Smart Structures and Systems, *Vol. 30, No. 6 (2022) 557-569*
https://doi.org/10.12989/sss.2022.30.6.557

Vibration-based structural health monitoring using CAE-aided unsupervised deep learning

Minte Zhang [1a], Tong Guo*[1,3], Ruizhao Zhu [1b], Yueran Zong [1c] and Zhihong Pan [2d]

[1] *School of Civil Engineering, Southeast University, Nanjing, Jiangsu, People's Republic of China*
[2] *School of Architecture and Civil Engineering, Jiangsu University of Science and Technology, Zhenjiang, Jiangsu, People's Republic of China*
[3] *The Centre for BIM Studies, Smart City and Sustainable Development Academy, Chongqing, China*

(*Received April 10, 2022, Revised August 31, 2022, Accepted September 8, 2022*)

Abstract. Vibration-based structural health monitoring (SHM) is crucial for the dynamic maintenance of civil building structures to protect property security and the lives of the public. Analyzing these vibrations with modern artificial intelligence and deep learning (DL) methods is a new trend. This paper proposed an unsupervised deep learning method based on a convolutional autoencoder (CAE), which can overcome the limitations of conventional supervised deep learning. With the convolutional core applied to the DL network, the method can extract features self-adaptively and efficiently. The effectiveness of the method in detecting damage is then tested using a benchmark model. Thereafter, this method is used to detect damage and instant disaster events in a rubber bearing-isolated gymnasium structure. The results indicate that the method enables the CAE network to learn the intact vibrations, so as to distinguish between different damage states of the benchmark model, and the outcome meets the high-dimensional data distribution characteristics visualized by the t-SNE method. Besides, the CAE-based network trained with daily vibrations of the isolating layer in the gymnasium can precisely recover newly collected vibration and detect the occurrence of the ground motion. The proposed method is effective at identifying nonlinear variations in the dynamic responses and has the potential to be used for structural condition assessment and safety warning.

Keywords: damage identification; on-site test; structural health monitoring; unsupervised deep learning; vibration assessment

1. Introduction

Buildings are the foundation of a city and provide human shelter. Maintaining and strengthening buildings has become an increasingly prominent problem to ensure the safety of people's lives and properties. Engineers have been attempting to measure building responses and excitations in order to monitor and evaluate the health state of existing structures due to the rapid development of computational science and physical sensor technology. Unlike traditional non-destructive tests, structural health monitoring (SHM) is a continuous on-site inspection method that records various real-time crucial safety parameters of the structure, such as location, deformation, stress, load, temperature, and accelerations. These parameters can be used to form the structures' on-site dynamic model. Engineers can determine the causes of the damage by modifying the previously defined model to fit the on-site monitored response. This allows them to make recommendations for structural maintenance and reinforcement. The need to detect the damage state of operational structures has prompted the development of various vibration-based methods.

Initially, a vibration-based damage identification method for SHM was proposed for aerospace and mechanic assessment. Civil engineers began to study vibration-based methods for civil infrastructure in the 1980s. Doebling *et al*. (1998) comprehensively reviewed the methods used between the 1980s and 1990s and concluded the fundamental idea to identify damage from structural vibration responses, i.e., every damage introduced to the structure may change its intrinsic characteristics, such as natural frequency, damping, and mode vector, all of which are closely related to structural responses. Researchers were inspired by this idea and used several signal processing methods to determine the structural parameters, which can be classified into three categories: frequency domain methods (Jalali and Rideout 2022), modal domain methods (Yi *et al*. 2019), and time-frequency domain methods (Pan *et al*. 2018).

Frequency-domain methods calculate the modal parameters of the monitored structure based on frequency response function (FRF). Reynolds and Pavic (2003) determined the natural frequencies and modal damping ratios of a long-span concrete floor by measuring and calculating the FRF data during the modal test. By installing a false floor, the damage was introduced. They revealed the influence of the damage on stiffness and damping of the test structure by comparing the modal parameters of the floor

*Corresponding author, Professor,
E-mail: guotong@seu.edu.cn
[a] Ph.D. Student, E-mail: zmint@seu.edu.cn
[b] Ph.D. Student, E-mail: ruizhaozhu@outlook.com
[c] Postgraduate Student, E-mail: zongyueran1997@163.com
[d] Professor, E-mail: Zhhpan@tom.com

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

before and after the installation, and finally evaluated its damage state by updating the finite element (FE) model of the structure (Yi *et al*. 2019). Furthermore, to make the best of the modal information recognized through frequency-domain methods, many researchers have used modal-domain methods (Bayissa and Haritos 2007). With the modal properties of the structure determined, the damage state can be easily quantified using various parameters such as the modal flexibility and modal strain energy, both of which can reveal the intrinsic features of the damage based on measured response data. Wang *et al*. (2018b) proposed a modified modal strain energy method to realize the damage location of asymmetric buildings after conducting a series of experimental tests and FE analysis. Unfortunately, neither of the methods described above considers the time-series characteristics of the vibrations. To overcome this limitation, some time-frequency domain methods (Tibaduiza Burgos *et al*. 2020), such as the Wavelet Transform (WT), Wavelet Package Transform (WPT) and Hilbert-Huang Transform (HHT) were developed to process real-time vibration signals (Wang *et al*. 2018c). These modern signal processing methods have been demonstrated to be effective in structural vibration analysis. Hera and Hou (2004) used the WT method to analyze vibration signals acquired from the ASCE-Benchmark Phase-I test, and it was verified that the WT time-frequency spectrum of the signals can clearly represent abnormal points when damage occurs. Qu and Lian (2012) employed the HHT method when identifying damage from the vibration response of electric transmission towers. The testing results demonstrated that the HHT method can detect the intrinsic differences before and after damage occurs.

All of the traditional damage identification methods mentioned above require the creation of an analytical model using FE methods, which is necessary for engineers when comparing the modal parameters of theoretical and measured values. However, the drawbacks of these physical model-based methods are apparent. First, it is a significant challenge to precisely simulate the real structure with FE methods because of the uncertainties of the in-service structures, such as the material properties and loadings on floors (Erdogan *et al*. 2014). Furthermore, the influence of ambient noise may also lead to different modal identification results, some of which are mistakes, though all of the results are within the allowable errors in FE models. In recent years, with the development of artificial intelligent, data-driven methods for modeling structural features using real-measured data is trending on structural damage identification (Flah *et al*. 2021). Data-driven methods can directly extract features from monitored vibration data. Based on the pattern recognition and machine learning (ML) theory, the dynamic signatures of the structure can be learned from mass vibration datasets (Gentile *et al*. 2019).

In general, ML establishes an analytical model and optimizes its structure by training its network with numerous datasets so as to extract their common features. Following extensive training and learning from previous experiences, the model not only has good distinguishing ability, but it can also predict the unhappened occasions using the experiences accumulated from the training datasets. During past decades, civil engineers have analyzed various structural response data with the support vector machine (SVM) (Na *et al*. 2022), artificial neural network (ANN) (Goyal and Pabla 2016), genetic network (GN) (Tiachacht *et al*. 2018), principal component analysis (PCA) (Yan *et al*. 2005), random forest (Smarra *et al*. 2020), etc. Unfortunately, traditional machine learning methods have some limitations when identifying structural damages: (1) Overcoming the influences of occasional noise is a significant challenge, which limits the application scene of the methods; (2) simple linear network structure can only extract shallow information, ignoring many nonlinear features of the datasets; (3) the methods always involve complicated feature pre-extraction before learning. The feature pre-extraction not only consumes more computational resources, but it also separates the original datasets and damage features of the structure because it only focuses on local details. On the other hand, deep learning with multi-layer neural networks has been investigated extensively over the years and has been shown to be effective in solving the problems listed above.

