to noise contamination and can only successfully detect severe levels of damage [14]. Environmental effects have been found to have a significant effect on modal-based methods [15]. Such approaches have been found to perform well in controlled laboratory or theoretical conditions but are extremely difficult to implement in the field. In many cases, temperature variations can have a greater effect on the dynamic behaviour of the bridge than the presence of damage [15]. This can result in false damage identification or in some cases obscure the detection of real damage [16, 17]. The effect can be reduced through data normalisation and statistical training methods such as principal component analysis [18].

Data-driven methods which incorporate artificial intelligence (AI) have been found to improve the robustness of this technique, but the drawbacks of modal-based analysis are still prevalent and have led to the investigation of alternative methods of damage detection. AI methods such as adaptive neuro-fuzzy interference systems (ANFIS) have been found to provide a high degree of accuracy for structural response and when coupled with interval modelling can effectively extract damage indicators [19]. ANFIS was found to identify damage within 0.03 s of occurrence in a series of finite element models but, as with many damage detection methods, it has not been applied in the field. A method which was successfully trialled in experimental laboratory conditions was the use of 1D Convolutional Neural Networks (CNNs) for real-time damage detection [20]. Large-scale experiments were carried out on a grandstand simulator at Qatar University, in cases of a single damage location the CNN correctly identified the damage location in all 18 test cases. For field applications, a low-cost monitoring system is made feasible by the simple structure of the CNN and the inexpensive computational demands lends to its suitability for field SHM applications. A downfall of the system is that the data for damaged structure are required to train the CNN which infers it as currently unsuitable for field applications as it is difficult to obtain training data for bridge damage scenarios. Alternative AI methods include the use of auto-encoder-based frameworks in deep neural networks which can provide a solution for damage detection in non-linear cases [21]. The framework was tested in both numerical and experimental conditions utilising pattern recognition of

model information; the output stiffness reduction parameters provided the damage indicator using a regression model. The technique performed well in both numerical and experimental conditions which allowed for environmental variability that is encountered on site. Other vibration-based analyses include the interpolation damage detection method (IDDM), where the damage index is defined in terms of mode shapes estimated from the frequency response functions [22]. The main drawback of this method is the assumptions that are made when the data for an intact structure are unavailable, making the method unsuitable for some field applications.

Recently, a symbiotic data-driven approach has been developed based on clustering analysis which reduces the raw vibration data into representative sets with the capabilities for real-time monitoring [18]. This method is particularly useful for applying SHM to ageing structures as it does not require baseline data. In this study, the dynamic cloud clustering algorithm was used; this assesses large data sets from multisensory systems. A cluster validity index then evaluates the quantitative descriptive measure of cluster compactness, thereby reducing the possibility of a false positive outlier. Cluster analysis can be seen as an alternative approach because it does not require a prior baseline to perform feature discrimination, which is useful for health assessment of aged structures or for post-accident/post-retrofitting situations. Recent studies in damage detection have provided real-time solutions using recursive principal component analysis in conjunction with time varying autoregressive modelling to successfully detect of instantaneous structural damage [23–25]. As a recursive model updating approach, the Kalman filter (EKF) has been used in combination with various regularisation methods to identify structural parameters and their changes in both vibration and vision-based data. Given its favourable performance in arithmetic robustness, identification accuracy and fast convergence EKF have been extensively applied to vibration-based measurement with effective outcomes [26–28]. When applied to vision-based measurements, it was found to perform well in small-scale testing but less suitable for large-scale structures subject to ambient vibrations [29, 30].

Overall, there is no absolute consensus on which vibration-based technique is most suitable for bridge SHM, as documented many of the methods have challenges in implementing in the field. The requirement of dense array of sensors for accurate modal curvatures at higher modes can often be prohibitive in field application due to access issues of power requirements and measurement resolution [31].

Displacement measurements provide a valuable insight into the structural conditions and in-service behaviour of bridges under operational and environmental loadings. Displacement has been used as a metric for bridge condition rating in numerous studies outlined in the following section. Analysis of monitored displacement values over

time can provide an insight in possible excessive loading or changes to structural behaviour, since displacements can be directly linked to structural stiffness and external loading. In a long-term analysis (multiple years), displacements can be used to create a pattern of structural response to temperature or vehicle loading. Extreme variance of measured response from its baseline response suggests that there has been a change to the structural properties of the monitoring subject [32].

Feng et al. [33] developed a vision sensor for multipoint displacement monitoring based on an advanced template matching algorithm. Feng and Feng [34] employed the vision sensor to verify the feasibility of output-only damage detection using vehicle-induced displacements and mode shape curvature index in a laboratory study. A 1.6 m simply supported steel beam was excited with hammer impacts at intact and damage states (20% section stiffness reduction). Damage location was accurately detected. However, the motion range of vertical displacements at the midspan was almost 30 mm (i.e. 15 mm amplitude). 15 mm deflection in 1.6 m is  $L/106$ , where  $L$  is the length of the span. In [35] proposed the deformation area difference (DAD) method using deflections, inclination angles and curvatures for condition assessment of bridges. The method resolves the problem of unknown initial structural conditions using numerical or theoretical models with known initial conditions as a reference system. The method is able to detect local stiffness reductions starting from 23.8% as validated using numerical and laboratory models with vision-based measurement [36]. The application of the DAD method was also demonstrated on a newly constructed bridge, where, of course, no damage was detected [37].

In [38], a displacement curve was used to detect damage to a cantilever beam structure. In [39], Zhang et al. used the displacement caused by a vehicle passing over a bridge structure as a modelling scenario for simulated damage detection.

This type of experiment, where changes to the displacement curvature based on repeated passes by a vehicle established the methodology of the laboratory work described in this paper. A test setup, where a bridge model is fully instrumented to determine displacement from a passing vehicle, is laid out by Catbas et al. [40]. This was developed upon in [41] where a series of Linear Variable Differential Transducers (LVDT) were used to detect damage under a roving sensor approach. Damage was detected under several scenarios in this study, with localised damage detected at multiple instances of location and severity. The research presented in this paper further enhances the development of roving sensor system for damage detection by replacing the cumbersome setup of LVDTs with a system of time-synchronised cameras for displacement measurement that was demonstrated by the authors in [42].

These studies all demonstrate the use of displacement measurements as a powerful means of assessing bridge condition through its performance. This study proposes a data-driven approach, where displacement measurements of a bridge at a healthy/undamaged state are used to derive its baseline conditions. A data-driven approach was selected for this study because they are less susceptible to noise in the data, increasing the likelihood that the approach developed here can be successfully deployed in the field. Additionally, the removal of the requirement for having detailed information about the structure allows for greater potential application on real world structures. The collected measurements are then compared against the baseline conditions for anomaly/damage detection. Displacement measurements, which are estimated using vision sensor, are used for training a selection of the most promising unsupervised learning algorithms found in the literature. These algorithms are (1) autoencoder, (2) local outlier factor, (3) kernel density estimation, (4) K-means nearest neighbours, and (5) Isolation Forest. These methods have also been applied in the literature for a diverse array of applications [43–47].

The selected algorithms provide a variety of approaches to unsupervised novelty/outlier detection, examples are probabilistic novelty detection (Kernel Density), nearest neighbour (K-means and Local Outlier Factor), Recursive partitioning (Isolation Forest) and reconstruction-based (Autoencoder). Further discussion on the types of novelty/outlier detection is presented in [48, 49]. These algorithms are chosen as they provided different examples of approaches for outlier detection, as it could not be estimated ahead of time which approach proves to be most successful. Unsupervised methods were chosen because this allowed greater applicability to real world bridges where it is not possible to be absolutely certain of bridge condition before inspection. The combination of unsupervised data analysis paired with a low-cost and robustly validated roving vision sensor will allow for the development of a practical solution for bridge monitoring that can be applied to a diverse field of bridge types. The performances of the techniques are assessed on a laboratory-scale model of a bridge structure from which displacement information was collected with a roving sensor technique.

## 2 Methodology

### 2.1 Roving sensor technique

Environmental effects, camera resolution, field of view, and image processing algorithms predominantly govern the accuracy of vision-based measurements. A complete derivation of the response of the bridge can rarely be obtained from a single camera even in the case of a short span bridge.

It is highly likely that multiple synchronised cameras are needed, which is often cost prohibitive. Alternatively, the number of cameras can be reduced, whilst a control vehicle is crossing the bridge multiple times, in the following arrangement.

- One camera records displacements of the reference target for each crossing.
- The focus or location of the other monitoring cameras are varied under each vehicle pass event (camera roving).

A structure's response at a target location over a crossing of a unit load (e.g. a vehicle) is referred to as time histories of target displacement or unit influence line [50]. Target displacements can be collected at any location on the bridge using vision-based measurement. Capturing the response of all targets, when, for example, the reference target reaches the highest displacement could give erroneous values for some targets that are located far from the load location. The duration of vehicle crossings may vary each time; therefore, any time-dependent response parameters should be excluded when defining a damage sensitive feature or damage factor. This research proposes to compute the range (i.e. peak to peak) of displacements of each target during vehicle crossings. Figure 1 illustrates the proposed approach on a continuous bridge. Parametric studies were performed by changing the boundary conditions of a scaled bridge model.

