Article | Received 29 March 2025; Accepted 19 June 2025; Published 30 June 2025
https://doi.org/10.55092/sc20250015

Working status prediction for a high-formwork support system using finite element model-informed deep learning model and GPT-aided method

Linlin Zhao[1,*], Jasper Mbachu[2], Siyu Liu[1] and Rongtian Zhang[1]

[1] College of Architecture and Civil Engineering, Beijing University of Technology, Beijing 100124, China

[2] Faculty of Society & Design, Bond University, Gold Coast 4226, Australia

* Correspondence author; E-mail: llzhao@bjut.edu.cn.

Highlights:

- GA-optimized FEM generates HFSS structural responses under varied conditions.
- CNN classifier achieves satisfactory accuracy in predicting HFSS working statuses.
- Experimental validation shows CNNs outperform SVM in HFSS status classification.
- RAG model with KG automatically generates SHM reports for HFSS.
- Multi-metric evaluation confirms RAG's superiority over GPT-4 in SHM reporting.

Abstract: A High-formwork support system (HFSS) is essential during the construction process for complex structures. A lack of an effective monitoring method occasionally causes the system's tragic collapse. In order to establish a reliable method for structural condition identification, experimental measurements on structures are always required. However, in practice, the experiments on sites are either expensive or difficult to conduct. A data-driven algorithm convolutional neural networks (CNNs) for working status monitoring of the HFSS structure is proposed in this study. First, a finite element (FE) model for the HFSS was developed and optimized by using the genetic algorithm (GA). Then, the optimal FE model was employed to produce structural response data for three statuses of the HFSS, such as normal working status, local instability, and fully unstable statuses. The generated dataset was adopted to train a CNNs classifier, which can correctly predict the working states of the structure. Finally, the potential of CNNs was validated on the experimental measurements derived from the HFSS, and the performance of CNNs and support vector machine (SVM) was compared. CNNs performed much better than SVM on the experimental dataset. Moreover, this study developed a Retrieval-Augmented Generation (RAG) model by leveraging a Generative Pre-Trained Transformer (GPT) to synthetically generate an SHM report to describe the structural condition of the given HFSS structures. A knowledge graph (KG) was also developed to enable comprehensive, reliable, informative SHM contexts. Multiple evaluation metrics were employed to assess the performance of the RAG model. The findings indicate that the RAG model could generate accurate and reasonable SHM reports for HFSS.

Keywords: finite element model updating; CNNs; high-formwork support system; working status prediction; GPT; knowledge graph

1. Introduction

A high-formwork support system (HFSS) is crucial during the project construction stage, especially in the case of complex projects. However, their importance is usually ignored, sometimes resulting in tragic collapses. Such collapses may cause substantial economic and life loss [1–3]. Appropriate safety management for HFSS is required. Especially, an effective monitoring method should be provided during its working. Monitoring and evaluating the working statuses of the structure is essential for a project's technological development and safety management. Therefore, this study proposes an effective monitoring approach for assessing the working status of high-formwork support systems (HFSS). This method utilizes structural response measurements, including pole displacements and structural vibrations, which contain valuable information about the structure's operational state.

Recently, machine learning (ML) technologies have been conveniently adopted to classify the working status of the structure by using the extracted information from structural responses [4,5]. ML is comprised of a wide range of mathematical models that are trained to conduct a specific task. Model training requires a great deal of reliable data, which indicates that a large number of measurements in various conditions are demanded. The main challenge for successful deep learning (DL) models is to attain the training dataset, especially for complex and large structures in which experiments are forbidden, either due to safety or cost.

In practice, it may require too much effort to perform experiments on structures to attain datasets, including different statuses of the structures. Data relating to specific states could be replaced with numerically produced ones from a physics-based digital twin of the structure. Such data can be exploited for the training of the deep learning model. Nevertheless, results derived from finite element models (FEM) often deviate from experimental findings. The established FEM is usually required to be optimized. Establishing proper values for model updating parameters constitutes a critical phase in the optimization procedure. The values of uncertain parameters could not be known at first, then, various procedures are adopted to explore the values that make the model forecasts best match the measurements [6,7]. Measurement data can be obtained based on experimental tests on a real structure. The study employs an optimized finite element model to minimize deviations between initial numerical predictions and experimental structural response data. Using the optimal FE model, both normal or abnormal working states' datasets could be produced without conducting experiments on the actual structure. This means that the costs and efforts of data attainment can be reduced by using such a FE model.

A FEM-informed CNN model is proposed in this study. A FEM of a high-formwork support system (HFSS) was developed and optimized. The working statuses of the HFSS have been simulated with the optimal FE model previously. The simulated datasets were then utilized to train a CNNs model that can be used to classify the working statuses of the HFSS. The trained CNNs classifier was then validated on experimental measurements from a nearly 8 m high full-scale HFSS to evaluate its performance gains. The main stages of the framework are listed below. The process is graphically depicted in Figure 1.

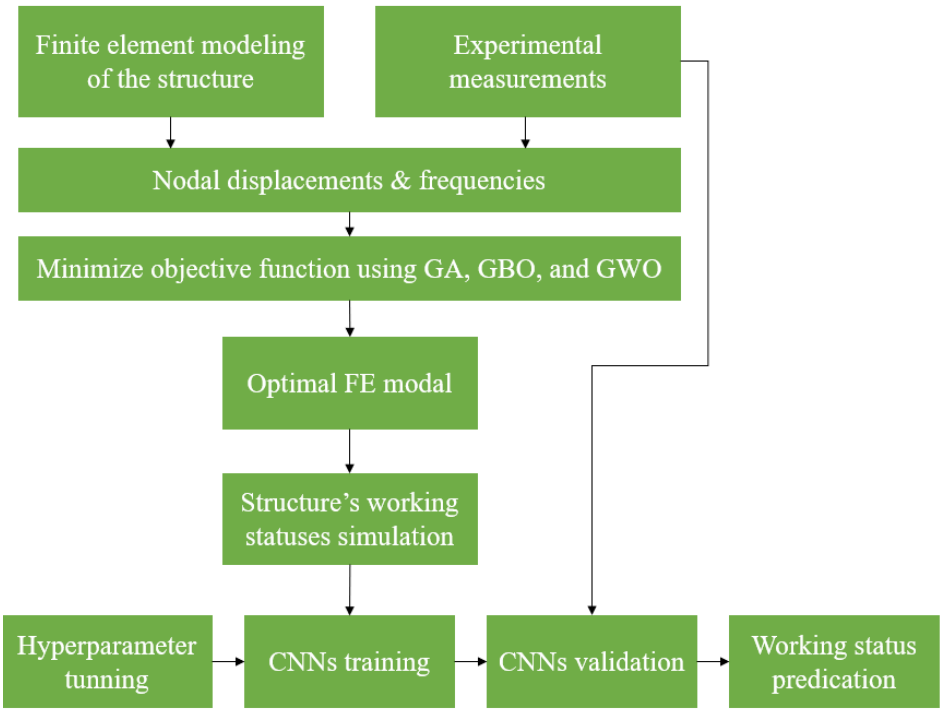


Figure 1. Workflow of the working status predication for HFSS.

(1) Build a FEM of the high-formwork support system and optimize the FEM based on actual experimental measurements from the structure.

(2) Use the optimal FEM simulation to generate learning data of different status scenarios of the structure.

(3) Deploy suitable CNNs and use the FEM simulated data to train the CNNs.

(4) Validate the CNNs classifier by using experimental measurements.

Current HFSS structural health monitoring (SHM) reports often depend excessively on subjective expertise rather than standardized protocols. This reliance on anecdotal evidence and collective experience, while valuable, results in non-quantitative data that challenges digital processing and lacks essential structural information. This highlights the need for generating more standard SHM reports that can provide essential, valuable, and practical information. In this study, several sections are included in the final report to gain an understanding of the working situation of the HFSS.

The integration of large language models (LLMs) is capable of simplifying paperwork and management duties, as LLMs can categorize documents, understand content, draft documents, and extract relevant information [8]. However, LLMs used in SHM domain and risk assessment are still at the infant stage. This study aims to develop and assess an SHM-specialized LLM using a Retrieval-Augmented Generation (RAG) framework. Specifically, the study compares the effectiveness of the developed model with GPT-4. The evaluation incorporates multiple quantitative measures: BLEU (assessing translation quality), ROUGE (evaluating summary content), along with vector-space metrics (cosine similarity and Euclidean distance). This multi-metric validation confirms the model's robustness in SHM applications.

The main contribution of the study lies below:

- Optimizing finite element model (FEM) using a genetic algorithm to generate structural response data for HFSS under different working statuses (normal working, local instability, fully unstable).
- Training a convolutional neural networks (CNNs) classifier with the data generated from

optimized FEM model to accurately predict the working statuses of HFSS.

- Validating the prediction accuracy of the CNNs classifier, which generated better results than support vector machines (SVM) in performance.
- Utilizing a knowledge graph (KG) as an external database alliance with GPT to form an RAG model, which could automatically generate accurate and informative structural health monitoring (SHM) reports.
- Validating the performance of the RAG model in generating SHM reports using various evaluation metrics such as BLEU, ROUGE, cosine similarity, and Euclidean distance.

The paper is organized as follows: Section 2 reviews relevant literature, while Section 3 elaborates on the finite element model updating methodology. Section 4 details the CNN training process using simulated data from optimized FEM and presents experimental validation results, including a comparative analysis with SVM. The subsequent section covers the development process of the RAG model for generating SHM report for the HFSS. The RAG model's validation is shown in Section 6. Finally, the conclusion is provided in Section 7.