The convolutional neural network (CNN) is a deep learning method that concentrates the multi-layer neural network and shares convolutional kernel. The method was initially studied by LeCun *et al*. (2015) to identify the hand-written numbers. Researchers can use 2D-CNN to classify and recognize different types of objects from 2D images. Based on its outstanding performance in the vision domain, some engineers began to identify structural damages from photos of structures. Cha and Choi (2017) trained the CNN with cracked concrete images and built a classifier with the optimized network to separate the damaged from the undamaged concrete surface images noised with different lighting and shadow environments. The CNN-based DL method was validated using existing concrete crack detection data. The method has also been applied to other non-destructive vision inspection tests, such as corrosion detection of steel structures (Khayatazad et al 2022), spall of masonry structures (Wang *et al*. 2018a), and splitting of bridge elastomeric bearings (Cui *et al*. 2021). However, most 2D-CNN-based vision methods can only detect superficial construction anomalies, and these static images cannot reveal the dynamic characteristics of the structure. This limitation can be overcome by using 1D-CNN-based signal classification methods. Abdeljaber *et al*. (2017) established vibration signal classifiers with 1D-CNN, trained the network with ASCE-Benchmark Phase I dataset and tested the classification performance of the network with the one-dimensional convolutional kernel. At Qatar University, an experimental grandstand frame was established and monitored, and its acceleration response was excited by the modal shaker. The test results demonstrated that the network could identify various damage models through vibration data and the damage can be located from the multi-sensor network. The neural network with the convolutional kernel can classify vibration data into several categories of pre-labeled damage state, but these datasets labeled for network training require manual classification, which is practically impossible in real-time

damage identification. Without the classification of the training dataset, the unsupervised learning method is ideal for real-time structural damage identification based on vibrations because it can automatically extract hidden features from a multitude of unlabeled data. Autoencoder (AE), which is composed of an encoder for data compression and a decoder for data decompression, is one of the most popular unsupervised learning methods. AE employs a deep-layer neural network to self-adaptively learn sharing features of the vibrations. Jiang *et al*. (2021) built the AE for structural damage diagnosis using multi-layer ANN and discussed the space distributions of experimental vibration data under different damage cases.

In recent years, many algorithms have been proposed and tested on vibration based SHM, but research on civil structures under real-time ambient or earthquake excitation vibration is still very limited. This paper is aimed to use an unsupervised deep learning method to identify and diagnose civil structures with online monitored accelerations. The remainder of this paper is organized as follows. First, the unsupervised DL method, as well as the CAE network architecture, data preprocessing method, training strategy and techniques for data analysis are introduced. Then, the routine for training and verifying the unlabeled monitoring data through CAE network are studied with the two real on-site tests. The ASCE Benchmark experiment-based test is conducted to verify the CAE network trained with steel frame experimental vibrations. With the help of the t-SNE method, the test result will reveal how the damage features embody in random vibrations data reconstruction through the autoencoder network with the convolutional kernel. On top of that, another test applies the method on a rubber bearing-isolated gymnasium structure to detect the damage and instant seismic events through vibration data. The relationship between the reconstruction error and real-time monitoring vibrations is also compared and discussed. Finally, the result of the test and potential future work are discussed.

2. Methodology

In this study, the convolutional autoencoder (CAE) is used on vibration-based unsupervised deep learning. The CAE-based unsupervised damage identification method for vibration data evaluation is shown in Fig. 1. First, the input data preprocessed from the vibration of the monitored structure is compressed by the convolutional network (Conv) and reconstructed by the transposed convolutional network (TransConv). The original and reconstructed data are then compared and scored to assess the vibration.

2.1 Data preprocessing

The acceleration vibration data collected by N sensors can be expressed as follows

$$\boldsymbol{Dd} = [\boldsymbol{u}_1 \quad \boldsymbol{u}_2 \quad \ldots \quad \boldsymbol{u}_N] \tag{1}$$

and for each sensor, a dataset containing N_a acceleration data is generated. Considering the effectiveness of stationary random vibration signals, several l-long segments are sampled from the N_a-long dataset, where l is greater than twice the sampling frequency of sensors. Two sampling approaches are used, i.e., sequential sampling and random window sampling. To begin, the dataset is sequentially divided into N_g independently repeated segments. Furthermore, to eliminate the uncertainty of random data and enlarge the number of training samples for the DL network, a random sampling window function RandWin (l, seed) is defined, as is shown in Fig. 2. For example, to sample a segment with the function, a random floating numbers seed $\in[0, N_a\text{-}l]$ is generated with the random number function to locate the starting point of the window. The function is repeatedly called to obtain another N_r segments. Each sampled segment $\mathbf{u}_{i,j}$ is normalized and arranged in a random order. Finally, $(N_g + N_r)$ acceleration signal segments for the DL test are generated, which can be written as

$$\begin{aligned} \boldsymbol{u}_i &= [\boldsymbol{u}_{i,1} \quad \boldsymbol{u}_{i,2} \quad \ldots \quad \boldsymbol{u}_{i,(N_g+N_r)}] \\ \boldsymbol{u}_{i,j} &= [a_1 \quad a_2 \quad \ldots \quad a_l] \end{aligned} \tag{2}$$

where $\mathbf{u}_{i,j}$ denotes the jth signal segment collected from sensor-i and $\mathbf{u}_i$ denotes a subset of sensor-i recording data. The RandWin function can also be used to generate representative datasets from massive online monitoring vibration data.

2.2 Unsupervised deep learning method with CAE

After preprocessing, the dynamic response data of the monitored building can be defined as follows

$$x = [x_1, x_2, x_3, \ldots x_i, \ldots x_n] \tag{3}$$

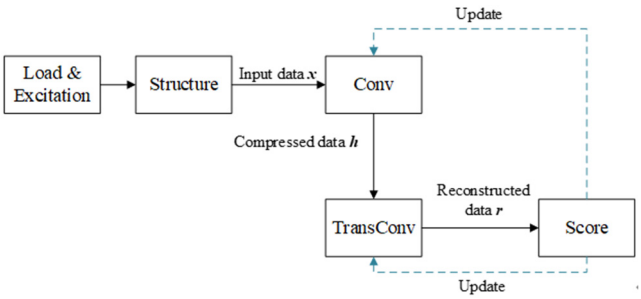


Fig. 1 Procedure of the damage identification method

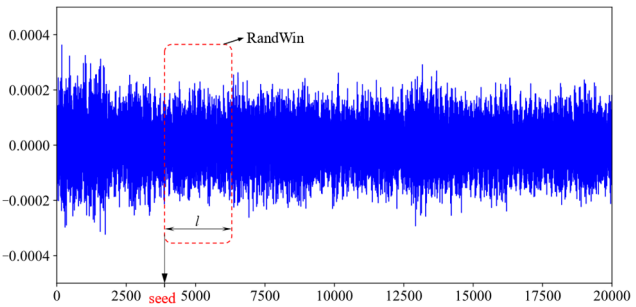


Fig. 2 Random window sampling

where $\boldsymbol{x}$ represents the input vibration signal for network training and x_i represents the ith structural vibration data. The CAE includes a convolutional encoder and a trans-convolutional decoder.

The encoder uses convolutional kernels to under-sample the input data $\boldsymbol{x}$ layer by layer

$$x_i^k = \sigma\left(\beta_i^k \cdot \boldsymbol{Conv}\left(x_i^{k-1}\right) + a_i^k\right) \tag{4}$$

where x_i^k represents the ith step data sampling in kth convolutional neural network, **Conv** is the convolution function, β_i^k is the weight of convolutional calculation and $\sigma(\cdot)$ is the activation function that employs the **LeakyReLU** function (Paszke *et al*. 2019), which can be expressed as follows.

$$\sigma(x) = \begin{cases} x, x > 0 \\ \alpha x, x \le 0 \quad \alpha \in (0,1) \end{cases} \tag{5}$$

The input data tensor $\boldsymbol{x}$ is compressed into a shorter tensor $\boldsymbol{h}$, which is unsampled and reconstructed by the decoder. Similar to the encoder, the decoder employs trans-convolutional kernels to under-sample the input data $\boldsymbol{x}$ layer by layer

$$x_i^k = \sigma\left(\beta_i^k \cdot \boldsymbol{TransConv}\left(x_i^{k-1}\right) + a_i^k\right) \tag{6}$$

where **TransConv** is the trans-convolution function. It up-samples the compressed data and recovers it into tensor $\boldsymbol{r}$ using the transposed convolutional kernel (LeCun *et al*. 2015). The network is trained to reduce deviation between $\boldsymbol{x}$ and $\boldsymbol{r}$. The mean squared error is defined as the loss function.

2.3 Network structure and training strategy

The DL network is established using the CAE method described in Section 2.2. A defined intact vibration dataset is used to train the DL network, whose network structure is shown in Fig. 3. The DL network is programmed on the PyTorch platform. 1D-CNN is established with nn.Conv1d() command and 1D-transposed-CNN is established with nn.ConvTranspose1d() command. Up-sampling and down-sampling are both packaged with nn.Sequential() to realize function of auto-encoder. The loss function is nn.MSELoss() and training optimizing algorithm is optim.Adamax. To accelerate the training process, the CUDA toolkit is used to boost GPU calculation (Paszke *et al*. 2019).

