**Smart Structures and Systems**, *Vol. 29, No. 1 (2022) 53-62*
https://doi.org/10.12989/sss.2022.29.1.053


# Data anomaly detection for structural health monitoring using a combination network of GANomaly and CNN

Gaoyang Liu [†1,2], Yanbo Niu [†1,3], Weijian Zhao [1], Yuanfeng Duan [1] and Jiangpeng Shu[*1]

[1] *College of Civil Engineering and Architecture, Zhejiang University, Hangzhou 310058, P.R. China*
[2] *Center for Balance Architecture, Zhejiang University, Hangzhou 310058, P.R. China*
[3] *The Architectural Design & Research Institute of Zhejiang University Co. Ltd., Hangzhou 310058, P.R. China*

(*Received April 5, 2021, Revised July 28, 2021, Accepted July 29, 2021*)

**Abstract.** The deployment of advanced structural health monitoring (SHM) systems in large-scale civil structures collects large amounts of data. Note that these data may contain multiple types of anomalies (e.g., missing, minor, outlier, etc.) caused by harsh environment, sensor faults, transfer omission and other factors. These anomalies seriously affect the evaluation of structural performance. Therefore, the effective analysis and mining of SHM data is an extremely important task. Inspired by the deep learning paradigm, this study develops a novel generative adversarial network (GAN) and convolutional neural network (CNN)-based data anomaly detection approach for SHM. The framework of the proposed approach includes three modules : (a) A three-channel input is established based on fast Fourier transform (FFT) and Gramian angular field (GAF) method; (b) A GANomaly is introduced and trained to extract features from normal samples alone for class-imbalanced problems; (c) Based on the output of GANomaly, a CNN is employed to distinguish the types of anomalies. In addition, a dataset-oriented method (i.e., multistage sampling) is adopted to obtain the optimal sampling ratios between all different samples. The proposed approach is tested with acceleration data from an SHM system of a long-span bridge. The results show that the proposed approach has a higher accuracy in detecting the multi-pattern anomalies of SHM data.

**Keywords:** convolutional neural network; data anomaly detection; generative adversarial network; Gramian angular field; long-span bridge; structural health monitoring

## 1. Introduction

With the rapid development of advanced sensing techniques, a substantial number of large-scale civil infrastructures are equipped with sophisticated structural health monitoring (SHM) systems (Gatti 2019, Heinlein *et al*. 2019, Li and Ou 2016, Xu 2018), aiming to ensure the safe and stability of the structure. After long-term monitoring of structures, these SHM systems have collected a large volume of data including wind, temperature, displacement, strain, acceleration and so on. Such information can be employed to evaluate the structural performance, judge whether the structure is damaged, and predict the future behavior of the structure. However, it is noteworthy that massive data in SHM systems generally present various types of anomalies caused by a series of factors like environment interference, sensor faults and system malfunctions (Huang *et al*. 2016, Park *et al*. 2017, Yuen and Mu 2012). These anomalies have a great impact on the structural performance assessment, because it is difficult to distinguish whether these anomalies are caused by structural damage or by incorrect data. Therefore, to overcome this problem, it is essential to identify and remove anomalies from SHM data, which requires data processing methods to have high recognition accuracy and high calculation efficiency.

To date, many development traditional anomaly detection methods have been utilized in the field of SHM (Kerschen *et al*. 2005, Park *et al*. 2018, Yang and Nagarajaiah 2014). For instance, Yi *et al*. (2016) employed a statistical process control approach, called as the cumulative sum (CUSUM) chart, for detecting persistent shifts in global positioning system (GPS) monitoring. Hernandez-Garcia and Masr (2014) used three latent-variable methods, i.e., principal component analysis (PCA), independent component analysis (ICA), and modified independent component analysis (MICA), for identifying faulty sensors in a three-dimensional truss and an actual cable-supported bridge measurement. Gul and Catbas (2009) proposed a modified Mahalanobis distance-based outlier detection method for detecting different types of structural changes on different test structures. Rao *et al*. (2014) put forward a combined method based on the PCA and an adaptive differential evolution method to improve the performance of sensor fault isolation, and numerical investigations were carried out to verify the effectiveness of the proposed method. Ding *et al*. (2016) proposed a latent correlation-based anomaly detection approach that effectively identifies potential anomalies from a large number of correlative isomerous monitoring data. Note that

*Corresponding author, Ph.D., Professor,
E-mail: jpeshu@zju.edu.cn
† These authors contributed equally to this work and should be considered co-first authors



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

the traditional anomaly detection approaches are inefficient. Meanwhile, the uniqueness and the robustness of the features extracted by the traditional approaches are not good enough. The large variations of extracted features from massive SHM data make the traditional data anomaly detection approaches prone to be over-processed or under-processed.

Nowadays, artificial intelligence (AI) techniques have been increasingly applied for solving some typical problems such as image recognition (He and Li 2019), audio recognition (Kong *et al*. 2020), noise reduction (Isogawa *et al*. 2018), data defect (Tian *et al*. 2020), *et al*. Compared with the traditional methods mentioned above, AI-based method is capable of detecting multiple anomaly data fast and automatically with a higher accuracy. Therefore, it has become a popular research topic in SHM. In particular, deep learning artificial neural networks offer the advantage of selecting useful features, which have been paid more attention to in the field of SHM. For example, Xu *et al*. (2019) developed a modified fusion convolutional neural network (CNN) method for identifying fatigue cracks in bridge girders, and proved that the identification errors of the proposed method are smaller than those of the regular CNN. Bao *et al*. (2019) put forward a novel data anomaly detection approach based on computer vision and deep neural network. In this approach, the original time series signals are first transformed into image vectors, and these image vectors were then input into a deep neural network for training and classifying data pieces. Experimental verification was that the proposed approach can effectively detect multi-pattern anomalies of SHM data and obtain a moderate detection accuracy of 87.0%. Following this, Tang *et al*. (2019) proposed a CNN-based data anomaly detection approach that learned and fused multiple graphical features based on time and frequency response information of SHM data. Their study validated that the proposed CNN-based approach can detect anomaly from SHM data with a mean overall accuracy of 93.5%. Subsequently, Ni *et al*. (2020) designed a one-dimensional CNN for extracting the essential features of input signals and detect the anomaly of SHM data. They demonstrated that the proposed approach is more effective than the approach that converts time series signals into deep learning-based image vectors. It is noteworthy that imbalanced time series are common in practical engineering, where the number of normal samples is far larger than that of abnormal samples. CNN is difficult to obtain high classification accuracies for class-imbalanced problems, because it is based on a class-balanced hypothesis (Yin and Gai 2015). The development of generative adversarial network (GAN) offers a new perspective for overcoming this problem (Jiang *et al*. 2019). The basic principle of GAN is that it generates prototypical samples via a generator with random samples that satisfy a definite distribution (e.g., Gaussian distribution) (Jiang *et al*. 2019). Some GAN-based networks, such as AnoGAN (Thomas *et al*. 2017), BiGAN (Mutlu and Alpaydın 2020), and GANomaly (Akcay *et al*. 2019), have been employed in the field of image recognition. These approaches can distinguish abnormal images based on the distribution difference between the normal and abnormal, and are suitable for processing imbalanced samples. However, in the field of SHM, GAN-based approaches for anomaly detection are rarely seen.

