Smart Structures and Systems, *Vol. 31*, *No. 1* (*2023*) *45-56*
https://doi.org/10.12989/sss.2023.31.1.045

Automated structural modal analysis method using long short-term memory network

Jaehyung Park [1a], Jongwon Jung [1b], Seunghee Park [2c] and Hyungchul Yoon[*1]

[1] *Department of Civil Engineering*, *Chungbuk National University*,
1 Chungdae-ro, *Seowon-gu*, *Cheongju-si*, *Chungcheongbuk-do*, *28644 Republic of Korea*
[2] *School of Civil*, *Architectural Engineering & Landscape Architecture*, *Sungkyunkwan University*,
2066 Seobu-ro, *Jangan-gu*, *Suwon-si*, *Gyeonggi-do*, *16419 Republic of Korea*

(*Received December 28*, *2021*, *Revised August 10*, *2022*, *Accepted August 30*, *2022*)

Abstract. Vibration-based structural health monitoring is used to ensure the safety of structures by installing sensors in structures. The peak picking method, one of the applications of vibration-based structural health monitoring, is a method that analyze the dynamic characteristics of a structure using the peaks of the frequency response function. However, the results may vary depending on the person predicting the peak point; further, the method does not predict the exact peak point in the presence of noise. To overcome the limitations of the existing peak picking methods, this study proposes a new method to automate the modal analysis process by utilizing long short-term memory, a type of recurrent neural network. The method proposed in this study uses the time series data of the frequency response function directly as the input of the LSTM network. In addition, the proposed method improved the accuracy by using the phase as well as amplitude information of the frequency response function. Simulation experiments and lab-scale model experiments are performed to verify the performance of the LSTM network developed in this study. The result reported a modal assurance criterion of 0.8107, and it is expected that the dynamic characteristics of a civil structure can be predicted with high accuracy using data without experts.

Keywords: deep learning; long short-term memory (LSTM); modal analysis; structural health monitoring; system identification

1. Introduction

Many civil infrastructures around the world have been constructed more than 50 years ago, and several issues have been reported in these structures because of deterioration. The West Seattle Bridge in Seattle, United States, was closed on March 23, 2020, owing to a crack in the girder caused by deterioration (Lindblom 2020). The Hammersmith Bridge, located over River Thames in the United Kingdom, was closed on April 10, 2019, owing to corrosion (Collins *et al*. 2019). On August 1, 2007, bridge 9340 in Minneapolis, Minnesota, collapsed because of deterioration, causing 13 deaths and 150 injuries (Haberman 2014). On August 14, 2018, a 200-m section of the Morandi Bridge in Genova, Italy, collapsed for the same reason, causing dozens of vehicles to crash and 35 deaths (Harlan and Selk 2018). These accidents illustrate how the deterioration of bridges can wreak havoc on society.

At present, safety inspections for bridges are performed primarily through exterior inspection by experts. In Korea, the Special Act on the Safety Control and Maintenance of Establishments specifies periodic safety inspections (Special Act on the Safety Control and Maintenance of Establishments 2021). These inspections are conducted by a person with experience and expertise in examining the inherent risk factors in a particular establishment with the naked eye or using inspection tools. This inspection requires considerable time and cost, and the results of the inspection may be subjective.

One of the most popular methods for ensuring the integrity of a structure in an objective manner is vibration-based structural health monitoring (SHM). SHM diagnoses the integrity of the structure by measuring the vibration response of the structure, such as dis-placement and acceleration. Traditional SHM was developed using wired sensors such as linear variable displacement transducers (LVDTs) and accelerometers (Sun *et al*. 1995). To overcome the limitations of wired SHM systems, wireless SHM systems have been developed (Spencer *et al*. 2004, Lynch 2006, Shen *et al*. 2021). Further, unmanned aerial vehicles and cameras have recently been utilized to improve the efficiency of SHM systems (Basharat *et al*. 2005, Yoon *et al*. 2016, Khuc and Catbas 2017, Feng and Feng 2018, Yoon *et al*. 2018, Jung *et al*. 2019, Lydon *et al*. 2019, Spencer *et al*. 2019, Nasimi and Moreu 2021).

System identification is a method to determine the dynamic characteristics of a structural system by measuring the vibration response of the structure, such as displacement and acceleration response. The peak picking method is the

*Corresponding author, Ph.D., Associate Professor,
E-mail: hyoon@cbnu.ac.kr
[a] Graduate Student, E-mail: pjcomputer@naver.com
[b] Associate Professor, E-mail: jjung@chungbuk.ac.kr
[c] Professor, E-mail: shparkpc@skku.edu

Copyright © 2023 Techno-Press, Ltd.

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

most common approach for identifying structural systems. The peak picking method intuitively determines the natural frequency and mode shape by extracting the peak point from the frequency response function (FRF). However, the result of the peak picking method may be different for each person owing to the subjective judgment of the person selecting the peak point with the naked eye. To address this problem, in 1985, Juang and Pappa developed the eigen-system realization algorithm, which identifies the system using the singular value of the Hankel Matrix (Juang and Pappa 1985), and, in 2001, Bricker *et al*. developed a frequency domain decomposition algorithm that identifies a system using the singular value of the power spectral density (PSD) (Brincker *et al*. 2001). In addition, the Ibrahim time domain algorithm and the random decrement algorithm have also been developed (Ibrahim and Pappa 1982, Cole 1973). However, most of these system identification algorithms are not intuitive like peak picking methods or fully automated, requiring users to specify system orders or singular values in advance.

Recent technologies, including artificial intelligence and deep learning, may provide an opportunity to overcome the limitations of the peak picking method. Recently, several studies have applied deep learning techniques to resolve engineering problems in the construction field. For example, various methods have been developed for detecting damages using deep learning (Cha *et al*. 2017, Zhang *et al*. 2017, Nex *et al*. 2019, Ye *et al*. 2019, Deng *et al*. 2020, Lopez-Pacheco *et al*. 2020, Qi *et al*. 2021). In addition, Narazaki *et al*. (2020) and Lee et. al. (2020) introduced a system based on deep learning that can automatically classify bridge components, and Lim and Yoon (2022) introduced a vision-based dynamic load estimation for pedestrian bridges using deep learning.

Few studies have applied deep learning techniques to system identification. Kim and Sim developed an automated system for implementing the peak picking method using a convolutional neural network (CNN) (Kim and Sim 2019), and Jin *et al*. (2021) has introduced fully automated peak picking method using Faster R-CNN. However, the proposed method requires a specific format for the input data, and, thus, a preprocessing stage is required for scaling the FRF. In addition, the method uses only the amplitude of the FRF, whereas the phase of the FRF can provide crucial information as well.

To overcome the limitations of the previous study, this study aims to develop an automated peak picking method using long short-term memory (LSTM), an architecture of a recurrent neural network (RNN). The proposed method directly uses the sequential data of the FRF as the input of the LSTM network without requiring any preprocessing procedure. In addition, the proposed method aims to increase the accuracy by using the phase information as well as the amplitude of the FRF. A theoretical background for the peak picking method and LSTM network is briefly introduced, and the configuration of the proposed LSTM-based peak picking network is explained. Finally, the results for the simulation-based experiments as well as the lab-scale experiment are discussed.

2. Background

Long short-term memory is an architecture type of the RNN, which has a characteristic recurrent structure, where the output of the layer is reconnected to the input value of the same layer. RNNs are widely used for classifying sequential data such as voice recognition (Hochreiter and Schmidhuber 1997). The fundamental structure of an RNN network is illustrated in Fig. 1.