2.4 Techniques for data assessment and visualization

2.4.1 Damage state assessment indicators

Two indicators are used to assess the reconstruction ability of the CAE network. Formula (7) is the mean square error that describes the discrete error of the input and output data. Formula (8) is the Pearson correlation coefficient that describes the linear correlation between the input and output data.

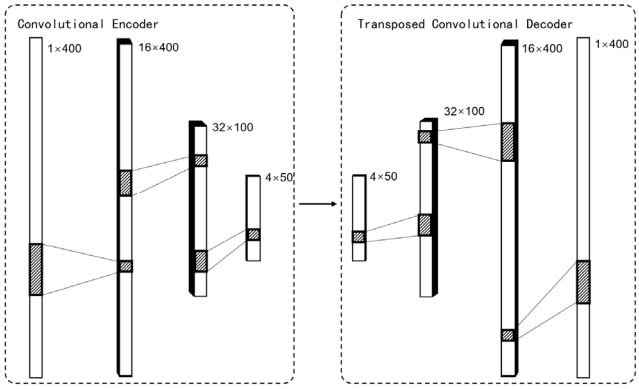


Fig. 3 CAE network structure

$$\mathrm{MSE} = \frac{\sum(\boldsymbol{x} - \boldsymbol{r})^2}{n} \tag{7}$$

$$\mathrm{Cor} = \frac{\sum_{i=1}^{n}(x_i - \bar{x})\,(r_i - \overline{r})}{\sqrt{\sum_{i=1}^{n}(x_i - \bar{x})^2}\sqrt{\sum_{i=1}^{n}(r_i - \overline{r})^2}} \times 100 \tag{8}$$

2.4.2 t-SNE method

It is assumed that $\alpha = [\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots \mathbf{x}_i \ldots, \mathbf{x}_N] \in R^D$ is a dataset in D-dimensional space, whereas $\beta = [\boldsymbol{y}_1, \boldsymbol{y}_2, \ldots \mathbf{y}_i \ldots, \mathbf{y}_N] \in R^D$ is in d-dimensional space, with $d << D$. To map data in high dimensional space into low dimensional space, the t-distributed stochastic neighborhood embedding (t-SNE) method is used to represent a long vibration data in the second -dimension space (Leon-Medina *et al*. 2020). The method considers similarities between elements in α set and retains these correlations when mapping them into lower dimensions. The similarities between data $\boldsymbol{x}_i$ and $\boldsymbol{x}_j$ can be described with the Gaussian distribution

$$p_{j|i} = \frac{exp\left(-\left\|\boldsymbol{x}_i - \boldsymbol{x}_j\right\|^2 / 2\sigma_i^2\right)}{\sum_{k \ne i} \exp\left(-\left\|\boldsymbol{x}_i - \boldsymbol{x}_j\right\|^2 / 2\sigma_i^2\right)} \tag{9}$$

where σ_i represents the bandwidth of the Gaussian distribution kernel and $p_{j|i}$ is the value of gauss probability between the ith and jth data, increasing when $\boldsymbol{x}_i$ and $\boldsymbol{x}_j$ become similar. $p_{j|i} = 0$ when $i = j$. The symmetrized expression of p_{ij} is defined as follows

$$p_{ij} = \frac{p_{i|j} + p_{j|i}}{2N} \tag{10}$$

Then, q_{ij}, the distance between $\boldsymbol{y}_i$ and $\boldsymbol{y}_j$ in d-dimensional space, can be quantified using the single-degree-of-freedom Student's t distribution

$$q_{ij} = \frac{\left(1 + \left\|\boldsymbol{y}_i - \boldsymbol{y}_j\right\|^2\right)^{-1}}{\sum_{k \ne l}^{k,l=1toN}(1 + \|\boldsymbol{y}_k - \boldsymbol{y}_l\|^2)^{-1}} \tag{11}$$

Similarly, $q_{ij} = 0$ when $i = j$. The mapped d-dimensional dataset β is obtained by optimizing the K-L divergence between p and q

$$L(\boldsymbol{Y}) = KL(p||q) = \sum_i \sum_j p_{ij}\, log \frac{p_{ij}}{q_{ij}} \tag{12}$$

and the gradient of $L(Y)$ is as follows.

$$\frac{\partial L}{\partial \boldsymbol{y}_i} = 2\sum_j \left(p_{j|i} - q_{j|i} + p_{i|j} - q_{i|j}\right)\left(\boldsymbol{y}_i - \boldsymbol{y}_j\right) \tag{13}$$

The optimized $\boldsymbol{y}_i$ and β is accessed using the gradient descent algorithm.

3. Experimental study for damage detection – a benchmark model

3.1 Introduction to ASCE-Benchmark experiment

The ASCE-Benchmark model, a widely used structural experimental model, was first proposed by IASC and ASCE in 2002 to assist different civil engineers and SHM data analysts in testing various structural damage identification methods (Das and Saha 2018). Using a group of shared vibration data obtained from the benchmark test model, researchers worldwide can study different physical feature and structural identification indicators under the same experimental background. The benchmark model was studied in two phases between 1999 and 2002. The Phase-I response data came from a simulated 120 DOF FE model (Johnson *et al*. 2004), while the Phase-II response data came from a real experimental steel frame model (see Fig. 4), including steel beams, steel columns and braces (Das and Saha 2018). Damage of the experiment structure was introduced by removing braces and loosening bolts at the column-beam joints, descriptions of damage simulated are shown in Table 2. With 15 acceleration sensors, the IASC-ASCE SHM group accessed real-time monitored data from the model under three excitation cases, i.e., ambient vibration, impact hammer and dynamic shaker.

The benchmark model is a steel frame with 4-story, 3.6 m height and 2×2 bay, and additional masses are placed on the top of each floor. The total masses for stories 1-4 are 4000 kg, 4140 kg, 4000 kg and 3000 kg, respectively. There are 15 sensors (sensor-1~15) placed between floors to record structural response. The location information of each sensor is shown in Table 1. The sampling frequency of the sensor is 200 Hz. In this study, data generated by the ambient vibration excitation model is used for the unsupervised damage identification test. Table 2 displays five different structural damage simulation cases. The model in Case 1 is undamaged, while the models in Cases 2 to 4 are damaged.

3.2 Network training

The acceleration vibration data is collected from 15 sensors, with each sensor generating a dataset containing 60000 acceleration data. In this test, 300 acceleration signal segments for DL test are generated, with 150 independent segments divided in average using the method in section 2.1, and another 150 segments sampled using the RandWin

Table 1 Description of sensors

Sensor No.	Location	Orientation
1	ground floor, west side	North/South
2	ground floor, center	East/West
3	ground floor, east side	North/South
4	1st floor, west side	North/South
5	1st floor, center	East/West
6	1st floor, east side	North/South
7	2nd floor, west side	North/South
8	2nd floor, center	East/West
9	2nd floor, west side	North/South
10	3rd floor, east	North/South
11	3rd floor, center	East/West
12	3rd floor, west side	North/South
13	4th floor, west side	North/South
14	4th floor, west side	East/West
15	4th floor, west side	North/South

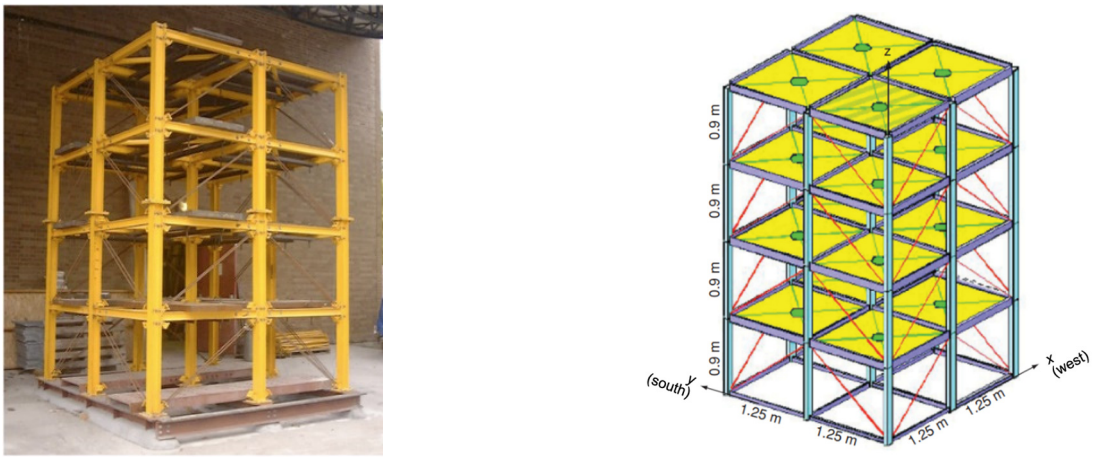


Fig 4 ASCE SHM Benchmark steel frame model

Table 2 Description of test cases

Test cases	Damage description
1	Undamaged structure
2	First floor braces in one bay on southeast corner removed
3	First & fourth floor braces in one bay on southeast corner removed
4	All east side braces removed
5	Case 4 + bolts loosened on all floors at both ends of beam on east face, north side

function. Using the network structure described in section 2.3, the network is trained with case 1 dataset. The learning rate is 0.0005 and the training epoch is 10000.