2. Literature review

2.1. Finite element model updating

FEM is an efficient tool in the monitoring of structures, forecasting of service life, damage detection, and decision of maintenance options. Hence, the reliability and accuracy of the FE models are essential. The errors and uncertainties related to the model must be quantified. Modelling uncertainties typically encompasses three primary components: parameter-level uncertainties, model structure limitations, and code-related factors [9]. In general, uncertainty in model parameters is caused by wrong assumptions, including material properties or geometric properties of structure [10]. Model structure uncertainties primarily originate from mechanical properties and physical behavior characteristics. These uncertainties stem from multiple modeling approximations, including but not limited to: (1) improper connection modeling, (2) boundary condition misrepresentation, (3) mass distribution inaccuracies, and (4) structural simplification misassumptions [11–13]. Uncertainties in structure might additionally originate from inappropriate behavioral representations (nonlinear/linear simplifications) and erroneous assumptions of structure geometry and loads [14]. Furthermore, code-related errors typically emerge from technical model uncertainties and numerical uncertainties, predominantly caused by hardware deficiencies or software imperfections [9].

Finite element model updating (FEMU) fundamentally constitutes an optimization procedure aimed at reducing discrepancies between computational predictions and the real behavior from the actual structure. This process can be characterized as solving either an optimization problem or a statistical inference task. The development of FEMU approaches usually uses the structure's actual response derived from static and/or dynamic experiments on the structure. An optimization objective is formulated by minimizing the discrepancy between numerical predictions and experimental measurements. These datasets include static datasets [15], structural dynamic datasets [16,17], or a combination of the two datasets [18,19].

Iterative finite element model updating algorithms employ structural parameter modification schemes coupled with optimization techniques to achieve realistic and reasonable results [20]. These

algorithms fundamentally solve optimization problems to generate the global optimum. Genetic algorithm (GA) and particle swarm optimization (PSO) algorithms have demonstrated particular effectiveness, but there are several algorithms also available in FEMU. Most of them are inspired by nature, including grey wolf [21], harmony search [22], gravitational search [23], colliding bodies [24], and simulated annealing [25].

The present FEMU process comprises four key steps: (1) identification of critical updating parameters, (2) experimental configuration for structural testing, acquisition and processing of response measurements, (3) formulation of optimization criteria, and (4) selection of appropriate search algorithms. Commercially available software packages were adopted to create and optimize the model. Moreover, a full-scale high-formwork support system was built to provide experimental measurements that are essential during the FE updating process. In addition, several optimization algorithms including GA, gradient-based optimizer (GBO) and grey wolf optimizer (GWO) were compared to obtain the best results.

2.2. Large language models (LLMs) and Retrieval-Augmented Generation (RAG)

Large language models (LLMs) are trained on extensive datasets, developing the capability to recognize intricate patterns, discern nuanced variations, and derive meaningful semantic representations from textual data [26,27]. LLMs show remarkable efficiency in language processing works including classification, summarization, sentiment analysis, text generation, and machine translation [28]. Despite GPT could understand the questions and present common-sense information, it exhibits limitations in comprehending domain-specific knowledge. Additionally, they are prone to generating factually inaccurate outputs, a phenomenon commonly referred to as “hallucination”. LLMs usually come short of capturing factual information, while knowledge graphs (KGs) consist of structural knowledge that clearly stores rich factual information. Knowledge graphs could benefit LLMs, as they offer a structured context that improves LLMs capability to understand and produce more accurate and relevant responses. As such, it is advantageous to integrate LLMs and KGs and leverage their advantages.

The Retrieval-Augmented Generation (RAG) framework significantly improves large language models’ information retrieval and precise knowledge application capabilities [40]. As an emerging paradigm, RAG concentrates on enhancing the abilities of LLMs by connecting them to an external knowledge database. Large language models usually store a large number of information but struggle to accurately manipulate specific domain knowledge. By dynamically accessing domain-specific knowledge bases during inference, the RAG model effectively incorporates relevant external information, thereby significantly enhancing relevance, accuracy, efficiency, and depth of its responses. Afzal *et al.* [29] and Siriwardhana *et al.* [30] addressed RAG as allowing retrieval of information from an external database, and easing hallucination problems in construction education. Taiwo *et al.* [31] developed a RAG model using contract documents, demonstrating enhancement of GPT-4’s performance in processing contractual documents. Contract documents were processed into the structured database and embedding vectors. The proposed RAG framework leverages semantic similarity matching within its knowledge base to generate contextually appropriate responses. Comparative evaluations demonstrated superior response quality and enhanced relevance over baseline GPT-4 implementations, while simultaneously mitigating hallucination phenomena.

Several LLMs are available, but they do not properly solve the unique problems in the structural health monitoring field including hallucinations, domain-specific knowledge, data and interoperability. This study employed RAG structure to better suit the structural health monitoring domain. Previous studies on specialized LLMs in the construction field mainly demonstrated the potential of these models but did not describe how to develop a model specifically tailored to the structural health monitoring domain. Prieto *et al.* [32] and Uddin *et al.* [33] collectively proved that the absence of domain-specific knowledge in training data directly correlates with increased rates of erroneous outputs and hallucinatory generations in LLMs.

This study aims to address these gaps by building a RAG model, using domain-specific database focusing on structural health monitoring for high formwork. The RAG model retrieves information from the database that incorporates a domain-specific knowledge graph. The proposed RAG model is designed to produce relevant and accurate information suited to structural health monitoring tasks. The assessment incorporates multiple quantitative indicators to systematically evaluate the performance of the developed RAG model.

3. Finite element model updating

The mathematical formulations of the finite element (FE) model updating for a high-formwork support system with semi-rigid joints and initial defects are described in this section. Experiments were set up on a real high-formwork support system for the FE model updating. The flow chart of model updating is shown in Figure 2.

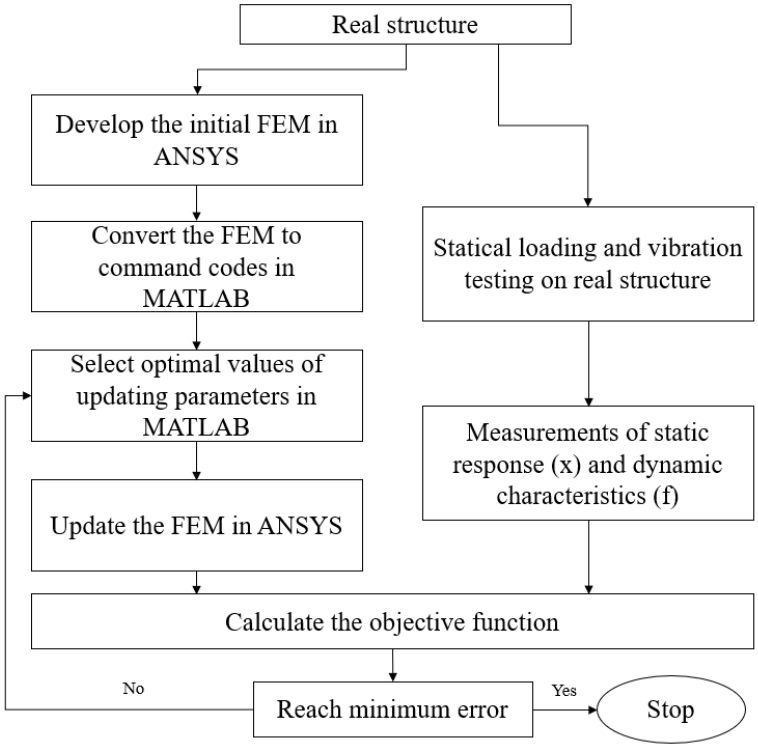


Figure 2. The new technique of FE model updating using MATLAB and Ansys.

Model updating procedures can be summarized as follows.

Step 1: Build the initial FE model;

Step 2: Select the updating parameters θ based on sensitivity analysis;

Step 3: Attain the experimental results and compute the structural response: displacements of standing poles and frequencies of the structure;

Step 4: Establish objective function;

Step 5: Use the optimization algorithm to minimize the objective function;

Step 6: Analyze and evaluate the updated FE model.

3.1. Local stiffness for semi-rigid joints and initial defects of the structure

The joints are semi-rigid, meaning that, depending on the fixity factor, they can rotate to some degree. The fixity factor ranges from 0 to 1, 0 indicating a totally hinged joint, and 1 denoting an entirely rigid joint. In this study, the semi-rigid joint is adopted instead of a perfectly rigid joint. This makes the FE model represent the real situation of the structure. There is a direct connection among Young's modulus, initial stiffness, and stiffness matrix. Hence, stiffness matrices are updated by adjusting Young's modulus and the initial stiffness of the joints.

In addition, the initial defects of the structure should be considered, especially for structures like high-formwork support systems, as they are not rigidly connected like a real structure. Based on the building code for high-formwork support structures, the error of the perpendicularity for the structure must be less than 1/500 H with H indicating the height of the structure. Moreover, the initial bending of the standing poles of the structure must be less than 1/1000 L, with L indicating the length of the standing pole. Hence, the developed FE model used the first-order bulking mode shapes as the initial defects of the poles.

3.2. Selection of updating parameters

When updating an FE model, it is important to choose updating parameters that have a significant impact on the response data. Based on previous studies, the elastic moduli, Poisson's ratio, and density of the standing pole of the structure, real constant of the diagonal brace, and the initial stiffness of the semi-rigid joints were chosen as the updating model parameters. The updating model parameters are presented in Table 1. LB and UB denote lower and upper bounds for each parameter. The initial parameters for the structure are Young's modulus E = 210 GPa, Poisson's ratio ν = 0.3 and density ρ = 7850kg/m^3. Nominal damping is utilized as constant ζ = 0.01 at all frequencies leading to Rayleigh coefficients α_1 = 0.01998 and α_2 = 1.9998.