RNN has a vanishing gradient problem, which is a problem that occurs when a large amount of data is trained using the gradient descent method (Bengio *et al*. 1994). Several methods have been proposed to resolve the vanishing gradient problem. Hochreiter (1998) proposed an LSTM structure consisting of multiple gates, such as the cell state, forget gate, input gate, and output gate in a single-layer structure using only a hyperbolic tangent activation function to solve the vanishing gradient problem. Cho *et al*. (2014) proposed the structure of a gated recurrent unit that combines the forget and input gates of the LSTM as one gate and the cell state and hidden state as another. Both the structures of RNN solved the vanishing gradient problem such that decision-making for long sequential data could be processed without facing the vanishing gradient problem. In this study, the basic structure of the LSTM, illustrated in Fig. 2, was modified and used, and detailed explanations of the network will be discussed in a later section.

3. System development

An overview of the proposed system is presented in Fig. 3. First, using dynamic response data such as displacement or acceleration, the cross-power spectral density (CPSD) and Power Spectral Density (PSD) are calculated. Next, the Frequency Response Function (FRF) is calculated using the PSD and CPSD. The amplitude of the FRF, $|H_{xf}|$, and phase of the FRF, φ, can be obtained from the complex form of the FRF. These two parameters are considered as

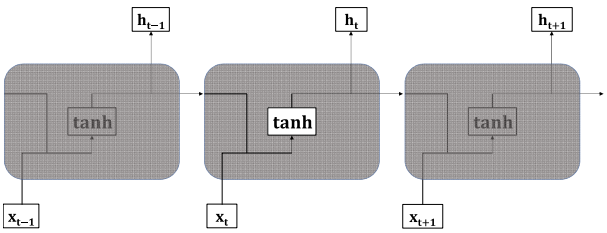

Fig. 1 Recurrent Neural Network (RNN) Module

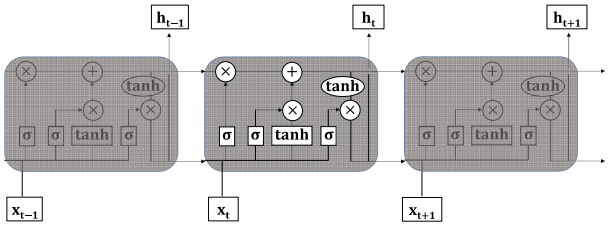

Fig. 2 Long Short-Term Memory (LSTM) Module

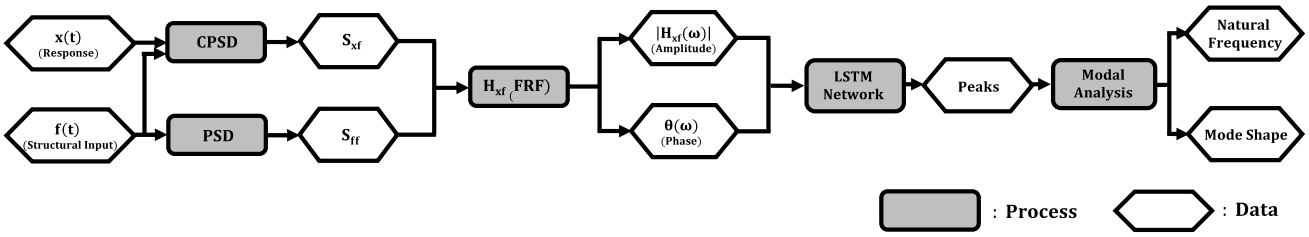

Fig. 3 System overview

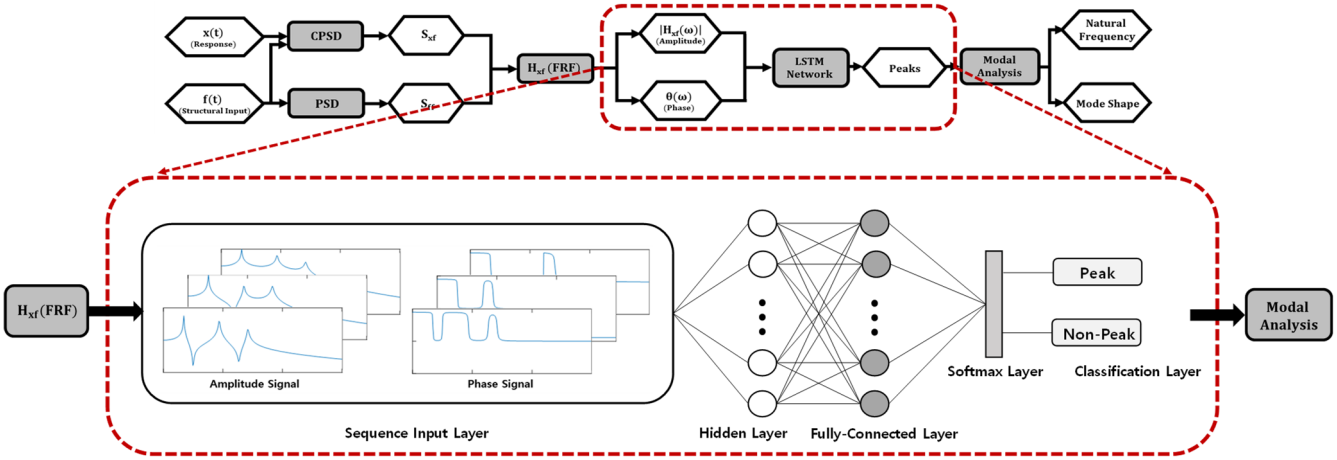

Fig. 4 LSTM network configuration

the input of the LSTM network, which calculates the location of the peak points. Finally, the natural frequencies and mode shapes are calculated.

3.1 Frequency response function

The theoretical FRF can be calculated from the equation of motion of the particular system using Eq. (1). However, the system identification method is a method used to estimate the dynamic system; thus, the FRFs are estimated using the dynamic response, not the numerical model. The dynamic response of the structure can be obtained by attaching sensors such as accelerometers and LVDTs to the structure or by capturing a video of the structural motion and using computer vision methods. Using excitation and the measured response, the auto-PSD of the excitation, $S_{ff}(\omega)$, and the CPSD between the input and output, $S_{xf}(\omega)$, can be calculated. Finally, the FRF of the system can be obtained by combining these two parameters, as in Eq. (2)

$$H_{xf}(\omega) = \frac{p}{k - m\omega^2 + ic\omega} \tag{1}$$

$$H_{xf}(\omega) = \frac{S_{xf}(\omega)}{S_{ff}(\omega)} \tag{2}$$

where p is the magnitude of the input, k is the stiffness of the structure, m is the mass of the structure, ω is the forcing frequency, c is the damping coefficient of the structure, and i is an indeterminate satisfying $i^2 = -1$. where S_{ff} is the auto-PSD of the input and S_{xf} is the CPSD between the input and output. If the structure is a multi-degree-of-freedom (MDOF) system with n DOFs, H_{x_if}, the FRF for the ith DOF can be expressed as Eq. (3)

$$H_{x_if}(\omega) = \frac{S_{x_if}(\omega)}{S_{ff}(\omega)}, \qquad (i = 1,2,\cdots,n) \tag{3}$$

$$\text{Amplitude of FRF}: \ \left|H_{xf}(\omega)\right| = \sqrt{\text{Imag}^2 + \text{Real}^2} \tag{4}$$

$$\text{Phase of FRF}: \quad \theta(\omega) = \tan^{-1}\left(\frac{\text{Imag}}{\text{Real}}\right) \tag{5}$$

where S_{x_if} is the CPSD between the input and x_i. The amplitude and phase of the FRFs can be obtained using Eqs. (4) and (5) and are used as the input data of the LSTM network. The length of sequential data in the FRF depends on the number of points used in the fast Fourier transform (NFFT). Although the total number of points used in the fast Fourier transform is NFFT, owing to the Nyquist rate, the valid length of the FRF sequential data becomes $\frac{1}{2}N_{\text{FFT}} + 1$.

