Smart Structures and Systems, *Vol. 33, No. 2 (2024) 93-103*
https://doi.org/10.12989/sss.2024.33.2.093

Deep learning-based anomaly detection in acceleration data of long-span cable-stayed bridges

Seungjun Lee [1c], Jaebeom Lee*[2], Minsun Kim [1c], Sangmok Lee [3b] and Young-Joo Lee [1a]

[1] *Department of Civil, Urban, Earth, and Environmental Engineering, Ulsan National Institute of Science and Technology (UNIST), 50, UNIST-gil, Eonyang-eup, Ulju-gun, Ulsan, Republic of Korea*
[2] *Interdisciplinary Materials Measurement Institute, Korea Research Institute of Standards and Science (KRISS), 267, Gajeong-ro, Yuseong-gu, Daejeon, Republic of Korea*
[3] *Dam Safety Management Center, Korea Water Resources Corporation (K-water), 200, Sintanjin-ro, Daedeok-gu, Daejeon, Republic of Korea*

(*Received March 2, 2023, Revised December 20, 2023, Accepted December 28, 2023*)

Abstract. Despite the rapid development of sensors, structural health monitoring (SHM) still faces challenges in monitoring due to the degradation of devices and harsh environmental loads. These challenges can lead to measurement errors, missing data, or outliers, which can affect the accuracy and reliability of SHM systems. To address this problem, this study proposes a classification method that detects anomaly patterns in sensor data. The proposed classification method involves several steps. First, data scaling is conducted to adjust the scale of the raw data, which may have different magnitudes and ranges. This step ensures that the data is on the same scale, facilitating the comparison of data across different sensors. Next, informative features in the time and frequency domains are extracted and used as input for a deep neural network model. The model can effectively detect the most probable anomaly pattern, allowing for the timely identification of potential issues. To demonstrate the effectiveness of the proposed method, it was applied to actual data obtained from a long-span cable-stayed bridge in China. The results of the study have successfully verified the proposed method's applicability to practical SHM systems for civil infrastructures. The method has the potential to significantly enhance the safety and reliability of civil infrastructures by detecting potential issues and anomalies at an early stage.

Keywords: anomaly patterns; deep neural network (DNN); feature extraction; sensor data; structural health monitoring (SHM)

1. Introduction

Structural health monitoring (SHM) has been widely adopted as one of the safety management methods for civil infrastructures such as cable-stayed bridges (Sohn *et al.* 2003). It aims to detect the structural problems by measuring structural responses including vibration, displacement, and environmental information; thus, several studies have been conducted to develop reliable sensing and data analysis techniques (Li *et al.* 2016). Data acquisition of structural responses is a crucial step in SHM. Therefore, vibration-, strain-, or displacement-based sensing techniques have been developed frequently. For example, various types of sensors, including fiber optic sensors (Ansari 2007), piezoelectric wafer active sensors (Mei *et al.* 2019), wireless data acquisition systems (Straser *et al.* 1998), and computer vision-based systems (Lee *et al.* 2017) have been studied, and they are sometimes complementally applied to infrastructures as they have advantages and disadvantages.

Meanwhile, the identification of structural damages from the measured sensor data has become a significant issue as well. Sohn *et al.* (2001) suggested an auto-regressive-based pattern recognition in a statistical view, and Manson (2002) developed a feature-based damage detection technique using a principal component analysis that can identify damages by filtering temperature effects out. Arangio and Bontempi (2015) suggested an SHM model for cable-stayed bridges using Bayesian neural networks. Lee *et al.* (2018) suggested a probabilistic method to detect the unusual vertical deflection of railway bridges. Xin *et al.* (2018) used a combinatory method of a Kalman filter, an autoregressive integrated moving average model, and a generalized autoregressive conditional heteroskedasticity process to the SHM problem. Lee *et al.* (2019) proposed a method to integrate a finite element model with sensor data for detecting the structural problems using Bayesian inference. Bao *et al.* (2019) suggested a deep learning-based SHM method using computer vision sensor data. Lee *et al.* (2022) suggested a computer-vision sensing based probabilistic monitoring of time-history deflections of railway bridges, and Son *et al.* (2022) proposed a combinational identification method of three efficient techniques, including statistical analysis,

*Corresponding author, Ph.D.,
E-mail: jblee@kriss.re.kr
[a] Ph.D., Associate Professor
[b] Ph.D., PE
[c] Ph.D. Student

ISSN: 1738-1584 (Print), 1738-1991 (Online)

clustering, and neural network models to detect the damaged cable in a cable-stayed bridge. These studies aimed to develop new methods to identify structural damage from the measured sensor data to enhance the accuracy of SHM.

Despite the rapid development of sensing and data analysis techniques, SHM occasionally suffers difficulty in damage detection owing to the functional degradation of devices and harsh environmental loads, which may result in measurement errors, missing data, or outliers (Posenato *et al*. 2010). This data quality problem may result in inaccurate assessments; thus, data filtering is important for a successful SHM. To overcome this issue, Kullaa (2011) suggested a framework to distinguish the effects of sensor faults, environmental loads, and structural damage. Yi *et al*. (2016) proposed a data transformation method for reducing the risk of false alarms and missed detections of a bridge deflection. Smarsly and Law (2014) studied a sensor fault detection problem for wireless SHM systems. Xu *et al*. (2019) suggested a methodology for diagnosing sensor faults based on the principle that the structural responses at symmetric locations should be similar.

Substantial efforts have been made in research aimed at alleviating the effects of sensor faults. However, despite these developments, the effective integration of such advancements into practical SHM systems for civil structures poses challenges in the following aspects: (1) in large-scale civil structures, handling faults in multiple sensors becomes particularly challenging due to the extensive time and resources required to test every possible sensor combination with the conventional approach (Yi *et al*. 2017); (2) detection of multiple types of sensor anomalies verified through simulation and laboratory experiments may face difficulties in practical application due to differences with field monitoring (Kullaa 2013, Chang *et al*. 2017, Fu *et al*. 2019); and (3) the process of extracting features in deep learning-based anomaly detection framework that utilize raw time-series data from multi-channel sources can be computationally intensive (Bao *et al*. 2019, Oh *et al*. 2020, Ni *et al*. 2020).

In this study, a deep learning-based classification method, to detect anomalies in data acquired from multiple sensors quickly and accurately, was proposed. The proposed method includes three distinct phases: (1) scaling of data acquired from multiple sensors; (2) time domain and frequency domain feature extraction and selection process; and (3) deep learning-based classification model training for multiple anomaly pattern detection. To evaluate the performance of the proposed method, we participated in the 1st International Project Competition for Structural Health Monitoring (IPC-SHM; Bao *et al*. 2021) which gives the acceleration data from a real long-span highway bridge. In the competition, the proposed method achieved promising results in terms of quick and accurate anomaly detection, which highlights the potential of deep learning techniques in the field of anomaly detection.