## 2.2 Image processing algorithm

The algorithm used to calculate displacement of the test structures is composed of three elements. First, a scale factor calculation is performed using the dimension correspondences method. This method involves selecting a known physical dimension in the view frame of the camera and measuring it in pixels. A simple conversion is then performed to convert any measurements from pixels to engineering units. This method has been used in several studies and has been

applied in laboratory and field trials [51, 52]. The next step is feature extraction, where high contrast key points known as features are detected in the image of the structure. Feature extraction was done using the Speeded Up Robust Features (SURF) [53] algorithm. SURF uses integral images to create a low computation cost-scale space. The scale space is calculated with the formula:

$$L(x, y, \sigma) = G(x, y, \sigma) * (I(x, y)), \quad (1)$$

where  $*$  is the convolution operator,  $I(x, y)$  is the input image,  $\sigma$  is a scale parameter and  $G(x, y, \sigma)$  is the Gaussian blur operator, which can be expressed as:

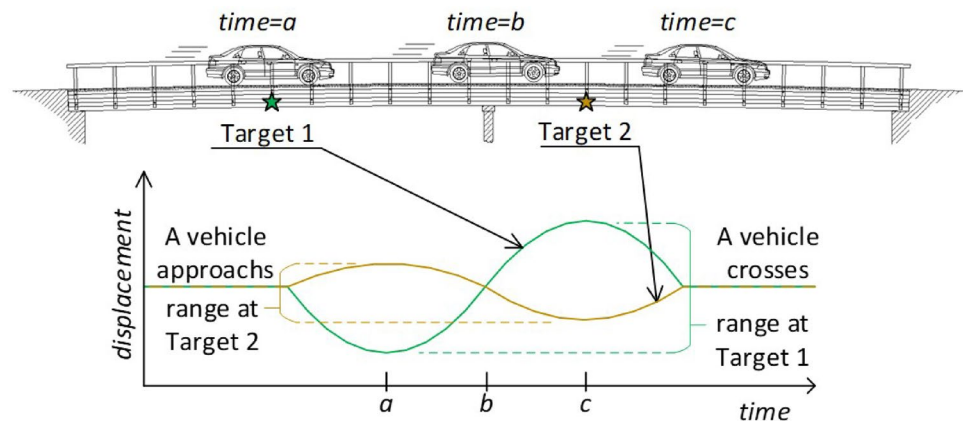
$$G(x, y, \sigma) = \frac{1}{2\pi\sigma^2} e^{-(x^2+y^2)/2\sigma^2}. \quad (2)$$

The original image is then resized to half size and the Laplacian of Gaussians between the scaled and original image is found. SURF applies a box filter to the integral images to approximate this calculation. The determinant of the Hessian matrix  $H$  for each key point is used to detect blob structures which can be then converted into feature descriptors. For a point  $x = (x; y)$  in an image  $Im$ ,  $H(x; \sigma)$  is given by:

$$H(x, \sigma) = \begin{bmatrix} \tau_{xx}(x, \sigma) & \tau_{xy}(x, \sigma) \\ \tau_{yx}(x, \sigma) & \tau_{yy}(x, \sigma) \end{bmatrix}, \quad (3)$$

where  $\tau_{xx}(x, \sigma)$  is the convolution of the Gaussian second-order derivative  $\frac{\partial^2}{\partial x^2} g(\sigma)$  with the image  $Im$  in point  $x$ , this also applies to  $\tau_{xy}(x, \sigma)$  and  $\tau_{yy}(x, \sigma)$ . The box filter of size  $9 \times 9$  approximates a Gaussian with  $\sigma = 1.2$  and represents the lowest level (highest spatial resolution) for blob-response maps. The use of box filters and integral images mean that SURF can apply filters of varying size to the image efficiently. Non-maximum suppression is applied to determine the location and scale of key points, with scale space interpolation of the maxima of the determinant of the Hessian

**Fig. 1** Estimating displacement range at target locations

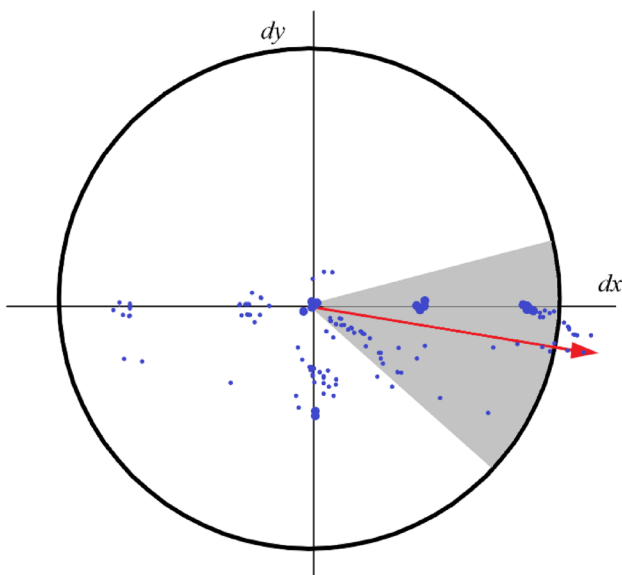


matrix done according to the method proposed by Ref. [54]. The orientation of the key points is assigned by calculating the Haar wavelets in horizontal and vertical directions for a neighbourhood of size 6 around the location of the point, where  $s$  is the scale the feature was detected in. The Haar wavelet of a region around a point is the sum of the pixel intensities around that point paired with the difference between these sums. This is used as a means of speeding up representations of image regions as direct calculations of each pixel in the region would be computationally expensive. This method was first proposed by [55]. These responses are then weighted by an adequate Gaussian ( $\sigma=2s$ ) with the dominant orientation estimated by calculating the sum of responses within a sliding window of size  $60^\circ$ , as shown in Fig. 2.

A region of  $20s$  is then created and split in  $4 \times 4$  square sub-regions. In these regions, SURF computes Haar wavelet responses at  $5 \times 5$  evenly spaced sample points. The responses are weighted with a Gaussian and the wavelet responses in horizontal  $w_{av_x}$  and vertical  $w_{av_y}$  directions for each sub-region are summed to form the first entry of the feature descriptor. The information on polarity of intensity change is given by extracting the sum of the absolute values of the responses. By concatenating the descriptor vector  $v$

$$v = \left( \sum w_{av_x}, \sum w_{av_y}, \sum |w_{av_x}|, \sum |w_{av_y}| \right), \quad (4)$$

from each sub-region, a float-vector descriptor of length 64 is formed. This descriptor can then be used to track key points throughout a series of images. Once the features have been detected, they are matched in subsequent frames



**Fig. 2** Sliding window used in the SURF algorithm to determine key point orientation

of the video by use of the RANdom SAmple Consensus (RANSAC) [56] method of matching features. The basic principle of RANSAC involves drawing a random uniform set of points  $m$  (smallest number of matches required) from the dataset. A least squares line is fitted to this random set and any points outside a distance  $dist$  are judged to be outliers. If there are enough points inside the line (inliers), then there is a good fit. This process is repeated  $N$  times and the best fit from these tests is used with the fitting error applied as the criteria for decision.  $N$  is computed by solving the following equation:

$$N = \frac{\log(1 - P)}{\log(1 - \epsilon)^\rho}, \quad (5)$$

where  $\epsilon$  is the probability that a point is an outlier,  $P$  is the desired probability that we get a good fit and  $\rho$  is the number of points in a sample. A homography between the randomly selected matching features  $H$  is calculated, and the  $H$  with the highest number of inliers is selected as the final  $H$ . The final  $H$  is applied to the dataset of matched points and the outliers are removed before transform estimation is performed. This matching process is repeated throughout the video and displacement for the structure over time is calculated. The accuracy of this algorithm was verified in laboratory and field trials by the authors in [57].

### 2.3 Details of camera hardware

The cameras used in this were modified GoPro Hero 4 [58] action cameras. These cameras were chosen as they are portable, resistant to adverse weather effects, have wireless functionality and offer an inexpensive and high-resolution solution for image capture. A standard GoPro lens would have too short a focal length to be viable for bridge monitoring, a modified GoPro [59] allows attachment of C or F-mount lenses to the GoPro cameras. A Computer [60]  $\frac{1}{2}$ " 25–135 mm F1.8C-Mount lens was attached to the GoPros for the laboratory trial detailed in this research.

The footage from each GoPro was synchronised through use of a Syncbac [69], an accessory that can be attached to the extension port of the GoPro to allow for embedding of timecode metadata into each frame. Analysis of this metadata allows for synchronisation of recordings obtained by the system using a solution developed by the authors in C++ in Microsoft Visual Studio. The Syncbac sends live timecode data via Radio Frequency (RF), with a range of 30–60 m. The videos for the laboratory trial were captured in 1080p.

### 3 Experimental setup

#### 3.1 Laboratory model setup

A laboratory model of a two-span bridge was developed in the Experimental Design and Monitoring (EDM) laboratory of Civil Infrastructure Technologies for Resilience and Safety (CITRS) at University of Central Florida (UCF). The bridge has two 300 cm main continuous spans supported by three steel frame sections as shown in Fig. 2. The bridge deck is 1.2 m wide and 600 cm long of steel plate construction with the thickness of 3.18 mm. The steel deck is supported by two  $25 \times 25 \times 3.2$  mm girders spaced 0.61 m from each other. The girders are denoted as A and B as illustrated in Fig. 3. The connection sets with four M6 bolts and 3.18-mm-thick plates are used to connect the girders and the deck. A small-scale toy truck is employed as the moving load on the bridge in this study.

The supports of the bridge are varied during the experiment to simulate the change of boundary condition, replicating common real-life bridge conditions. The undamaged case, i.e. baseline, in this experiment is the case that all the supports are rollers. Four damage cases are designated by changing the supports of the Girder B. Two lanes were predefined on the deck: one was close to Girder A (lane 1) and the other was close to Girder B (lane 2). The dimensions and layout of the left span are shown in Fig. 6.

Figure 6, these dimensions are mirrored for the right span of the structure. The truck ran on the lane 2 travelling

longitudinally from left to right to simulate the moving load. Four cameras were placed next to the structure as shown in Fig. 4.

To measure the displacements at the strategic locations, measurement nodes are defined on the girders. The measurement nodes are bolts or similar distinctive objects on the underside of the bridge that could be reliably and consistently tracked throughout the course of the video. The nodes were labelled from 1 to 16, as laid out in Fig. 5. The dimensions and layout of the left span are shown in Fig. 6, these dimensions are mirrored for the right span of the structure.

#### 3.2 Measurement collection strategy

Each camera is set up to measure one node in each run, as such the measurement of all the nodes cannot be captured in one run. This technique is known as a roving vision sensor setup to accommodate all the measurement nodes. The process for testing is detailed in the flowchart in Figs. 7 and 8.

As shown in Fig. 5, Node 1 was set as the reference node (Ref. Node), and it is measured in each run. The triangles in Fig. 5 represent the cameras used in each run in addition to the reference camera that remained fixed at Node 1. For Test 1, this indicates Nodes 2–4, Test 2 consists of Nodes 5–7 and so on until all nodes have been monitored. Vertical displacements of the Ref. Node for the first crossing at no damage condition are shown in Fig. 9a.