3.3 Result and discussion

3.3.1 t-SNE result

A 400-long acceleration segment can be seen as a point in 400-dimensional space. The t-SNE method in section 2.4.2 is used to map the benchmark data segment into 2-dimensional vector. The vectors of sensors 1 to 15 are separately drawn in Fig. 5 and colored with different test cases.

It is obvious that the collected vibration data can locally reflect the damage state in the corresponding monitor location. In terms of an independent monitor location, the structural response at different times shares similar intrinsic features under the same operation state. If the physical feature of the structure (such as damping ratio, stiffness, mass, and geometry) alters, the structural vibration response and intrinsic features at that location will also change due to local performance degradation. Accordingly, the location of each vector (mapping from high-dimensional space) will deviate. The more severe damage introduced to the structure, the further these vectors' locations deviate.

According to the space distribution of vibration data collected from each monitoring station in different damage states shown in Fig. 5, structural response data under Case 2 (i.e., the slight damage model) has similar feature to Case 1 (i.e., the undamaged model). Some of the data collected from monitoring stations are close to the damage (such as sensor-4 and 6 stations) and those in higher stories (such as sensor-9, 12 and 15 stations) can be nonlinearly classified clearly between two damage cases. In addition, data collected in Cases 3 to 5 have a clear clustering boundary when compared to data collected in Cases 1 to 2, triggering pronounced distinctions between damage model cases.

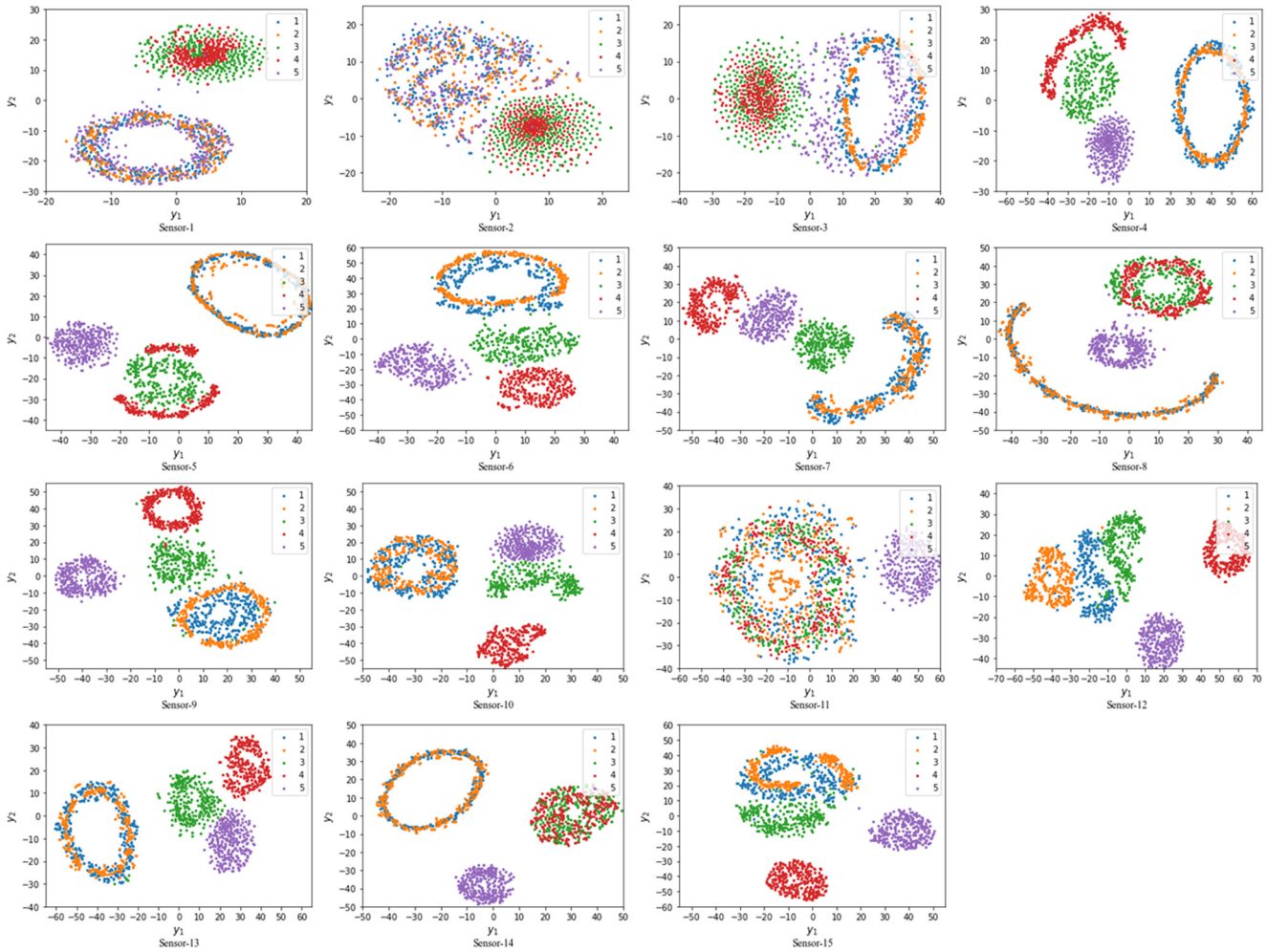


Fig. 5 Low-dimensional mapping of dataset in 15 sensors using t-SNE

In a nutshell, 1) monitored response data from same damage cases have a similar embedding feature; 2) the Euclidean space distribution of the experiment data changes as structural damage severity increases; and 3) the sensitivity of each case's response data increases as the distances between the damaged region and monitoring station increases. Based on these findings, the deep learning method can extract features from healthy state data and clarify damaged data with different nonlinear physical changes.

3.3.2 CAE-based data reconstruction result

After preprocessing the dataset of Cases 1 to 5, the convolutional auto-encoder network with Case 1 dataset is trained to learn the features of undamaged structure and test the reconstruction assessment indicators (MSE and Cor) of each cases' dataset. The average MSE and Cor between input and output datasets are shown in Tables 2 and 3. In general, structural response datasets collected from each monitoring station under Case 1 can be accurately reconstructed by the optimized CAE network. Therefore, the network can compress and decompress the undamaged response data. To reveal the relationship between two indicators and signal reconstruction patterns, four typical original signals and corresponding CAE reconstructed signals are chosen, which are plotted in Fig. 6. A Low MSE and a High Cor value indicate that the reconstruction ability is excellent. Detailed discussions of the CAE experiment results are as follows.

(1) As the damage state develops, the reconstruction error for vibration response datasets increases, which meets the conclusion about the Euclidean space distribution of vibration datasets under different damage cases.

(2) MSE represents the deviation and discreteness between the original and reconstructed data, whereas Cor represents the linear correlation between them, revealing the precision of the intrinsic features extracted by CAE. Noticeably, some of the response datasets have less apparent Cor changes than MSE when the test case develops from the undamaged case to damaged Cases 4 and 5. In other words, data from some cases with a high MSE value also maintain a high Cor value. As shown in Fig. 6, though the second and third vibration curves have similar MSE values (i.e., 0.5542 and 0.5266), the difference in Cor value between the two curves is relatively large. In terms of CAE reconstruction performance, the former reconstructed data have more discrete error but maintain a good linear relation, whereas the latter are more likely to comprise overfitted nonlinear details but maintain less discrete error. Therefore, both MSE and Cor should be considered when quantifying the completeness for the undamaged feature of the test data.

(3) The average MSE and Cor of the 15 sensors datasets change as the damage state progresses. When the test structure encounters slight damage, the discreteness of the reconstructed data may increase sharply, but its linear trends remain relatively complete. As the damage worsens, CAE can detect fewer undamaged physical features, resulting in more apparent variation on reconstructed data.