2. Proposed method for data anomaly detection

This study aims to suggest a classification model that categorizes the sensor data according to the anomaly patterns, which may occur due to sensor faults or harsh environmental loads. The proposed method for constructing the classification model is depicted in Fig. 1. First, data scaling is conducted to adjust the scales of the raw data, which may have different magnitudes and ranges. The next step is the selection and extraction of informative variables (i.e., features) such as statistical properties to reduce the data complexity, and the extracted features are used as inputs for the classification model. In this study, the classification model is constructed as a deep neural network (DNN), which is capable of dealing with complex nonlinear problems.

2.1 Data scaling

To ensure a reliable data analysis, it is important to address the issue of different magnitudes, ranges, and units in data obtained from various sources, which can make model training difficult (Bakar *et al*. 2006). Therefore, data scaling is a crucial pre-processing step. In this study, standardization, which is one of the popular methods for data scaling, has been adopted. Standardization involves rescaling the data to have a zero mean and a unit variance, and is defined as

$$d_{i,standardized} = \frac{d_i - \mu_d}{\sigma_d} \tag{1}$$

where d_i and $d_{i,standardized}$ are the i^{th} data before and after standardization, respectively, and μ_d and σ_d are the mean and standard deviation of the given dataset, respectively. By using standardization as a data scaling method, we can

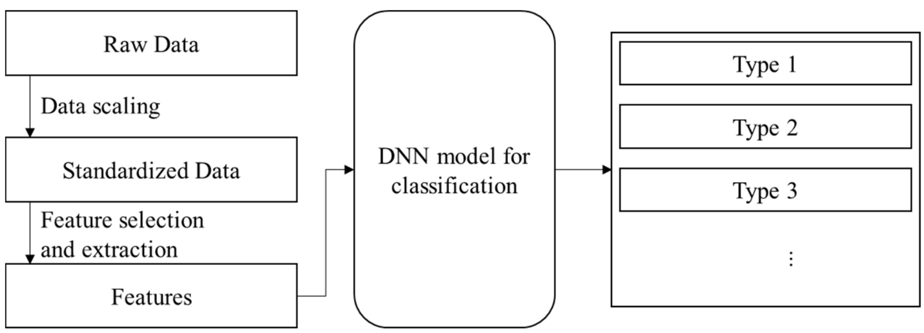


Fig. 1 Overview of the proposed method

ensure that the data is on the same scale, which facilitates the comparison of data across different sensors, and can improve the accuracy and reliability of the data analysis. This pre-processing step is crucial in our study to detect anomaly patterns in sensor data for structural health monitoring and could be beneficial in other fields where data from different sources need to be compared and analyzed.

2.1 Feature selection and extraction for dimension reduction

In the process of constructing a predictive model, let us consider the output variable (i.e., y) as a function of a set of input variables (i.e., t_1, t_2, t_3), which can be expressed as follows

$$y = 2 \times t_1 + t_2 - 0.4 \times t_3 \tag{2}$$

It is suggested that, although additional inputs t_4 and t_5 are available, the accuracy of the model may be optimized by exclusively restricting the training data to t_1, t_2, and t_3. This selective approach, known as feature selection (Khalid *et al*. 2014, Zebari *et al*. 2020), involves the strategically choosing variables that exhibit a direct and substantial correlation with the output variable. The rationale behind feature selection is its ability to reduce dataset complexity and decrease the dimensionality of the input space, thereby enhancing the predictive model's effectiveness. In addition to feature selection, the introduction of a novel variable, denoted as t_{new} according to Eq. (3), contributes to a more precise characterization of the relationship between the input variables and the output. This process, referred to as feature extraction (Khalid *et al*. 2014, Zebari *et al*. 2020), involves constructing t_{new} such that the output y is twice the value of t_{new}.

$$t_{new} = t_1 + 0.5 \times t_2 - 0.2 \times t_3 \tag{3}$$

The integration of both feature selection and extraction becomes essential in the data pre-processing phase. These techniques play a pivotal role in improving the predictive model's accuracy. By carefully selecting the most relevant input variables and deriving new features that more accurately represent their relationship with the output, it is possible to significantly reduce data redundancy and ensure that only the most pertinent variables are included. This strategic application of feature selection and extraction is fundamental in developing an efficient predictive model that maintains a high degree of accuracy.

In the field of civil infrastructure monitoring, such as a long-span cable-stayed bridge, the acceleration data collected may contain massive amounts of information. However, using raw data directly to train a model may increase computational costs without significantly improving the model accuracy. Even though many deep learning methods, such as a convolutional neural network, are known to cover the feature selection and extraction parts, those are crucial steps in identifying the latent variables that are most relevant to the output, especially when the given data is not sufficient. This study proposes selecting several features in both the time and frequency domains, which have been widely recognized as informative variables in previous research (Altın and Er 2016). To calculate these features, equations listed in Tables 1 and 2 are used on the i^{th} raw sensor data x_i. By conducting feature selection and extraction, the resulting model is more efficient, accurate, and enables effective detection of potential anomalies or problems in civil infrastructure.

After conducting feature selection to identify the relevant variables for the output, the next crucial step is feature extraction. In many studies, a principal component analysis (PCA) has been adopted for this purpose (Abdi and Williams 2010). The PCA method uses a linear combination of the selected variables, such as t_{new} in Eq. (3), to generate principal components (PCs) that are designed to be orthogonal to one another (i.e., the inner product is zero). The informativeness of each PC is determined by the magnitude of the corresponding eigenvalue. By selecting the PCs with high eigenvalues, the dimensionality of the input data can be reduced, resulting in a more efficient and effective model.

2.3 Construction of deep neural network for the classification problem

Deep neural networks, which have been widely applied to various engineering problems owing to their capability to deal with complex nonlinear problems (Kung and Diamantaras 1990), have been adopted to solve the classification problem in this study. The DNN is a network with multiple hidden layers having multiple nodes, which are connected as depicted in Fig. 2. When N_{11}, N_{12}, and N_{13} are the input nodes, the output node of the next layer (i.e., N_{21}) can be calculated as $N_{21} = f_{act}(w_0 + w_1 N_{11} + w_2 N_{12} + w_3 N_{13})$, where f_{act} is the activation function. The performance of the DNN depends on the proper design of the activation function, and the number of layers and nodes.