The response pattern at Node 1 is similar to other nodes located on the left span (Nodes 1–8 in Fig. 3) Vertical displacement histories of the nodes on the right span (Nodes

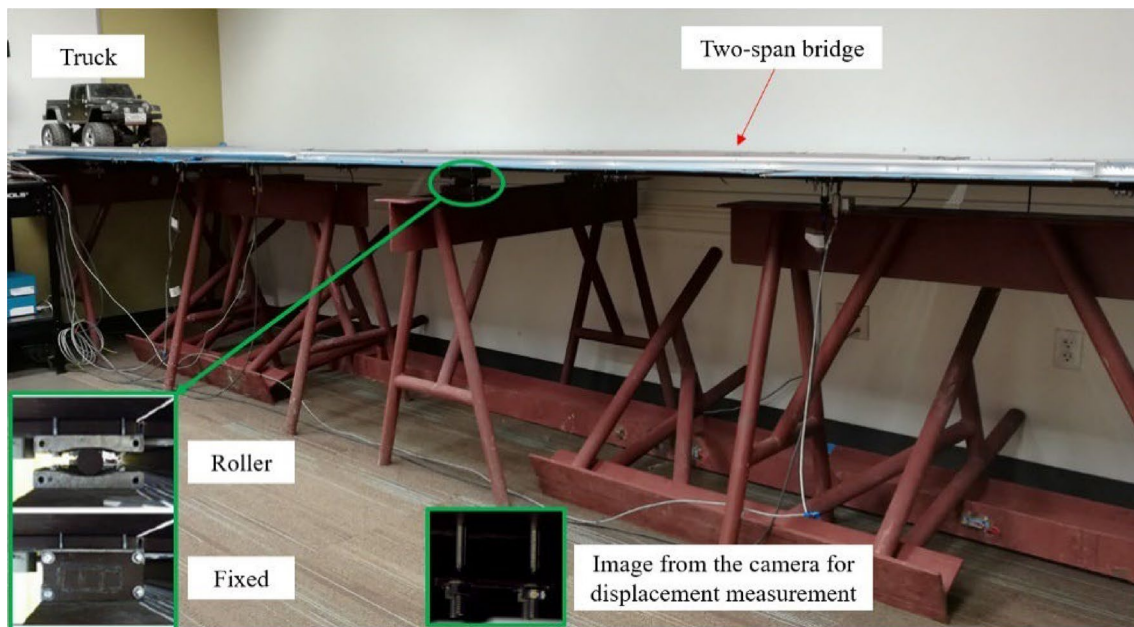
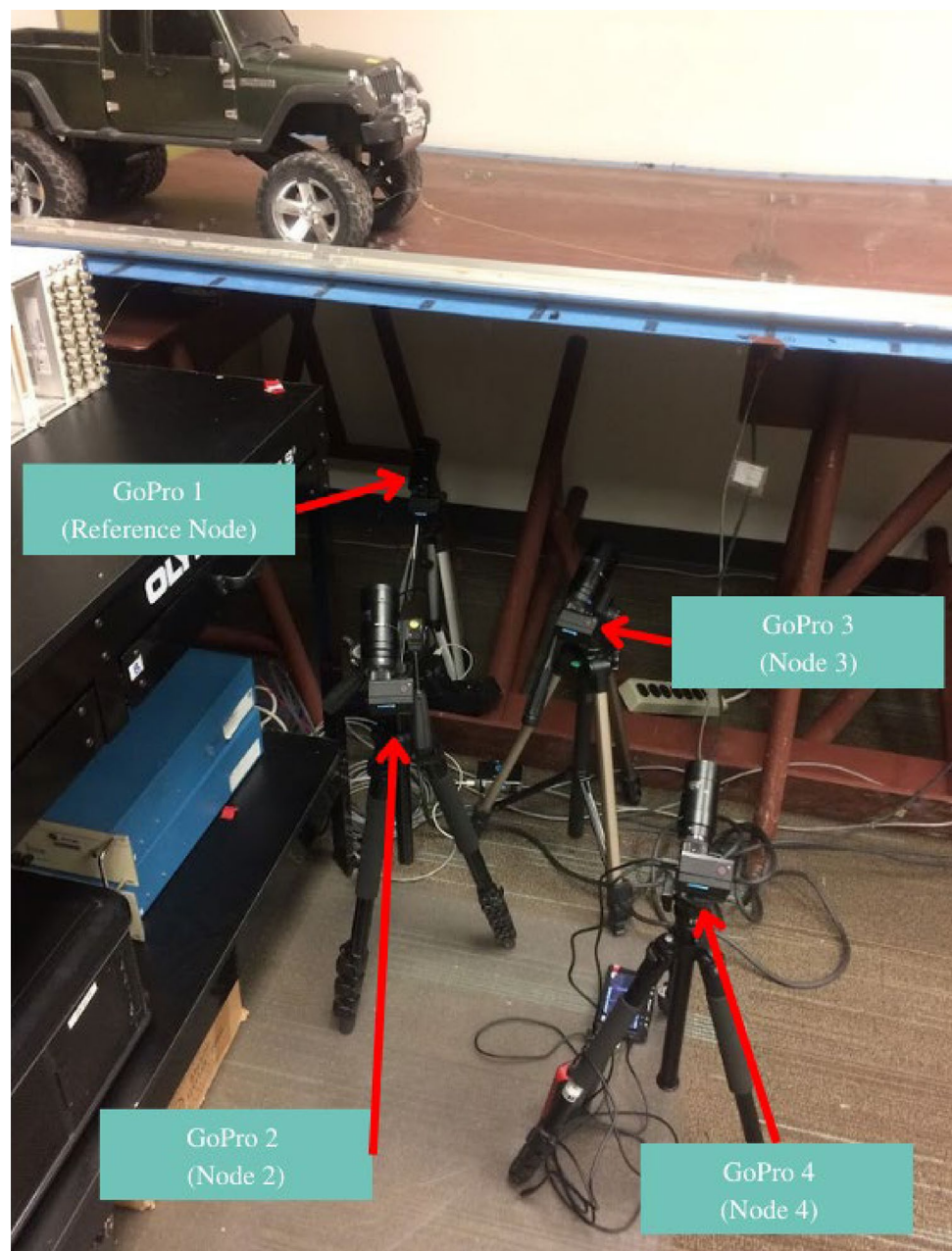


Fig. 3 Two span model bridge

**Fig. 4** Camera setup for monitoring



9–16 in Fig. 2) are opposite to the ones on the left support. When the vehicle is on the left span, the right span lifts up, and vice versa. The raw response measurements contain both static component (bridge deflection) and dynamic component (bridge vibration induced by the vehicle). The static response can be isolated by processing the raw response with an adequate high-pass filter. The conversion of the response measurement from time domain to frequency domain reveals fundamental frequencies. The power spectrum density (PSD) plot is used to set a suitable high-pass filter. The lowest frequency component, which is 0.098 Hz, presents duration of the vehicle crossing. The crossing lasts approximately 10 s. The frequency range of the dynamic

response is above 4 Hz (see Fig. 3b). Thus, the high-pass frequency is set to 1 Hz. The resulting signal is subtracted from the raw measurements leaving only the component of the static response. During the trials, it was not possible to ensure that all runs start and end at the same time. They also vary slightly in their duration. For these reasons, selecting the range of vertical displacements of each node as the damage factor (DF) is the best choice.

### 3.3 Data generation and analysis

The filtered displacement ranges are used as a basis for a training dataset for all damage detection techniques. This

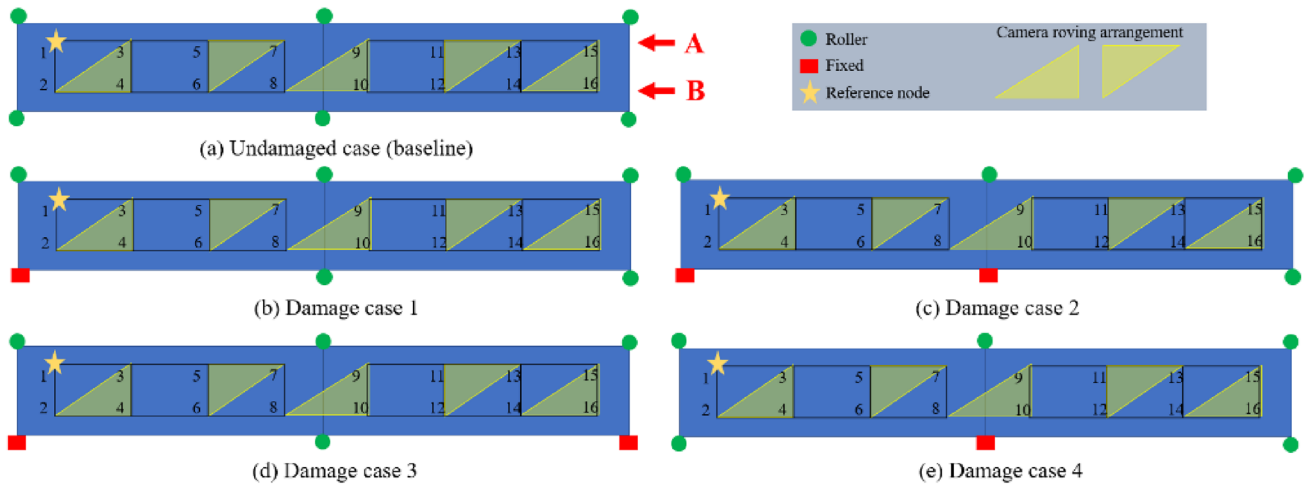
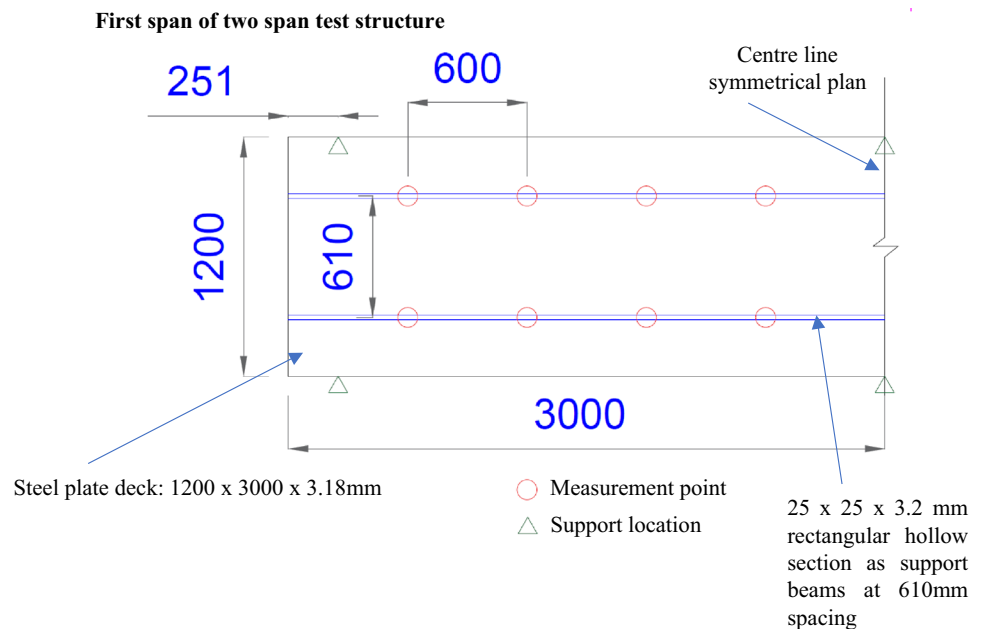


Fig. 5 Experimental cases and configurations (Plan)

Fig. 6 Details of left span of model bridge structure all dimensions show in mm



was done using the variance between displacement ranges between runs 1–5 in the undamaged scenario as boundary conditions to generate randomised displacement scenarios for all nodes on the bridge that would resemble baseline undamaged behaviour. This meant every generated scenario had a displacement for each node that did not below the minimum displacement of the node recorded in the earlier runs, nor did the generated displacement exceed the maximum displacement. This is repeated 50,000 times to make a dataset that encompassed baseline behaviour of the bridge. The dataset is split into 40,000 training examples and 10,000 validation examples. The 4 damage scenarios are also used as candidates for scenario generation in the test dataset, with 8000 baseline scenarios and 2000 damaged scenarios

generated. The nodes are split in four layouts, or node groupings. This involves reducing the number of nodes used for analysis of the structure. The analysis of smaller node groupings is carried out because if acceptable results can be gathered using a lower amount of monitoring locations, it would mean a more streamlined data gathering process. The node groupings were used for the data from the scenarios, these are laid out in Table 1.