Therefore, to quantify the structural damage, Score (MSE, Cor) is defined. Steps for CAE-based multi-station monitoring and vibration damage assessment are as follows. First, structural vibration response data are collected from each station. After preprocessing the dataset, the data from each station are input into the CAE network for feature extraction and reconstruction. The MSE and Cor values are calculated after comparing the input and output data. Finally, the Score value is obtained from each distributed sensor. The average Score from sensors 1 to 15 is used to comprehensively assess test structure. The Scores of the Benchmark model in Cases 1 to 5 are proved to be coincident with the initially defined damage state, as shown

Table 3 Average MSE values of the reconstructed data

Sensor No.	Case 1	Case 2	Case 3	Case 4	Case 5
1	0.1297	0.1645	0.4667	0.5521	0.1966
2	0.2259	0.2540	0.4895	0.5331	0.2687
3	0.1023	0.1839	0.4731	0.5617	0.3824
4	0.0803	0.1729	0.5519	0.6115	0.6372
5	0.0829	0.0886	0.5554	0.5717	0.6212
6	0.0656	0.2611	0.3899	0.5575	0.6035
7	0.1020	0.2434	0.5304	0.5497	0.5379
8	0.0447	0.0632	0.5749	0.5726	0.6488
9	0.0957	0.3642	0.3529	0.4717	0.5770
10	0.0788	0.1167	0.3383	0.4044	0.4811
11	0.0933	0.1857	0.1845	0.1699	0.8875
12	0.1285	0.2086	0.1890	0.4221	0.5819
13	0.1123	0.1878	0.5065	0.5678	0.5726
14	0.0642	0.0966	0.5538	0.5762	0.6521
15	0.0843	0.5321	0.2584	0.4070	0.4921
MSE_{avg}	**0.0994**	**0.2082**	**0.4277**	**0.5019**	**0.5427**

Table 4 Average Cor values of the reconstructed data

Sensor No.	Case 1	Case 2	Case 3	Case 4	Case 5
1	99.15	98.93	93.97	85.76	98.53
2	97.45	96.79	86.45	82.24	96.43
3	99.46	99.46	93.35	83.31	96.41
4	99.67	99.70	92.98	85.63	69.80
5	99.66	99.57	92.07	90.98	73.19
6	99.78	99.84	96.92	87.36	76.53
7	99.47	99.44	93.21	90.25	85.48
8	99.90	99.84	81.56	81.68	65.55
9	99.53	99.47	97.39	93.21	78.84
10	99.71	99.33	96.89	93.25	86.37
11	99.55	98.90	99.09	99.05	65.19
12	99.18	98.15	98.25	92.45	78.16
13	99.36	99.38	94.26	86.56	81.49
14	99.79	99.63	86.35	82.49	63.44
15	99.66	99.43	98.23	93.40	86.57
Cor_{avg}	**99.42**	**99.19**	**93.40**	**88.51**	**80.13**

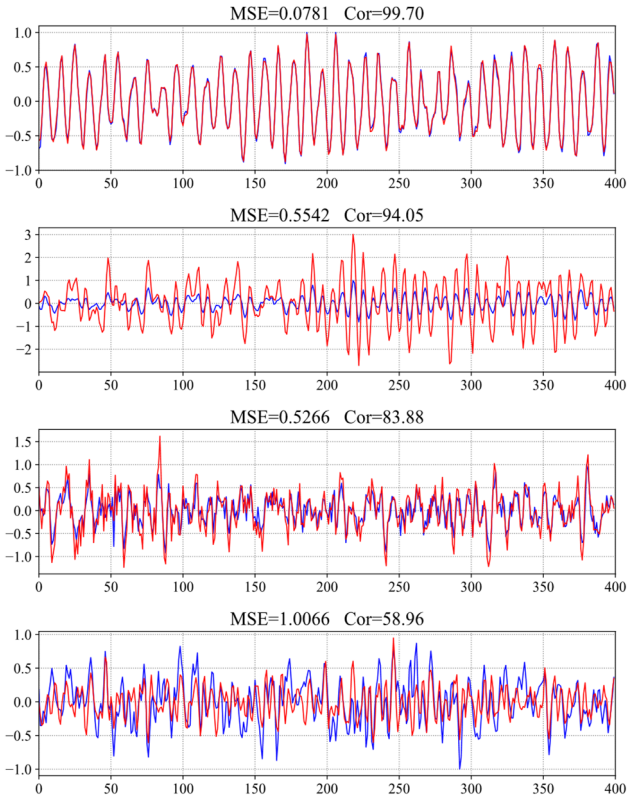


Fig. 6 Vibration patterns of different MSE and Cor values

in Fig. 7. According to the damage description in Table 2, the undamaged structure can maintain an average Score of 90.03. After removing the first-floor braces in one bay on the southeast corner, the Score dropped to 79.61 on average. With more braces removed, the stiffness of the steel frame changes evidently, and the Score decreases to 49.66 after all the east side braces are removed. In addition, the loosening of bolts at the beams ends also negatively affected the Score value to 42.54.

4. Demonstrating study – a rubber bearing-isolated structure

The proposed unsupervised deep learning method is applied in a rubber bearing isolated gymnasium to identify the earthquake and daily vibration damage, and the

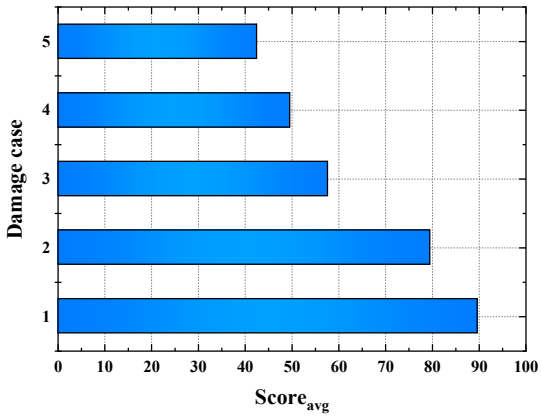


Fig. 7 Average scores in five damage cases

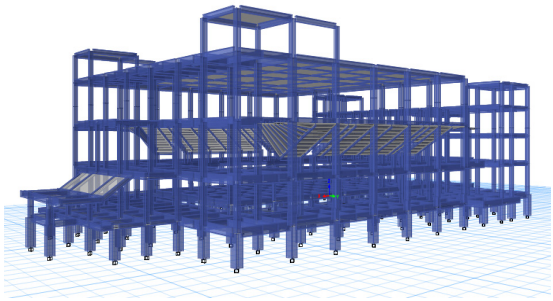
performance of the DL network is tested and verified.

4.1 Introduction to the rubber bearing-isolated gymnasium

The gymnasium is located in Suqian, Jiangsu province, China, and its main structure is a 7-story reinforced concrete space frame. It is 85.1 m long and 51.2 m wide. The building roof is a steel spatial grid. The building was constructed in 2001, when the Chinese Code for Seismic Design of Buildings (CCSDB) stipulated that the seismic fortification intensity of the building site was 8 degrees. However, due to the influence of the fracture zone, the revised 2010 CCSDB has adjusted the seismic fortification intensity here to 8.5 degrees. That is, the original structure form is insufficient to resist strong ground motions. Therefore, the base-isolation technology has been applied to improve its seismic performance. Fig. 8(a) shows the photos of the gymnasium with viscous dampers and rubber bearings. A 3D model of the gymnasium is developed using the ETABS finite element program, as shown in Fig. 8(b).

Fig. 9 shows the arrangements of the rubber bearings on the ground floor. Each column is cut in half to sandwich the bearing between the top column and bottom halves. Besides, the viscous dampers are placed near the corner columns to control the displacement of the isolating layer in the X and Y directions. The mechanical parameters of the bearings are shown in Table 5. The damping exponent α and the coefficient C of the viscous damper are 0.25 and