Contrary to the regression problem where the final output values are continuous, the outputs for the classification problem are discrete. Therefore, the activation function, which connects the final hidden layer and the output layer, can be designed as the softmax function, $\sigma_{softmax}$

$$\sigma_{softmax}(N_{fi} | N_{f1}, N_{f2}, \ldots, N_{f(k-1)}, N_{fk}) = \frac{e^{N_{fi}}}{\sum_{j=1}^{k} e^{N_{fj}}} \tag{4}$$

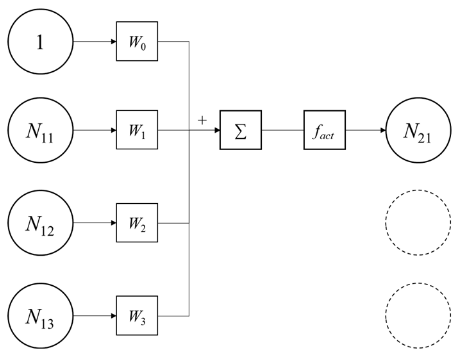


Fig. 2 Calculation of the output node (i.e., N_{21}) from the input nodes (i.e., N_{11}, N_{12}, N_{13}) in DNN

Table 1 Time domain features

Feature:	Equation:
Mean	$T_{mean} = \left(\frac{\sum_{i=1}^{n} x_i}{n}\right)$
Standard deviation	$T_{std} = \sqrt{\frac{\sum_{i=1}^{n}(x_i - T_{mean})^2}{n-1}}$
Root mean square	$T_{rms} = \sqrt{\frac{\sum_{i=1}^{n}(x_i)^2}{n}}$
Square root of the amplitude	$T_{sra} = \left(\frac{\sum_{i=1}^{n}\sqrt{\lvert x_i\rvert}}{n}\right)^2$
Skewness	$T_{skew} = \frac{\sum_{i=1}^{n}(x_i - T_{mean})^3}{(n-1){T_{std}}^3}$
Kurtosis	$T_{kurt} = \frac{\sum_{i=1}^{n}(x_i - T_{mean})^4}{(n-1){T_{std}}^4}$
Shape factor	$T_{sf} = \frac{T_{rms}}{(\frac{1}{n})\sum_{i=1}^{n}\lvert x_i\rvert}$
Crest factor	$T_{cf} = \frac{max\lvert x_i\rvert}{T_{rms}}$
Impulse factor	$T_{if} = \frac{max(\lvert x_i\rvert)}{\left(\frac{1}{n}\right)\sum_{i=1}^{n}\lvert x_i\rvert}$
Clearance factor	$T_{clf} = \frac{max\lvert x_i\rvert}{T_{sra}}$
Skewness factor	$T_{skf} = \frac{T_{sk}}{T_{rms}^3}$
Kurtosis factor	$T_{kuf} = \frac{T_{sk}}{T_{rms}^4}$

Table 2 Frequency domain features

Feature:	Equation:
Mean frequency	$F_{mf} = \frac{\sum_{i=1}^{k} p_i}{k}$
Frequency center	$F_{fc} = \frac{\sum_{i=1}^{k} f_i \cdot p_i}{\sum_{i=1}^{k} p_i}$
Root mean square frequency	$F_{rmsf} = \sqrt{\frac{\sum_{i=1}^{k} f_i^2 \cdot p_i}{\sum_{i=1}^{k} p_i}}$
Power spectrum standard deviation	$F_{stdf} = \sqrt{\frac{\sum_{i=1}^{k}(f_i - F_{fc})^2 \cdot p_i}{\sum_{i=1}^{k} p_i}}$
Stabilization factor #1 of wave shape	$F_{sf1} = \sqrt{\frac{\sum_{i=1}^{k} f_i^4 \cdot p_i}{\sum_{i=1}^{k} f_i^2 \cdot p_i}}$
Stabilization factor #2 of wave shape	$F_{sf2} = \frac{\sum_{i=1}^{k} f_i^2 \cdot p_i}{\sqrt{\sum_{i=1}^{k} p_i \sum_{i=1}^{k} f_i^4 \cdot p_i}}$
Frequency domain skewness	$F_{skf} = \frac{\sum_{i=1}^{k}(f_i - F_{fc})^3 \cdot p_i}{\sigma^3 \cdot k}$
Frequency domain kurtosis	$F_{kuf} = \frac{\sum_{i=1}^{k}(f_i - F_{fc})^4 \cdot p_i}{\sigma^4 \cdot k}$
Frequency domain Coefficient of variation	$F_{fcv} = \frac{\sigma}{F_{fc}}$
Root mean square ratio	$F_{rmsr} = \frac{\sum_{i=1}^{k}\sqrt{(f_i - F_{fc})} \cdot p_i}{\sqrt{\sigma} \cdot k}$

* p_i is the power spectrum density; k is the number of spectrum lines; f_i is the frequency value of the i^{th} spectrum line; $\sigma = \sqrt{\sum_{i=1}^{k} (f_i - F_{fc})^2 \cdot p_i / k}$ (Fleming and Egeseli 1980).

where N_{fi} is the i^{th} node in the final hidden layer, which has k-number of nodes. The resultant value is between zero and one, which can be regarded as the likelihood, therefore, the output node with the highest likelihood can be selected as the classification result.