By the adoption of sensitivity analysis, the measurements' impact on the parameters was quantitatively assessed. Each parameter varied by 10%, and the changes in sensitivity and responses were computed accordingly. In this study, displacements of standing poles and frequencies of the structure were chosen as structural responses for FE model updating. Hence, 27 responses were regarded: the displacements of 20 points on standing poles as shown in Figure 3, and the first seven frequencies. Figure 4a presents the sensitivity analysis results**.** The results denote that the displacements D1, D4, D6, D10, D11, D13, D15, D17 and D18 have lower sensitivity (less than 2) compared with others, and these measurements will no longer be regarded in subsequent model updating. As shown in Figure 4b, in the horizontal plane of the figure, the updating parameters are displayed against the first seven frequencies. The height of the columns indicates the scaled sensitivity of each frequency to each of the parameters.

Hence, in this study, the displacements D2, D3, D5, D7, D8, D9, D12, D14, D16, D19 and D20 and the first three nature frequencies were adopted as the structure response data and structural measurements from the experiment. The sensitivities are all positive, as a rise in stiffness usually leads to an increase in all the frequencies.

Table 1. Updating parameters for the optimization of the FE model.

Components		Standing poles & horizontal ledgers	Diagonal brace	Joints
Young's Modulus E (GPa)	Initial value	206	206	~
	LB	185.4	~	~
	UB	226.6	~	~
Density ρ (kg/m^3)	Initial value	7800	7800	~
	LB	7020	~	~
	UB	8580	~	~
Poisson's ratio ν	Initial value	0.3	0.3	~
	LB	0.27	~	~
	UB	0.33	~	~
Geometric dimension	OD (mm)	48	48	~
	Thickness (mm)	3.2	2.5	~
Real constant r (mm^2)	Initial value	450	357	~
	LB	420	320	~
	UB	480	380	~
Initial stiffness (kN·m/rad)	Initial value	~	~	65
	LB	~	~	58.5
	UB	~	~	71.5

OD = outer diameter, LB = lower bound, UB = upper bound.

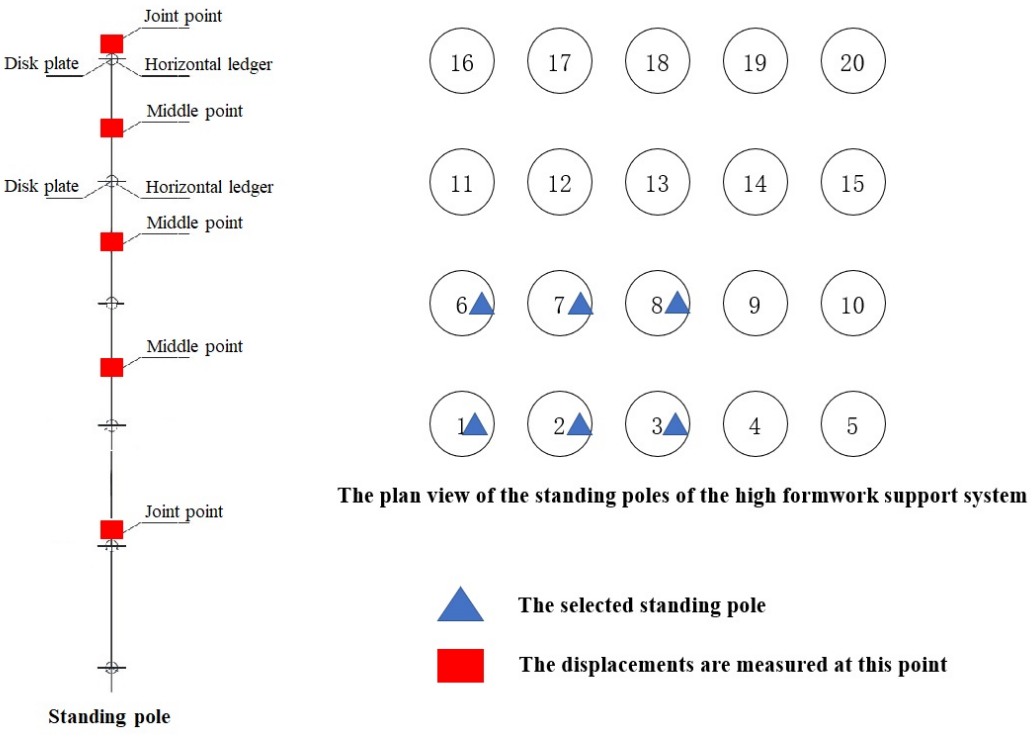
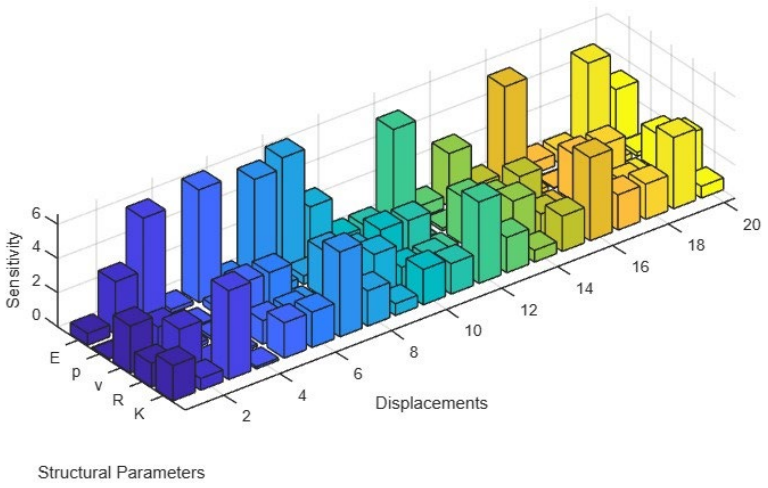


Figure 3. Arrangement of measuring points.

(a)



(b)

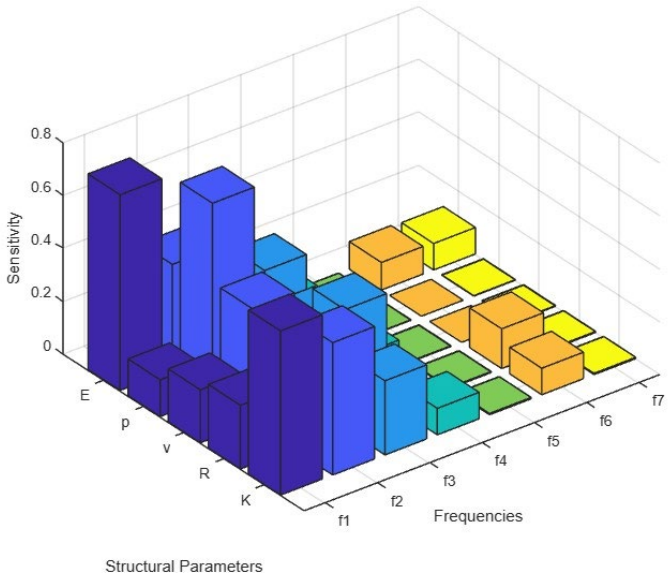


Figure 4. Sensitivity analysis**. (a)** displacement sensitivity; **(b)** frequency sensitivity.

3.3. Objective function definition

The selected model parameters should be updated. Firstly, suitable initial values are assigned to the model parameter. Then, their values are updated by employing an optimization process. To update all the selected parameters, a suitable objective function should be defined. Experimental tests should be carried out, and the structural responses, such as acceleration signals and displacements, must be collected. The objective function is built depending on displacement and frequency, as presented in Equation 1.

$$\min R(\theta) = \min \left(\sum_{i=1}^{m} \left(\frac{x_i' - x_i}{x_i} \right)^2 + \sum_{j=1}^{n} \left(\frac{f_j' - f_j}{f_j} \right)^2 \right) \quad (1)$$

Where x_i and x_i' indicate the displacement measurement and computed displacement, respectively; f_j and f_j' indicate the frequency measurement and computed frequency, respectively.

3.4. Experiment set-up

The experimental set-up of this work is presented in this paragraph. The set-up involves a five-storey high-formwork support system which is composed of steel pipes, including standing poles, horizontal ledgers, and diagonal brace bolted on joints to form a complete truss structure, as shown in Figure 5. The height of the system is nearly 8 m, and the lift height is 1.5 m. The length of the system is nearly 5 m, and the transverse space between two neighbor standing poles is 1.2 m. The width of the system is nearly 4 m, and the longitudinal distance between two neighbor standing poles is 0.9 m.

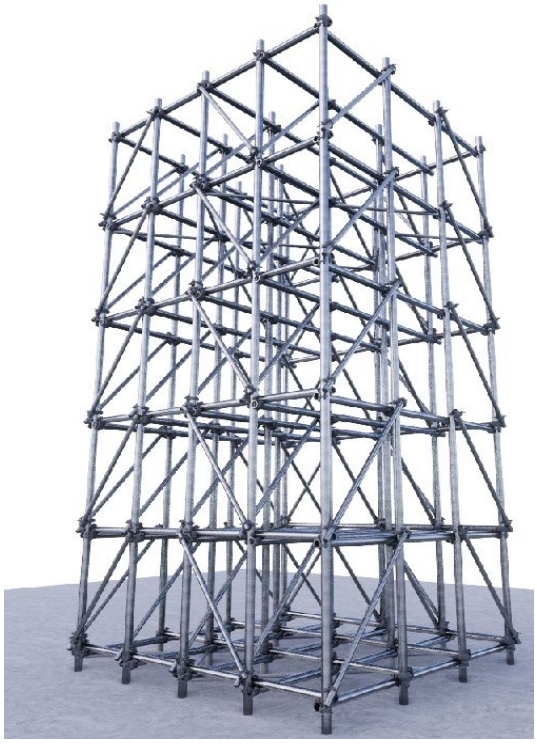


Figure 5. The full-scale of the high-formwork support system.