3.2 LSTM network configuration

The configuration of the automated peak picking method using LSTM is presented in Fig. 4. First, the amplitude and phase information of the FRF are connected to the input layer in the form of sequential data. Next, these two pieces of information are delivered to the LSTM and fully connected layers, respectively. The calculated values

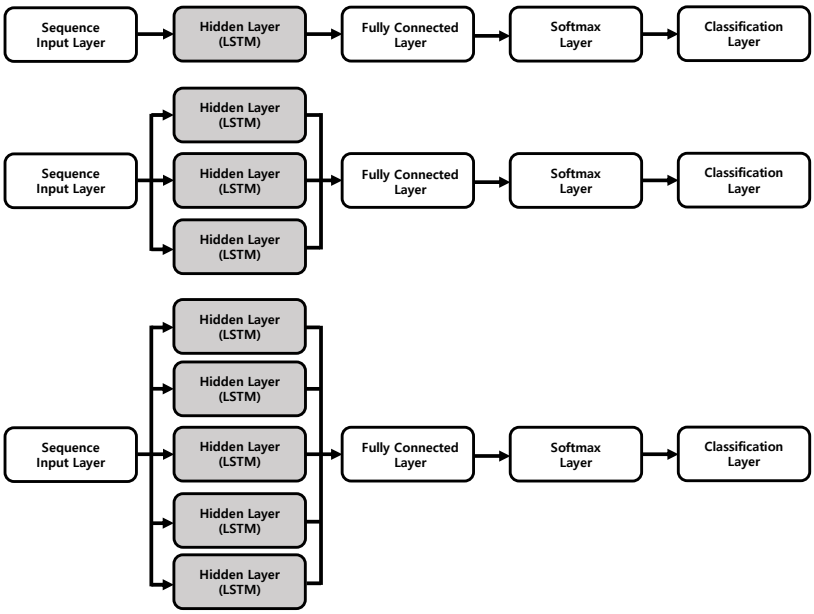

Fig. 5 Configuration of LSTM network used for comparison: (a) single LSTM; (b) three layers in parallel; (c) five layers in parallel

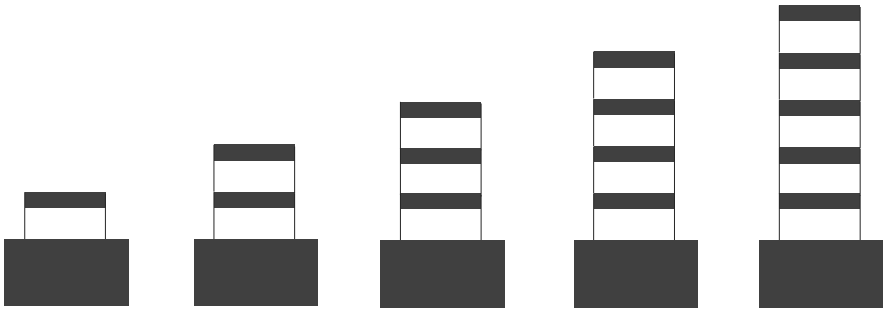

Fig. 6 Numerical models used for collecting the training data

are then processed in the softmax layer so that the sum of the values becomes 1, and the classification layer classifies each point as either a peak or a nonpeak point.

The performance of the LSTM network is highly dependent on the number of hidden units. The number of hidden units corresponds to the amount of information stored in between the sequences (hidden state). In general, as the number of hidden units increases, more complex models can be expressed; thus, the accuracy of the network improves. However, a network with a large number of hidden units requires more training time. Furthermore, if the number of hidden units is unconditionally large, the accuracy will no longer improve; the accuracy might decrease owing to the overfitting of the training data. Therefore, an appropriate number of hidden units should be chosen through trial and error. In this study, five different networks using different numbers of hidden units (20, 50, 100, 150, 200) were designed, and their performances were compared.

Even with the same number of hidden units, the performance of the LSTM network can vary according to the configuration of the LSTM (hidden) layers. The result obtained from the LSTM network can be different depending on the number of hidden layers. In this study,three different networks were constructed using different LSTM layouts, as illustrated in Fig. 5. The first layout uses a single hidden layer, the second uses three hidden layers in a parallel structure, and the third layout uses five layers in a parallel structure.

Numerical models were generated to collect training data for the automated peak picking method using LSTM. For the properties of the numerical model, random values between 3 and 4 kg for mass, 2,000 and 16,000 kN for stiffness, and 0.4% and 2% for damping ratio were used for each DOF. A band-limited white noise with a noise power of 0.2 was selected as the ground excitation. For each of the single DOF, 2DOF, ..., and 5DOF systems, 6,000 numerical models were generated, leading to a total of 30,000 models. For all the models, it was assumed that all the DOFs were fully observable.

The equation of motion for the model can be expressed as follows (6)

$$[\boldsymbol{M}]\ddot{x} + [\boldsymbol{C}]\dot{x} + [\boldsymbol{K}]x = f \tag{6}$$

where $[\boldsymbol{M}], [\boldsymbol{C}],$ and $[\boldsymbol{K}]$ are the mass, damping, and

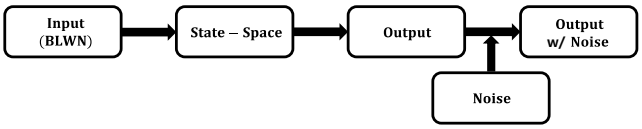

Fig. 7 Configuration of the dynamic system

stiffness matrices, respectively, of the nDOF system of size (n × n) and x and f are the displacement and load vectors of size (n × 1), respectively, as follows

$$[M] = \begin{bmatrix} m_{11} & m_{12} & \cdots & m_{1n} \\ m_{21} & m_{22} & \cdots & m_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ m_{n1} & m_{n2} & \cdots & m_{nn} \end{bmatrix},$$

$$[C] = \begin{bmatrix} c_{11} & c_{12} & \cdots & c_{1n} \\ c_{21} & c_{22} & \cdots & c_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ c_{n1} & c_{n2} & \cdots & c_{nn} \end{bmatrix},$$

$$[K] = \begin{bmatrix} k_{11} & k_{12} & \cdots & k_{1n} \\ k_{21} & k_{22} & \cdots & k_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ k_{n1} & k_{n2} & \cdots & k_{nn} \end{bmatrix}$$

$$x = \begin{Bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{Bmatrix}, \quad f = \begin{Bmatrix} f_1 \\ f_2 \\ \vdots \\ f_n \end{Bmatrix}$$

Once the equation of motion is generated, the state-space representation of the system can be obtained as follows

$$\dot{\boldsymbol{z}} = \boldsymbol{A}_s \cdot \boldsymbol{z} + \boldsymbol{B}_s \cdot f \tag{7}$$

$$\boldsymbol{y} = \boldsymbol{C}_s \cdot \boldsymbol{z} + \boldsymbol{D}_s \cdot f \tag{8}$$

where $z = \begin{Bmatrix} x \\ \dot{x} \end{Bmatrix}$ is a state-space vector and y is an observation vector composed of the values that the user wants to obtain as the output. A_s is a state matrix, B_s is an input matrix, C_s is an output matrix, and D_s is a feed through matrix. In this study, y is defined as $y = x$ because the displacement is the only information used in the proposed system. The response of the structure was calculated using the state-space model with Simulink, as illustrated in Fig. 7. Random noise was added to the output to represent the field data and extreme conditions.