(a) Photos of the isolated gymnasium



(b) Structural analysis model

Fig. 8 Photos and 3D model of the gymnasium

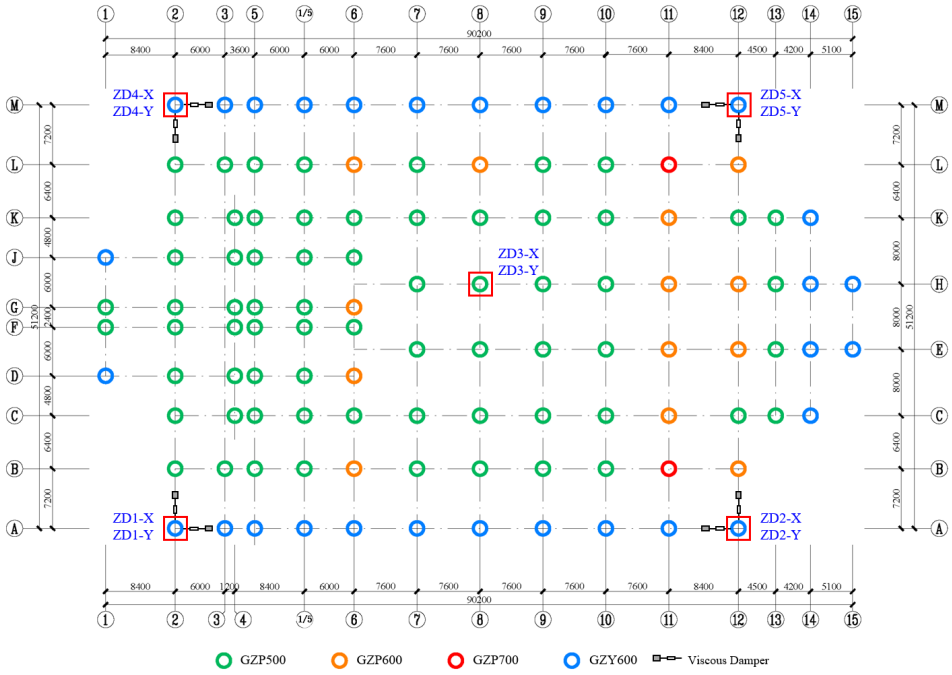


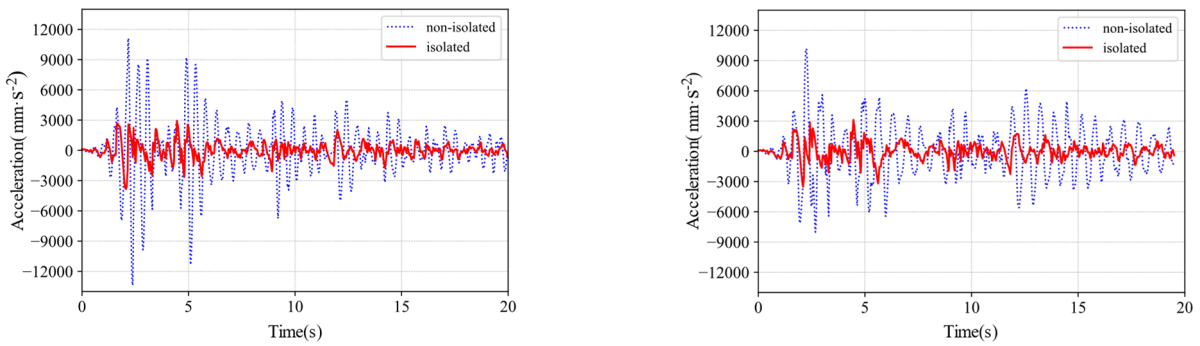
Fig. 9 Layout of bearings and accelerometers in the isolating layer

Table 5 Mechanical parameters for bearings

Item	Rubber bearing			Lead-rubber bearing
	GZP500	GZP600	GZP700	GZY600
Bearing diameter (mm)	500	600	700	600
Lead core diameter (mm)	/	/	/	120
Equivalent stiffness (kN/m)	767	921	1078	1703
Lead core stiffness (kN/m)	/	/	/	11951
Yield force (kN)	/	/	/	96

300(kN·m/s), respectively. As shown in Fig. 10, the peak acceleration response at the top story of the structure FE model excited by El-Centro ground motion significantly decreases after seismic isolation, for the X-direction peak acceleration decreases from over 13377 mm/s^2 to 3843 mm/s^2 and the Y-direction peak acceleration decreases from over 10183 mm/s^2 to 3524 mm/s^2. Due to the isolation effect, the three fundamental periods of the structure increase from T_1 = 0.758 s, T_2 = 0.661 s and T_3 = 0.644 s to T_1 = 2.279 s, T_2 = 2.265 s and T_3 = 2.142 s.

According to the ground motion filtration theory, which considers the seismic-isolation as a process of wave filtering, the original earthquake ground motion may attenuate into a relatively less-destructive motion by the isolation layer. Therefore, to maintain the normal in-service performance of the isolation devices on the ground floor, 10



(a) X-direction (b) Y-direction

Fig. 10 Acceleration curves of non-isolated and isolated structures

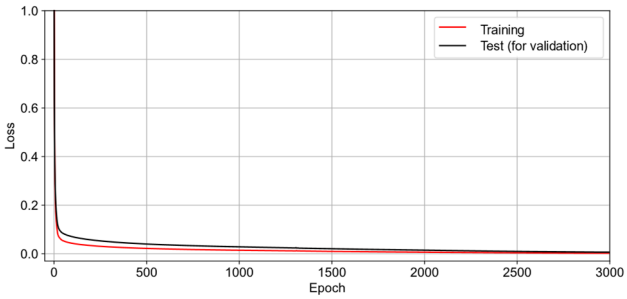


Fig. 11 Loss curve of the training process

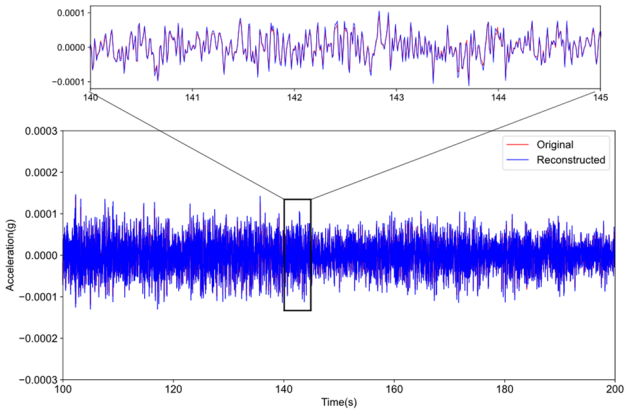


Fig. 12 Original and reconstructed vibrations

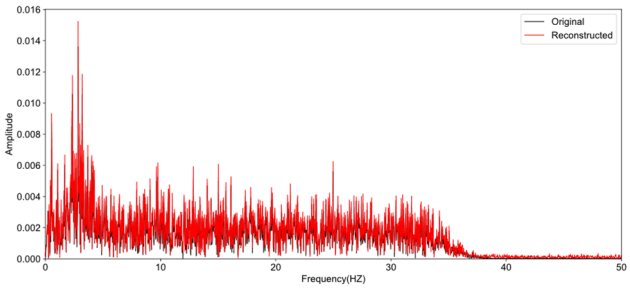


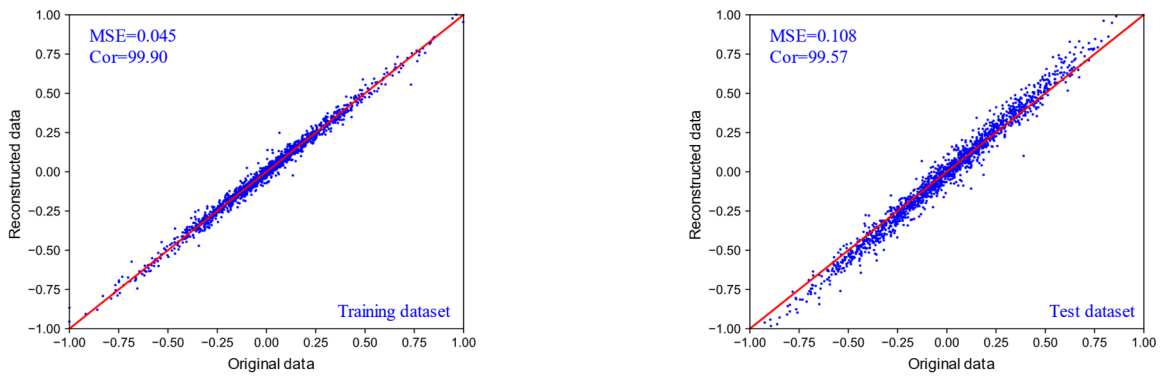
Fig. 13 Spectra of original and reconstructed vibrations

accelerometers have been installed above the bearings. Fig. 9 shows the sensor layout, which measure vibrations in two horizontal directions at five points. All measured acceleration signals are recorded at a sampling frequency of 100 Hz, and these real-time signals are analyzed by the structural assessment system.

4.2 Network training and test with daily vibrations

In this case, the acceleration data collected on August 16, 2021 from the ZD1 measuring point in the X-direction was used to train and test the CAE network designed in Section 2.3. Vibrations from 14:00 to 15:00 are used to generate the training dataset, while vibrations from 20:00 to 21:00 are used to generate the test dataset. Both datasets are generated using the preprocess method in Section 2.1, which divides vibration data in each hour into 900 segments using the RandWin function. Train the network until the training loss is less than 0.05. The loss curves for the training and test datasets are shown in Fig. 11. The loss value of the test dataset for verification keeps the same downward trend as the loss value of the training dataset. Fig. 12 compares the original and reconstructed vibration signals in 20:00 and Fig. 13 depicts the Fourier spectral of the two vibration signals. The features of the signal are accurately extracted and recovered in both the time domain and frequency domains.