### 3.4 Unsupervised algorithms description

This section gives an overview of the five algorithms chosen for analysis of the training dataset in this study. An autoencoder is an unsupervised neural network frequently used for



### BASELINE DATA COLLECTION

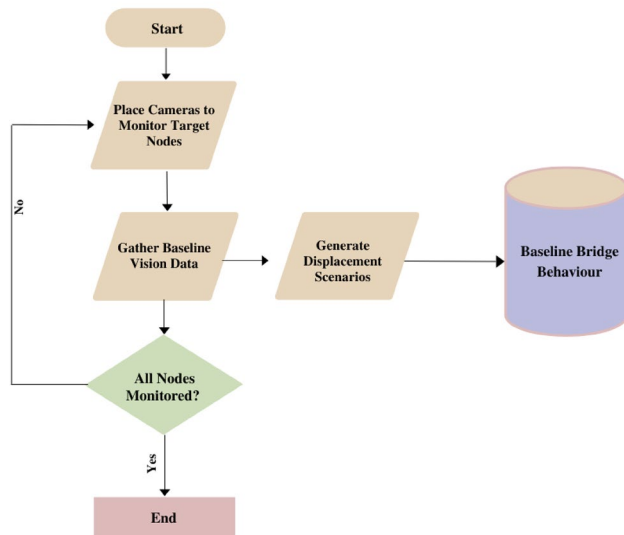


Fig. 7 Flowchart of baseline data collection methodology

### DAMAGE SCENARIO DATA COLLECTION

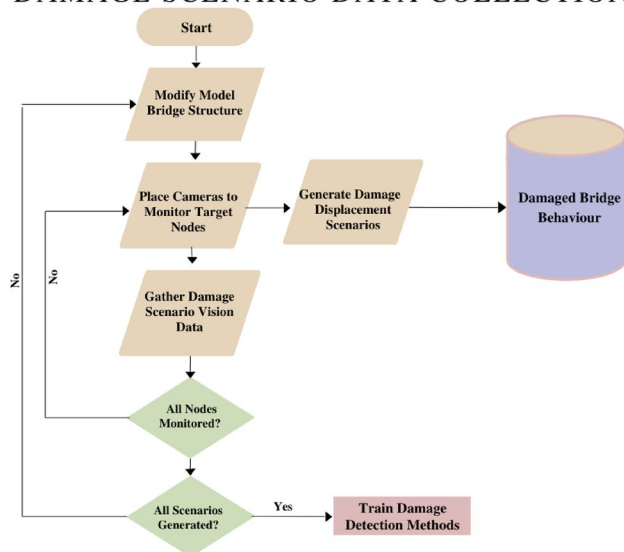


Fig. 8 Flowchart of damage scenarios data collection methodology

anomaly detection [61–63]. It consists of an encoder and a decoder. The encoder maps the input to a lower-dimensional space, and the decoder maps the encoded data back to the input. If an auto-encoder is trained to recognise a certain type of input, such as bridge baseline (normal) conditions, any deviation from this output has a high reconstruction error, thus indicating structural damage. The Local Outlier Factor is an anomaly detection technique, which assign each object a degree of being an outlier [64]. It is local in that the degree depends on how isolated the object is with respect to the surrounding neighbourhood. This technique allows

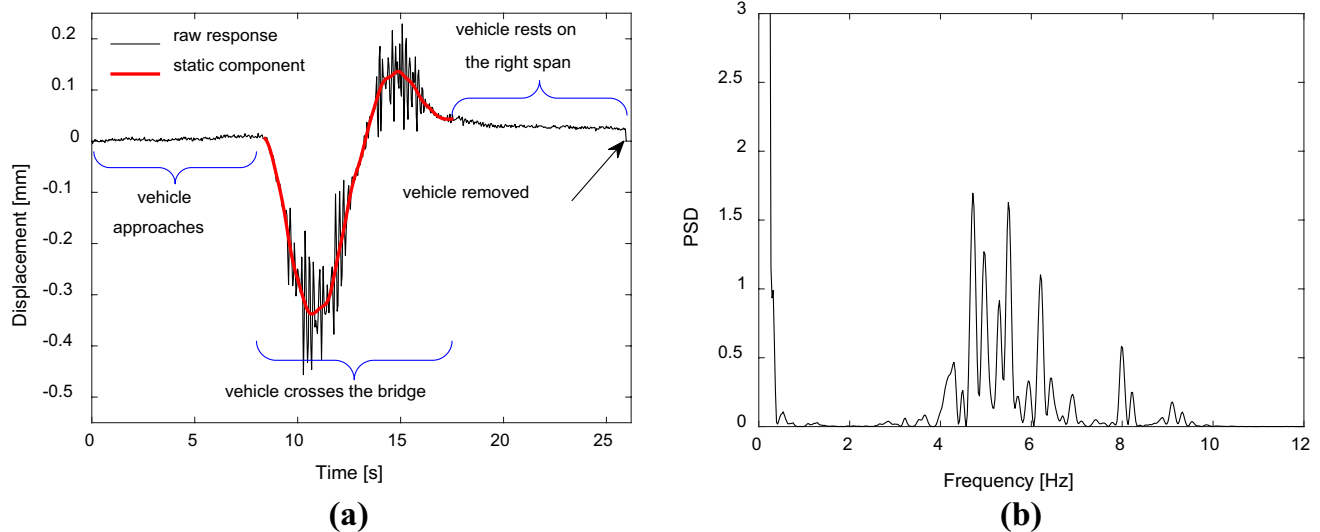
detecting outliers that can otherwise not be detected with existing approaches [43]. The Kernel Density Estimation is the process of estimating an unknown probability density function using a kernel function [65]. Whilst a histogram counts the number of data points in somewhat arbitrary regions, a kernel density estimate is a function defined as the sum of a kernel function on every data point. This estimate is then used to detect outliers. The K-means nearest neighbours technique [66] attempts to group data points into clusters, by first choosing initial random centroids of the number of clusters to be created. After initialisation, K-means consists of looping between the two other steps. The first step assigns each sample to its nearest centroid. The second step creates new centroids by taking the mean value of all the samples assigned to each previous centroid. The differences between the old and the new centroids are computed, and the algorithm repeats these last two steps until this value is less than the defined threshold, which allows undamaged and damaged scenarios to be grouped separately and therefore to be identified at inference time. For this research, the mini batch K-means implementation is used [67]. The Isolation Forest is an ensemble algorithm [68]. Isolation Forest ‘isolates’ observations by randomly selecting a feature and then randomly selecting a split value between the maximum and minimum values of the selected feature. Since recursive partitioning can be represented by a tree structure, the number of splittings required to isolate a sample is equivalent to the path length from the root node to the terminating node. This path length, averaged over a forest of such random trees, is a measure of normality and our decision function. Random partitioning produces noticeably shorter paths for anomalies. Hence, when a forest of random trees collectively produces shorter path lengths for particular samples, they are highly likely to be anomalies (or damage events).

### 3.5 Unsupervised algorithms training details

This section details the parameters, and training details for the five algorithms chosen for analysis of the training dataset in this study.

#### 3.5.1 Kernel density

For kernel density training, a Gaussian kernel with a bandwidth of 1 is used to create the kernels. The kd-tree [69] algorithm with a leaf size of 40 was the tree method employed. The kd-tree algorithm splits a dataset using  $k$  hyperplanes to split the datapoints. The leaf size determines size of the boxes that are generated by these planes. A smaller leaf size would mean that very few points must be checked if they are inside an area, but this results in a deeper tree and longer construction and traversal time. 40 was chosen after a randomised grid search to find the



**Fig. 9** Node 1 displacement time history (a) and its PSD plot (b)

optimal parameter. For the kd-tree, the Minkowski distance between points with a  $p$  value of 2, which means the distance was equivalent to the Euclidean distance, was used as the distance metric to select outliers. Kernel Density will be referred to as KDE for the results discussion.

### 3.5.2 Autoencoder

The autoencoder was trained for 100 epochs with early stopping, the optimal loss for each node grouping varied from 50 to 90 epochs with a small number requiring the full 100 epochs. A threshold for classification of undamaged vs damaged scenarios was then calculated by comparing the root mean square error for the predictions on the test set and determining the optimal decision boundary. If the DF generated by the autoencoder from the supplied displacements was outside the boundary conditions determined in the training phase, the scenario was labelled as damaged. This method was successful at identifying the presence of damage, but it was incapable of localising where damage occurred on the structure. The auto-encoder will be referred to as AE for the results discussion. The layers for AE are shown in Fig. 10.