The structure was imposed static force by the instruments, as shown in Figure 6. Applying concentrated loads on the top of the structure uses four hydraulic jacks. The concentrated loads can be evenly transferred to every standing pole via the adoption of the reaction frame and distribution beams. Only triggered 2 of the 4 jacks can apply local loading on the structure. During the experiment, the load was applied to the structure in multiple steps. At every load step, 10 kN was applied to the structure. As the loading approaches the structure's ultimate bearing capacity, the loading increment should be reduced at each step. When the strain and displacement of the monitored standing poles were no longer developing, the next step loading was carried out. Until the displacements of the monitored standing poles still increase but the jacks are unloaded, the structure reaches its ultimate bearing capacity.

Moreover, an electrodynamic shaker was used as an excitation source. It may then excite the structure with impacts. A load sensor was employed to record the imposed impact forces. The acceleration history for experimental measurements during the testing phase is recorded at the accelerometer sensors. In this study, the frequencies of the structure were measured in X, Y, and Z directions. Hence, the sensors were installed in three directions. The deployment of the acceleration sensors is shown in Figure 7. A time-force curve for multiple impacts measured at the load cell sensor

and recorded accelerations is shown in Figure 8. The sampling frequency was selected at 0–100 Hz for all measurements. Based on the experimental loading steps, the frequencies of the structure can be measured several times. Due to the changes in stiffness matrix of the structure, the frequencies are different at each loading step.

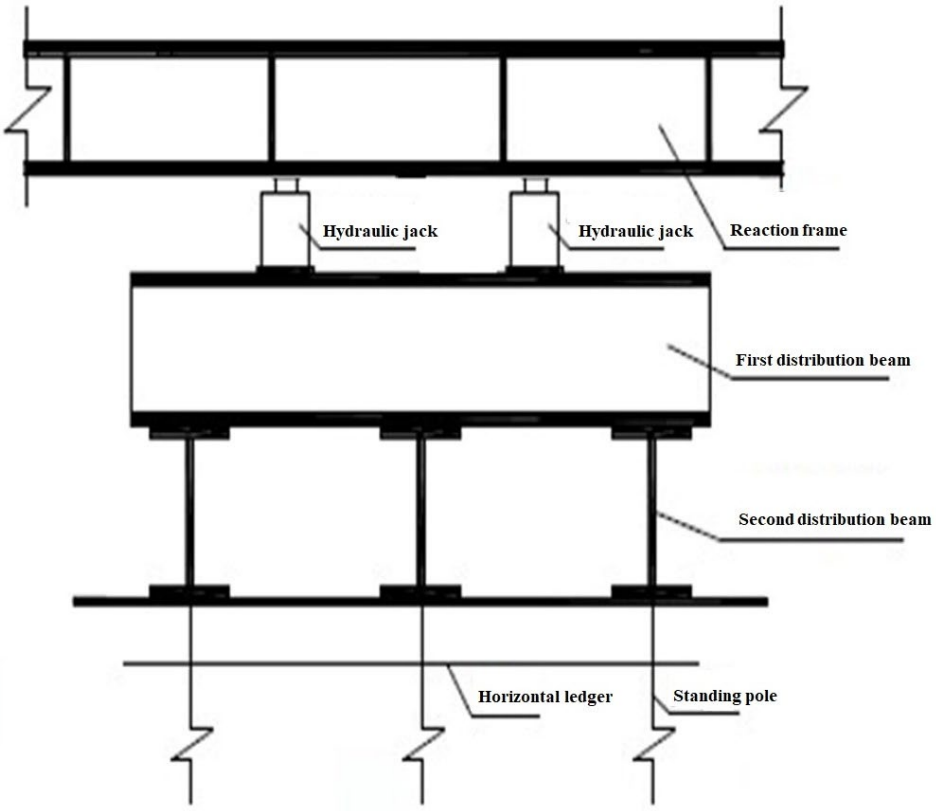


Figure 6. The sketch of the loading equipment.

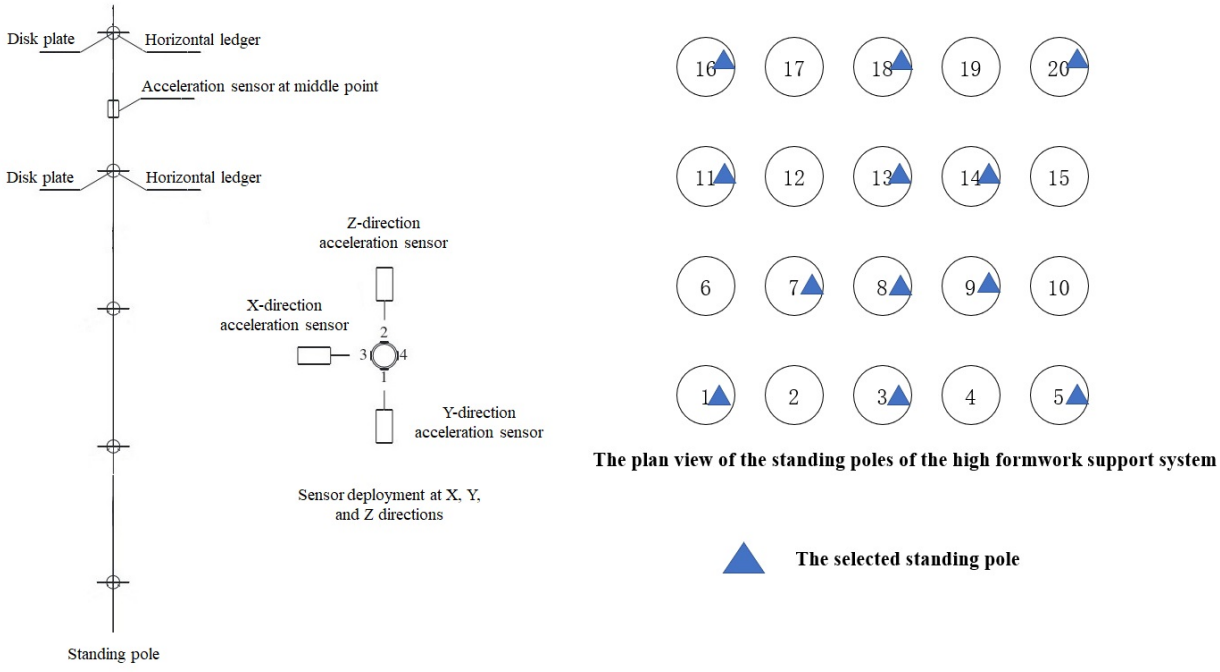
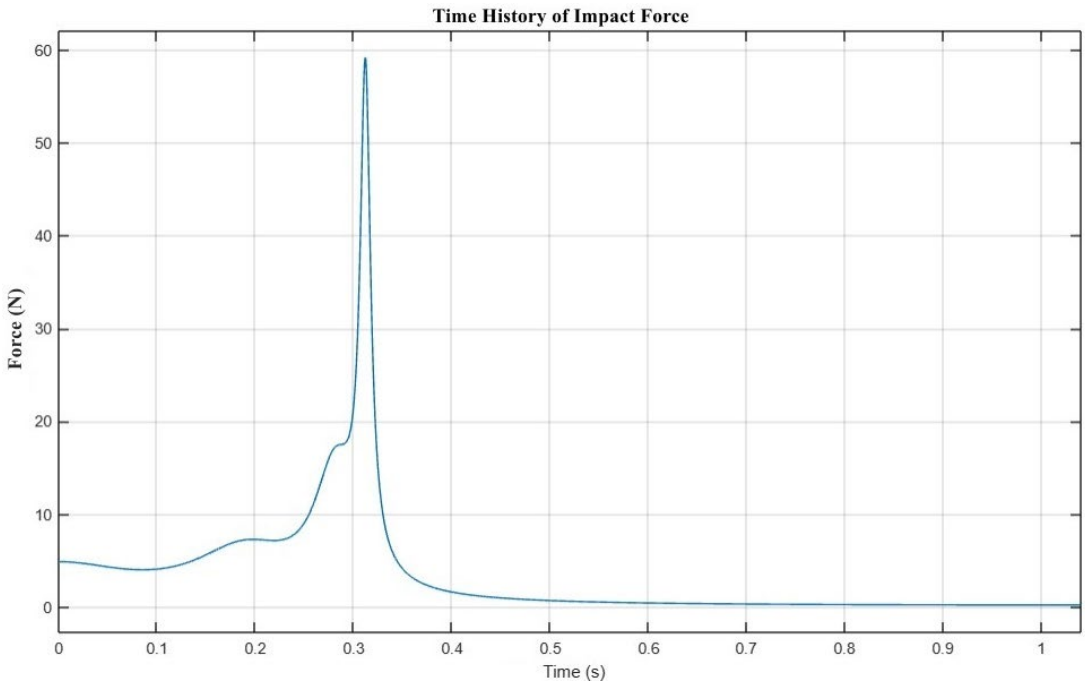


Figure 7. The arrangement of acceleration sensors.