3.3 Modal analysis

The proposed method performs modal analysis using the peak points estimated from the LSTM networks. The peak picking method is an approach used to estimate the natural frequency and mode shape by extracting the peak point from the FRF of the particular system. The FRF is a quantitative measure of the output spectrum of a dynamic system with respect to the input. An FRF for an SDOF system can be expressed as follows

$$H_{xf}(\omega) = \frac{p}{k - m\omega^2 + ic\omega} \tag{9}$$

where p is the magnitude of the input, k is the stiffness of the structure, m is the mass of the structure, ω is the forcing frequency, c is the damping coefficient of the structure, and i is an indeterminate satisfying $i^2 = -1$. The amplitude and phase of an FRF, respectively, $\left|H_{xf}(\omega)\right|$ and $\theta(\omega)$, can be obtained using the real and imaginary parts of the FRF, which can be expressed as Eqs. (2) and (3). "Real" and "Imag" represent the real and imaginary parts of the FRF, respectively.

The natural frequency ω_n of a particular system can be obtained by selecting the peak point from the FRF plot, where the denominator is minimized in Eq. (9). In addition, the phase of the FRF dropped by 180° at the location of the natural frequency. If the system is expanded to a MDOF system consisting of n DOFs, the plots of amplitude and phase for the FRF are as presented in Fig. 8.

For an nDOF system, n FRFs $(H_{x_1f}, H_{x_2f}, \cdots, H_{x_if}, \cdots, H_{x_nf})$ can be obtained by assuming that the observability and controllability are equal to the number of DOFs. In Fig. 8, the natural frequency for each

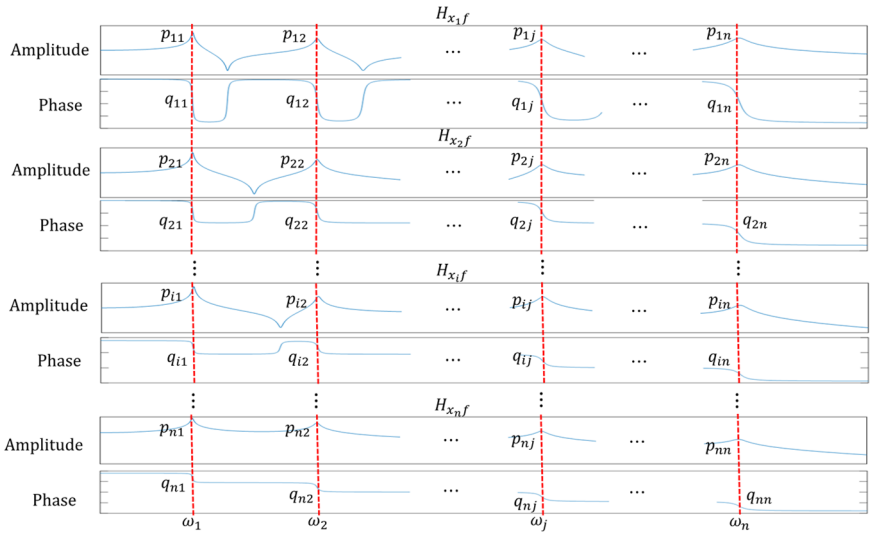

Fig. 8 Peak points in the amplitude and phase of frequency response functions

mode ω_j $(j = 1, 2, \cdots, n)$ can be estimated by selecting the peak points p_{ij} from the amplitudes of the FRFs, where p_{ij} indicates the jth peak point in the ith FRF. The x- and y-coordinate components of p_{ij} are ω_j and $\left|H_{x_if}(\omega_j)\right|$, respectively. The natural frequency for each mode ω_j can also be estimated from the phase of the FRFs, where q_{ij} indicates the jth point of the phase in the ith FRF. The x- and y-coordinate components of q_{ij} are ω_j and $\theta_i(\omega_j)$, respectively. As mentioned before, the modal matrix φ_{ij} can be estimated using the relative magnitude $\left|H_{x_if}(\omega_j)\right|$ and relative phase difference $\theta_i(\omega_j)$, as expressed by Eqs. (8) and (9)

$$\varphi_{ij} = \begin{bmatrix} \alpha_{11}\frac{p_{11}}{\beta_1} & \alpha_{12}\frac{p_{12}}{\beta_2} & \cdots & \alpha_{1j}\frac{p_{1j}}{\beta_j} & \cdots & \alpha_{1n}\frac{p_{1n}}{\beta_n} \\ \alpha_{21}\frac{p_{21}}{\beta_1} & \alpha_{22}\frac{p_{22}}{\beta_2} & \cdots & \alpha_{2j}\frac{p_{2j}}{\beta_j} & \cdots & \alpha_{2n}\frac{p_{2n}}{\beta_n} \\ \vdots & \vdots & \ddots & \vdots & \cdots & \vdots \\ \alpha_{i1}\frac{p_{i1}}{\beta_1} & \alpha_{i1}\frac{p_{i2}}{\beta_2} & \cdots & \alpha_{ij}\frac{p_{ij}}{\beta_j} & \cdots & \alpha_{in}\frac{p_{in}}{\beta_n} \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ \alpha_{n1}\frac{p_{n1}}{\beta_1} & \alpha_{n2}\frac{p_{n2}}{\beta_2} & \cdots & \alpha_{nj}\frac{p_{nj}}{\beta_j} & \cdots & \alpha_{nn}\frac{p_{nn}}{\beta_n} \end{bmatrix} \quad (10)$$

α_{ij} is a value representing the sign of the modal matrix of the ith mode and the jth DOF and can be simplified in the form of Eq. (9) by classifying it into in-phase and out-of-phase values

$$\alpha_{ij} = \begin{cases} 1, & \text{if } \left|\theta_i(\omega_j) - \theta_{ref}(\omega_j)\right| = 0 \text{ or } 360° \\ -1, & \text{if } \left|\theta_i(\omega_j) - \theta_{ref}(\omega_j)\right| = 180° \end{cases} \quad (11)$$

β_i is the normalization constant of the ith modal matrix. In general, the value of β_i is set such that the largest value of the ith modal matrix becomes 1

$$\beta_i = \max_{j=1,2\ldots,n} \left|H_{x_if}(\omega_j)\right| \quad (12)$$

4. Validation test

4.1 Simulation-based validation test

To validate the performance of the proposed method, the results of the proposed LSTM-based automatic modal analysis method were compared with those of the manual peak picking method using the naked eye and ERA. As mentioned in the Introduction, the ERA identifies a system using the singular value of the Hankel matrix. In the ERA, the singular values are truncated to select an appropriate order for the realized model such that it contains only the physical modes. In this study, the best singular value was selected manually by considering the system order.

A total of 7,000 FRF data (1,400 each for the SDOF, 2DOF, ⋯, 5DOF systems) were collected as test data. The numerical model and simulation procedure were the same as described in Section 3.3, and a total of 250 FRF data (50 each for the SDOF, 2DOF, …, 5DOF systems) were collected for manual peak picking using the naked eye. The manual peak picking test was conducted by 14 experts.

Four experimental cases were conducted by adding different levels of noise to the FRF: Case 1 has no noise, where peaks can be identified easily with the naked eye; Case 2 has a low level of noise, where peaks can be identified with some difficulty through the naked eye; Case 3 has a medium noise level, where peaks are difficult to identify with the naked eye; and Case 4 has an extremely high noise level, where peaks are almost impossible to identify with the naked eye.