**Table 1** Node groupings for analysis

Scenario name	Node groupings
All Nodes	(1,2,3,4,5,6,7,8,9,10,11,12,13,14,15,16)
Node Octets	(1,2,3,4,5,6,7,8), (9,10,11,12,13,14,15,16)
Node Quartets	(1,2,3,4), (5,6,7,8), (9,10,11,12), (13,14,15,16)
Pairs	(1,2), (3,4), (5,6), (7,8), (9,10), (11,12), (13,14), (15,16)

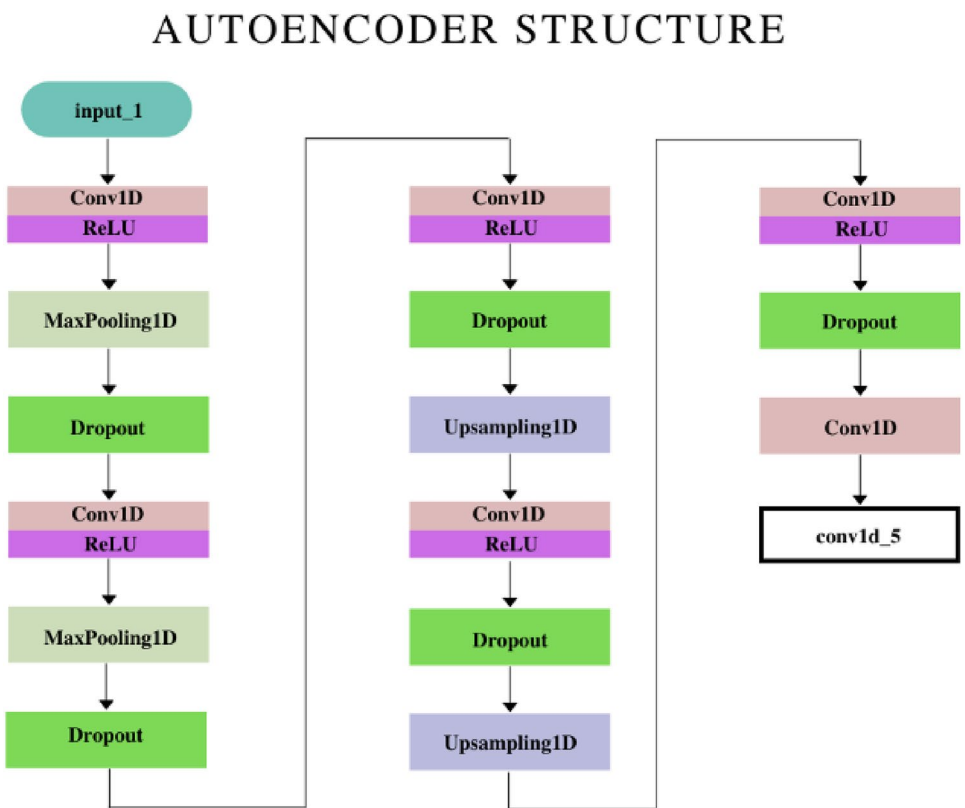
### 3.5.3 Isolation forest

For the training of the isolation forest structure, the number of estimators was set to 100, whilst the contamination ratio was set to 10% of the dataset according to the recommendations in the original paper. The number of estimators for an isolation forest is the number of tree structures used to construct the forest. The contamination ratio controls the threshold for the decision function when a scored data point should be considered an outlier. Isolation forest will be referred to as IF for the results discussion.

### 3.5.4 K-means nearest neighbours

For the training of the K-means nearest neighbours' algorithm, the optimal number of initialisations ( $n_{init}$  was set to be 20 following a randomised grid search [70]). The number of initialisations is used to avoid K-means falling into local minima. In K-means, the initial placement of centroid plays a very important role in its convergence. Sometimes the initial centroids are placed in a such a way that during consecutive iterations of K-means the clusters keep on changing drastically and even before the convergence condition may occur, the maximum number of iterations is reached, resulting in incorrect cluster allocations. Hence, the clusters obtained in such may not be correct.  $N_{init}$  is used to overcome this problem.  $N_{init}$  controls how many times the k-means algorithm is run on the training dataset, as if a non-optimal initial position for the centroids is randomly chosen, there is a chance that training could end before the best solution is found.  $N_{init}$  does not determine the number of clusters used, merely how many times initialisation/training is performed. In

Fig. 10 Auto-encoder structure



the case of this research, the clusters were initialised 20 times and the best performing solution was chosen from the 20 passes over the dataset. For each different set of points, a comparison is made about how much distance did the clusters move, i.e. if the clusters travelled small distances than it is highly likely that we are closest to ground truth/best solution. The points which provide the best performance and their respective run along with all the cluster labels are returned. This is then used to determine which cluster each datapoint belongs to. Because there are 5 different scenarios present in testing, a total of 5 clusters were used. This meant that if the algorithm converged well, damage could be not only identified, but localised on the beam structure. K Means nearest neighbours will be referred to as KM for the results discussion.

3.5.5 Local outlier factor

For the local outlier factor algorithm, five clusters were also chosen in an attempt to identify and localise damage on the beam structure. The Ball-Tree [71] algorithm was used for selecting the clusters with the Minkowski distance with a value for  $p = 2$  between data points and the cluster centroids used to detect outliers in the dataset. The local outlier factor will be referred to as LOF for the results discussion.

4 Results and discussion

4.1 Analysis of all nodes

The results for the analysis of all nodes with a combination of all scenarios are shown in Fig. 11. AE, KM, IF and KDE all performed quite well on the data, detecting damage in the general case for AE, IF and KDE and by type in the case of KM. KM and KDE had perfect scores in all scenarios. There was a 2% false positive rate for IF in all scenarios, and a variance of 0.5% of false negatives and positives for AE in each scenario. The LOF failed completely to detect damage on the beam, returning a F1-score of 0.06 across all scenarios. False positive is when the algorithm predicts damage when none is present, whilst false negatives are when the algorithm predicts no damage when damage is present. Recall is the proportion of damage scenarios in the testing set that were correctly detected by the algorithm ( $Recall = \frac{True\ Positives}{True\ Positives + False\ Negatives}$ ) and precision is the proportion of damage predictions that were actually correct ( $Precision = \frac{True\ Positives}{True\ Positives + False\ Positives}$ ). The results from all nodes analysis are shown in Table 2.

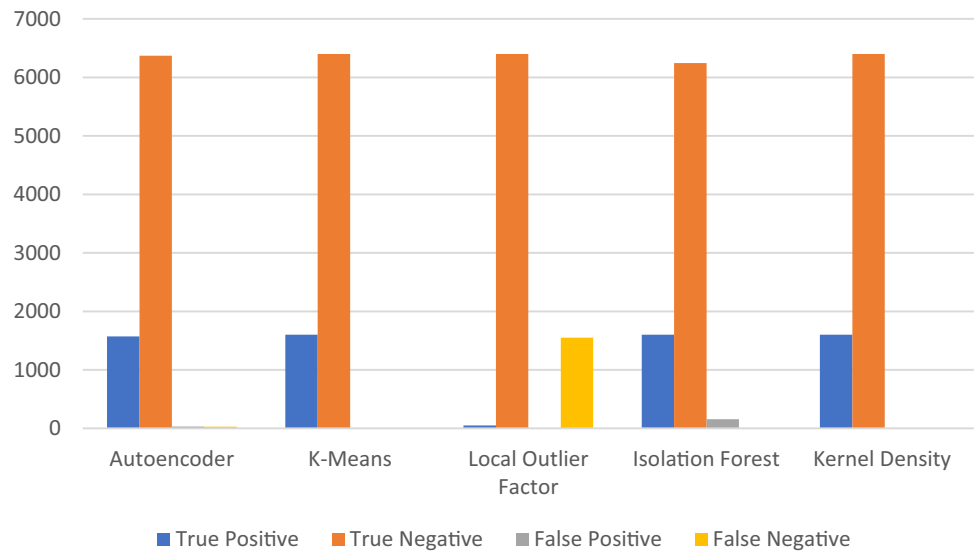
4.2 Node Octets analysis

The results for the node octets with a combination of all scenarios are shown in Fig. 12. Figure 12 demonstrates

that AE, KM and KDE continue to work well across all scenarios with LOF still unable to detect damage at any acceptable level. There is a greater level of false positive

readings from IF analysis in comparison to all nodes analysis. This is more significant for nodes 1–8 compared to Nodes 9–16 as can be seen in Fig. 13. The results from

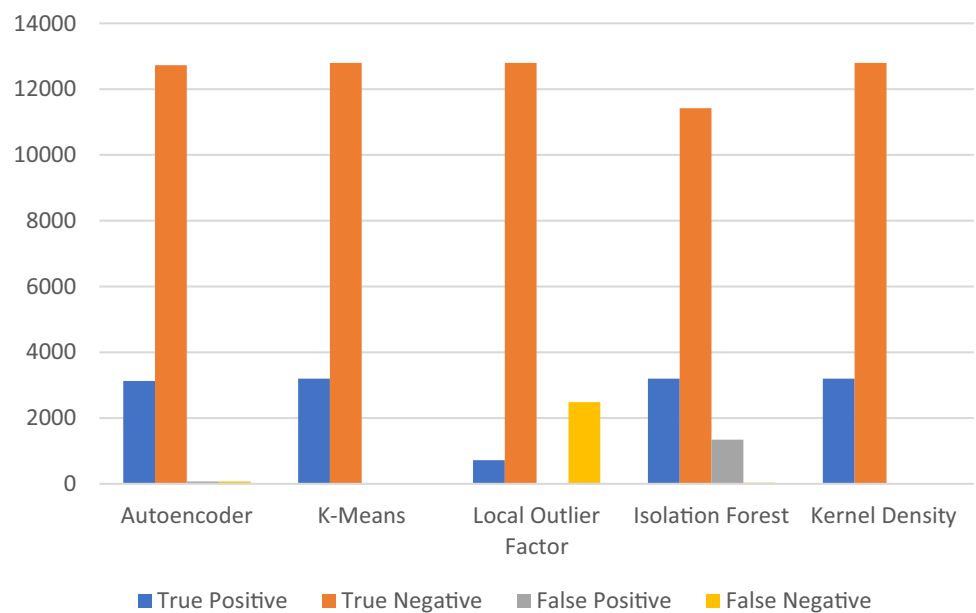
**Fig. 11** Combined scenarios for all nodes analysis



**Table 2** Results from all nodes analysis on test dataset

Damage detection technique	True Positive (%)	True Negative (%)	False Positive (%)	False Negative (%)	Precision	Recall	F1-Score
Autoencoder	19.64	79.63	0.38	0.36	0.98	0.98	0.98
K Means	20	80	0	0	1	1	1
Local outlier factor	0.63	80	0	19.38	1	0.03	0.06
Isolation forest	20	78.06	1.94	0	0.91	1	0.95
Kernel density	20	80	0	0	1	1	1

**Fig. 12** Combined scenarios for Node Octets analysis



all algorithms on the Node Octets analysis are shown in Table 3.

### 4.3 Node quartets analysis

The results for the node quartets analysis with a combination of all scenarios are shown in Fig. 14. KM and KDE continue to show perfect scores for this collection of scenarios/nodes, with AE close behind in third with precision and recall of 0.96. An interesting feature for the quartet scenarios is the improvement of LOF in terms of *F1*-score with a value of 0.74 for the 4 node groupings compared to 0.06 and 0.37 for all and octet analysis. Figure 11 The results from all algorithms on the node quartet analysis are shown in Table 4.