(a)



(b)

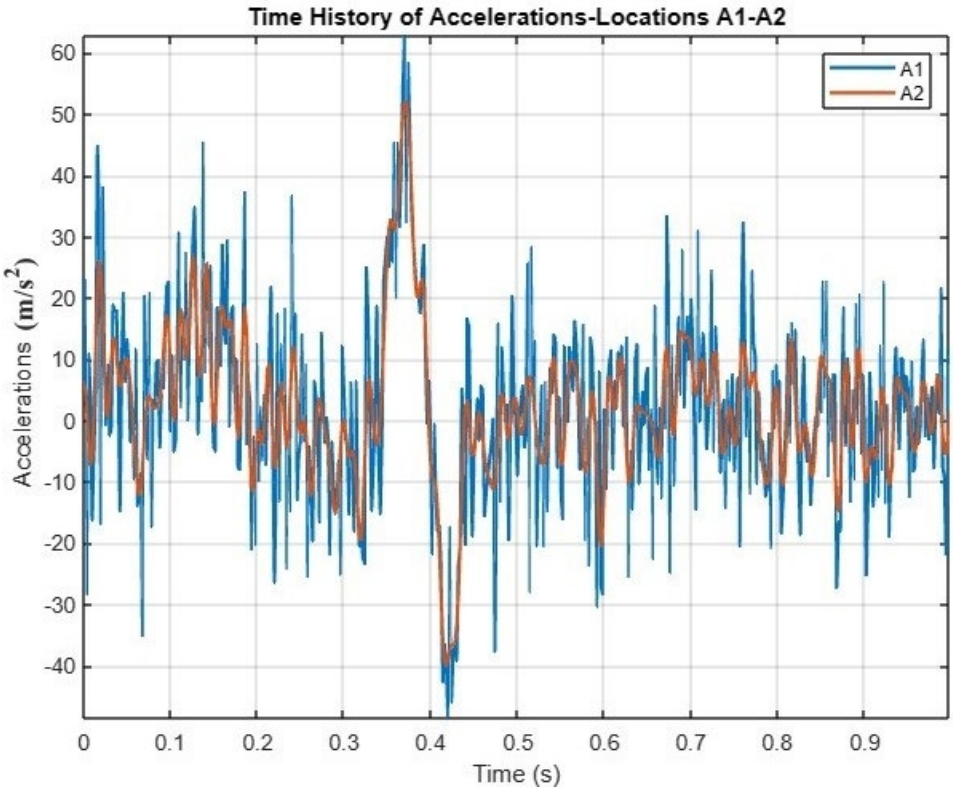


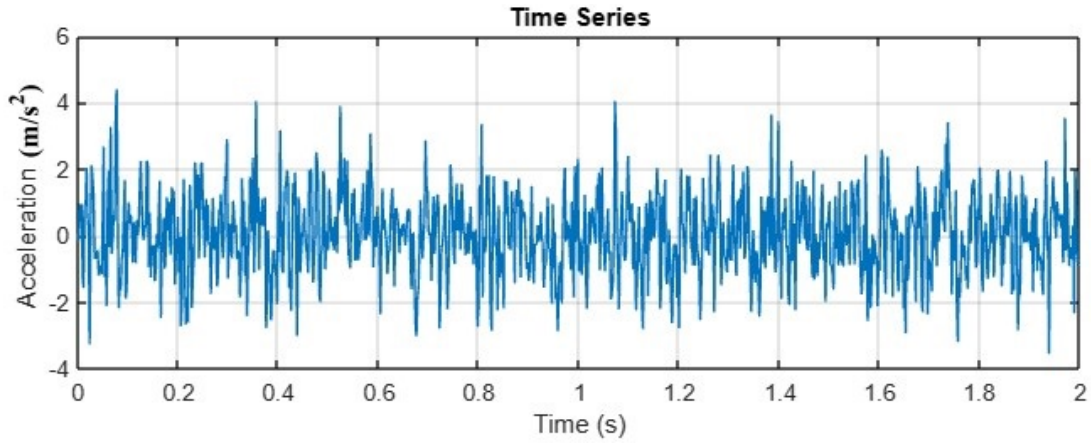
Figure 8. Measured force and acceleration during multiple impacts. **(a)** Recorded time history of impact force; **(b)** recorded time history of acceleration at points A1 and A2.

Displacement meters, total-station instruments, a levelling instrument, and a 3D laser scanner were used to measure the displacements of standing poles. A detailed explanation of using a laser scanner to measure the displacements of the standing poles on a high-formwork support system was described in [34].

3.4.1. Frequency and mode shapes

The frequencies and mode shapes can be derived from the on-site acceleration measurements by using a power spectral density (PSD) technique. The ability of PSD to identify modal parameters including frequency and mode shape has been applied and verified in many studies. In this study, PSD has been applied on acceleration signals to identify the frequency and mode shapes of the structure. Based on the theoretical background in [35,36], the PSD value always has a sharp change around the frequency and achieves a local maximum value at the natural frequency. The time history of the vibration response signal for a typical measurement point is shown in Figure 9a. The power spectral density curve of the measurement point is attained, as shown in Figure 9b. It is obvious that three peaks are occurring at 3.024, 3.152, and 4.826 Hz, respectively.

(a)



(b)

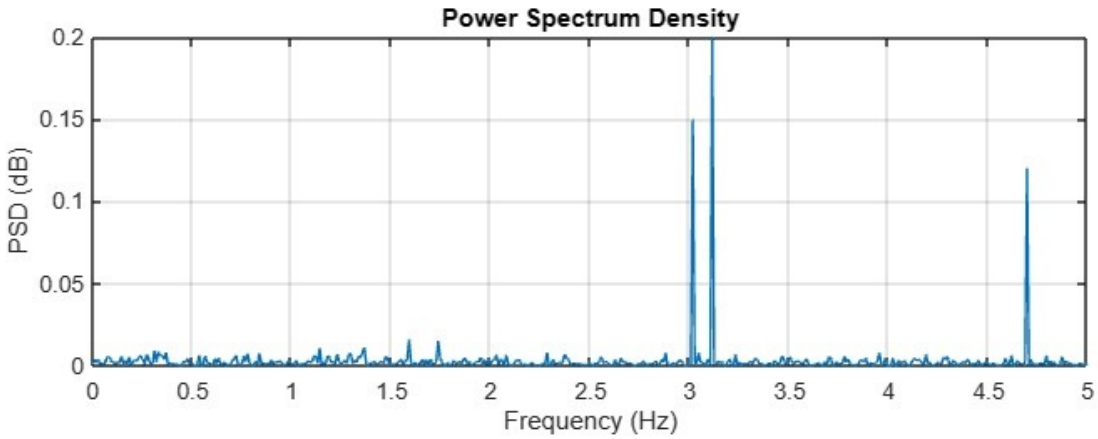


Figure 9. Acceleration & PSD. **(a)** Time history of acceleration; **(b)** Power spectrum density.

3.4.2. The simulation of abnormal working conditions of the structure

The structure can work properly referred to as the normal working status. To evaluate the performance of the proposed approach for the classification of the abnormal working status of the structure, two main types are considered: local instability status and fully instable status. The first case, referred to as local instability, indicates that local load was applied on the real structure to make some standing poles reach

their ultimate bearing capacity. The second case, referred to as fully instable, denotes that the structure reaches its ultimate bearing capacity. The goal is to evaluate the accuracy of the classification of the deep learning (DL) model.

The simulation of the two cases might demand a complex FE analysis. However, the effect of the cases can be observed via the structure's frequency and mode shape as they diverge from the normal working status. A comparison of the first four mode shapes of the structure for normal working status and local instability status is shown in Figure 10. The first 4 modes of the structure still remain as bending and torsion in a normal working state, while that of the local instability case stays away from rigid element stiffness. In addition, time response can also provide evidence to verify the local instability case. A direct comparison in acceleration time series between responses of the normal working status and local instability status is shown in Figure 11. The time responses with reduced stiffness diverge from the normal working state.

In this study, in order to attain the experimental measurements of the structure, a full-scale high-formwork support system was built in the lab. At first, local forces have been applied on the structure to attain the measurements for the status of local instability. Then, the standing poles that reach their ultimate bearing capacity have been replaced by new ones. When adequate measurements were obtained for status of local instability, full load was applied to collect the measurements until the structure reached its ultimate bearing capacity.