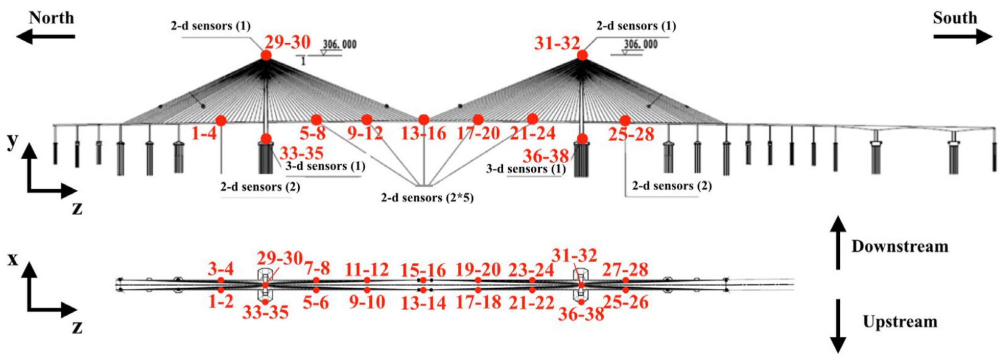
3. Application example

3.1 Problem description

The acceleration data from the Su-Tong Yangtze River Highway Bridge (SYRHB) from January 1 to January 31, 2012, were used for this example, provided by the 1st International Project Competition for Structural Health Monitoring (IPC-SHM; Bao *et al*. 2021). The bridge has two side spans of 300 m each, a main span of 1088 m, and two 306 m-high towers. The dataset contains one month of acceleration data obtained from 38 sensors attached to the long-span cable-stayed bridge in China, with a sampling frequency of 20 Hz, as shown in Fig. 3. The raw time series measurements are divided into 1-hour intervals, resulting in 744 time series measurements for each sensor over one month, for a total of 744 × 38 datasets. The dataset has seven anomaly patterns, including *Normal*, *Missing*, *Minor*, *Outlier*, *Square*, *Trend*, and *Drift*, as listed in Table 3. Time domain and frequency domain examples of each data pattern are depicted in Fig. 4. As shown in Table 3, 48.02% of the dataset is classified as *Normal* patterns; however, more than half show signs of anomalies, with the *Trend* pattern being the most prominent at 20.44%. The *Missing* and *Square* anomalies are also notable, each accounting for approximately 10% of the dataset. In contrast, *Outlier* and *Drift* patterns are less frequent, at 1.9% and 2.4% respectively, indicating an imbalance in the distribution of labelled datasets. These anomaly patterns may result from environmental variations or sensor faults, which can hinder reliable warnings for structural damage or accidents. Thus, identifying and removing data with unexpected anomalies are crucial for structural condition monitoring.

3.2 Classifier model construction

To address this problem, the proposed method (Fig. 1) was applied to construct the classification model for the seven anomaly patterns in the given dataset of the target bridge. At first, 12 features in the time domain and 10

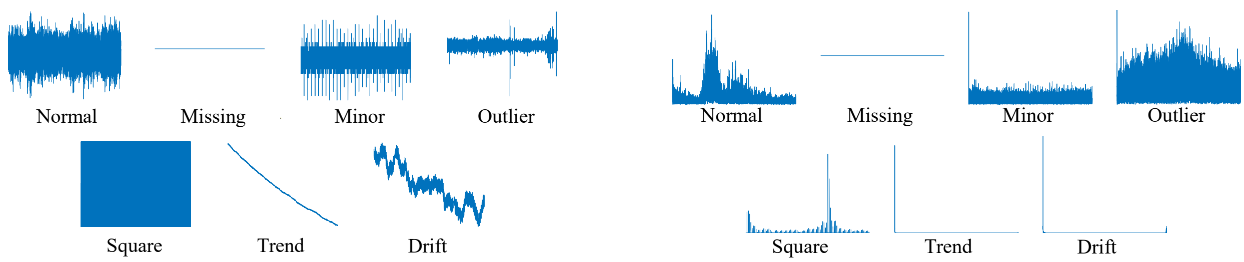


x-direction:	2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 29, 31, 33, 36	y-direction:	1, 3, 5, 7, 9, 11, 13, 15, 17, 19, 21, 23, 25, 27, 35, 38	z-direction:	30, 32, 34, 37

Fig. 3 Location of the sensors attached to a target long-span cable-stayed bridge

Table 3 Seven anomaly patterns

Anomaly patterns:	Description:	Quantity in dataset
Normal	The time response is a normal oscillation curve whereas the frequency response is peak-like.	13575 (48.02%)
Missing	Majority/all of the time response is missing, which makes the time and frequency response zero.	2942 (10.41%)
Minor	Relative to normal sensor data, the amplitude is very small in the time domain.	1775 (6.28%)
Outlier	One or more outliers appear in the time response.	527 (1.86%)
Square	The time response is like a square wave.	2996 (10.60%)
Trend	The data has an obvious trend in the time domain and has an obvious peak value in the frequency domain.	5778 (20.44%)
Drift	The vibration response is non-stationary, with random drift.	679 (2.40%)

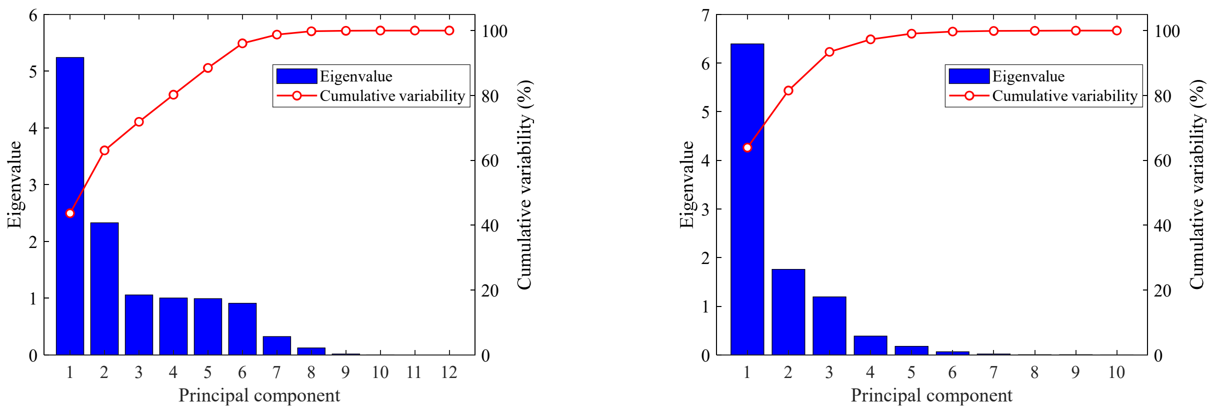


(a) Time domain (b) Frequency domain

Fig. 4 Examples of each data anomaly pattern in the dataset

features in the frequency-domain (as listed in Tables 1 and 2) were derived from the standardized sensor data. PCA was conducted for the time- and frequency-domains, respectively, to find the informative coordinates for each domain. In the time domain, the linear combinations of the features and corresponding eigenvalues were estimated as shown in Fig. 5(a). Because the percent variability of PCt1 to PCt6 accounts for 96.06%, these six PCs were extracted, which resulted in the dimension reduction from 12- to 6-dimensional spaces. In the frequency domain, the top three PCs (i.e., PCf1 to PCf3) were extracted as the new input features because they accounted for 93.44% of the percent variability, as depicted in Fig. 5(b). Therefore, a total of nine PCs were used as inputs for the classification model.