Furthermore, a statistical analysis is conducted to compare the reconstruction error between the training and test datasets. Fig. 14 presents the input and the output normalized data of the CAE network. The reconstructed accuracy of the test dataset is more discrete than that of the training dataset, which is also reflected in MSE and Cor. The training dataset's MSE and Cor are 0.045 and 99.90, respectively, while the test dataset's values are 99.57 and 0.108, respectively. Given the exact data reconstruction ability of the CAE network, it is ideal to use the network to learn the intrinsic feature of the daily monitored data. Under operational conditions, the monitoring system may randomly sample newly collected data to update network parameters, thereby mitigating the effect of the environment and temperature. The network can then be practiced on earthquake detection and assessment.



(a) Training dataset (b) Test dataset

Fig. 14 Statistical analysis of data reconstructed by CAE network

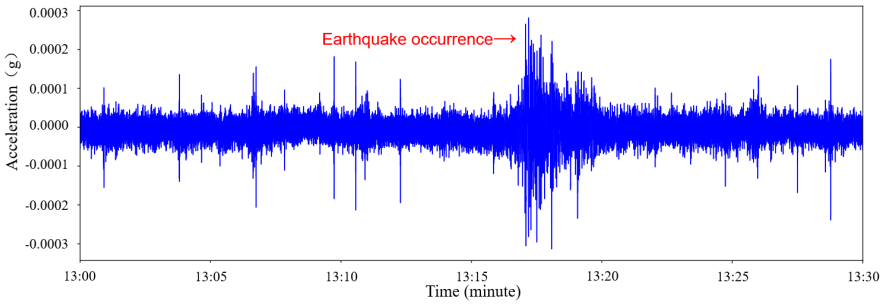


Fig. 15 Acceleration of the earthquake events

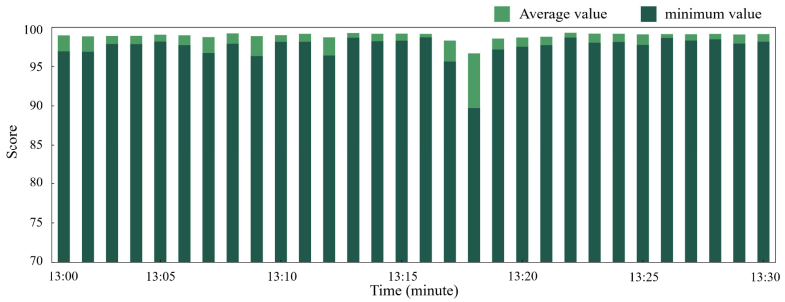


Fig. 16 Safety score of the acceleration

4.3 Earthquake event detection and assessment

4.3.1 Earthquake records and assessment results

On October 24, 2021, at around 13:11, a magnitude-6.2 earthquake with a depth of 66.8 km was occurred in Yilan County, Taiwan. The ground motion also spread to many coastal provinces on the Chinese mainland. The vibration monitoring and assessment system in the Suqian's gymnasium also recorded the response of the earthquake event and represented the variation in its scoring program. Fig. 15 depicts the vibration signal collected by the X-direction sensor in ZD1 measuring point from 13:00 to 13:30 on October 24, 2021, and Fig. 16 is a bar chart that represents the mean and minimum score value in each minute during that time.

4.3.2 Discussion

According to the vibration plotted in Fig. 15 and score result in Fig. 16, the patterns of acceleration change clearly when the earthquake occurs, and the instant score value also decreases as the ground motion develops. There was a 5-minute time lag between the reported earthquake time in Taiwan and the time identified by CAE, which was caused by the wave travelling from the Circum-Pacific seismic belt to Suqian. The score fluctuated significantly between 13:15 and 13:25, which decreased from over 99 to 96.68. The instantaneous score touched the lowest point at 89.72, satisfying the fluctuation characteristic of the vibrations.

When the ground motion arrived, the intrinsic features of the acceleration changed drastically when compared to the normal in-service accelerations used to train the CAE network. Without enough information for feature extraction and reconstruction, the network is hard to picture the wave based on previously learned experiences. During daily operation, the safety score of all sensors fluctuates between 90 and 100, and the accelerometers above the bearing solators always record stationary signals generated by ambient vibration. When the earthquake arrives, the original ambient signals will become nonstationary. Some nonlinear features caused by ground motion and the dynamic response of the lead rubber bearings trigger a downward trend in the safety score. The local ground motion signal has a high impact on the safety score, which decreases as the magnitude, peak acceleration values and ground soil reactions changes. In this case, the weak earthquake motion elicited a remarkable but limited change in the CAE reconstruction value. Therefore, when the structure receives a score of 90 or higher, it is considered to be in the safety state, and when it receives a score between 80 and 90, it is considered to be in the slight damaged state.

In addition, the gymnasium is located in the downtown of the city, where the surrounding noise level is relatively high. Some unexpected acceleration changes, peaking at 0.0002g, were also recorded. These are mainly triggered by the outside pass-by effect of heavy trucks.

5. Conclusions

A time-series signal feature extraction and reconstruction method using one-dimensional convolutional autoencoder (CAE) based unsupervised deep learning technology is proposed to identify and assess the damage directly from responses of the structure. A preprocessing method is regulated to generate datasets for DL network training and verification from massive on-site monitoring vibrations. The proposed CAE network uses 1D-convolutional kernel to extract features from time-

correlated vibrations effectively and recover the input data adaptively with transposed 1D-convolutional kernel. Two indicators, MSE and Cor, are used to quantify the reconstructive ability of the network, enabling the SHM system to score the damage and earthquake events based on vibrations. The assessment performance of the network has been verified through two SHM test.

First, the experimental ambient vibration data from ASCE-Benchmark Phase II is used to test the damage identification performance of the CAE network. The test result shows that the unsupervised DL method can learn the features of a healthy structure just from intact vibrations to optimize the network parameters. The optimized network can distinctively reconstruct data from different damage cases that are distributed nonlinearly in Euclidean space. Two indicators are ideal for quantifying different damage cases in which the assessment result meets the designed damage state on the experimented steel frame.

Additionally, the CAE network is practiced on an earthquake health monitoring system in a rubber bearing-isolated gymnasium located in Suqian, China. According to the test on it, with the increase of the number of iterations, the loss function of the test dataset with 900 segments and the training dataset with 900 segments dropped and became stable synchronously, which consequently revealed that the perfectly trained network can distill features from newly collected daily accelerations data. The vibration records of an earthquake event on October 24, 2021 also demonstrated that the method can precisely detect the occurrence of the ground motion and represent its analysis on a score value.

In conclusion, the above result demonstrates the reliable performance of the proposed unsupervised DL approach, which can identify different damages and detect disaster events from unlabeled vibrations in the SHM system under real-engineering conditions. In the future work, multi-layer input data-based CAE will be considered to fit more complicated structure, such as long-span bridges and spatial buildings.

Acknowledgments

The authors gratefully acknowledge the Ministry of Science and Technology of the People's Republic of China (No.2018YFE0206100). The research was financially supported by the Jiangsu Provincial Department of Science and Technology under (No.BE2019107).

References

Abdeljaber, O., Avci, O., Kiranyaz, S., Gabbouj, M. and Inman, D. J. (2017), "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

Bayissa, W.L. and Haritos, N. (2007), "Structural damage identification in plates using spectral strain energy analysis", *J. Sound Vib.*, **307**(1-2), 226-249. https://doi.org/10.1016/j.jsv.2007.06.062

Cha, Y.J. and Choi, W. (2017), "Vision-based concrete crack detection using a convolutional neural network", In: *Dynamics of Civil Structures*, Springer, Volume 2, Cham, pp. 71-73.