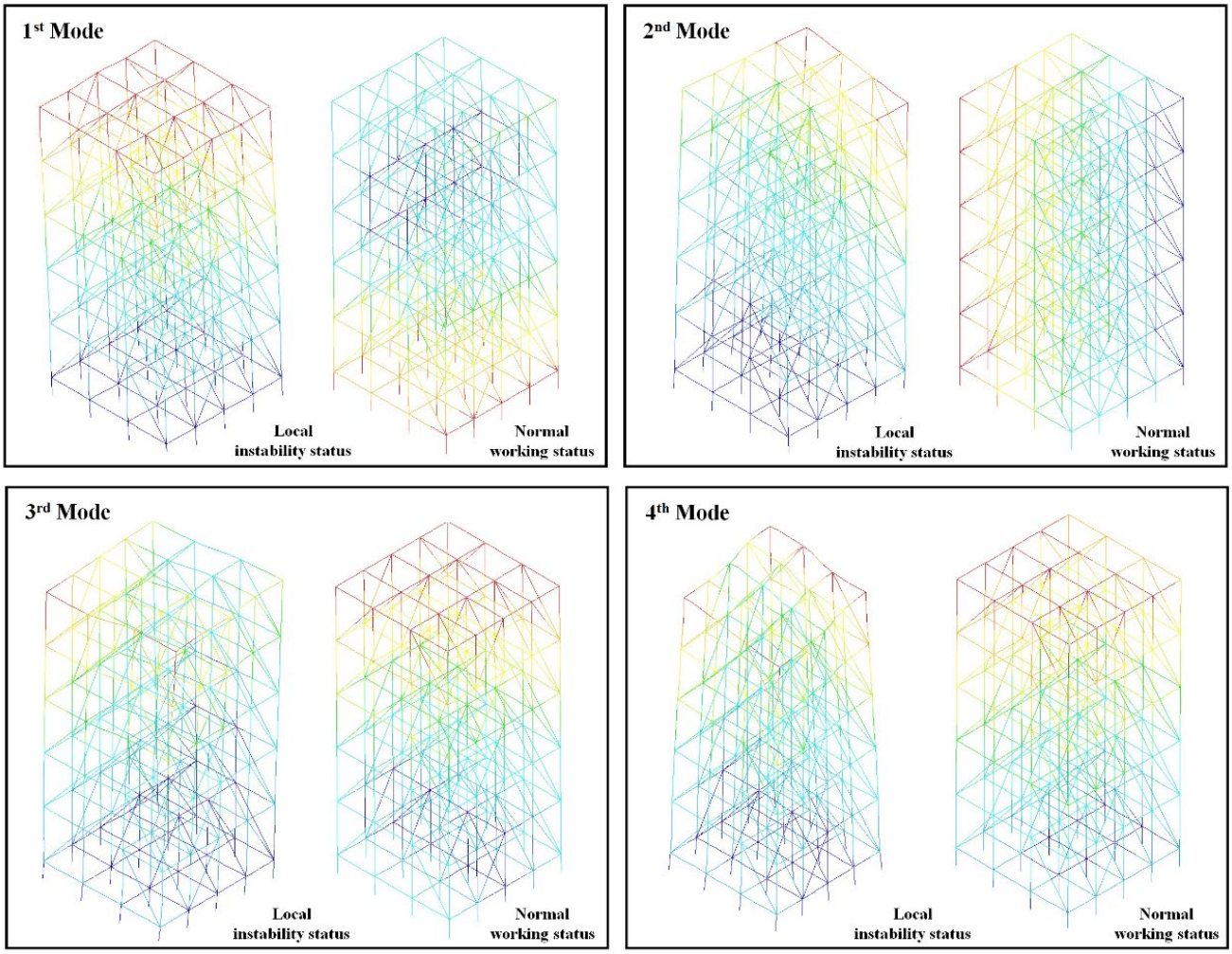


Figure 10. The first four modes of the HFSS for normal working status and local instability status.

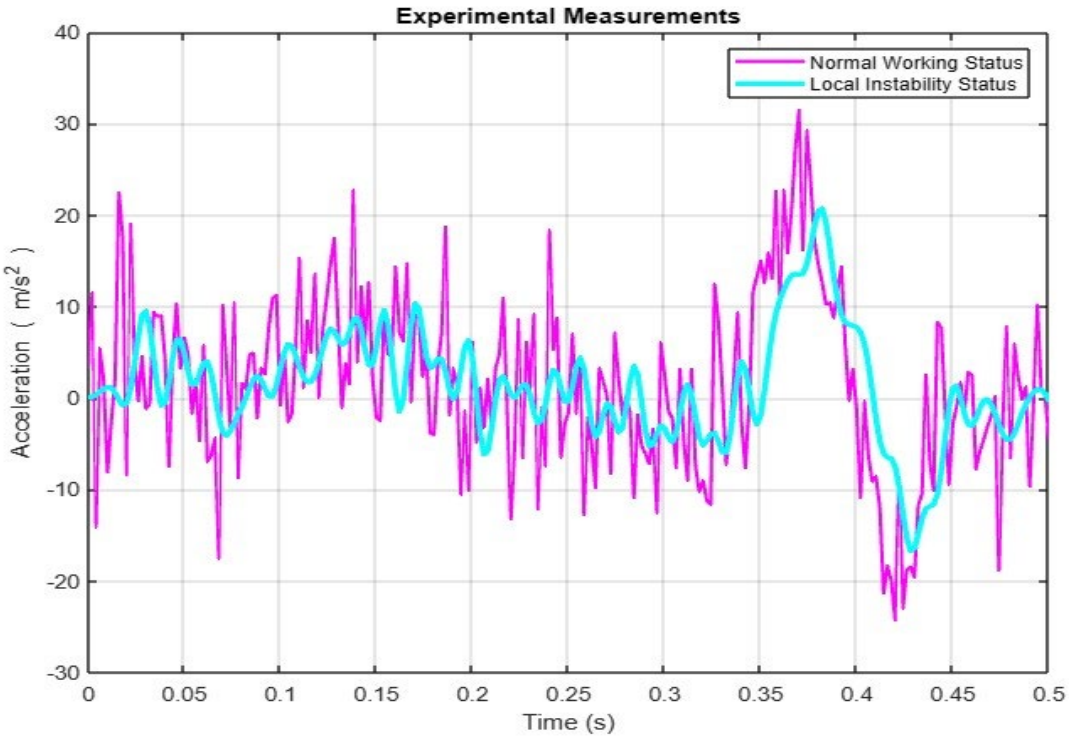


Figure 11. Comparison of measured acceleration time responses for the normal working status and local instability status of the high-formwork support system.

3.5. A comparison of the results derived from different optimal algorithms

The proposed objective function was minimized using three optimization algorithms: gradient-based optimizer (GBO), genetic algorithm (GA), and grey wolf optimizer (GWO). The statistical results of FEMU when using different optimization algorithms are summarized in Table 2. It can be observed that the performance of GA is the most suitable and reasonable. The first four frequencies and corresponding mode shapes of the structure derived from FEM completely agree with the measured modal data. At the same time, the statistical results of FEMU using GBO and GWO provide a satisfactory result for displacements of standing poles. However, some challenges still exist in having a good correlation between the computed and experimental frequencies. On the whole, the process of FEMU using GA algorithm is the most efficient compared to those other algorithms. The convergence curves of the objective function values using GA, GBO and GWO are shown in Figure 12. It clearly shows that the convergence of GA and GBO is better than that of GWO, which indicates that these technologies can generate more acceptable results. In addition, the convergence speed of the GA is much faster than GBO. The GBO requires more iterations to converge.

A comparison between the forecasting modes from the FE model optimized by GA and the test modes is displayed in Figure 13. The pairing between predicted and test modes seems good from the diagonal dominance of modal assurance criterion (MAC) which arrays up to the eighth mode. The frequencies computed by the optimal and nominal FEMs, and test frequencies are summarized in Table 3. A MAC value of 1 denotes a strong correlation between mode shapes, while a value close to 0 indicates no correlation. In fact, MAC values should be close to 1, which suggests that modes are matching.

Table 2. Statistical results of FEMU from different optimization algorithms.

Mode		1	2	3	4
Measured frequencies (Hz)		3.024	3.152	4.826	5.397
GA	mean f (Hz)	3.139	3.272	5.009	5.602
	std f	1.474E-02	3.081E-02	2.461E-02	2.783E-02
GBO	mean f (Hz)	3.211	3.347	5.125	5.732
	std f	2.845E-02	3.675E-02	3.724E-02	3.897E-02
GWO	mean f (Hz)	3.236	3.373	5.164	5.775
	std f	2.547E-02	3.112E-02	5.012E-02	3.962E-02

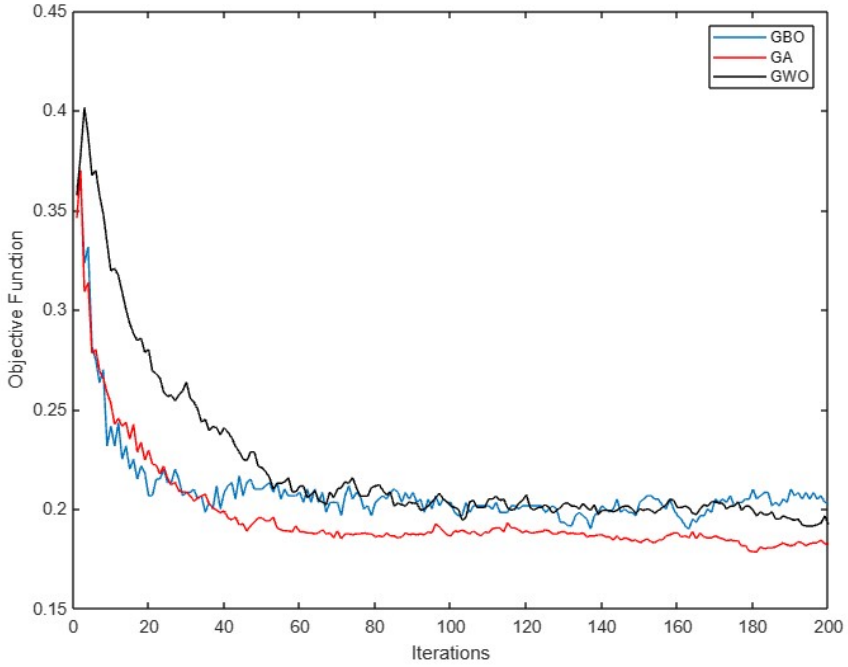


Figure 12. The convergence curves of the objective function values of the three optimization algorithms.