Inspired by the above research, this study developed a new combined GANomaly and CNN based method for detecting multi-pattern anomalies from a large amount of SHM data. In particular, this article encoded time series into two-dimensional image via Gramian Angular Field (GAF) (Yang *et al*. 2020). In addition, it is remarkable that time-domain features may not be sufficient to distinguish some types of anomalies well, such as “Trend” and “Drift” (Tang *et al*. 2019). Hence the frequency domain characteristics of the original signal were introduced in this article. Finally, the multi-pattern anomalies of SHM data from an actual long-span bridge are detected based on the proposed GANomaly-CNN model.

## 2. Proposed network structure

The purpose of this study is to detect multiple types of anomalies in SHM data. In particular, class-imbalance is a common problem in SHM data. To address this problem and obtain accurate detection results, a novel data anomaly detection method is developed. This method mainly includes three modules (as presented in Fig. 1): (1) A three-channel input including time response image, frequency response image and GAF image is built; (2) A GANomaly network is introduced to solve the class-imbalanced issue; (3) Different types of anomalies are distinguished by a designed CNN network. In the first module, to better distinguish multiple anomaly types, extra frequency domain features are introduced via using fast Fourier transform (FFT) method. Meanwhile, the original time series are encoded into two-dimensional vector images by using GAF method. In addition, a two-stage of hybrid sampling strategy called MUS1 and MUS2 is adopted. In the second module, only normal samples are considered in the training stage. All different types of anomalies fall into one class. Such a class of samples can be discriminated by higher anomaly scores compared to normal samples in the testing stage. In the third module, a training set consisting of labeled samples is fed into a CNN network, and then multiple types of anomalies can be detected via the trained CNN.

### *2.1 Three-channel input establishment and multistage sampling*

In this study, abnormal samples include six types, i.e., missing, minor, outlier, square, trend and drift. Considering that it may be difficult to distinguish multiple anomaly types only by using time response information. A three-channel input including time response image, frequency response image and GAF image is established. Particularly, GAF encodes time series into image by polar coordinates based matrix and it can preserve absolute temporal correlation (Barra *et al*. 2020). To construct the GAF image, the original time series $X = \{x_1,\ x_2,\ \cdots,\ x_n\}$ needs to be rescaled via the following equation

$$\tilde{x}_i = \frac{(x_i - max(X)) + (x_i - min(X))}{max(X) - min(X)} \tag{1}$$

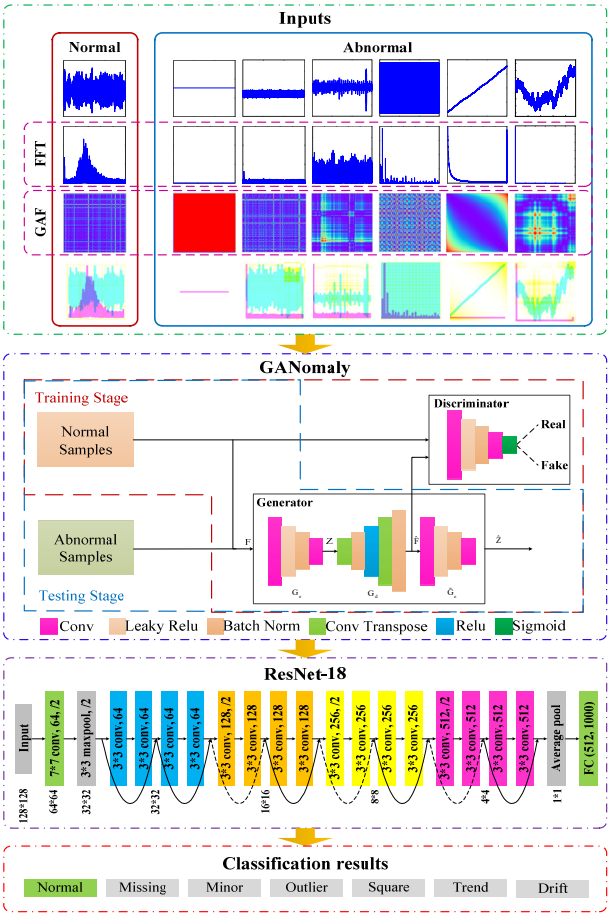


Fig. 1 Framework of the proposed data anomaly detection method

The rescaled series is represented by $\tilde{X} = \{\tilde{x}_1, \tilde{x}_2, \cdots, \tilde{x}_n\}$. Then, angular cosine and the time stamp are used to encode the rescaled data into polar coordinates, and the equation is given by

$$\begin{cases} \theta_i = arccos(\tilde{x}_i), & \tilde{x}_i \in \tilde{X}, \\ r_i = \dfrac{t_i}{N}, & \text{with } t_i \le N \end{cases} \tag{2}$$

where $t_i$ is the time stamp and $N$ is a constant factor to regularize the span of the polar coordinate system.

GAF can generate two images by different equations, i.e., Gramian Angular Summation Field (GASF) and Gramian Angular Difference Field (GADF), as follows

$$\text{GASF} = \begin{bmatrix} cos(\theta_1+\theta_1) & \cdots & cos(\theta_1+\theta_n) \\ cos(\theta_2+\theta_1) & \cdots & cos(\theta_2+\theta_n) \\ \vdots & \ddots & \vdots \\ cos(\theta_n+\theta_1) & \cdots & cos(\theta_n+\theta_n) \end{bmatrix} = \tilde{X}' \cdot \tilde{X} - \sqrt{I-\tilde{X}^2}' \cdot \sqrt{I-\tilde{X}^2} \tag{3}$$

$$\text{GADF} = \begin{bmatrix} sin(\theta_1+\theta_1) & \cdots & sin(\theta_1+\theta_n) \\ sin(\theta_2+\theta_1) & \cdots & sin(\theta_2+\theta_n) \\ \vdots & \ddots & \vdots \\ sin(\theta_n+\theta_1) & \cdots & sin(\theta_n+\theta_n) \end{bmatrix} = \sqrt{I-\tilde{X}^2}' \cdot \tilde{X} - \tilde{X}' \cdot \sqrt{I-\tilde{X}^2} \tag{4}$$

Fig. 2 shows the process for transforming a time series to the GADF and GASF images using Eqs. (3) and (4). Note that the derived matrix via Eqs. (3) and (4) actually represents a heat-map, whose values range from 0 (blue) to 1 (red). In this study, the grayscale images are used as input to the network.