Fig. 6 depicts the classification process for the given data. Among the seven anomaly patterns, the *Missing* type has NaN values or continuously repeated values in the sensor data where the time and frequency response are zero; thus, it can be priorly classified by finding a dataset that has several NaN values or repeated constant values. To classify the dataset with the remaining six anomaly patterns, a DNN structure for the classification problem was constructed, as depicted in Fig. 6. For nine inputs (i.e., PCt1–PCt6 and PCf1–PCf3) and six outputs (i.e., anomaly patterns), the DNN model was designed to have four hidden layers with 64-128-128-64 number of hidden nodes. A hyperbolic



(a) (b)

Fig. 5 Principal component analysis results for (a) time domain features; and (b) frequency domain features

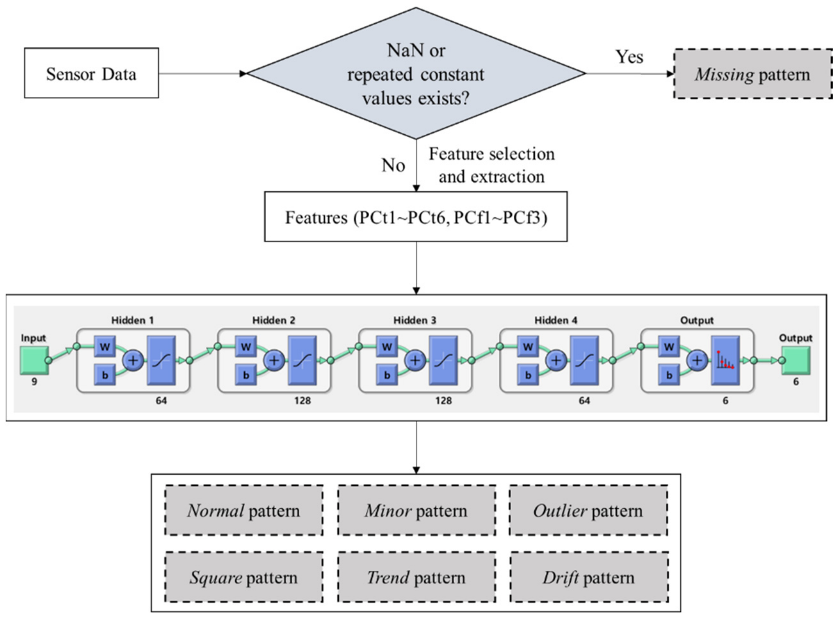


Fig. 6 Diagram of the classification process

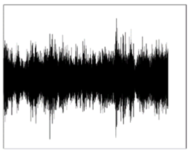
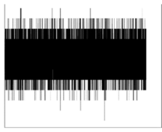
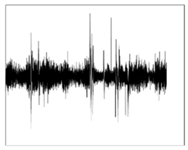
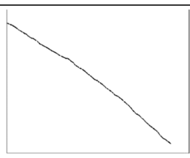
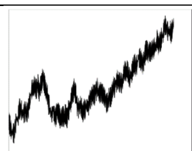
tangent sigmoid transfer function was employed as an activation function in the hidden layer, and a softmax function was used in the output layer. To enhance the reliability of the DNN model, cross-validation was conducted (70%, 15%, and 15% of the data were used for model training, validation, and testing, respectively).

As a result, the output vector consists of six elements between zero and one, which denotes the likelihood of classifying each anomaly pattern. Table 4 lists the examples of model outputs. For example, the classification model provides six by one vector for Data #1 with the highest value for *Normal*, which indicates that the given data is likely a *Normal* pattern. Indeed, Data #1 is a *Normal* pattern; thus, the classification result is correct. Similarly, in the case of Data #6, the model gives the highest value for the *Drift* pattern, which is also correct.

3.3 Model performance for anomaly pattern detection

Fig. 7 depicts a confusion matrix, which describes the performance of a classification model for training, validation, test, and overall dataset, respectively. From this matrix, the percentages for true positive (*TP*), true negative (*TN*), false positive (*FP*), and false negative (*FN*) can be found. For example, 8315 training data with true label of *Normal* were classified as *Normal*, which is the case of *TP* in terms of *Normal* pattern. Similarly, 6541 (i.e., 903+305+2037+2915+381) number of training data with true label of other than *Normal* were classified as not *Normal*, which is the case of *TN* in terms of *Normal* pattern. That is, the two cases of *TP* and *TN* are correct. However, 71 (i.e., 56+15) number of training data with true label of *Normal* were classified as *Normal*, which is called *FP*, and 71 (i.e., 38+33) number of training data were categorized as *Normal* even though the true labels were not *Normal*, which is *FN*. In the confusion matrix, green text in the gray areas

Table 4 Examples of model outputs

		Data #1:	Data #2:	Data #3:
Data plot:				
True label:		*Normal*	*Minor*	*Outlier*
Model Output	*Normal*	**0.9993**	0.2356	0.1408
	Minor	2.05×10^{-7}	**0.7500**	8.20×10^{-7}
	Outlier	9.09×10^{-5}	0.0134	**0.8583**
	Square	0.0006	3.72×10^{-6}	0.0007
	Trend	1.03×10^{-6}	0.0003	9.68×10^{-6}
	Drift	4.44×10^{-8}	0.0006	7.97×10^{-5}
		Data #4:	Data #5:	Data #6:
Data plot:				
True label:		*Square*	*Trend*	*Drift*
Model Output	*Normal*	0.0001	3.15×10^{-8}	4.21×10^{-5}
	Minor	2.65×10^{-8}	3.33×10^{-9}	4.50×10^{-6}
	Outlier	4.72×10^{-5}	1.31×10^{-9}	8.00×10^{-5}
	Square	**0.9997**	3.48×10^{-7}	6.05×10^{-9}
	Trend	4.07×10^{-6}	**0.9999**	0.0617
	Drift	6.30×10^{-5}	3.57×10^{-7}	**0.9382**

indicates the percentage of correct classifications for each class, calculated as *TP* divided by (*TP* + *FN*) for the right area and *TP* divided by (*TP* + *FP*) for the bottom area. Red text in the gray areas denotes the percentage of misclassified instances, considering *FP* and *FN* in those calculations.

From the given confusion matrix, the performance of the classification model can be assessed using four indices of *Precision*, *Recall*, *Accuracy*, and *F1-score*

$$Precision = \frac{TP}{(TP+FP)} \tag{5}$$

$$Recall = \frac{TP}{(TP+FN)} \tag{6}$$

$$Accuracy = \frac{(TP+TN)}{(TP+TN+FP+FN)} \tag{7}$$

$$F1-score = \frac{2\times(Precision \times Recall)}{(Precision + Recall)} \tag{8}$$

High *Precision* means that the ratio of *FP* is low, and high *Recall* indicates that the *FN* case is few. *Accuracy* denotes the ratio of the correct model output, and *F1-score* is a popular index for performance assessment that considers *Precision* and *Recall*. The values for these indices are within zero to one; thus, closer to one indicates better performance for all indices. These four indices of performance assessment for six anomaly patterns, summarized in Tables 5 to 8.