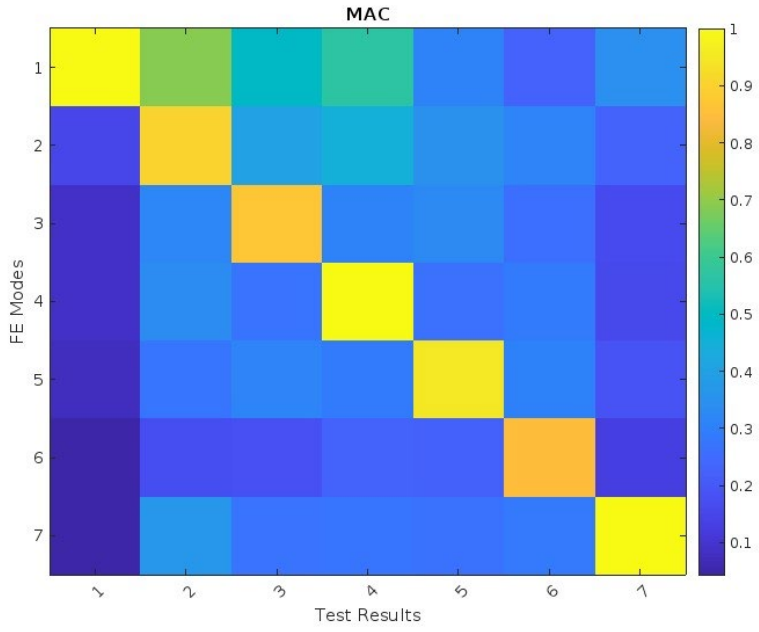


Figure 13. MAC table after optimization.

Table 3. Natural frequencies (Hz) and errors (%) between the FEMs and measurements.

Mode	Measured	FEM (before updating)		FEM (after updating)	
	Frequency	Frequency	Error	Frequency	Error
1	3.024	3.325	9.95	3.056	1.05
2	3.152	3.512	11.42	3.252	3.17
3	4.826	4.931	2.17	4.801	-0.51
4	5.397	5.724	6.05	5.459	1.14
5	8.143	8.825	8.38	8.171	0.34
6	10.051	9.625	-4.24	9.613	-4.36
7	11.331	11.760	3.79	11.465	1.18
$\sum\|Error\|/8$			6.57		1.68

4. FE model data training CNNs and experimental validation

4.1. FEM data training CNNs

In this section, the learning data generation derived from FE models and learning procedures of CNNs are discussed. The network is trained on FEM produced data in two types, nominal and optimal FE produced data, in order to show performance gains of optimal FEM generated data on forecasting the structural statuses.

The optimal FEM discussed in section 3 was adopted to generate training data for CNNs. The FE equation of motion is presented in Equation 2.

$$\begin{gathered} \mathrm{M}(\rho)\mathrm{A} + \mathrm{C}(\mathrm{K},\mathrm{M},\alpha_1,\alpha_2)\mathrm{V} + \mathrm{K}(\mathrm{E},\mathrm{I})\mathrm{U} = \mathrm{F} \\ \mathrm{C} = \alpha_1\mathrm{M} + \alpha_2\mathrm{K} \end{gathered} \tag{2}$$

Where U, V, and A are the displacement, velocity, and acceleration vector respectively. K, C, and M represent stiffness matrices of the structure, damping, and global mass that relate to the model parameters of elasticity modulus E, damping α_1, α_2, and density ρ. The displacements of standing pole and structural frequencies used for CNNs learning were simulated for arbitrary load cases by using Algorithm 1. The status labels Y were the categories of the working statuses of the structure in this work.

Algorithm 1 Finite element model data generation algorithm

Input: Number of load cases n and upper (up) and lower (lo) statistical bounds for each quantity
Output: n displacements and n number of labeled acceleration vectors

1. **for** i =1:n **do**
2. define status y according to FE model→ $\boldsymbol{Y}$
3. sample E→ $\boldsymbol{K} = \boldsymbol{K}(E)$
4. Solve $\boldsymbol{KX} = \boldsymbol{F}$
5. **return** displacement X
6. sample $\rho \rightarrow \boldsymbol{M} = \boldsymbol{M}(\rho)$
7. Solve $\boldsymbol{MA} + \boldsymbol{KX} + \boldsymbol{CV} = \boldsymbol{F}$
8. **return** $\boldsymbol{A}$ *and* $\boldsymbol{Y}$

End

$$\mathrm{Train}_{\mathrm{set}} = \{(\mathrm{F}_1,\mathrm{U}_1,\mathrm{Y}_1),(\mathrm{F}_2,\mathrm{U}_2,\mathrm{Y}_2),\ldots,(\mathrm{F}_n,\mathrm{U}_n,\mathrm{Y}_n)\} \tag{3}$$

Data generated from nominal FEM was employed with the purpose of showing the difference in the CNNs results. The nominal and optimal FEMs were developed using the ANSYS software package on an Intel i7-8700 PC with 64GB of RAM. The number and properties of produced data are shown in

Table 4. Cases were simulated for all the three statuses (normal working, local instability, and fully instable) of the structure generated from both nominal and optimal FEMs.

Table 4. FE produced data for the CNNs.

FE model	Optimal	Nominal	Optimal	Nominal	Optimal	Nominal
Case	NWS	NWS	LIS	LIS	FIS	FIS
Number of cases	1000	1000	1000	1000	1000	1000

NWS = normal working status, LIS = local instability status, FIS = fully instability status.

A total of 3000 cases were generated for all the various statuses of the structure. Each status includes 1000 cases. For 3000 input-output pairs extracted from the FEM simulator, in CNNs training and testing, 3/4 of them are used for training, while the remaining 1/4 are utilized for testing in order to assess the classification performance of the proposed CNNs model.

The training results for data produced from nominal and optimal FEMs are summarized in Table 5. The CNNs model trained from the optimal FEM dataset achieves the best forecasting results with 89.4% global accuracy, which indicates that CNNs can adequately learn the characters and perform well on this problem. For the CNNs trained by nominal FEM datasets, the prediction results are significantly different with only 65.7% accuracy. It proved that the data derived from optimized FEM can really improve the performance of a CNN classifier.

Table 5. Training results for data derived from optimal and nominal FEM.

	Optimal FEM data	Nominal FEM data
Number of learning cases	3000	3000
Mean learning accuracy	89.4%	65.7%

Using the above datasets, a CNNs classifier was well-trained. There are four independent convolutional filters in the CNNs. The structure of the network is displayed in Figure 14 and Table 6. The network architecture includes the number of neurons, filters, and filter length, which were selected over various parameters. The simple and powerful network with the fewest parameters which achieved good training results on the learning data, was selected. The learning results are summarized in Table 7 in terms of mean accuracy for each status of the structure. Model checkpoints were used to avoid overfitting. The learning scores denote that the CNNs can classify the three statuses of the structure using the optimal FEM generated data with scores exceeding 80%. The training process is shown in Figure 15. The trained CNNs were finally used to forecast the status of the structure.

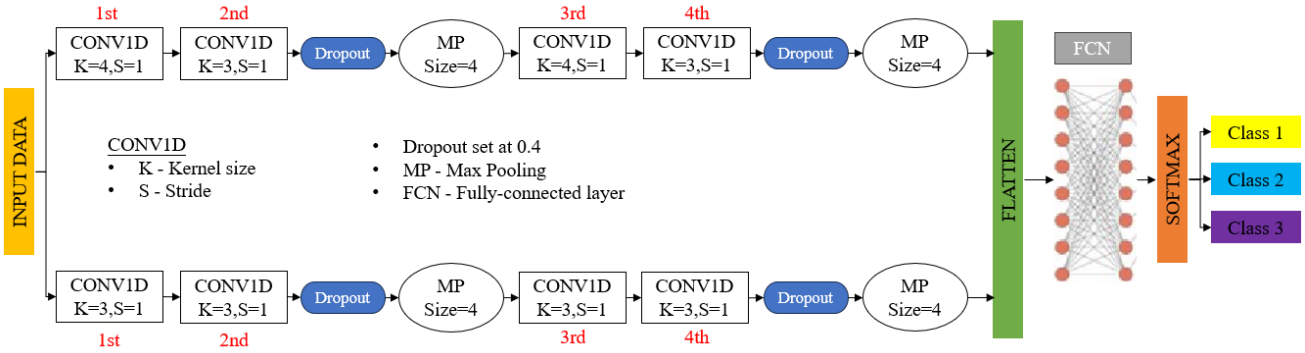


Figure 14. The proposed CNNs classifier architecture.

Table 6. The architecture of the proposed CNNs.

Hyperparameter	Predefined value
Activation	Relu, softmax
Loss	Categorical cross entropy
Optimizer	Adam
Down-sampling	Max-pooling
Overfitting algorithm	Early stopping
Learning rate	0.001
Max training epochs	100
Metrics	Accuracy
Validation split	0.2

Table 7. The learning and validation results for CNNs classifier.

Case	NWS	LIS	FIS
Number of learning cases	1000	1000	1000
Number of validation cases	200	200	200
Mean learning accuracy	91.3%	87.7%	89.2%
Mean validation accuracy	88.1%	83.8%	87.6%