FFT method is employed to obtain frequency response information. Then, the frequency response signal is converted into grayscale image and input to the network. Thus far, a three-channel input is derived.

In addition, to address class-imbalance issue in SHM data, a multistage sampling method is introduced. In anomaly detection, normal data usually plays a dominant role in the whole dataset. The increasing ratio of examples between majority and minority classes as well as the number of minority classes had a negative effect on performance of the resulting classifiers. Hybrid sampling method can simply rebalance the datasets by taking constant numbers of samples from different classes. Nevertheless, the number of samples from different classes being equal did not always correspond to the optimal results. Therefore, a two-stage of hybrid sampling called MUS1 and MUS2, is adopted to effectively relieve severe imbalance of training samples and achieve the optimal sampling ratios.

In the MUS1 stage, the sample proportions are equal within minority classes and equal within majority classes by hybrid sampling (Xia *et al*. 2019). The fractions of majority classes $\mu_{\text{maj}}$ and minority classes $\mu_{\text{min}}$ are defined by

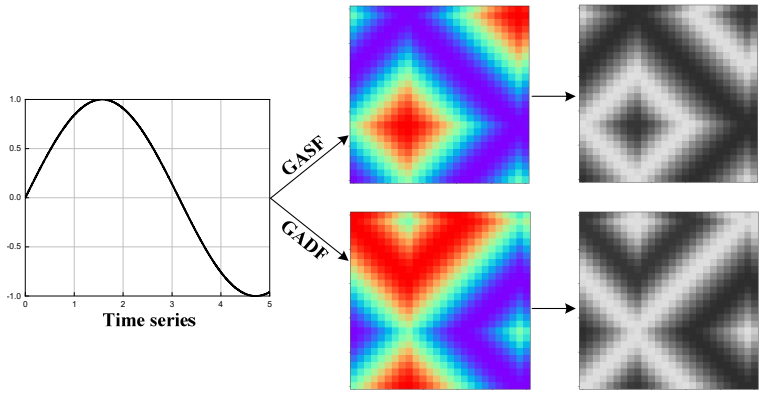


Fig. 2 An example of encoding time series as image

$$\begin{cases}\mu_{\min} = \dfrac{|\{i = \{1,\ 2,\ \cdots,\ N\}:\ C_i \text{ is minority}\}|}{N} \\ \mu_{\text{maj}} = \dfrac{|\{i = \{1,\ 2,\ \cdots,\ N\}:\ C_i \text{ is majority}\}|}{N}\end{cases} \tag{5}$$

where $N$ is the total number of classes, and $C_i$ is a set of the examples in class $i$. The sample proportions of majority classes $p_{\text{maj}}$ and minority classes $p_{\min}$ are

$$\begin{cases}p_{\text{maj}} = \dfrac{\mu_{\text{maj}}}{\sum_{i\in\{1,\ldots,N\}:\ C_i \text{ is majority}} \mu_{maj}^{i} + \sum_{i\in\{1,\ldots,N\}:\ C_i \text{ is minority}}^{\Sigma} \mu_{min}^{i}} \\ p_{\min} = \dfrac{\mu_{\min}}{\sum_{i\in\{1,\ldots,N\}:\ C_i \text{ is majority}} \mu_{maj}^{i} + \sum_{i\in\{1,\ldots,N\}:\ C_i \text{ is minority}}^{\Sigma} \mu_{min}^{i}}\end{cases} \tag{6}$$

After the first stage, the second stage is applied to obtain the MUS2 dataset that has a smaller imbalance ratio. The hybrid sampling process is repeated to make the sample proportions in the multiple classes able to be interpolated linearly. The difference between consecutive pairs of classes is of constant order by the value of sample proportion. The constant $\rho$ was defined by the following equation, and it was used in MUS2 as listed

$$\rho = \frac{p_{\text{maj}} - p_{min}}{N - 1} \tag{7}$$

Eq. (6) is substituted into Eq. (7), the proportion constant $\rho$ can be obtained. The sample proportion $p_i$ of class $i$ in reverse order is given by

$$p_i = p_{min} \tag{8}$$

### 2.2 GANomaly

For class-imbalance problem, the dataset-oriented method (i.e., multistage sampling) is an available scheme. Beyond this, GANomaly has also been proved to be an effective method for processing class-imbalance problem (Jiang *et al*. 2019). Therefore, it is introduced to detect abnormal samples from normal samples involving imbalanced dataset, which consists of a generator and a discriminator. An encoder-decoder-encoder three-sub-network is employed in the design of the generator. Two encoders $G_e$ and $\tilde{G}_e$ learn to get the input representations of original images $X$ and regenerated images $\hat{X}$ respectively. The decoder $G_d$ tries to reconstruct $\hat{X}$ at the same time. The corresponding process of the generator is the following:

a) With the use of convolutional layers followed by batch-norm and leaky ReLU activation, respectively, the encoder $G_e$ downscales $x$ into latent vector $z$ via $z = G_e(x)$;
b) The decoder $G_d$ adopts convolutional transpose layers, ReLU activation layers and batch-norm layers. This method upscales the latent vector $z$ to reconstructs the output, defined as $\hat{x}$, where $\hat{x} = G_d(z)$;
c) The last encoder $\tilde{G}_e$ has the same architecture as $G_e$ but with different parametrization. $\tilde{G}_e$ downscales $\hat{x}$ to obtain its feature representation $\hat{z}$ via $z = \tilde{G}_e(\hat{x})$.
d) The discriminator $D$ is based on the standard discriminator introduced in DCGAN (Radford *et al*. 2016). It is applied to distinguish whether input data is real or fake.

The objective function $L$ combines three loss functions

$$L = w_{adv}L_{adv} + w_{con}L_{con} + w_{enc}L_{enc} \tag{9}$$

$$\begin{cases}L_{adv} = \|f(x) - f(\hat{x})\|_2 \\ L_{con} = \|x - \hat{x}\|_1 \\ L_{enc} = \|z - \hat{z}\|_2\end{cases} \tag{10}$$

where $L_{adv}$, $L_{con}$, $L_{enc}$ are adversarial loss, contextual loss and encoder loss respectively and $f$ is a function that outputs an intermediate layer of the discriminator $D$.