### 4.4 Node pairs analysis

The results for the node pairs with a combination of all scenarios are shown in Fig. 15. There is a drop in accuracy for all algorithms for the node pairs analysis, this could be caused by insufficient information being present to capture an accurate portrayal of the model bridge response. The previously high performing KDE and KM algorithms suffered a large drop on the pairing of Nodes 9–10 for KDE and Nodes 13–14 for KM, this was common across all damage scenarios and is presented in Fig. 16. This drop in accuracy meant that AE was the top performing method on the node pairs scenario, LOF placed second in terms of *F1*-Score for the node pairings, this potentially means that if limited monitoring points are available, it can be used as a viable method of general damage detection. The improvement in *F1*-score for the LOF for the node pairs could be due to the relative density of datapoints around the nodes being quite high when they are split into small

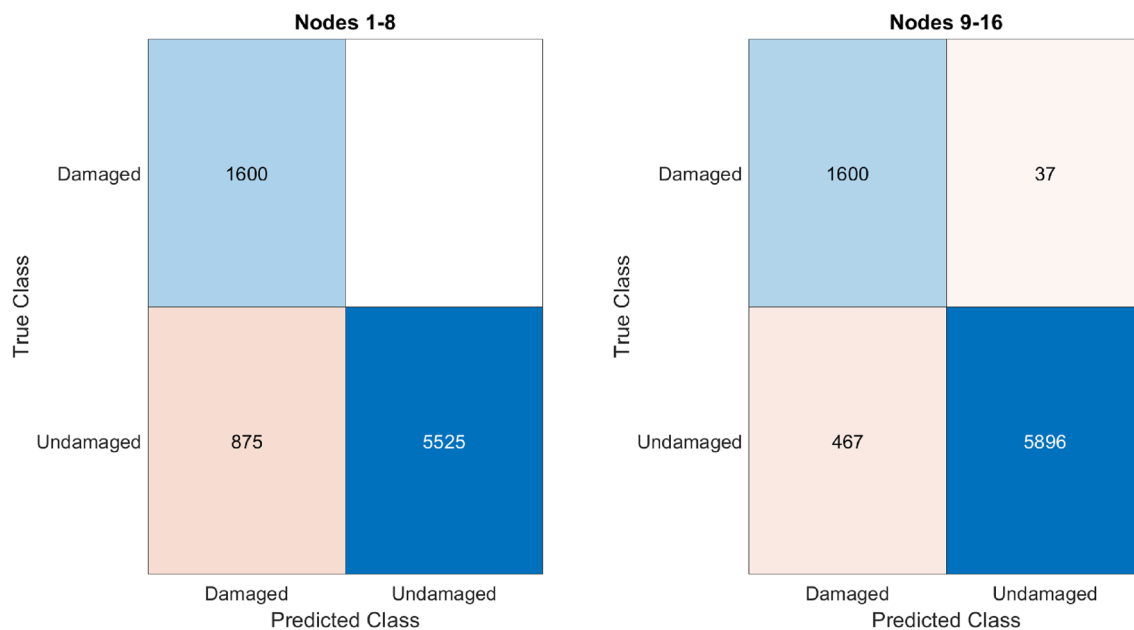
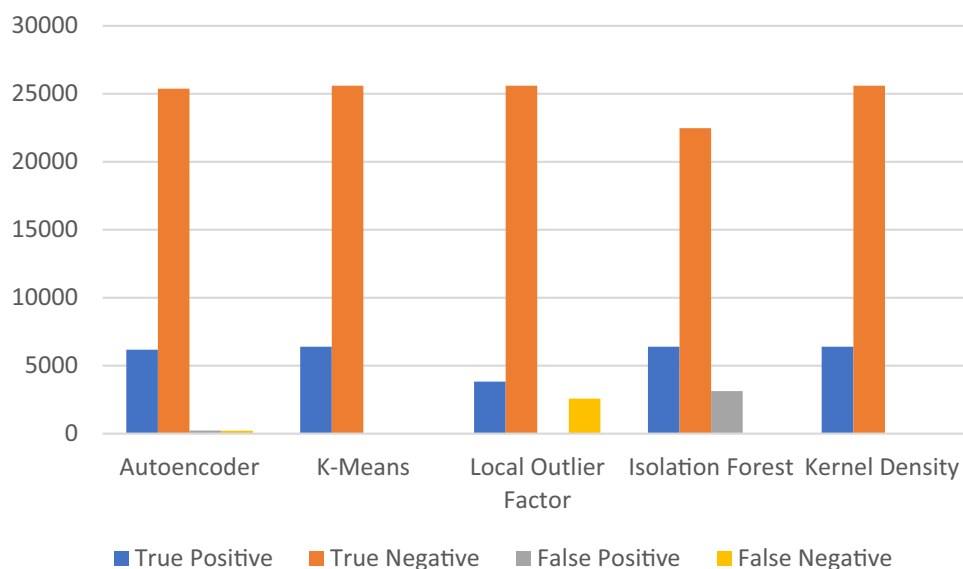


Fig. 13 Confusion matrix for isolation Forest Node octets analysis

Table 3 Results from Node Octets analysis on test dataset

Technique	True Positive (%)	True Negative (%)	False Positive (%)	False Negative (%)	Precision	Recall	<i>F1</i> -Score
Autoencoder	19.56	79.55	0.45	0.44	0.97	0.97	0.97
K Means	20	80	0	0	1	1	1
Local outlier factor	4.49	80	0	15.51	1	0.22	0.37
Isolation forest	20	71.38	8.39	0.23	0.70	0.98	0.82
Kernel density	20	80	0	0	1	1	1

**Fig. 14** Combined scenarios for Node Quartet analysis**Table 4** Results from Node Quartet analysis on test dataset

Scenario Type	True Positive (%)	True Negative (%)	False Positive (%)	False Negative (%)	Precision	Recall	F1-Score
Autoencoder	19.32	79.30	0.70	0.68	0.96	0.96	0.96
K Means	20	80	0	0	1	1	1
Local outlier factor	11.95	80	0	8.05	1	0.59	0.74
Isolation forest	20	70.22	9.78	0	0.67	1	0.80
Kernel density	20	80	0	0	1	1	1

subsets, meaning that the outliers would be more obviously over a threshold, hence improving accuracy. As discussed above, the principal idea behind LOF is that points with local densities lower than their neighbours can be considered outliers. The results from the node pairings are presented in Table 5.

#### 4.5 Running time comparison

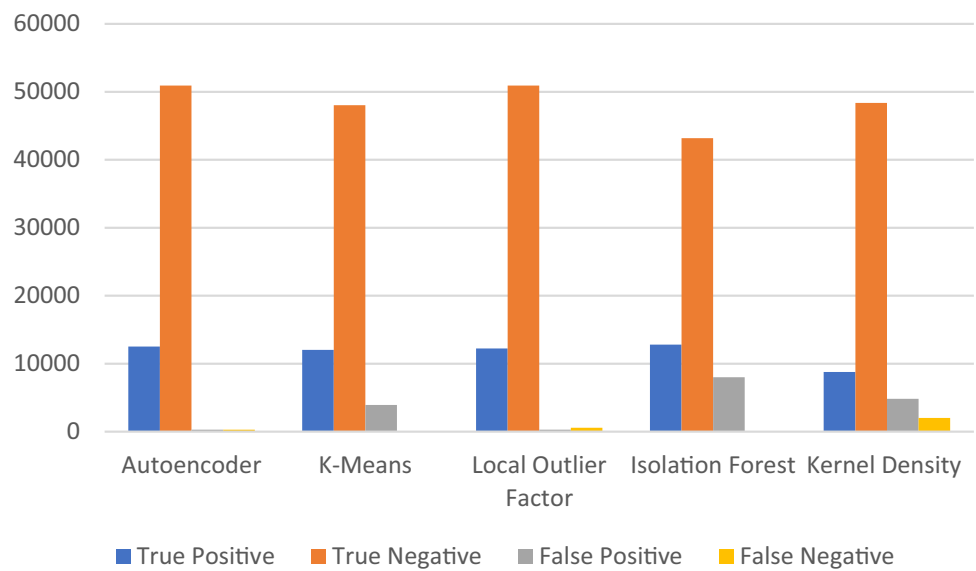
Another consideration to be taken into account when selecting between two unsupervised algorithms that have broadly similar performance is computational running time. All the algorithms used were deployed in scikit-learn or TensorFlow using a GPU accelerated Docker container on an NVIDIA GTX 970 card. The running time to train each algorithm for the node pairs analysis for each algorithm is given in Table 6. For inference, the running times for all algorithms were quite similar for all algorithms, with marginal differences between each approach.

## 5 Conclusions and future work

The results show that a computer vision-based roving sensor setup can be employed to establish a bridge baseline condition, which then can serve to check the newly collected measurements for anomaly events/to detect changes to bridge conditions. The laboratory study has shown that the sensor roving technique is viable as a means of data collection for vision-based monitoring. The technique can lead to a reduction in monitoring costs for real-life scenarios. The following conclusions are drawn from the results of the laboratory tests:

- From the comparison of damage detection algorithms, the K Means nearest neighbours' method performed best as it allowed for localised damage detection on the span of the bridge whilst also having a very short training time (2.91 s) compared to the auto-encoder and kernel density estimation, the next best performing algorithms. Whilst this was extremely accurate for all, octet and quartet node

**Fig. 15** Combined scenarios for node pairs analysis

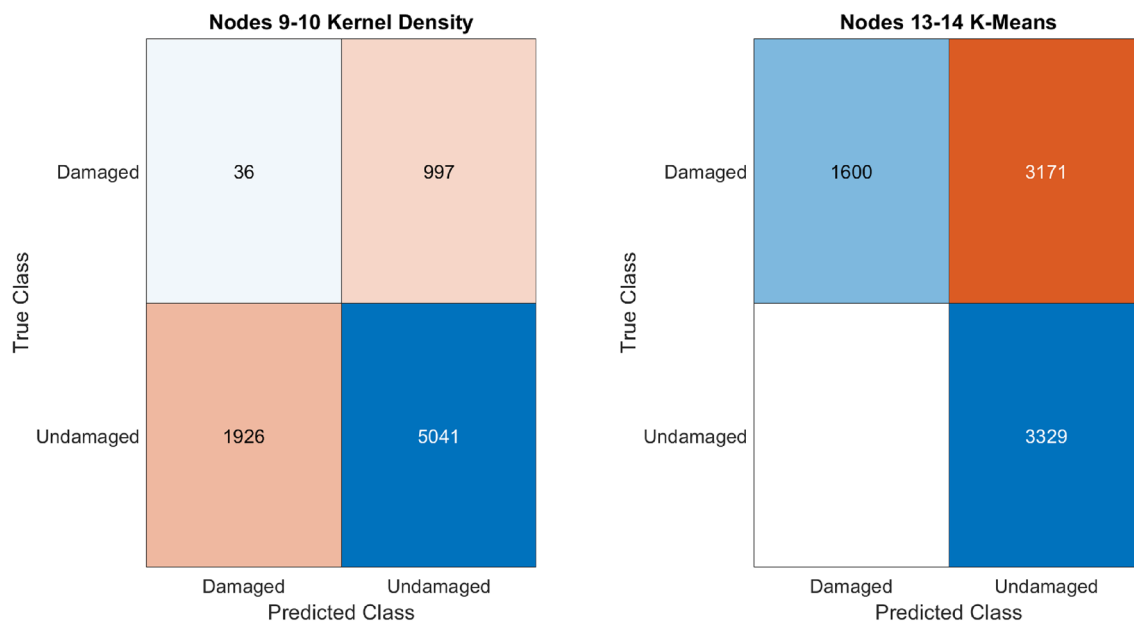


analysis, if only the node pairings are available, K-means suffers a notable loss in accuracy.