An example of the estimated natural frequencies and the mode shapes for a 3DOF system using the proposed method is presented in Fig. 9. The systems for each case are independent, showing the result for different dynamic systems. The simulated FRF data are represented by the red solid lines, the locations of the theoretical natural frequencies are represented by the blue dashed lines, and the locations of the estimated peak points are indicated by black circles. As indicated in Fig. 9, the natural frequencies and mode shapes were estimated precisely for data without noise, whereas certain errors were indicated for the data with higher noise levels. However, the proposed method

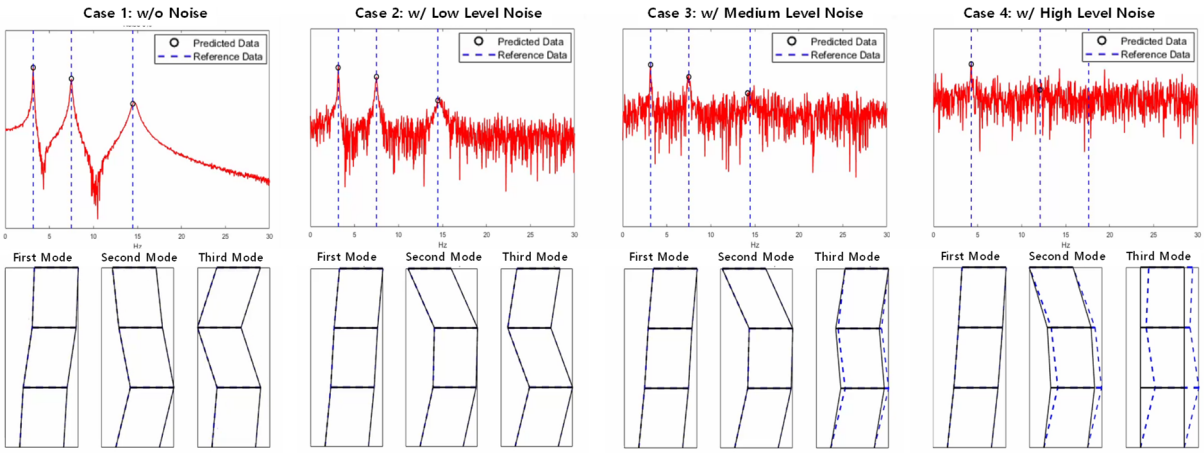

Fig. 9 Estimated natural frequencies (top) and the mode shape (bottom) using the proposed method

could at least determine the first two modes with high noise level, which could not be detected using manual peak picking.

To quantitatively compare the performance of the proposed method, the modal assurance criterion (MAC) value was selected as the performance metric. The MAC value is a statistical indicator of the correlation between the test modal vector and the compatible analytical modal vector. The MAC value was bounded between 0 and 1. If it is close to 0, it indicates no consistent correspondence, and if it is close to 1, it represents a consistent correspondence. The MAC value can be obtained as follows

$$MAC(i,j) = \frac{\left|\{\varphi_A\}_i^T\{\varphi_X\}_j\right|^2}{(\{\varphi_A\}_i^T\{\varphi_A\}_i)\left(\{\varphi_X\}_j^T\{\varphi_X\}_j\right)} \tag{13}$$

where $\{\varphi_A\}_i$ is a compatible modal vector (mode i), $\{\varphi_X\}_j$ is the test modal vector (mode j), $\{\varphi_A\}_i^T$ is the transpose of $\{\varphi_A\}_i$, and $\{\varphi_X\}_j^T$ is the transpose of $\{\varphi_X\}_j$. In this study, only the MAC diagonals were used as performance metrics.

The MAC values for the 3DOF system using the proposed method and the manual peak picking method are summarized in Table 1. The proposed method yielded a MAC value of 0.9897 without any noise, which gradually decreased at a higher noise level. The MAC value for manual peak picking (0.9748) and ERA (0.9940) were also high enough without any noise but decreased more steeply with higher noise levels. For the FRF with high noise level, the MAC value for proposed method indicated was 0.8008, which was significantly higher than that of manual peak picking (0.6824) or ERA (0.7017). From these results, it can be concluded that the proposed method could not only automate the procedure of the peak picking method but also increase the accuracy compared to the naked eye method when the noise level is high.

The MAC values for the 5DOF system using the proposed method, manual peak picking method, and ERA are summarized in Table 2. The proposed method indicated a MAC value of 0.8772 without any noise, which gradually decreased at a higher noise level, similar to that of 3DOF. The MAC value for manual peak picking (0.8384) and ERA (0.8458) were also high without any noise, but it decreased more steeply at higher noise levels. In the extreme case of the high noise level, the MAC value obtained using the proposed method was 0.7494, while the MAC values for manual peak picking using the naked eye and ERA showed 0.3904 and 0.6151 respectively. Compared to the 3DOF system, the difference between the MAC values of the proposed method and that of the manual peak picking or ERA was even higher. When the system gets more complex, the human eyes get more confused resulting higher error, but the proposed LSTM network did not affect by the complexity significantly as much as the human did.

4.2 Experimental validation test

A lab-scale validation test was conducted to determine the applicability of the proposed LSTM-based modal analysis method. A three-story building model was used as the validation structure and was mounted on a shaking table. The specifications of the building model are indicated in Fig. 10. The material of the floor was steel, and SUS304 was used for the columns. The lab-scale validation test was conducted for five cases by changing the input of the shaking table: band-limited white noise with noise powers of 0.010, 0.013, 0.015, 0.017, and 0.020.

To measure the response of the structure, vision-based displacement measurement system proposed by Yoon *et al*.

Table 1 MAC values for the simulation validation test: 3DOF systems

Noise level	Method	MAC values			
		1st Mode	2nd Mode	3rd Mode	Avg.
Case 1 w/o Noise	Proposed method	0.9787	0.9945	0.9960	0.9897
	Manual peak picking	0.9932	0.9882	0.9431	0.9748
	ERA	1.0000	0.9997	0.9824	0.9940
Case 2: w/ Low level noise	Proposed method	0.9752	0.9632	0.9321	0.9568
	Manual peak picking	0.9992	0.9783	0.5814	0.8530
	ERA	0.9994	0.9130	0.6657	0.8594
Case 3: w/ Medium level noise	Proposed method	0.9397	0.8225	0.7229	0.8284
	Manual peak picking	0.9646	0.7755	0.5035	0.7479
	ERA	0.9924	0.7009	0.5557	0.7497
Case 4: w/ High level noise	Proposed method	0.9896	0.7334	0.6794	0.8008
	Manual peak picking	0.9060	0.6651	0.4760	0.6824
	ERA	0.9638	0.5660	0.5754	0.7017
Avg.	Proposed method	0.9393	0.8714	0.8072	**0.8726**
	Manual peak picking	0.9658	0.8518	0.6260	**0.8145**
	ERA	0.9889	0.7949	0.6948	**0.8262**