During the testing stage, the encoder loss $L_{enc}$ is used for scoring the abnormality of a given image. The anomaly score $A(x)$ for a test sample $x$ is defined as

$$A(x) = \left\|G_E(x) - E\left(G(x)\right)\right\|_2 \tag{11}$$

For an anomaly sample $A(x)$ will be much higher compared to a normal sample, which help us to identify the anomalies. The procedure above should be enough to solve anomaly detection in the scenario of unsupervised learning, which aims at distinguishing between normal and abnormal samples. For a supervised learning task, the probability of a normal sample $p(x)$ is given by a sigmoid function (Han and Moraga 1995).

$$p(x) = 1 - \frac{1}{1 + e^{-2\gamma A(x)}} \tag{12}$$

where $\gamma$ is a parameter for the slope of the function.

### 2.3 CNN

A residual neural network, i.e., ResNet-18 (He *et al*. 2016), is established and employed for the automated classification of different types of anomalies. This network consists of five convolutional layers (see Fig. 1 for further details). The classification layers consist of the full connection layer and the softmax layer, and each classification label corresponds to one output type.

The probability of the normal class obtained by CNN should be multiplied by $p(x)$ in Eq. (12), which could relieve over-fitting to the majority in an imbalanced dataset.

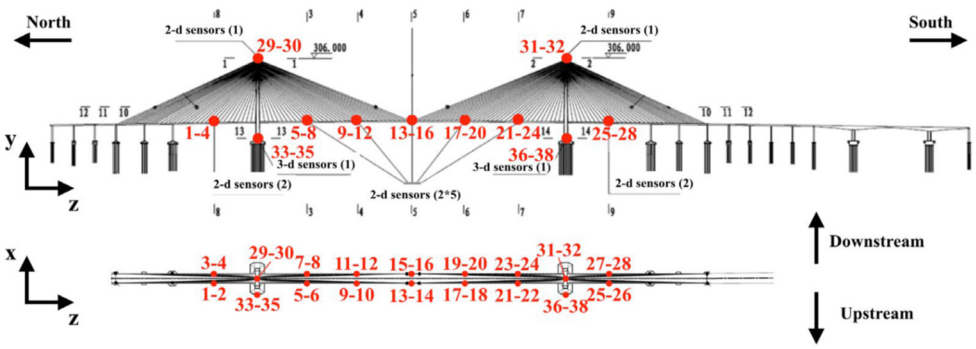


Fig. 3 The position of accelerometers on the deck and towers

Table 1 Description of each data pattern

| No. | Anomaly patterns | Description |
|---|---|---|
| 0 | Normal | The time response is normal oscillation curve; frequency response is peak-like (may differ between bridges) |
| 1 | Missing | Most/all of the time response is missing, which makes the time and frequency response zero |
| 2 | Minor | Relative to normal sensor data, the amplitude is very small in the time domain |
| 3 | Outlier | One or more outliers appear in the time response |
| 4 | Square | The time response is like a square wave |
| 5 | Trend | The data has an obvious trend in the time domain and has an obvious peak value in the frequency domain |
| 6 | Drift | The vibration response is non-stationary, with random drift |

## 3. Application example

### *3.1 Dataset description*

A dataset that consists of one-month of acceleration data for a long-span cable-stayed bridge is provided by the first International Project Competition for Structural Health Monitoring (IPC-SHM) (Bao *et al*. 2021). There are 38 sensors, whose locations are illustrated in Fig. 3. The sampling frequency is 20 Hz. Table 1 describes the characteristics of the normal data and the six classes of anomalies. A dataset with labels for the time series by hour is also provided. For the one-month (31 days) of data for 38 sensors, the dimension of the labeled dataset is 744 × 38.

### *3.2 Data pre-processing*

The pre-processing of the raw SHM data involves time-frequency transform and time series to image encoding. In this study, FFT is first employed to obtain power spectral density (PSD) functions of original data. Meanwhile, the time series is transformed into two-dimensional images by using GAF approach (only GASF is adopted). Then, a series of composite images including three channels are generated by stacking time response image (channel 1), frequency response image (channel 2) and GASF image (channel 3). The size of each image is 128 × 128 × 3. Fig. 4 depicts partial inputs of all patterns. Among them, the upper left corner of each sub image represents a single-channel time response image, the upper right corner represents a single-channel frequency response image, the lower left corner represents a single-channel GAF image, and the lower right corner represents a three-channel mixed image.

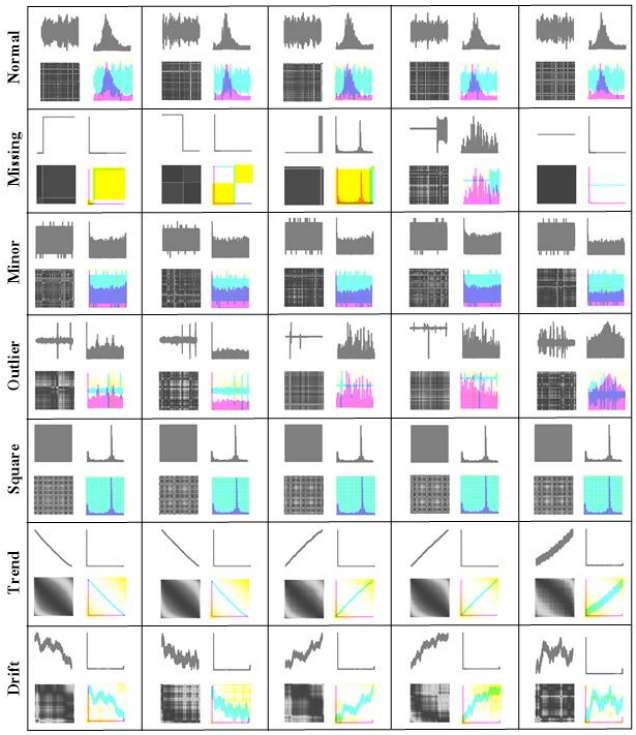


Fig. 4 Examples of data set grouped by class

In addition, Fig. 5 presents the distribution ratios of different data patterns. It can be observed that the normal type has a large proportion, and then trend, square, missing, minor, drift and outlier in order. Note that an imbalanced data set usually causes under-fitting to the minor classes and over-fitting to the major classes (Tang *et al*. 2019).