- If limited monitoring technology is present on site (i.e. only a small group of nodes can be captured), the local outlier factor could be considered a viable method as it delivers comparable accuracy (0.96 *F1-Score*) to the highest scoring method which was the auto-encoder (0.97 *F1-Score*), with a significantly shorter training time (1.17 s vs 5740 s).
- The automated and unsupervised nature (data-driven approach) of the results analysis means that, if paired with an accurate method for load evaluation as demonstrated by the authors in [72], real-time damage detec-

tion can be implemented. The accuracy of the proposed system could also be improved by feeding in live data in a continuous training methodology similar to the work in [73].

The future research can include the additional laboratory trials to obtain measurements from a large number of undamaged scenarios to validate if the boundary conditions set out in the initial runs are feasible. Other scenarios, such as multiple vehicles, different axle loads, and less severe damages, could also be explored to determine the accuracy of the proposed system.



**Fig. 16** Selected node pairs results

**Table 5** Results from node pairs analysis on test dataset

Scenario Type	True Positive (%)	True Negative (%)	False Positive (%)	False Negative (%)	Precision	Recall	F1-Score
Autoencoder	19.56	79.55	0.45	0.44	0.97	0.97	0.97
K Means	18.89	75.05	6.12	0.02	0.75	0.99	0.85
Local outlier factor	19.11	79.57	0.43	0.89	0.97	0.95	0.96
Isolation forest	20	67.49	12.51	0	0.62	1	0.76
Kernel density	13.73	75.59	7.54	3.14	0.64	0.81	0.72

**Table 6** Training time for node pairs analysis

Technique	Training time (s)	Inference time (s)
Autoencoder	5740	9.73
K Means	2.91	0.94
Local outlier factor	1.17	0.73
Isolation forest	4.69	1.56
Kernel density	765	5.79

Ideally field trials of the system should be carried out to determine the measurement accuracy of a range of scenarios outside a controlled laboratory environment. This would involve careful planning as there are some challenges when performing such trials in the field, particularly being able to reliably place cameras to monitor multiple points on the structure. This may not be feasible in all locations, so selected nodes may be chosen to be paired with the reference node to provide analysis of bridge displacements. Additionally, it may not be possible to obtain a reference vehicle for multiple crossings of the bridge at repeated instances. It is possible that if a bridge is known to be accessed by similar vehicles repeatedly, e.g. if the bridge is on a bus route, then knowledge of the bus timetable could allow for repeated data gathering opportunities or a railway bridge. If these concerns can be overcome, being able to obtain very accurate displacements for multiple nodes of the bridge could provide a detailed reference of bridge baseline conditions and allow for early damage detection of the bridge.

**Funding** This research was supported by Engineering and Physical Sciences Research Council (Grant EP/S036695/1).

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes

were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

## References

- Phares BM, Washer GA, Rolander DD, Graybeal BA, Moore M (2004) Routine highway bridge inspection condition documentation accuracy and reliability. *J Bridg Eng* 9(4):403–413. [https://doi.org/10.1061/\(ASCE\)1084-0702\(2004\)9:4\(403\)](https://doi.org/10.1061/(ASCE)1084-0702(2004)9:4(403))
- Mainline (2013) Maintenance, renewal and improvement of rail transport infrastructure to reduce economic and environmental impacts. In: Deliverable D1.1: Benchmark of New Technologies to Extend the Life of Elderly Rail Infrastructure European Project, Luleå, Sweden: 7<sup>th</sup>. Sweden
- BBC. Northern Ireland floods: More than 100 people rescued - BBC News
- Telegraph. UK weather: Bridge collapses and roads washed away as flood warnings continue to midnight
- Sigurdardottir DH, Glisic B (2015) On-site validation of fiber-optic methods for structural health monitoring: Streicker Bridge. *J Civ Struct Heal Monit* 5(4):529–549. <https://doi.org/10.1007/s13349-015-0123-x>
- García D, Tcherniak D (2019) An experimental study on the data-driven structural health monitoring of large wind turbine blades using a single accelerometer and actuator. *Mech Syst Signal Process* 127:102–119. <https://doi.org/10.1016/j.ymssp.2019.02.062>
- Sarrafi A, Mao Z, Niezrecki C, Poozesh P (2018) Vibration-based damage detection in wind turbine blades using Phase-based Motion Estimation and motion magnification. *J Sound Vib* 421:300–318. <https://doi.org/10.1016/j.jsv.2018.01.050>
- Kozin F, Natke HG (1986) System identification techniques. *Struct Saf* 3(3–4):269–316. [https://doi.org/10.1016/0167-4730\(86\)90006-8](https://doi.org/10.1016/0167-4730(86)90006-8)
- Shi ZY, Law SS, Zhang LM (2000) Structural damage detection from modal strain energy change. *J Eng Mech* 126(12):1216–1223. [https://doi.org/10.1061/\(ASCE\)0733-9399\(2000\)126:12\(1216\)](https://doi.org/10.1061/(ASCE)0733-9399(2000)126:12(1216))
- Chen J, Xu YL, Zhang RC (2004) Modal parameter identification of Tsing Ma suspension bridge under Typhoon Victor: EMD-HT method. *J Wind Eng Ind Aerodyn* 92(10):805–827. <https://doi.org/10.1016/j.jweia.2004.04.003>