Table 2 MAC values for the simulation validation test: 5DOF systems

Noise level	Method	MAC values					
		1st Mode	2nd Mode	3rd Mode	4th Mode	5th Mode	Avg.
Case 1 w/o Noise	Proposed method	0.9984	0.9837	0.9574	0.8225	0.6241	0.8772
	Manual peak picking	0.9886	0.9878	0.9341	0.8998	0.3816	0.8384
	ERA	0.9981	0.9741	0.9175	0.7212	0.6182	0.8458
Case 2: w/ Low level noise	Proposed method	0.9977	0.9536	0.907	0.6925	0.6012	0.8304
	Manual peak picking	0.9349	0.9369	0.7069	0.4503	0.2417	0.6541
	ERA	0.9989	0.9573	0.7895	0.5368	0.5358	0.7637
Case 3: w/ Medium level noise	Proposed method	0.9946	0.9454	0.7451	0.6214	0.6226	0.7858
	Manual peak picking	0.7138	0.7152	0.348	0.2895	0.2768	0.4687
	ERA	0.9892	0.7664	0.5300	0.4381	0.5209	0.6489
Case 4: w/ High level noise	Proposed method	0.9915	0.8614	0.6354	0.6249	0.6337	0.7494
	Manual peak picking	0.5779	0.6286	0.2694	0.2579	0.218	0.3904
	ERA	0.9588	0.6399	0.4194	0.4907	0.5665	0.6151
Avg.	Proposed method	0.9956	0.9360	0.8112	0.6903	0.6204	**0.8107**
	Manual peak picking	0.8038	0.8171	0.5646	0.4744	0.2795	**0.5879**
	ERA	0.9588	0.6399	0.4194	0.4907	0.5665	**0.6151**

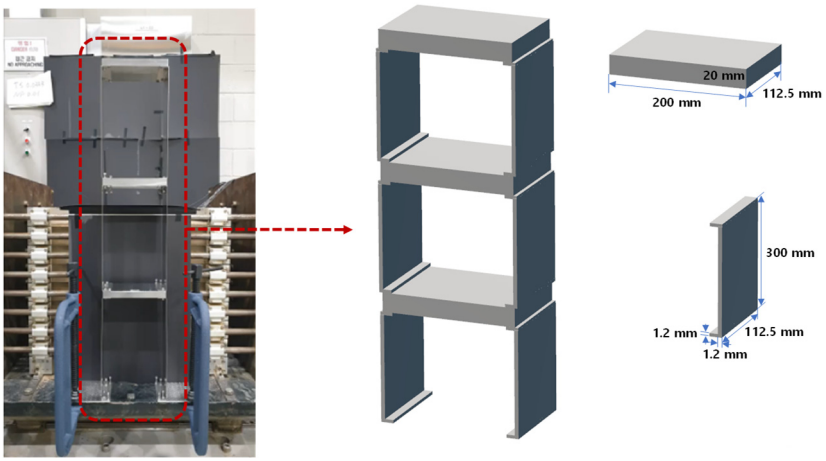

Fig. 10 Three-story building model for the lab-scale experimental test

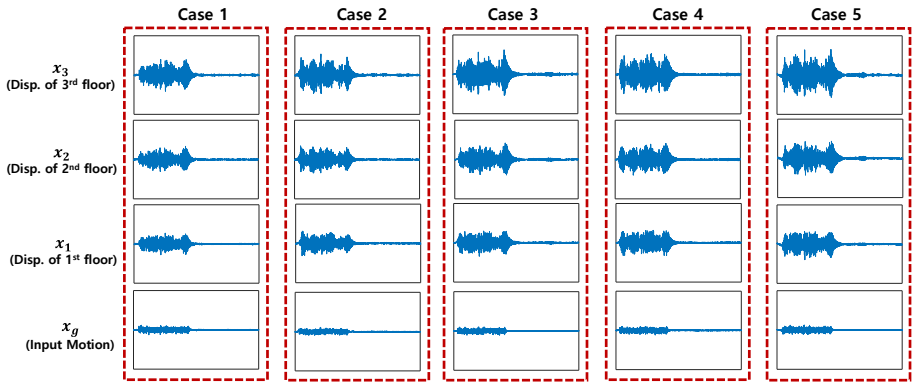

Fig. 11 Measured displacement using the vision-based displacement system

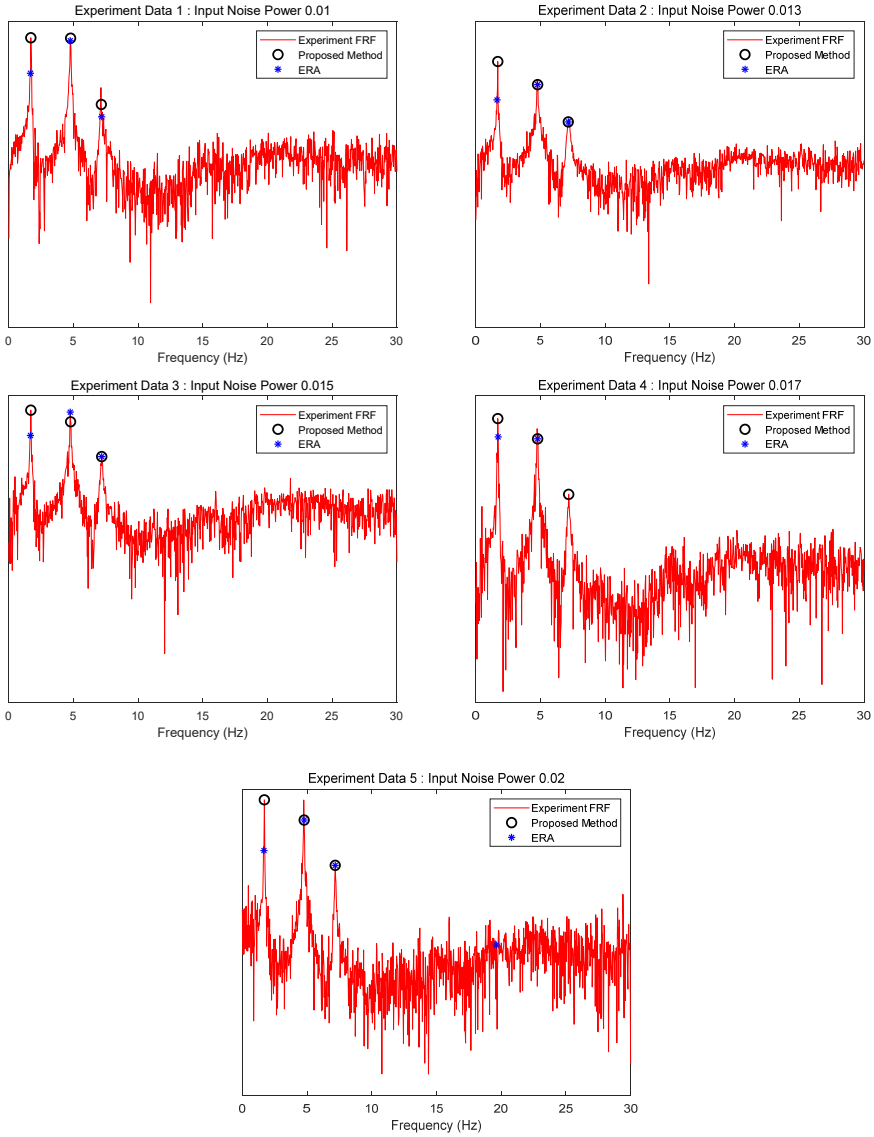

Fig. 12 Lab-scale validation test result

The natural frequencies and the mode shapes were calculate using the proposed method, and MAC values were calculated by comparing the mode shapes with the reference values. Also, to compare the performance of the proposed method with another modal analysis method, MAC values were calculated using the eigen-system realization algorithm (ERA). Similar to the simulation-based validation test, the best singular value was selected manually by considering the system order.

The reference natural frequencies and mode shapes were calculated based on the measured geometry and material properties of the structural components. However, due to the error during the fabrication and assembly process, the modal property of the structure might be different with the complete structure. Therefore, the reference values were calibrated by applying the Finite Element (FE) model updating technique. The FE model updating techniques use genetic algorithm, setting the objective function as below.

$$Minimize: W = \sqrt{\sum_{i=1}^{n} \frac{(\overline{\omega_i} - \omega_i)^2}{n}} \tag{14}$$

where $\overline{\omega_i}$ is the ith natural frequency of the FE model, ω_i is the estimated ith natural frequency using the experimental data.