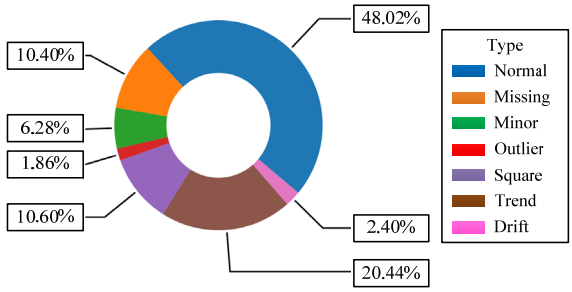


Fig. 5 Distribution ratios of different types of samples

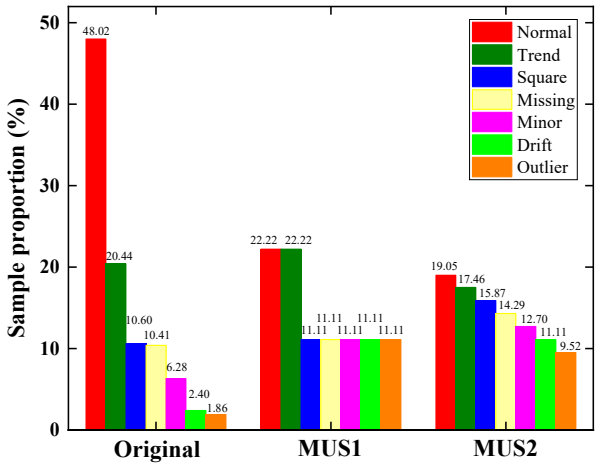


Fig. 6 Proportional distribution of three sampling schema

Multistage Sampling can relieve severe class-imbalance of training samples (He *et al*. 2016). Fig. 6 illustrated the proportional distribution of three sampling methods. The IPC-SHM dataset is first split into training set and validation set in a stratified fashion according to separate classes. Notice that the dataset is split before hybrid sampling, in order to ensure same distribution between test set and original dataset. Otherwise, the trained model might simply memorize specific data points and cause over-fitting and poor generalization to the test data.

## 4. Results

### *4.1 Evaluation metrics*

To verify the efficiency of the proposed GANomaly-CNN approach, the following metrics are utilized, i.e., precision, sensitivity (recall), accuracy and $F_1$ score.

$$\begin{cases} \text{precision} = \dfrac{\text{TP}}{\text{TP+FP}}, \quad \text{sensitivity} = \dfrac{\text{TP}}{\text{TP+FN}} \\ \text{accuracy} = \dfrac{\text{TP+TN}}{\text{TP+FN+FP+TN}} \\ F_1 = 2\dfrac{\text{precision} \cdot \text{sensetivity}}{\text{precision} + \text{sensetivity}} \times 100\% \end{cases} \tag{13}$$

where TP (true positive) is the number of positive samples predicted to be positive. FN (false negative) is the number of positive samples predicted to be negative. FP (false positive) is the number of negative samples predicted to be positive, and TN (true negative) is the number of negative samples predicted to be negative. Above four numbers also construct the confusion matrix, which is one of the evaluation indexes of various classification models. Additionally, the area under curve (AUC) of the receiver operating characteristic (ROC) is also employed to evaluate the overall performance of the proposed approach. In ROC, the true positive rate (TPR) is as a function of the false positive rate (FPR).

### *4.2 Performance evaluation*

In this section, the detection results were presented via employing the proposed network. Considering that only time response image may not be able to distinguish some types of anomalies well, frequency response image and GASF image were also introduced. Furthermore, a comparative study was conducted to investigate the performance of the following four cases with different input conditions via using the proposed network.

1) Case 1: raw time response image;
2) Case 2: frequency response image;
3) Case 3: GASF image;
4) Case 4: a mixed image of the first three cases.

Fig. 7 indicates the confusion matrix for different cases. It can be seen that the overall performance of each case is well with high accuracy. The comparison shows that the result of case 4 outperforms the other three cases by using total accuracy as the evaluation indicator. In cases 1-3, it is noteworthy that the outlier patterns have a relatively low precision, especially for case 3 (27.6%), which is caused by an enormous amount of misclassified samples. Most of these misclassified samples are from normal and minor. Nevertheless, for case 4, each pattern achieves a higher precision. That is to say, the three-channel superimposed image as input performs better than that of single input. Additionally, for the pattern "Missing," all four cases achieved high-level accuracies.

Tables 2 and 3 present F1 scores of different data types under four different cases. Based on a given sampling method, it can be found that F1 scores of the pattern outlier for cases 1-3 are low, but for case 4, the performance has a significant improvement. For drift pattern, the performance of case 3 is optimal. That is to say, GAF image is more suitable for detecting drift pattern than time domain image and frequency domain image. In addition, frequency domain image is better than time domain image for identifying trend and drift patterns, which is consistent with the previous study (Tang *et al*. 2019). Notice that $F_1$ scores of the pattern missing in all cases are high, especially for case 4, the value reaches 100%. Based on a given case, it can be seen that MUS2 method outperforms the other two sampling methods. Meanwhile, MUS1 is better that original sampling method. This means multistage sampling is effective for solving class-imbalance problem. The same conclusions are intuitively presented in Figs. 8 and 9.