The proposed deep learning-based anomaly detection model achieves an accuracy of 98.00% across the entire dataset. This high level of accuracy is complemented by notable *Precision* and *Recall* rates for various patterns: *Normal* at 99.12% and 99.17%, *Minor* at 95.32% and 92.64%, and *Trend* at 97.62% and 97.69%, respectively. The model demonstrates its exceptional capability in identifying *Square* patterns with perfect *Precision* and an impressive *Recall* of 99.97%. Furthermore, the *F1-score*s for most patterns are impressive, with *Normal* achieving 99.14% and *Square* reaching an outstanding 99.98%. Despite these achievements, the classifier shows a relatively lower *F1-score* of 84.75% for the *Drift* pattern compared to other anomaly patterns. As indicated by the confusion matrix in Fig. 7, the classifier often incorrectly predicts between *Trend* and *Drift* patterns. This issue may stem from the similarity in features extracted from the time and

Fig. 7 Confusion matrices of the constructed classification model

Table 5 Performance assessment results for training data

Anomaly patterns:	*Precision*:	*Recall*:	*Accuracy*:	*F1-score*:
Normal	0.9915	0.9915	0.9905	0.9915
Minor	0.9545	0.9290	0.9925	0.9416
Outlier	0.8640	0.9385	0.9954	0.8997
Square	1.0000	1.0000	1.0000	1.0000
Trend	0.9769	0.9782	0.9911	0.9775
Drift	0.8562	0.8448	0.9911	0.8504

Table 6 Performance assessment results for training data

Anomaly patterns:	*Precision*:	*Recall*:	*Accuracy*:	*F1-score*:
Normal	0.9900	0.9917	0.9897	0.9908
Minor	0.9442	0.9312	0.9916	0.9376
Outlier	0.9186	0.9294	0.9959	0.9240
Square	1.0000	0.9976	0.9997	0.9988
Trend	0.9824	0.9777	0.9922	0.9800
Drift	0.8495	0.8778	0.9922	0.8634

Table 7 Performance assessment results for training data

Anomaly patterns:	*Precision*:	*Recall*:	*Accuracy*:	*F1-score*:
Normal	0.9906	0.9922	0.9903	0.9914
Minor	0.9570	0.9082	0.9919	0.9319
Outlier	0.8864	0.9630	0.9959	0.9231
Square	1.0000	1.0000	1.0000	1.0000
Trend	0.9670	0.9700	0.9875	0.9685
Drift	0.8304	0.8158	0.9875	0.8230

Table 8 Performance assessment results for training data

Anomaly patterns:	*Precision*:	*Recall*:	*Accuracy*:	*F1-score*:
Normal	0.9912	0.9917	0.9904	0.9914
Minor	0.9532	0.9264	0.9923	0.9396
Outlier	0.8767	0.9409	0.9956	0.9077
Square	1.0000	0.9997	1.0000	0.9998
Trend	0.9762	0.9769	0.9907	0.9766
Drift	0.8508	0.8443	0.9907	0.8475

Table 9 Summary of the methods submitted to the IPC-SHM

Rank:	*Title*:	*Data Scaling Methods*:	*Feature Engineering Methods*:	*Classification Algorithm*:	*Reference*:
1	Data anomaly detection for structural health monitoring of bridges using shapelet transform	Shapelet Transform	Shapelet Transform	Random Forest	Arul and Kareem (2022)
2	SHM data anomaly classification using machine learning strategies: A comparative study	Standardization	Feature Selection	CNN-based ensemble method	Chou *et al*. (2022)
3	Data anomaly detection for structural health monitoring using a combination network of GANomaly and CNN	Gramian Angular Field Encoding & GANomaly	-	CNN	Liu *et al*. (2022)
4	**“THIS PAPER”**	**Standardization**	**Time and Frequency Domain Feature Selection & PCA**	**DNN**	-
5	Convolutional neural network-based data anomaly detection considering class imbalance with limited data	Gray Scaling	-	CNN	Du *et al*. (2022)
6	A semi-supervised interpretable machine learning framework for sensor fault detection	Standardization	Time and Frequency Domain Feature Selection	XGBoost & SHAP	Martakis *et al*. (2022)
7	Detection of multi-type data anomaly for structural health monitoring using pattern recognition neural network	-	Time Domain Feature Selection	Pattern Recognition Neural Network	Gao *et al*. (2022)
8	CNN based data anomaly detection using multi-channel imagery for structural health monitoring	-	Time, Spectrogram Channel, and Probability Density Function based Feature Selection	CNN	Shajihan *et al*. (2022)
9	Data abnormal detection using bidirectional long-short neural network combined with artificial experience	-	Time and Frequency Domain Feature Selection	RNN	Yang *et al*. (2022)
10	Data anomaly detection using ensemble deep convolutional neural networks	-	-	-	-

frequency domains used in training. To enhance the classifier’s performance, augmenting the dataset with additional data for the *Drift* pattern may be crucial. Currently, the *Drift* pattern constitutes only about 2.40% of the dataset. By addressing this issue of data imbalance and scarcity, we expect an improvement in the classifier’s ability to accurately identify and categorize anomalies.

4. Review of the IPC-SHM

Totally thirty methods were submitted to the IPC-SHM (Bao *et al*. 2021), and ten methods were awarded as summarized in Table 9. Our model achieved fourth place out of thirty methods in terms of model accuracy, presentation, and report quality. In comparing our approach to those of other participants, we found that each method has its own strengths; thus, we summarize the findings from the competition.

First, we observed that data scaling methods, such as standardization, were frequently used among the top performers. This suggests that the data scaling can be a crucial factor in achieving high model accuracy, particularly when working with data that has a wide range of values or varying distributions. Scaling the data can help to ensure that all data are treated equally and can prevent certain features from dominating the model’s predictions.

Second, we found that feature engineering methods were also necessary for achieving good results, despite not being required in many deep learning algorithms, such as CNN. This may be due to the fact that the given data was not sufficient to train the features, emphasizing the importance of feature engineering. In the data analysis competition, many participants used techniques, such as time and frequency domain feature selection, PCA, or transformation, to help identify and highlight the most important features for analysis. The use of feature engineering techniques may have been particularly important where the data is complex and high-dimensional.

Finally, while many participants utilized CNN, we found that simpler models, such as random forest and DNN, were also effective in achieving high model quality. These findings suggest that while CNN is a powerful algorithm, simpler models may be more appropriate, and that careful consideration of feature engineering and data scaling methods can be crucial. Moreover, simpler models may be easier to interpret and more efficient to train than complex models, making them a more practical choice.

5. Conclusions

This study proposes a classification method to detect anomaly patterns in the acceleration data of long-span cable-stayed bridges. The proposed method involves several steps, including the scaling of raw acceleration data, the estimation of informative features in the time and frequency domains, and the extraction of principal components as the linear combinations of these features. The deep neural network model was employed for classification, and the results showed that the proposed method accurately detected the data anomalies. The practical applicability of the proposed method was demonstrated using actual data from a long-span cable-stayed bridge in China, and the results showed its effectiveness in real-world SHM systems of civil infrastructures. Although this approach achieves high accuracy, it encounters challenges in distinguishing between *Trend* and *Drift* patterns. This issue highlights the need for further data augmentation, especially for underrepresented patterns such as *Drift*, which currently comprises only 2.40% of the dataset. Addressing this data imbalance and scarcity is anticipated to enhance the classifier's ability to accurately identify and categorize anomalies.