The results of the lab-scale test using the proposed method and the ERA are presented in Fig. 12. The simulated FRF data are represented by red solid lines, the locations of the estimated peak points using the proposed method are represented by black circles, and the locations of the estimated peak points using the ERA are represented by blue stars. The results for the natural frequencies using the proposed method and ERA are presented in Table 3. The proposed method was able to estimate all three modes accurately with minimum MAC value of 0.9830. However, the ERA method could not estimate the 3rd mode for case 4

Table 3 Experimental validation test results

Noise level	Method	MAC values			
		1st Mode	2nd Mode	3rd Mode	Avg.
1	Proposed method	0.9995	0.9878	0.9914	**0.9929**
	ERA	0.9959	0.9842	0.9315	**0.9705**
2	Proposed Method	0.9997	0.9879	0.9830	**0.9902**
	ERA	0.9941	0.9881	0.9539	**0.9787**
3	Proposed Method	0.9996	0.9881	0.9838	**0.9905**
	ERA	0.9915	0.9803	0.9335	**0.9684**
4	Proposed Method	0.9994	0.9894	0.9935	**0.9941**
	ERA	0.9820	0.9611	0.0076	**0.6502**
5	Proposed Method	0.9998	0.9870	0.9869	**0.9912**
	ERA	0.9944	0.9862	0.9752	**0.9853**
Avg	**Proposed Method**	**0.9996**	**0.9880**	**0.9877**	**0.9918**
	ERA	**0.9916**	**0.9800**	**0.7603**	**0.9106**

indicating the MAC value of 0.0076. The averaged MAC value for case 4 was 0.9941 for the proposed method, while the ERA method showed 0.6502. Even for other cases where ERA could estimate all three modes (e.g., case 1, 2, 3, and 5), the proposed method showed higher MAC values (0.9929, 0.9902, 0.9905, and 0.9912 respectively) compared to the MAC values of ERA (0.9902, 0.9787, 0.9684, 0.9853). From the experimental test, it was shown that the proposed method can not only automate the process but can estimate the modal properties more accurately compared to the conventional method.

5. Conclusions

In this paper, a new method was presented for modal analysis using a long short-term memory (LSTM) network. The proposed method calculates the cross-power spectral density (CPSD) and PSD from the dynamic response of the structure, and the frequency response function (FRF) is calculated using the CPSD and PSD. Once the FRF is determined, the amplitude and phase of the FRF are connected to the input layer of the LSTM network. The LSTM network estimates the peak points of the FRF and delivers the results through the output layer. Finally, the natural frequencies and mode shapes of the system were calculated using the FRF.

To validate the performance of the proposed method, two validation tests were conducted: (1) simulation-based and (2) lab-scale. For the simulation test, the proposed method was able to not only automate the conventional modal analysis procedure but also estimate the mode shapes with higher accuracy. It indicated average MAC values of 0.8447 and 0.8107 for the three- (3DOF) and five-degree-of-freedom (5DOF) systems, respectively, whereas manual peak picking indicated MAC values of 0.8145 and 0.5879, respectively. To determine the applicability of the proposed method to a real-world problem, a lab-scale test was conducted for a three-story building. The proposed method was able to automatically estimate the modal properties of the structure, whereas the ERA method requires the user to select the singular value. Furthermore, the results obtained using the proposed method indicated higher MAC values as compared with those obtained using ERA.

The proposed method has two main contributions as compared with conventional modal analysis methods. (1) The proposed method is a fully automated process that does not require any manual procedure or prior information such as a singular value or a system order. (2) The proposed method can estimate the dynamic characteristics of the structural system with better accuracy as compared with the manual method or ERA, particularly when the data contain high-level noise. Therefore, it is expected that the results of the proposed system will be more objective and accurate, so that even nonexperts can estimate the dynamic characteristics of the structural system.

Acknowledgments

This work was supported by the National Research Foundation of Korea (NRF) grant funded by the Korea government (MSIT) (NRF-2021R1A4A3033128 and NRF-2022R1C1C1003012).

References

Basharat, A., Catbas, N. and Shah, M.A. (2005), "Framework for Intelligent Sensor Network with Video Camera for Structural Health Monitoring of Bridges", *Proceedings of the 3rd IEEE International Conference on Pervasive Computing and Communications Workshops*, Kauai, HI, USA, March.

BBC News (2019), "Bridge closed because cracks in pedestals", BBC, London, UK. https://www.bbc.com/news/uk-england-london-48395371

Bengio, Y., Simard, P. and Frasconi, P. (1994), "Learning Long-Term Dependencies with Gradient Descent Is Difficult", *IEEE Transact. Neural Networks*, **5**(2), 157-166. ttps://doi.org/10.1109/72.279181

Brincker, R., Zhang, L. and Andersen, P. (2001), "Modal

Identification of Output-Only Systems Using Frequency Domain Decomposition", *Smart Mater. Struct.*, **10**(3), 441-445. https://doi.org/10.1088/0964-1726/10/3/303

Cha, Y.-J., Choi, W. and Büyüköztürk, O. (2017), "Deep Learning-Based Crack Damage Detection Using Convolutional Neural Networks", *Comput.-Aided Civil Infrastr. Eng.*, **32**(5), 361-378. https://doi.org/10.1111/mice.12263

Cho, K., Merrienboer, B.V., Gulcehre, C., Bahdanau, D., Bougares, F., Schwenk, H. and Bengio, Y. (2014) "Learning Phrase Representations Using RNN Encoder–Decoder for Statistical Machine Translation", *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing* (*EMNLP*), Doha, Qatar, October.

Cole Jr, H.A. (1973), "On-line failure detection and damping measurement of aerospace structures by random decrement signatures", Report No. CR-2205, NASA, Washington D.C., USA.

Collins, J., Ashurst, D., Webb, J., Sparkes, P. and Ghose, A. (2019), "Guidance on hidden defects in bridges in the UK and Ireland", *Proceedings of the Institution of Civil Engineers-Bridge Engineering*, Vol. 172, No. 1, pp. 41-53.

Deng, L., Chu, H.H., Shi, P., Wang, W. and Kong, X. (2020), "Region-based CNN method with deformable modules for visually classifying concrete cracks", *Appl. Sci.*, **10**(7), 2528. https://doi.org/10.3390/app10072528

Feng, D. and Feng, M.Q. (2018), "Computer vision for SHM of civil infrastructure: From dynamic response measurement to damage detection–A review", *Eng. Struct.*, **156**, 105-117. https://doi.org/10.1016/j.engstruct.2017.11.018

Haberman, C. (2014), "A Disaster Brought Awareness but Little Action on Infrastructure", The New York Times, New York, NY, USA. https://www.nytimes.com/2014/03/03/us/a-disaster-brings-awareness-but-little-action-on-infrastructure.html

Harlan, C. and Selk, A. (2018), "'Should not happen': Italian officials look for answers after dozens dead in bridge collapse", The Washington Post, Washington D.C., USA. https://www.washingtonpost.com/news/worldviews/wp/2018/08/14/immense-tragedy-feared-as-highway-bridge-collapses-over-italian-city/

Hochreiter, S. (1998), "The Vanishing Gradient Problem During Learning Recurrent Neural Nets and Problem Solutions", *Int. J. Uncertain. Fuzz. Knowledge-Based Syst.*, **06**(02), 107-116. https://doi.org/10.1142/S0218488598000094