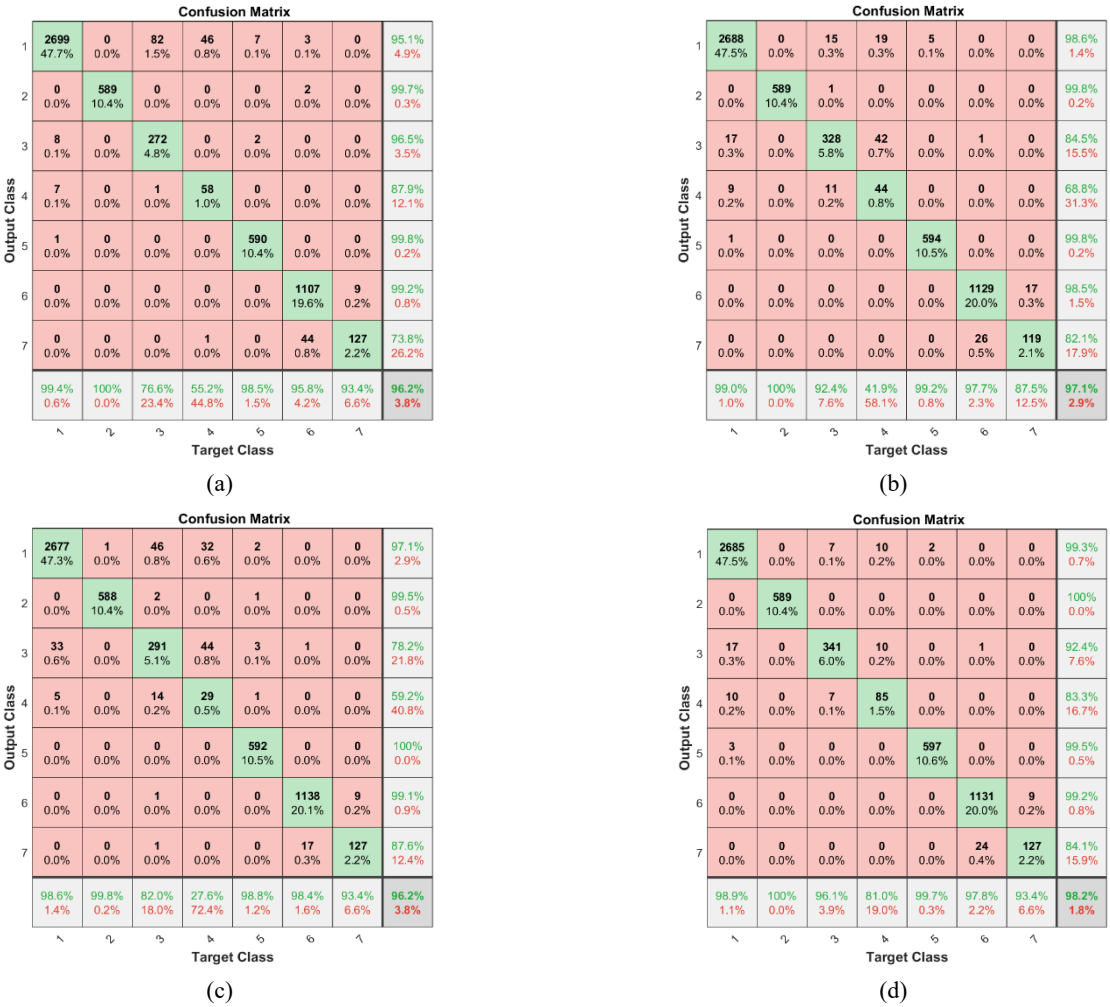


Fig. 7 Confusion matrix between detection result and actual anomaly distribution under different input conditions: (a) case 1; (b) case 2; (c) case 3; (d) case 4

The performance of the classifiers is also analyzed in terms of ROC curves, as shown in Fig. 10. Note that the closer the AUC area is to 1, the better the classification effect is. From Fig. 10, it can be found that the AUC areas of all patterns of each case are almost 1, indicating that the classification result performs well. In comparison, the performance of the outlier feature in case 3 is inferior, which is consistent with the analysis results from Fig. 7.

Additionally, compared with the previous studies, i.e., deep neural network (DNN) (Bao *et al*. 2019) and CNN (Tang *et al*. 2019), the proposed approach in this article has higher overall detection accuracy and $F_1$ score.

Table 2 Testing performance for cases 1 and 2. ($F_1$ score)

| | Case 1 | | | Case 2 | | |
|---|---|---|---|---|---|---|
| | Original | MUS1 | MUS2 | Original | MUS1 | MUS2 |
| Normal | 96.18 | 97.14 | 97.23 | 97.99 | 98.12 | 98.79 |
| Missing | 99.16 | 99.66 | 99.83 | 99.92 | 99.92 | 99.92 |
| Minor | 78.99 | 85.40 | 85.40 | 85.60 | 85.33 | 88.29 |
| Outlier | 56.44 | 67.02 | 67.84 | 38.22 | 48.65 | 52.07 |
| Square | 98.58 | 98.99 | 99.16 | 99.75 | 99.58 | 99.50 |
| Trend | 97.70 | 97.32 | 97.45 | 98.01 | 97.69 | 98.09 |
| Drift | 80.14 | 82.08 | 82.47 | 83.39 | 81.82 | 84.70 |
| Macro | 86.74 | 89.66 | 89.91 | 86.13 | 87.30 | 88.76 |
| Accuracy | 95.00 | 96.06 | 96.23 | 96.36 | 96.32 | 97.10 |

Table 3 Testing performance for cases 3 and 4 ($F_1$ score)

| | Case 3 | | | Case 4 | | |
|---|---|---|---|---|---|---|
| | Original | MUS1 | MUS2 | Original | MUS1 | MUS2 |
| Normal | 97.01 | 97.21 | 97.83 | 97.95 | 98.89 | 99.10 |
| Missing | 99.75 | 99.41 | 99.66 | 99.92 | 100.00 | 100.00 |
| Minor | 78.11 | 78.33 | 80.06 | 91.24 | 93.92 | 94.20 |
| Outlier | 30.38 | 36.48 | 37.66 | 71.55 | 81.77 | 82.13 |
| Square | 99.50 | 99.25 | 99.41 | 99.67 | 99.50 | 99.58 |
| Trend | 98.52 | 98.47 | 98.78 | 98.53 | 98.48 | 98.52 |
| Drift | 88.36 | 88.05 | 90.39 | 87.12 | 87.94 | 88.50 |
| Macro | 84.52 | 85.31 | 86.26 | 92.28 | 94.36 | 94.57 |
| Accuracy | 95.47 | 95.58 | 96.23 | 97.24 | 98.09 | 98.23 |