11. Nayeri RD et al (2009) Study of time-domain techniques for modal parameter identification of a long suspension bridge with dense sensor arrays. *J Eng Mech* 135(7):669–683. [https://doi.org/10.1061/\(ASCE\)0733-9399\(2009\)135:7\(669\)](https://doi.org/10.1061/(ASCE)0733-9399(2009)135:7(669))
12. Hart GC, Yao JTP (1977) System identification in structural dynamics. *J Eng Mech Div* 103(6):1089–1104
13. Yao JTP, Natke HG (1994) Damage detection and reliability evaluation of existing structures. *Struct Saf* 15(1–2):3–16. [https://doi.org/10.1016/0167-4730\(94\)90049-3](https://doi.org/10.1016/0167-4730(94)90049-3)
14. Talebinejad I, Fischer C, Ansari F (2011) Numerical evaluation of vibration-based methods for damage assessment of cable-stayed bridges. *Comput Civ Infrastruct Eng* 26(3):239–251. <https://doi.org/10.1111/j.1467-8667.2010.00684.x>
15. Peeters B, De Roeck G (2001) One-year monitoring of the Z24-Bridge: environmental effects versus damage events. *Earthq Eng Struct Dyn* 30(2):149–171. [https://doi.org/10.1002/1096-9845\(200102\)30:2%3c149::AID-EQE1%3e3.0.CO;2-Z](https://doi.org/10.1002/1096-9845(200102)30:2%3c149::AID-EQE1%3e3.0.CO;2-Z)
16. Moser P, Moaveni B (2011) Environmental effects on the identified natural frequencies of the Dowling Hall Footbridge. *Mech Syst Signal Process* 25(7):2336–2357. <https://doi.org/10.1016/j.ymsp.2011.03.005>
17. Wah WSL, Chen Y-T, Roberts GW, Elamin A (2017) Damage detection of structures subject to nonlinear effects of changing environmental conditions. *Procedia Eng* 188:248–255. <https://doi.org/10.1016/J.PROENG.2017.04.481>
18. Santos JP, Crémona C, Calado L, Silveira P, Orcesi AD (2016) On-line unsupervised detection of early damage. *Struct Control Heal Monit* 23(7):1047–1069. <https://doi.org/10.1002/stc.1825>
19. Zhu F, Wu Y (2014) A rapid structural damage detection method using integrated ANFIS and interval modeling technique. *Appl Soft Comput J* 25:473–484. <https://doi.org/10.1016/j.asoc.2014.08.043>
20. Abdeljaber O, Avci O, Kiranyaz S, Gabbouj M, Inman DJ (2016) Real-time vibration-based structural damage detection using one-dimensional convolutional neural networks. *J Sound Vib* 388:154–170. <https://doi.org/10.1016/j.jsv.2016.10.043>
21. Pathirage CSN, Li J, Li L, Hao H, Liu W, Ni P (2018) Structural damage identification based on autoencoder neural networks and deep learning. *Eng Struct* 172(May):13–28. <https://doi.org/10.1016/j.engstruct.2018.05.109>
22. Dilena M, Limongelli MP, Morassi A (2015) Damage localization in bridges via the FRF interpolation method. *Mech Syst Signal Process* 52–53:162–180. <https://doi.org/10.1016/J.YMSSP.2014.08.014>
23. Krishnan M, Bhowmik B, Hazra B, Pakrashi V (2018) Real time damage detection using recursive principal components and time varying auto-regressive modeling. *Mech Syst Signal Process* 101:549–574. <https://doi.org/10.1016/j.ymsp.2017.08.037>
24. Santhosh KV, Roy BK (2017) Online implementation of an adaptive calibration technique for displacement measurement using LVDT. *Appl Soft Comput* 53:19–26. <https://doi.org/10.1016/J.ASOC.2016.12.032>
25. Bhowmik B, Tripura T, Hazra B, Pakrashi V (2020) Real time structural modal identification using recursive canonical correlation analysis and application towards online structural damage detection. *J Sound Vib* 468:115101. <https://doi.org/10.1016/j.jsv.2019.115101>
26. Zhang C, Gao Y-W, Huang J-P, Huang J-Z, Song G-Q (2020) Damage identification in bridge structures subject to moving vehicle based on extended Kalman filter with  $l_1$ -norm regularization. *Inverse Probl Sci Eng* 28(2):144–174. <https://doi.org/10.1080/17415977.2019.1582650>
27. Yang JN, Lin S, Huang H, Zhou L (2006) An adaptive extended Kalman filter for structural damage identification. *Struct Control Heal Monit* 13(4):849–867. <https://doi.org/10.1002/stc.84>
28. Meiliang W, Smyth AW (2007) Application of the unscented Kalman filter for real-time nonlinear structural system identification. *Struct Control Heal Monit* 14(7):971–990. <https://doi.org/10.1002/stc.186>
29. Santos CA, Costa CO, Batista J (2016) A vision-based system for measuring the displacements of large structures: simultaneous adaptive calibration and full motion estimation. *Mech Syst Signal Process* 72–73:678–694. <https://doi.org/10.1016/j.ymsp.2015.10.033>
30. Cha YJ, Chen JG, Büyükköztürk O (2017) Output-only computer vision based damage detection using phase-based optical flow and unscented Kalman filters. *Eng Struct* 132:300–313. <https://doi.org/10.1016/j.engstruct.2016.11.038>
31. Casas JR, Moughty JJ (2017) Bridge damage detection based on vibration data: past and new developments. *Front Built Environ* 3:4. <https://doi.org/10.3389/fbuil.2017.00004>
32. Kromanis R, Kripakaran P (2014) Predicting thermal response of bridges using regression models derived from measurement histories. *Comput Struct* 136:64–77. <https://doi.org/10.1016/J.COMPSTRUC.2014.01.026>
33. Feng MQ, Fukuda Y, Feng D, Mizuta M (2015) Nontarget vision sensor for remote measurement of bridge dynamic response. *J Bridg Eng* 20(12):04015023. [https://doi.org/10.1061/\(ASCE\)BE.1943-5592.0000747](https://doi.org/10.1061/(ASCE)BE.1943-5592.0000747)
34. Feng D, Feng MQ (2016) Output-only damage detection using vehicle-induced displacement response and mode shape curvature index. *Struct Control Heal Monit* 23(8):1088–1107. <https://doi.org/10.1002/stc.1829>
35. Erdenebat D, Waldmann D, Scherbaum F, Teferle N (2018) The Deformation Area Difference (DAD) method for condition assessment of reinforced structures. *Eng Struct* 155:315–329. <https://doi.org/10.1016/j.engstruct.2017.11.034>
36. Erdenebat D, Waldmann D, Teferle N (2019) Curvature based DAD-method for damage localisation under consideration of measurement noise minimisation. *Eng Struct* 181:293–309. <https://doi.org/10.1016/j.engstruct.2018.12.017>
37. Erdenebat D, Waldmann D (2020) Application of the DAD method for damage localisation on an existing bridge structure using close-range UAV photogrammetry. *Eng Struct* 218:110727. <https://doi.org/10.1016/j.engstruct.2020.110727>
38. Dworakowski Z, Kohut P, Gallina A, Holak K, Uhl T (2016) Vision-based algorithms for damage detection and localization in structural health monitoring. *Struct Control Heal Monit* 23(1):35–50. <https://doi.org/10.1002/stc.1755>
39. Zhang Y, Lie ST, Xiang Z (2013) Damage detection method based on operating deflection shape curvature extracted from dynamic response of a passing vehicle. *Mech Syst Signal Process* 35(1–2):238–254
40. Khuc T, Catbas FN (2018) Structural identification using computer vision-based bridge health monitoring. *J Struct Eng (United States)* 144(2):04017202. [https://doi.org/10.1061/\(ASCE\)ST.1943-541X.0001925](https://doi.org/10.1061/(ASCE)ST.1943-541X.0001925)
41. Celik O, Terrell T, Gul M, Necati Catbas F (2018) Sensor clustering technique for practical structural monitoring and maintenance. *Struct Monit Maint* 5(2):273–295. <https://doi.org/10.12989/smm.2018.5.2.273>
42. Lydon D et al (2018) Development and field testing of a time-synchronized system for multi-point displacement calculation using low-cost wireless vision-based sensors. *IEEE Sens J* 18(23):9744–9754. <https://doi.org/10.1109/JSEN.2018.2853646>
43. Auskalnis J, Paulauskas N, Baskys A (2018) Application of local outlier factor algorithm to detect anomalies in computer network. *Elektron ir Elektrotechnika* 24(3):96–99. <https://doi.org/10.5755/J01.EIE.24.3.20972>

44. Laory I, Trinh TN, Smith IFC, Brownjohn JMW (2014) Methodologies for predicting natural frequency variation of a suspension bridge. *Eng Struct* 80:211–221. <https://doi.org/10.1016/j.engstruct.2014.09.001>
45. Del Buono F, Calabrese F, Baraldi A, Paganelli M, Guerra F (2022) Novelty detection with autoencoders for system health monitoring in industrial environments. *Appl Sci* 12(10):4931. <https://doi.org/10.3390/app12104931>
46. Smith J, Nouretdinov I, Craddock R, Offer C and Gammernan A (2014) Anomaly detection of trajectories with kernel density estimation by conformal prediction. pp 271–280
47. Kumar SG, Corrado SJ, Puranik TG and Mavris DN (2022) Application of isolation forest for detection of energy anomalies in ADS-B trajectory data. In: AIAA SCITECH 2022 Forum. <https://doi.org/10.2514/6.2022-2441>
48. Pimentel MAF, Clifton DA, Clifton L, Tarassenko L (2014) A review of novelty detection. *Signal Process* 99:215–249. <https://doi.org/10.1016/j.sigpro.2013.12.026>
49. Miljković D (2010) Review of novelty detection methods. In: MIPRO 2010 - 33rd Int. Conv. Inf. Commun. Technol. Electron. Microelectron. Proc., pp 593–598
50. Zaurin R, Catbas F (2010) Structural health monitoring using video stream, influence lines, and statistical analysis. *Struct Heal Monit* 10(3):309–332. <https://doi.org/10.1177/1475921710373290>
51. Feng D, Feng MQ (2015) Model updating of railway bridge using in situ dynamic displacement measurement under Trainloads. *J Bridg Eng* 20(12):04015019. [https://doi.org/10.1061/\(ASCE\)BE.1943-5592.0000765](https://doi.org/10.1061/(ASCE)BE.1943-5592.0000765)
52. Gamache R and Santini-Bell E (2009) Non-intrusive digital optical means to develop bridge performance information. In: Non-destructive testing in civil engineering, pp 1–6
53. Bay H, Ess A, Tuytelaars T, Van Gool L (2008) Speeded-up robust features (SURF). *Comput Vis Image Underst* 110(3):346–359. <https://doi.org/10.1016/j.cviu.2007.09.014>
54. Brown M and Lowe D (2002) Invariant features from interest point groups. In: Proceedings of the British machine vision conference 2002, pp 23.1–23. <https://doi.org/10.5244/C.16.23>
55. Papageorgiou CP, Oren M and Poggio T (1998) General framework for object detection. In: Proceedings of the IEEE International Conference on Computer Vision, pp 555–562. <https://doi.org/10.1109/iccv.1998.710772>
56. Hartley R, Zisserman A (2004) Multiple view geometry in computer vision. Cambridge University Press, Cambridge
57. 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. <https://doi.org/10.1016/j.ymsp.2018.11.015>
58. GoPro (2016) GoPro—refurbished HERO4 Black 4K Ultra HD Waterproof Camera. <https://shop.gopro.com/EMEA/refurbished/refurbished-hero4-black/CHDNH-B11.html>. (Accessed: 19-Jan-2018)
59. Back-Bone (2016) Ribcage AIR HERO4 Mod Kit Bundle | BACK-BONE. <https://www.back-bone.ca/product/ribcage-air-hero4-mod-kit/>. (Accessed: 19-Jan-2018)
60. Computar (2016) E5Z2518C-MP : Manual Iris: Megapixel Varifocal Lenses: ProductsMegapixel, FA, HD, Varifocal - : Computar. <https://computar.com/product/1115/E5Z2518C-MP>. (Accessed: 19-Jan-2018)
61. Hinton GE, Salakhutdinov RR (2006) Reducing the dimensionality of data with neural networks. *Science* (80–) 313(5786):504–507. <https://doi.org/10.1126/science.1127647>
62. Sakurada M and Yairi T (2014) Anomaly detection using autoencoders with nonlinear dimensionality reduction. In: Proceedings of the MLSDA 2014 2nd Workshop on Machine Learning for Sensory Data Analysis - MLSDA'14
63. Zhou C and Paffenroth RC (2017) Anomaly detection with robust deep autoencoders. In: Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, 2017, vol. Part F129685, pp 665–674. <https://doi.org/10.1145/3097983.3098052>
64. Breunig MM, Kriegel H-P, Ng RT and Sander J (2000) LOF. In: Proc. 2000 ACM SIGMOD Int. Conf. Manag. data - SIGMOD '00, pp 93–104. <https://doi.org/10.1145/342009.335388>
65. Rosenblatt M (1956) Remarks on some nonparametric estimates of a density function. *27(3):832–837*. <https://doi.org/10.1214/AOMS/1177728190>
66. Lloyd SP (1982) Least squares quantization in PCM. *IEEE Trans Inf Theory* 28(2):129–137. <https://doi.org/10.1109/TIT.1982.1056489>
67. Sculley D (2010) Web-scale k-means clustering. In: Proc. 19th Int. Conf. World Wide Web, WWW '10, pp 1177–1178. <https://doi.org/10.1145/1772690.1772862>
68. Liu FT, Ting KM and Zhou ZH (2008) Isolation forest. In: Proc. - IEEE Int. Conf. Data Mining, ICDM, pp 413–422. <https://doi.org/10.1109/ICDM.2008.17>
69. Friedman JH, Bentley JL and Finkel RA (1975) An algorithm for finding best matches in logarithmic expected time
70. Bergstra J, Ca JB, Ca YB (2012) Random search for hyper-parameter optimization Yoshua Bengio. *J Mach Learn Res* 13:281–305
71. Dolatshah M, Hadian A and Minaei-Bidgoli B (2015) Ball\*-tree: Efficient spatial indexing for constrained nearest-neighbor search in metric spaces. <https://doi.org/10.48550/arxiv.1511.00628>
72. Lydon D, Taylor SE, Lydon M, del Rincon JM, Hester D (2019) Development and testing of a composite system for bridge health monitoring utilising computer vision and deep learning. *SMART Struct Syst* 24(6):723–732. <https://doi.org/10.12989/sss.2019.24.6.723>
73. Tonioni A, Tosi F, Poggi M, Mattoccia S and DI Stefano L (2019) Real-time self-adaptive deep stereo

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