Hochreiter, S. and Schmidhuber, J. (1997), "Long short-term memory", *Neural Computat.*, **9**(8), 1735-1780. https://doi.org/10.1162/neco.1997.9.8.1735

Ibrahim, S.R. and Pappa, R.S. (1982), "Large Modal Survey Testing Using the Ibrahim Time Domain Identification Technique", *J. Spacecr. Rockets*, **19**(5), 459-465. https://doi.org/10.2514/3.62285

Jin, S.S., Jeong, S., Sim, S.H., Seo, D.W. and and Park, Y.S. (2021), "Fully automated peak-picking method for an autonomous stay-cable monitoring system in cable-stayed bridges", *Automat. Constr.*, **126**, 103628. https://doi.org/10.1016/j.autcon.2021.103628

Juang, J.-N. and Pappa, R.S. (1985), "An Eigensystem Realization Algorithm for Modal Parameter Identification and Model Reduction", *J. Guidance Control Dyn.*, **8**(5), 620-627. https://doi.org/10.2514/3.20031

Jung, H.J., Lee, J.H., Yoon, S. and Kim, I.H. (2019), "Bridge Inspection and condition assessment using Unmanned Aerial Vehicles (UAVs): Major challenges and solutions from a practical perspective", *Smart Struct. Syst., Int. J.*, **24**(5), 669-681. https://doi.org/10.12989/sss.2019.24.5.669

Khuc, T. and Catbas, F.N. (2017), "Completely contactless structural health monitoring of real-life structures using cameras and computer vision", *Struct. Control Health Monitor.*, **24**(1), 1852. https://doi.org/10.1002/stc.1852

Kim, H. and Sim, S.H. (2019), "Automated peak picking using region-based convolutional neural network for operational modal analysis", *Struct. Control Health Monitor.*, **26**(11). https://doi.org/10.1002/stc.2436

Korean Law Information Center (2016), Special Act on the Safety Control and Maintenance of Establishments; Ministry of Land, Infrastructure and Transport, Seoul, Korea. https://www.law.go.kr/LSW/eng/engLsSc.do?menuId=2&query=safety#liBgcolor6.

Lee, J.H., Park, J.J. and Yoon, H. (2020), "Automatic Bridge Design Parameter Extraction for Scan-to-BIM", *Appl. Sci.*, **10**(20), 7346. https://doi.org/10.3390/app10207346

Lim, J. and Yoon, H. (2022). "Vision-based Input-Output System identification for pedestrian suspension bridges", *Smart Struct. Syst., Int. J.*, **29**(5), 715-728. https://doi.org/10.12989/sss.2022.29.5.715

Lopez-Pacheco, M., Morales-Valdez, J. and Yu, W. (2020), "Frequency domain CNN and dissipated energy approach for damage detection in building structures", *Soft Comput.*, **24**(20), 15821-15840. https://doi.org/10.1007/s00500-020-04912-w

Lydon, D., Taylor, S.E., Lydon, M., Martinez del Rincon, J. and Hester, D. (2019), "Development and testing of a composite system for bridge health monitoring utilizing computer vision and deep learning", *Smart Struct. Syst., Int. J.*, **24**(6), 723-732. https://doi.org/10.12989/sss.2019.24.6.723

Lynch, J.P. (2006), "A summary review of wireless sensors and sensor networks for structural health monitoring", *Shock Vib. Digest*, **38**(2), 91-128. https://doi.org/10.1177/0583102406061499

Lindblom, M. (2020), "West Seattle Bridge closes tonight for repairs that could take months", The Seattle Times, Seattle, WA, USA. https://www.seattletimes.com/seattle-news/transportation/the-west-seattle-bridge-closes-tonight-for-major-repairs-that-could-take-several-weeks/

Narazaki, Y., Hoskere, V., Hoang, T.A., Fujino, Y., Sakurai, A. and Spencer Jr, B.F. (2020), "Vision-based automated bridge component recognition with high-level scene consistency", *Comput.-Aided Civil Infrastr. Eng.*, **35**(5), 465-482. https://doi.org/10.1111/mice.12505

Nasimi, R. and Moreu, F. (2021), "A methodology for measuring the total displacements of structures using a laser–camera system", *Comput.-Aided Civil Infrastr. Eng.*, **36**(4), 421-437. https://doi.org/10.1111/mice.12652

Nex, F., Duarte, D., Tonolo, F.G. and Kerle, N. (2019), "Structural building damage detection with deep learning: Assessment of a state-of-the-art CNN in operational conditions", *Remote Sensing*, **11**(23), 2765. https://doi.org/10.3390/rs11232765

Pastor, M., Binda, M. and Harčarik, T. (2012), "Modal Assurance Criterion", *Procedia Eng.*, **48**, 543-548. https://doi.org/10.1016/j.proeng.2012.09.551

Qi, Y., Kong, Q., Xiong, B. and Li, P. (2021), "A deep learning-based vision enhancement method for UAV assisted visual inspection of concrete cracks", *Smart Struct. Syst., Int. J.*, **27**(6), 1031-1040. https://doi.org/10.12989/sss.2021.27.6.1031

Shen, Y., Fu, W., Luo, Y., Yun, C., Liu, D., Yang, P., Yang, G. and Zhou, G. (2021), "Implementation of SHM system for Hangzhou East Railway Station using a wireless sensor network", *Smart Struct. Syst., Int. J.*, **27**(1), 19-33. https://doi.org/10.12989/sss.2021.27.1.019

Special Act on the Safety Control and Maintenance of Establishments, Act No. 17946, 16, Mar 2021. https://www.law.go.kr/LSW/eng/engLsSc.do?menuId=2&query=safety#AJAX

Spencer Jr, B.F., Ruiz-Sandoval, M.E. and Kurata, N. (2004),

"Smart Sensing Technology: Opportunities and Challenges", *Struct. Control Health Monitor.*, **11**(4), 349-368. https://doi.org/10.1002/stc.48

Spencer Jr, B.F., Hoskere, V. and Narazaki, Y. (2019), "Advances in computer vision-based civil infrastructure inspection and monitoring", *Engineering*, **5**(2), 199-222. Advances in computer vision-based civil infrastructure inspection and monitoring

Sun, F.P., Chaudhry, Z.A., Rogers, C.A., Majmundar, M. and Liang, C. (1995), "Automated real-time structure health monitoring via signature pattern recognition", *Smart Struct. Mater.*, **2443**, 236-247. https://doi.org/10.1117/12.208261

Ye, X.W., Jin, T. and Yun, C.B. (2019), "A review on deep learning-based structural health monitoring of civil infrastructures", *Smart Struct. Syst., Int. J.*, **24**(5), 567-586. https://doi.org/10.12989/sss.2019.24.5.567

Yoon, H., Elanwar, H., Choi, H., Golparvar-Fard, M. and Spencer, B.F. (2016), "Target-free approach for vision-based structural system identification using consumer-grade cameras", *Struct. Control Health Monitor.*, **23**(12), 1405-1416. https://doi.org/10.1002/stc.1850

Yoon, H., Shin, J. and Spencer, B.F. (2018), "Structural displacement measurement using an unmanned aerial system", *Comput.-Aided Civil Infrastr. Eng.*, **33**(3), 183-192. https://doi.org/10.1111/mice.12338

Zhang, A., Wang, K.C.P., Li, B., Yang, E., Dai, X., Peng, Y., Fei, Y., Liu, Y., Li, J.Q. and Chen, C. (2017), "Automated Pixel-Level Pavement Crack Detection on 3D Asphalt Surfaces Using a Deep-Learning Network", *Comput.-Aided Civil Infrastr. Eng.*, **32**(10), 805-819. https://doi.org/10.1111/mice.12297

HJ