Cheng-Zhong, Q. and Xu-Wei, L. (2012), "Damage identification for transmission towers based on HHT", *Energy Procedia*, **17**, 1390-1394. https://doi.org/10.1016/j.egypro.2012.02.257

Cui, M., Wu, G., Chen, Z., Dang, J., Zhou, M. and Feng, D. (2021), "Geometric attention regularization enhancing convolutional neural networks for bridge rubber bearing damage assessment", *J. Perform. Constr. Facil.*, **35**(5), 04021061. https://doi.org/10.1061/(ASCE)CF.1943-5509.0001634

Das, S. and Saha, P. (2018), "Structural health monitoring techniques implemented on IASC–ASCE benchmark problem: a review", *J. Civil Struct. Health Monitor.*, **8**(4), 689-718. https://doi.org/10.1007/s13349-018-0292-5

Doebling, S.W., Farrar, C.R. and Prime, M.B. (1998), "A summary review of vibration-based damage identification methods", *Shock Vib. Digest*, **30**(2), 91-105.

Erdogan, Y.S., Gul, M., Catbas, F.N. and Bakir, P.G. (2014), "Investigation of uncertainty changes in model outputs for finite-element model updating using structural health monitoring data", *J. Struct. Eng.*, **140**(11), 04014078. https://doi.org/10.1061/(ASCE)ST.1943-541X.0001002

Flah, M., Nunez, I., Ben Chaabene, W. and Nehdi, M.L. (2021), "Machine learning algorithms in civil structural health monitoring: a systematic review", *Arch. Computat. Methods Eng.*, **28**(4), 2621-2643. https://doi.org/10.1007/s11831-020-09471-9

Gentile, C., Ruccolo, A. and Canali, F. (2019), "Continuous monitoring of the Milan Cathedral: dynamic characteristics and vibration-based SHM", *J. Civil Struct. Health Monitor.*, **9**(5), 671-688. https://doi.org/10.1007/s13349-019-00361-8

Goyal, D. and Pabla, B.S. (2016), "The vibration monitoring methods and signal processing techniques for structural health monitoring: a review", *Arch. Computat. Methods Eng.*, **23**(4), 585-594. https://doi.org/10.1007/s11831-015-9145-0

He, W.Y., Zhu, S. and Ren, W.X. (2018), "Progressive damage detection of thin plate structures using wavelet finite element model updating", *Smart Struct. Syst., Int. J.*, **22**(3), 277-290. https://doi.org/10.12989/sss.2018.22.3.277

Hera, A. and Hou, Z. (2004), "Application of wavelet approach for ASCE structural health monitoring benchmark studies", *J. Eng. Mech.*, **130**(1), 96-104. https://doi.org/10.1061/(ASCE)0733-9399(2004)130:1(96)

Hsu, T.Y., Liu, C.Y., Hsieh, Y.M. and Weng, C.T. (2021), "Post-earthquake fast building safety assessment using smartphone-based interstory drifts measurement", *Smart Struct. Syst., Int. J.*, **29**(2), 287-299. https://doi.org/10.12989/sss.2022.29.2.287

Jalali, M.H. and Rideout, D.G. (2022), "Substructural damage detection using frequency response function based inverse dynamic substructuring", *Mech. Syst. Signal Process.*, **163**, 108166. https://doi.org/10.1016/j.ymssp.2021.108166

Jiang, K., Han, Q., Du, X. and Ni, P. (2021), "A decentralized unsupervised structural condition diagnosis approach using deep auto-encoders", *Comput.-Aided Civil Infrastr. Eng.*, **36**(6), 711-732. https://doi.org/10.1111/mice.12641

Johnson, E.A., Lam, H.F., Katafygiotis, L.S. and Beck, J.L. (2004), "Phase I IASC-ASCE structural health monitoring benchmark problem using simulated data", *J. Eng. Mech.*, **130**(1), 3-15. https://doi.org/10.1061/(ASCE)0733-9399(2004)130:1(3)

Khayatazad, M., Honhon, M. and De Waele, W. (2022), "Detection of corrosion on steel structures using an artificial neural network", *Struct. Infrastr. Eng.*, 1-12. https://doi.org/10.1080/15732479.2022.2069272

LeCun, Y., Bengio, Y. and Hinton, G. (2015), "Deep learning", *Nature*, **521**(7553), 436-444. https://doi.org/10.1038/nature14539

Leon-Medina, J.X., Anaya, M., Pozo, F. and Tibaduiza, D. (2020), "Nonlinear feature extraction through manifold learning in an electronic tongue classification task", *Sensors*, **20**(17), 4834. https://doi.org/10.3390/s20174834

MHURD-PRC (2010), Code for Seismic Design of Buildings, China Architecture & Building Press.

Na, S., Heo, S., Han, S., Shin, Y. and Roh, Y. (2022), "Acceptance Model of Artificial Intelligence (AI)-Based Technologies in Construction Firms: Applying the Technology Acceptance Model (TAM) in Combination with the Technology–Organisation–Environment (TOE) Framework", *Buildings*, **12**(2), 90. https://doi.org/10.3390/buildings12020090

Pan, H., Azimi, M., Yan, F. and Lin, Z. (2018), "Time-frequency-based data-driven structural diagnosis and damage detection for cable-stayed bridges", *J. Bridge Eng.*, **23**(6), 04018033. https://doi.org/10.1061/(ASCE)BE.1943-5592.0001199

Paszke, A., Gross, S., Massa, F., Lerer, A., Bradbury, J., Chanan, G., Killeen, T., Lin, Z., Gimelshein, N., Antiga, L. and Desmaison, A. (2019), "Pytorch: An imperative style, high-performance deep learning library", *Adv. Neural Inform. Process. Syst.*, **32**. https://proceedings.neurips.cc/paper/2019/hash/bdbca288fee7f92f2bfa9f7012727740-Abstract.html

Qu, C.Z. and Lian, X.W. (2012), "Damage identification for transmission towers based on HHT", *Energy Procedia*, **17**, 1390-1394. https://doi.org/10.1016/j.egypro.2012.02.257

Reynolds, P. and Pavic, A. (2003), "Effects of false floors on vibration serviceability of building floors. I: Modal properties", *J. Perform. Constr. Facil.*, **17**(2), 75-86. https://doi.org/10.1061/(ASCE)0887-3828(2003)17:2(75)

Smarra, F., Girolamo, G.D.D., Gattulli, V., Graziosi, F. and D'Innocenzo, A. (2020), "Learning models for seismic-induced vibrations optimal control in structures via random forests", *J. Optimiz. Theory Applicat.*, **187**(3), 855-874. https://doi.org/10.1007/s10957-020-01698-7

Tiachacht, S., Bouazzouni, A., Khatir, S., Wahab, M.A., Behtani, A. and Capozucca, R. (2018), "Damage assessment in structures using combination of a modified Cornwell indicator and genetic algorithm", *Eng. Struct.*, **177**, 421-430. https://doi.org/10.1016/j.engstruct.2018.09.070

Tibaduiza Burgos, D.A., Gomez Vargas, R.C., Pedraza, C., Agis, D. and Pozo, F. (2020), "Damage identification in structural health monitoring: A brief review from its implementation to the use of data-driven applications", *Sensors*, **20**(3), 733. https://doi.org/10.3390/s20030733

Wang, N., Zhao, Q., Li, S., Zhao, X. and Zhao, P. (2018a), "Damage classification for masonry historic structures using convolutional neural networks based on still images", *Comput.-Aided Civil Infrastr. Eng.*, **33**(12), 1073-1089. https://doi.org/10.1111/mice.12411

Wang, Y., Thambiratnam, D.P., Chan, T.H.T. and Nguyen, A. (2018b), "Damage detection in asymmetric buildings using vibration-based techniques", *Struct. Control Health Monitor.*, **25**(5), e2148. https://doi.org/10.1002/stc.2148

Wang, Z.C., Ren, W.X. and Chen, G. (2018c), "Time–frequency analysis and applications in time-varying/nonlinear structural systems: a state-of-the-art review", *Adv. Struct. Eng.*, **21**(10), 1562-1584. https://doi.org/10.1177/1369433217751969

Yan, A.M., Kerschen, G., De Boe, P. and Golinval, J.C. (2005), "Structural damage diagnosis under varying environmental conditions—part II: local PCA for non-linear cases", *Mech. Syst. Signal Process.*, **19**(4), 865-880. https://doi.org/10.1016/j.ymssp.2004.12.003

Yang, X.M., Yi, T.H., Qu, C.X., Li, H.N. and Liu, H. (2020), "Modal identification of high-speed railway bridges through free-vibration detection", *J. Eng. Mech.*, **146**(9), 04020107. https://doi.org/10.1061/(ASCE)EM.1943-7889.0001847

Yi, T.H., Yao, X.J., Qu, C.X. and Li, H.N. (2019), "Clustering number determination for sparse component analysis during output-only modal identification", *J. Eng. Mech.-ASCE*, **145**(1), 04018122. https://doi.org/10.1061/(ASCE)EM.1943-7889.0001557