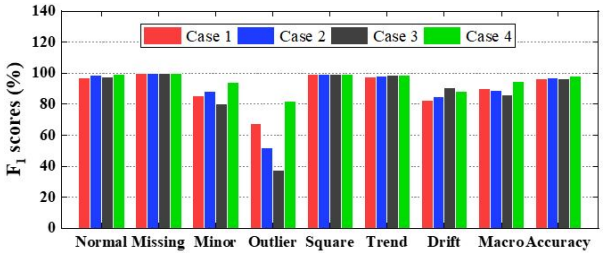


Fig. 8 A comparison of different cases based on MUS2 sampling method

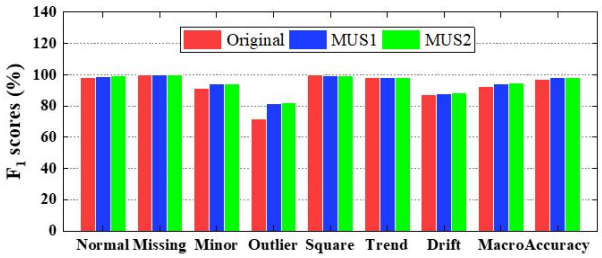


Fig. 9 A comparison of different sampling methods for case 4

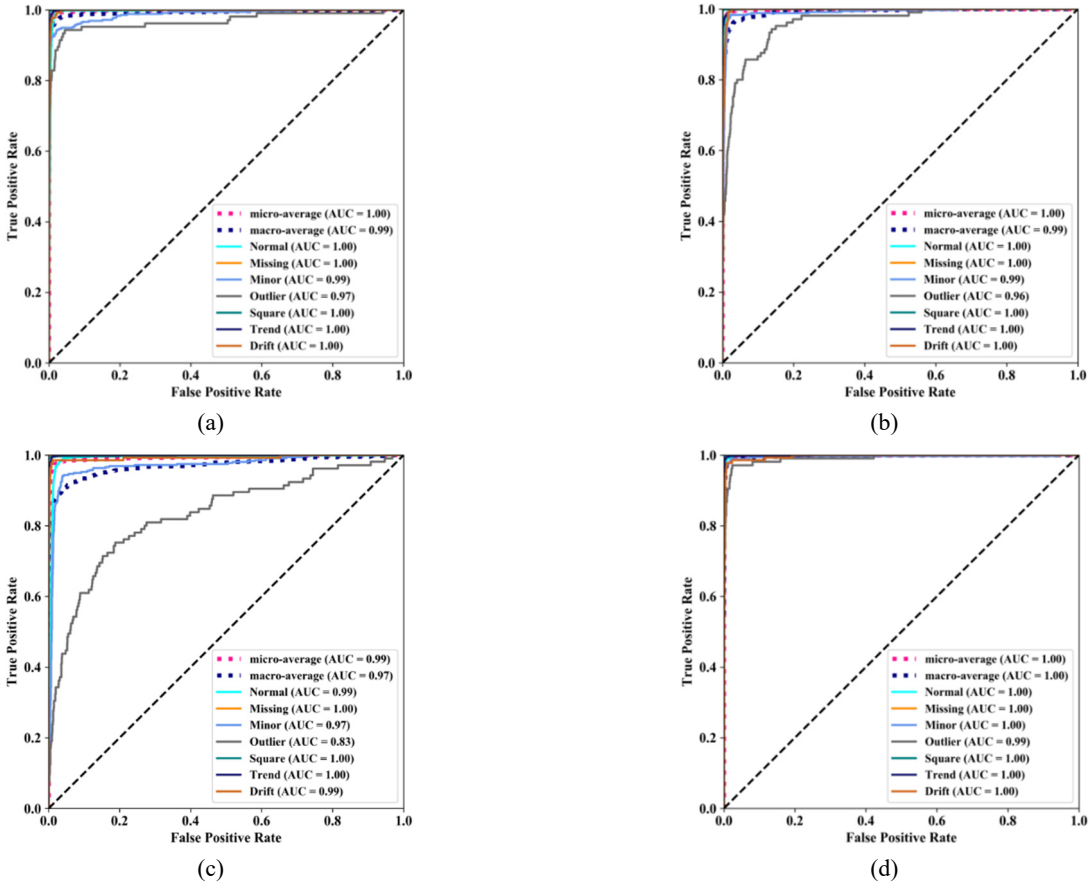


Fig. 10 ROC curves of the classifiers of each pattern under different input conditions: (a) case 1; (b) case 2; (c) case 3; (d) case 4

## 5. Conclusions

In this study, a novel GANomaly-Convolutional Neural Network (CNN) architecture is put forward to detecting multiple anomaly features for Structural Health Monitoring (SHM) data. In particular, the raw time series are encoded as two-dimensional image vectors by using Gramian Angular Field (GAF) method. Moreover, in order to better distinguish abnormal samples, extra frequency domain information is also introduced. According to the analysis results from SHM system of a real long-span bridge, the following conclusions are summarized:

- The acceleration data provided by IPC-SHM are utilized to verify the effectiveness of the proposed GANomaly-CNN model. The analysis results depict that this network architecture achieves a global detection accuracy of 98.2%, and $F_1$ scores for different patterns are 99.10%, 100.00%, 94.20%, 82.13%, 99.58%, 98.52%, and 88.50%. Compare with the former CNN-based approach, this approach achieves overall higher accuracy and $F_1$ score.
- By comparing the overall accuracy and $F_1$ score of anomalies under different input conditions (i.e., case 1, case 2, case 3 and case 4), it can be found that case 3 is more suitable for detecting trend and drift patterns than cases 1 and 2, but it is poor for outlier pattern. Meanwhile, a combination input including time domain image, frequency domain image and GAF image achieves a satisfactory result for all patterns. This implies that the proposed GANomaly-CNN based method with three-channel input information has a superior performance.
- Multistage Sampling methods, e.g., MUS1 and MUS2, are effective to solve the class-imbalance issue. Comparing the performance of MUS1 and MUS2 approaches for addressing class-imbalance issue, it can be found that MUS2 method is more effective than MUS1.

## Acknowledgments

The authors would like to thank the organizations of the International Project Competition for SHM (IPC-SHM 2020) ANCRiSST, Harbin Institute of Technology (China), and University of Illinois at Urbana-Champaign (USA) for their generously providing the invaluable data from actual structures. The authors also would like to thank the chairs of IPC-SHM 2020 Prof. Hui Li, and Prof. Billie F. Spencer Jr for their leadership on the competition.

The authors would like to gratefully acknowledge the National Natural Science Foundation of China (52108179), the China Postdoctoral Science Foundation (2021M692835, 2021M702866).

## References

Akcay, S., Atapour-Abarghouei, A. and Breckon, T.P. (2019), "Ganomaly: Semi-supervised anomaly detection via adversarial training", Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), 11363 LNCS, pp. 622-637.

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, 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

Barra, S., Carta, S.M., Corriga, A., Podda, A.S. and Recupero, D.R. (2020), "Deep learning and time series-To-image encoding for financial forecasting", *IEEE/CAA J. Automat. Sinica*, **7**(3), 683-692. https://doi.org/10.1109/JAS.2020.1003132

Buda, M., Maki, A. and Mazurowski, M.A. (2018), "A systematic study of the class imbalance problem in convolutional neural networks", *Neural Networks*, **106**, 249-259. https://doi.org/10.1016/j.neunet.2018.07.011

Ding, J., Liu, Y., Zhang, L., Wang, J. and Liu, Y. (2016), "An anomaly detection approach for multiple monitoring data series based on latent correlation probabilistic model", *Appl. Intell.*, **44**(2), 340-361. https://doi.org/10.1007/s10489-015-0713-7

Gatti, M. (2019), "Structural health monitoring of an operational bridge: A case study", *Eng. Struct.*, **195**, 200-209. https://doi.org/10.1016/j.engstruct.2019.05.102

Gul, M. and Catbas, F.N. (2009), "Statistical pattern recognition for Structural Health Monitoring using time series modeling: Theory and experimental verifications", *Mech. Syst. Signal Process.*, **23**(7), 2192-2204. https://doi.org/10.1016/j.ymssp.2009.02.013

Han, J. and Moraga, C. (1995), "The influence of the sigmoid function parameters on the speed of backpropagation learning", Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), **930**, 195-201. https://doi.org/10.1007/3-540-59497-3_175

He, T. and Li, X. (2019), "Image quality recognition technology based on deep learning", *J. Visual Commun. Image Represent.*, **65**, 102654. https://doi.org/10.1016/j.jvcir.2019.102654

He, K., Zhang, X., Ren, S. and Sun, J. (2016), "Deep residual learning for image recognition", *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pp. 770-778.