Additionally, this paper provides a comprehensive summary of the key findings and insights from the IPC-SHM competition. The analysis identified several critical factors that contribute to the success of high-performing models. Data scaling, such as standardization, emerged as a crucial element, especially when dealing with datasets that exhibit a wide range of values or varying distributions. This process ensures the reasonable treatment of all data and mitigates the dominance of specific features in model predictions. Furthermore, feature engineering, which is often overlooked in deep learning frameworks such as CNN, could be vital for model success. This highlights the importance of techniques such as time and frequency domain feature selection, PCA, or transformation in enhancing feature representation. Interestingly, not only advanced algorithms such CNN, RNN but also simpler models such as random forest and DNN were effective in achieving high model quality. This indicates that the choice of model should be tailored to the specific dataset and scenario, with careful consideration given to feature engineering, data scaling, and model complexity. These insights collectively offer a nuanced understanding of the various approaches in model optimization, proving invaluable for future research and application in the field.

Acknowledgments

This work was supported by the National Research Foundation of Korea (NRF) grant funded by the Korea government (MSIT) (2021R1C1C2008770), and the authors would like to thank the organizers of the 1st International Project Competition for SHM (IPC-SHM, 2020) for providing the invaluable data used in this paper.

References

Abdi, H. and Williams, L.J. (2010), "Principal component analysis", *Wiley Interdiscipl. Reviews*: *Computat. Statist*., **2**(4), 433-459. https://doi.org/10.1002/wics.101

Altın, C. and Er, O. (2016), "Comparison of different time and frequency domain feature extraction methods on elbow gesture's EMG", *Eur. J. Interdiscipl. Stud*., **2**(3), 35-44. https://doi.org/10.26417/ejis.v2i3.p35-44

Ansari, F. (2007), "Practical implementation of optical fiber sensors in civil structural health monitoring", *J. Intell. Mater. Syst. Struct*., **18**(8), 879-889. https://doi.org/10.1177/1045389X06075760

Arangio, S. and Bontempi, F. (2015), "Structural health monitoring of a cable-stayed bridge with Bayesian neural networks", *Struct. Infrastr. Eng*., **11**(4), 575-587. https://doi.org/10.1080/15732479.2014.951867

Arul, M. and Kareem, A. (2022), "Data anomaly detection for structural health monitoring of bridges using shapelet transform", *Smart Struct. Syst*., *Int. J*., **29**(1), 93-103. https://doi.org/10.12989/sss.2022.29.1.093

Bakar, Z.A., Mohemad, R., Ahmad, A. and Deris, M.M. (2006), "A comparative study for outlier detection techniques in data mining", In: *2006 IEEE Conference on Cybernetics and Intelligent Systems*, Bangkok, Thailand, June.

Bao, Y., Tang, Z., Li, H. and Zhang, Y. (2019), "Computer vision and deep learning–based data anomaly detection method for structural health monitoring", *Struct. Health Monitor*., **18**(2), 401-421. https://doi.org/10.1177/1475921718757405

Bao, Y., Li, J., Nagayama, T., Xu, Y., Spencer Jr, B.F. and Li, H. (2021), "The 1st International Project Competition for Structural Health Monitoring (IPC-SHM, 2020): A summary and benchmark problem", *Struct. Health Monitor*., **20**(4), 2229-2239. https://doi.org/10.1177/14759217211006485

Chang, C., Chou, J., Tan, P. and Wang, L. (2017), "A sensor fault detection strategy for structural health monitoring systems", *Smart Struct. Syst*., *Int. J*., **20**(1), 43-52. https://doi.org/10.12989/sss.2017.20.1.043

Chou, J.-Y., Fu, Y., Huang, S.-K. and Chang, C.-M. (2022), "SHM data anomaly classification using machine learning strategies: A comparative study", *Smart Struct. Syst*., *Int. J*., **29**(1), 77-91. https://doi.org/10.12989/sss.2022.29.1.077

Du, Y., Li, L., Hou, R., Wang, X., Tian, W. and Xia, Y. (2022), "Convolutional neural network-based data anomaly detection considering class imbalance with limited data", *Smart Struct. Syst*., *Int. J*., **29**(1), 63-75. https://doi.org/10.12989/sss.2022.29.1.063

Fleming, J.F. and Egeseli, E.A. (1980), "Dynamic behaviour of a cable-stayed bridge", *Earthq. Eng. Struct. Dyn*., **8**(1), 1-16. https://doi.org/10.1002/eqe.4290080102

Fu, Y., Peng, C., Gomez, F., Narazaki, Y. and Spencer Jr., B. (2019), "Sensor fault management techniques for wireless smart sensor networks in structural health monitoring", *Struct. Control Health Monitor*., **26**, e2362. https://doi.org/10.1002/stc.2362

Gao, K., Chen, Z.-D., Weng, S., Zhu, H.-P. and Wu, L.-Y. (2022), "Detection of multi-type data anomaly for structural health monitoring using pattern recognition neural network", *Smart Struct. Syst*., *Int. J*., **29**(1), 129-140. https://doi.org/10.12989/sss.2022.29.1.129

Khalid, S., Khalil, T. and Nasreen, S. (2014), "A survey of feature selection and feature extraction techniques in machine learning", In: *2014 Science and Information Conference*, London, UK, October. https://doi.org/10.1109/SAI.2014.6918213

Kullaa, J. (2011), "Distinguishing between sensor fault, structural damage, and environmental or operational effects in structural health monitoring", *Mech. Syst. Signal Process*., **25**(8), 2976-

2989. https://doi.org/10.1016/j.ymssp.2011.05.017

Kullaa, J. (2013), "Detection, identification, and quantification of sensor fault in a sensor network", *Mech. Syst. Signal Process.*, **40**(1), 208-221. https://doi.org/10.1016/j.ymssp.2013.05.007

Kung, S.Y. and Diamantaras, K.I. (1990), "A neural network learning algorithm for adaptive principal component extraction (APEX)", In: *International Conference on Acoustics, Speech, and Signal Processing*, Albuquerque, NM, USA, August.

Lee, J., Lee, K.-C., Cho, S. and Sim, S.H. (2017), "Computer vision-based structural displacement measurement robust to light-induced image degradation for in-service bridges", *Sensors*, **17**(10), 2317. https://doi.org/10.3390/s17102317

Lee, J., Lee, K.-C. and Lee, Y.-J. (2018), "Long-term deflection prediction from computer vision-measured data history for high-speed railway bridges", *Sensors*, **18**(5), 1488. https://doi.org/10.3390/s18051488

Lee, J., Lee, K.-C., Sim, S.H., Lee, J. and Lee, Y.-J. (2019), "Bayesian Prediction of Pre-Stressed Concrete Bridge Deflection Using Finite Element Analysis", *Sensors*, **19**(22), 4956. https://doi.org/10.3390/s19224956

Lee, J., Jeong, S., Lee, J., Sim, S.-H., Lee, K.-C. and Lee, Y.-J. (2022), "Sensor data-based probabilistic monitoring of time-history deflections of railway bridges induced by high-speed trains", *Struct. Health Monitor.*, **21**(6), 2518-2530. https://doi.org/10.1177/14759217211063424

Li, H.N., Ren, L., Jia, Z.G., Yi, T.H. and Li, D.S. (2016), "State-of-the-art in structural health monitoring of large and complex civil infrastructures", *J. Civil Struct. Health Monitor.*, **6**(1), 3-16. https://doi.org/10.1007/s13349-015-0108-9

Liu, G., Niu, Y., Zhao, W., Duan, Y. and Shu, J. (2022), "Data anomaly detection for structural health monitoring using a combination network of GANomaly and CNN", *Smart Struct. Syst., Int. J.*, **29**(1), 53-62. https://doi.org/10.12989/sss.2022.29.1.053

Manson, G. (2002), "Identifying damage sensitive, environment insensitive features for damage detection", *Proceedings of the 3rd International Conference on Identification in Engineering Systems*, Swansea, Wales, UK, April

Martakis, P., Movsessian, A., Reuland, Y., Pai, S.G.S., Quqa, S., Cava, D.G., Tcherniak, D. and Chatzi, E. (2022), "A semi-supervised interpretable machine learning framework for sensor fault detection", *Smart Struct. Syst., Int. J.*, **29**(1), 251-266. https://doi.org/10.12989/sss.2022.29.1.251

Mei, H., Haider, M.F., Joseph, R., Migot, A. and Giurgiutiu, V. (2019), "Recent advances in piezoelectric wafer active sensors for structural health monitoring applications", *Sensors*, **19**(2), 383. https://doi.org/10.3390/s19020383

Ni, F., Zhang, J. and Noori, M.N. (2020), "Deep learning for data anomaly detection and data compression of a long-span suspension bridge", *Comput.-Aided Civil Infrastr. Eng.*, **35**, 685-700. https://doi.org/10.1111/mice.12528

Oh, B.K., Glisic, B., Kim, Y. and Park, H.S. (2020), "Convolutional neural network–based data recovery method for structural health monitoring", *Struct. Health Monitor.*, **19**(6), 1821-1838. https://doi.org/10.1177/1475921719897571

Posenato, D., Kripakaran, P., Inaudi, D. and Smith, I.F. (2010), "Methodologies for model-free data interpretation of civil engineering structures", *Comput. Struct.*, **88**(7-8), 467-482. https://doi.org/10.1016/j.compstruc.2010.01.001

Shajihan, S.A.V., Wang, S., Zhai, G. and Spencer, B.F.J. (2022), "CNN based data anomaly detection using multi-channel imagery for structural health monitoring", *Smart Struct. Syst., Int. J.*, **29**(1), 181-193. https://doi.org/10.12989/sss.2022.29.1.181

Smarsly, K. and Law, K.H. (2014), "Decentralized fault detection and isolation in wireless structural health monitoring systems using analytical redundancy", *Adv. Eng. Software*, **73**, 1-10. https://doi.org/10.1016/j.advengsoft.2014.02.005

Sohn, H., Farrar, C.R., Hunter, N.F. and Worden, K. (2001), "Structural health monitoring using statistical pattern recognition techniques", *J. Dyn. Syst. Measur. Control*, **123**(4), 706-711. https://doi.org/10.1115/1.1410933

Sohn, H., Farrar, C.R., Hemez, F.M., Shunk, D.D., Stinemates, D.W., Nadler, B.R. and Czarnecki, J.J. (2003), "A review of structural health monitoring literature: 1996–2001", *Los Alamos National Laboratory*, USA.

Son, H., Yoon, C., Kim, Y., Jang, Y., Tran, L. V., Kim, S.-E., Kim, D.J. and Park, J. (2022), "Damaged cable detection with statistical analysis, clustering, and deep learning models", *Smart Struct. Syst., Int. J.*, **29**(1), 17-28. https://doi.org/10.12989/sss.2022.29.1.017

Straser, E.G., Kiremidjian, A.S., Meng, T.H. and Redlefsen, L. (1998), "A modular, wireless network platform for monitoring structures", *Proceedings of the International Modal Analysis Conference*, Santa Barbara, CA, USA, February.

Xin, J., Zhou, J., Yang, S.X., Li, X. and Wang, Y. (2018), "Bridge structure deformation prediction based on GNSS data using Kalman-ARIMA-GARCH model", *Sensors*, **18**(1), 298. https://doi.org/10.3390/s18010298E

Xu, X., Huang, Q., Ren, Y., Zhao, D.-Y. and Yang, J. (2019), "Sensor fault diagnosis for bridge monitoring system using similarity of symmetric responses", *Smart Struct. Syst., Int. J.*, **23**(3), 279-293. https://doi.org/10.12989/sss.2019.23.3.279

Yang, K., Jiang, H., Ding, Y., Wang, M. and Wan, C. (2022), "Data abnormal detection using bidirectional long-short neural network combined with artificial experience", *Smart Struct. Syst., Int. J.*, **29**(1), 117-127. https://doi.org/10.12989/sss.2022.29.1.117

Yi, T.H., Li, H.N., Song, G. and Guo, Q. (2016), "Detection of shifts in GPS measurements for a long span bridge using CUSUM chart", *Int. J. Struct. Stab. Dyn.*, **16**(04), 1640024. https://doi.org/10.1142/S0219455416400241

Yi, T.H., Huang, H.B. and Li, H.N. (2017), "Development of sensor validation methodologies for structural health monitoring: A comprehensive review", *Measurement*, **109**, 200-214. https://doi.org/10.1016/j.measurement.2017.05.064

Zebari, R., Abdulazeez, A., Zeebaree, D., Zebari, D. and Saeed, J. (2020), "A comprehensive review of dimensionality reduction techniques for feature selection and feature extraction", *J. Appl. Sci. Technol. Trends*, **1**(2), 56-70. https://doi.org/10.38094/jastt1224

HJ