Heinlein, S., Cawley, P. and Vogt, T. (2019), "Validation of a procedure for the evaluation of the performance of an installed structural health monitoring system", *Struct. Health Monitor.*, **18**(5-6), 1557-1568. https://doi.org/10.1177/1475921718798567

Hernandez-Garcia, M.R. and Masri, S.F. (2014), "Application of statistical monitoring using latent-variable techniques for detection of faults in sensor networks", *J. Intell. Mater. Syst. Struct.*, **25**(2), 121-136. https://doi.org/10.1177/1045389X13479182

Huang, L., Duan, S.W., Zhu, K.M. and Chen, W.G. (2016), "Study on deformation monitoring of subway station deep foundation construction", *Appl. Mech. Mater.*, **847**, 425-430. https://doi.org/10.4028/www.scientific.net/AMM.847.425

Isogawa, K., Ida, T., Shiodera, T. and Takeguchi, T. (2018), "Deep shrinkage convolutional neural network for adaptive noise reduction", *IEEE Signal Processing Letters*, *IEEE*, **25**(2), 224-228. https://doi.org/10.1109/LSP.2017.2782270

Jiang, W., Hong, Y., Zhou, B., He, X. and Cheng, C. (2019), "A GAN-based anomaly detection approach for imbalanced industrial time series", *IEEE Access*, **7**, 143608-143619. https://doi.org/10.1109/ACCESS.2019.2944689

Kerschen, G., De Boe, P., Golinval, J.C. and Worden, K. (2005), "Sensor validation using principal component analysis", *Smart Mater. Struct.*, **14**(1), 36-42. https://doi.org/10.1088/0964-1726/14/1/004

Kong, Q., Cao, Y., Iqbal, T., Wang, Y., Wang, W. and Plumbley, M.D. (2020), "Panns: Large-scale pretrained audio neural networks for audio pattern recognition", *IEEE/ACM Transactions on Audio Speech and Language Processing*, **28**, 2880-2894. https://doi.org/10.1109/TASLP.2020.3030497

Li, H. and Ou, J. (2016), "The state of the art in structural health monitoring of cable-stayed bridges", *J. Civil Struct. Health Monitor.*, **6**(1), 43-67. https://doi.org/10.1007/s13349-015-0115-x

Mutlu, U. and Alpaydın, E. (2020), "Training bidirectional generative adversarial networks with hints", *Pattern Recognition*, **103**, 107320. https://doi.org/10.1016/j.patcog.2020.107320

Ni, F.T., 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**(7), 685-700. https://doi.org/10.1111/mice.12528

Park, J., Kim, K. and Cho, Y.K. (2017), "Framework of automated construction-safety monitoring using cloud-enabled BIM and BLE mobile tracking sensors", *J. Constr. Eng. Manag.*, **143**(2), 05016019. https://doi.org/10.1061/(ASCE)CO.1943-7862.0001223

Park, S., Kim, S. and Choi, J.H. (2018), "Gear fault diagnosis using transmission error and ensemble empirical mode decomposition", *Mech. Syst. Signal Process.*, **108**, 58-72. https://doi.org/10.1016/j.ymssp.2018.02.028

Radford, A., Metz, L. and Chintala, S. (2016), "Unsupervised representation learning with deep convolutional generative adversarial networks", *Proceedings of the 4th International Conference on Learning Representations*, *ICLR 2016 - Conference Track Proceedings*, pp. 1-16.

Rao, A.R.M., Kumar, S.K. and Lakshmi, K. (2014), "A Sensor fault detection algorithm for structural health monitoring using adaptive differential evolution", *Int. J. Computat. Methods Eng. Sci. Mech.*, **15**(3), 282-293. https://doi.org/10.1080/15502287.2014.883239

Tang, Z., Chen, Z., Bao, Y. and Li, H. (2019), "Convolutional neural network-based data anomaly detection method using multiple information for structural health monitoring", *Struct. Control Health Monitor.*, **26**(1), 1-22. https://doi.org/10.1002/stc.2296

Thomas, S., Philipp, S., Sebastian, M.W., Ursula, S.-E. and Georg, L. (2017), "Unsupervised anomaly detection with generative adversarial networks to guide marker discovery", *Proceedings of International Conference on Information Processing in Medical Imaging*, pp. 146-157.

Tian, L., Fan, Y., Li, L. and Mousseau, N. (2020), "Identifying flow defects in amorphous alloys using machine learning outlier detection methods", *Scripta Materialia*, **186**, 185-189. https://doi.org/10.1016/j.scriptamat.2020.05.038

Xu, Y.L. (2018), "Making good use of structural health monitoring systems of long-span cable-supported bridges", *J. Civil Struct. Health Monitor.*, **8**(3), 477-497. https://doi.org/10.1007/s13349-018-0279-2

Xu, Y., Bao, Y., Chen, J., Zuo, W. and Li, H. (2019), "Surface fatigue crack identification in steel box girder of bridges by a deep fusion convolutional neural network based on consumer-grade camera images", *Struct. Health Monitor.*, **18**(3), 653-674. https://doi.org/10.1177/1475921718764873

Yang, Y. and Nagarajaiah, S. (2014), "Data compression of structural seismic responses via principled independent component analysis", *J. Struct. Eng.*, **140**(7), 04014032. https://doi.org/10.1061/(ASCE)ST.1943-541X.0000946

Yang, C.-L., Chen, Z.-X. and Yang, C.-Y. (2020), "Sensor classification using convolutional neural network by encoding multivariate time series as two-dimensional colored images", *Sensors*, **20**, 168. https://doi.org/10.3390/s20010168

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. Stabil. Dyn.*, **16**(4), 1-21. https://doi.org/10.1142/S0219455416400241

Yin, H. and Gai, K. (2015), "An empirical study on preprocessing high-dimensional class-imbalanced data for classification", *Proceedings - 2015 IEEE 17th International Conference on High Performance Computing and Communications*, *2015 IEEE 7th International Symposium on Cyberspace Safety and Security and 2015 IEEE 12th International Conference on Embedded Software and Systems*, New York, NY, USA, August, pp. 1314-1319. https://doi.org/10.1109/HPCC-CSS-ICESS.2015.205

Yuen, K.V. and Mu, H.Q. (2012), "A novel probabilistic method for robust parametric identification and outlier detection", *Probabil. Eng. Mech.*, **30**, 48-59. https://doi.org/10.1016/j.probengmech.2012.06.002