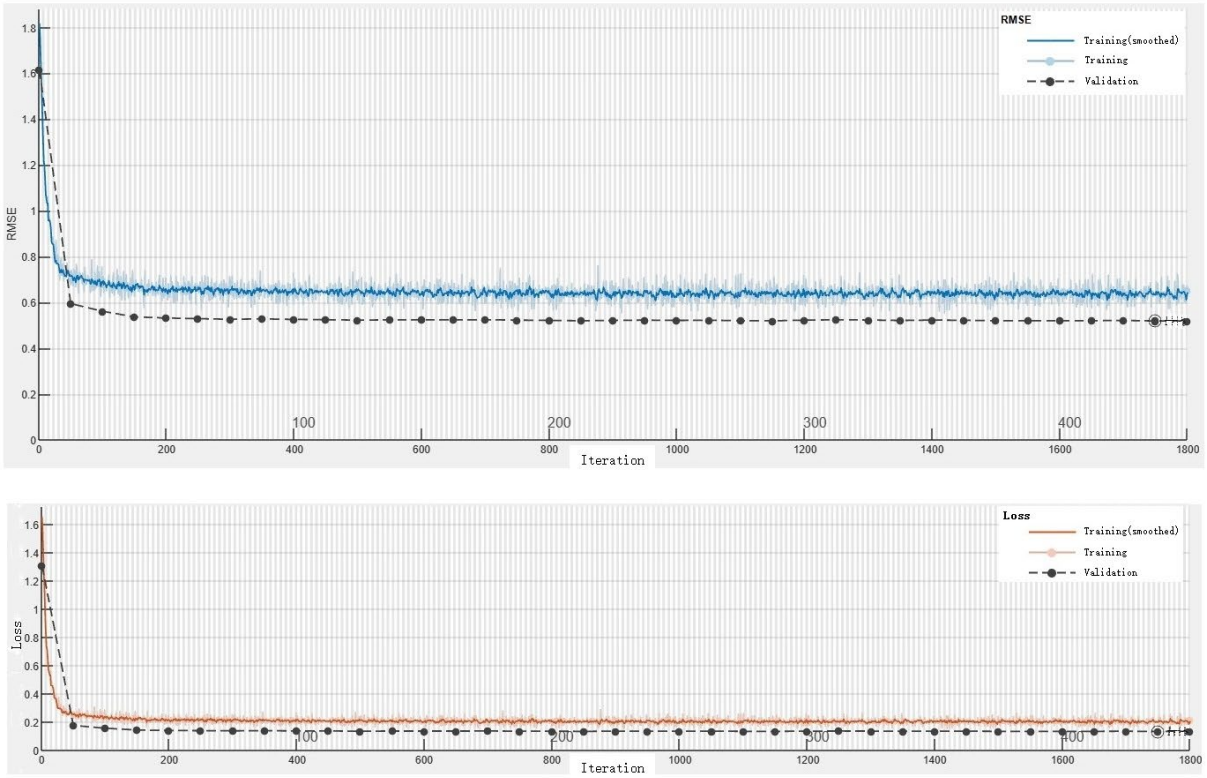


Figure 15. Accuracy and loss curves for the training process of the CNNs.

4.2. Experimental validation on trained CNNs

An experimental validation for the CNN's classifier is presented in this paragraph. Moreover, the performance of the CNNs and support vector machine (SVM) was also compared. Other neural networks, such as Long-Short Term Memory (LSTM) and Gated Recurrent Unit (GRU) [37], were also considered, but they did not show any better results compared with CNNs. Moreover, complicated

networks might cause an overfitting problem. The experimental dataset contains 300 measurements, and 100 measurements for each status of the structure. Class forecasting scores derived from CNNs and SVM are presented in Figure 16. The top charts indicate class scores from normal working status, while the middle and bottom charts are class prediction results from local instability status and fully instable status, respectively. The left charts show SVM forecasting results, and the right ones are CNNs. The class score ranges from 0 to 1, and the sum of the three classes is equal to 1.

(a) (b)

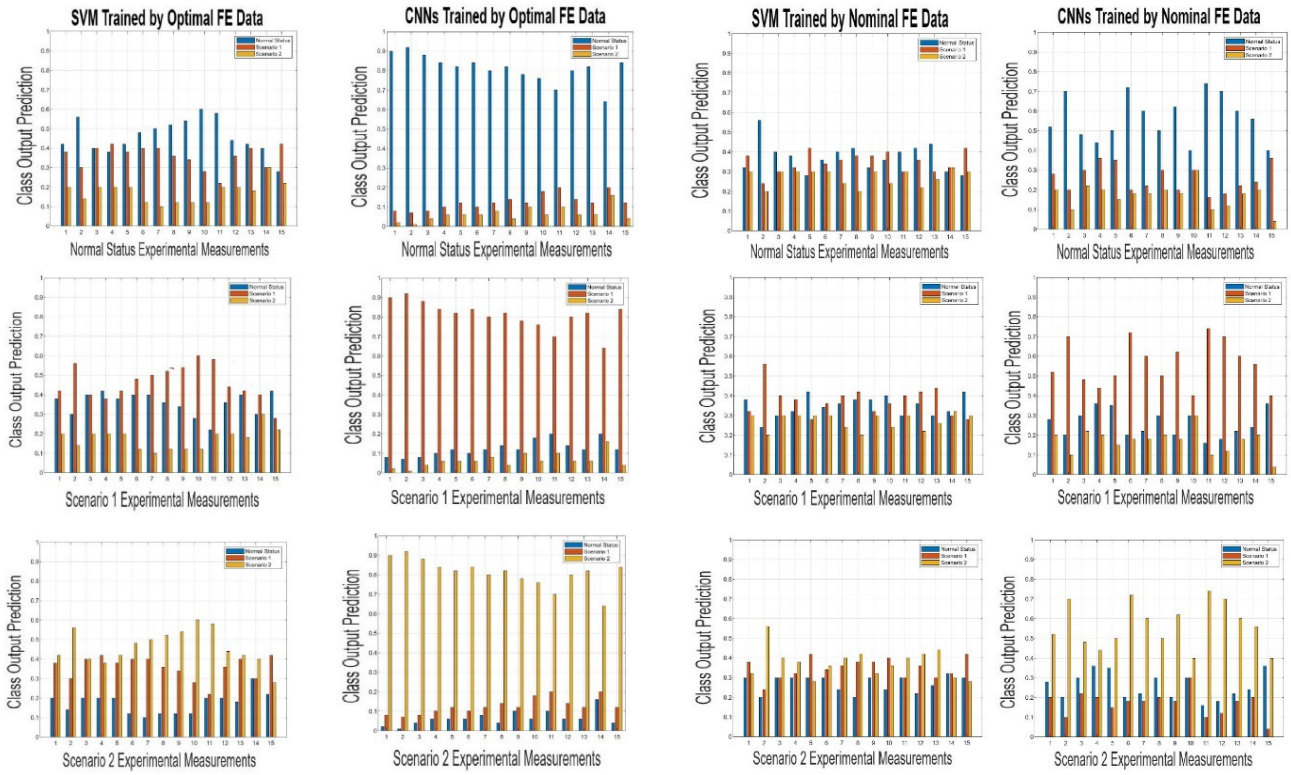


Figure 16. Class forecasting scores. **(a)** Status predictions on experimental measurement inputs for networks trained with optimal FE model; **(b)** Status predictions on experimental measurement inputs for networks trained with nominal FE model.

Only 20 measurements are displayed in Figure 16. There is a total of 300 measurements in confusion matrices, 100 of each set as displayed in Figure 17. Each class corresponds to 100 inputs. The representation of the elements in the confusion matrix is explained below. Diagonal elements indicate the numbers of accurately classified cases (together with their percent), while off-diagonal elements denote wrongly classified ones. The bottom diagonal element represents the overall precision (in green) and error (in red). Right columns indicate accuracy (in green) and error rate (in red). Bottom rows denote recall (in green) and false discovery rate (in red). The CNNs generated better scores for each class than SVM did. The multiple filter length can learn adequate features compared to a single filter classifier. The high validation accuracy indicates that the trained CNNs can minimize the difference between simulated and actual responses.

(a) (b)

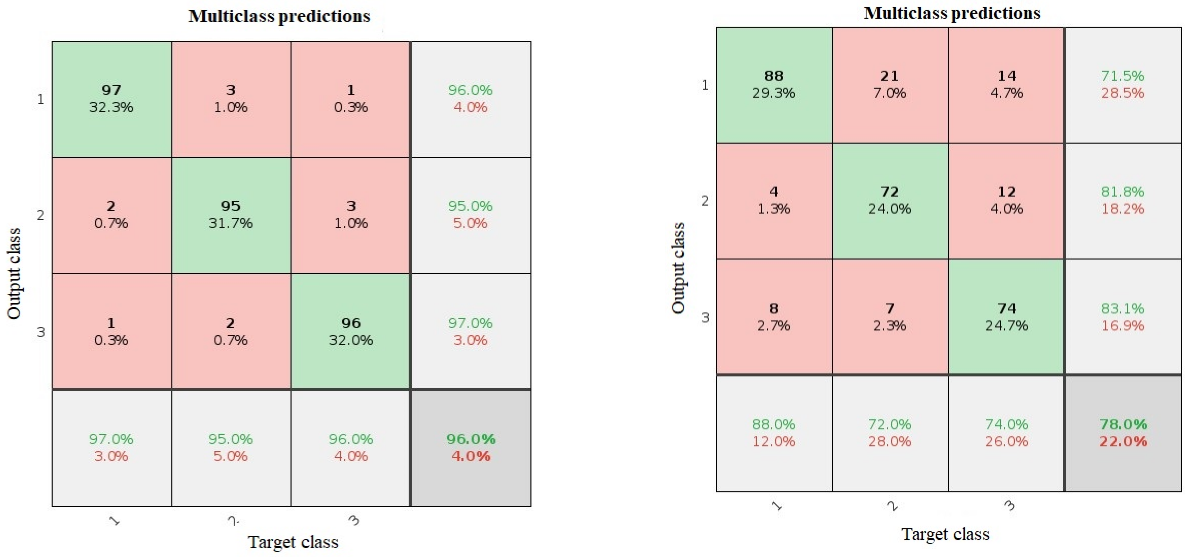


Figure 17. Confusion matrix from CNNs. **(a)** Confusion matrix for the multiclass predictions using CNNs trained by optimal FEM data; **(b)** confusion matrix for the multiclass predictions using CNNs trained by nominal FEM data.

Meanwhile, SVM for both optimal and nominal FEM data show poor performance in all cases. As shown in Figure 18, the SVM has a problem forecasting the class in a decisive way, as the scores are below 0.5 for most cases. The results indicate that only the CNNs trained with optimal FEM data could generate acceptable results while the SVM trained with nominal FE data generate results that show high inconsistency with the real status.

(a) (b)

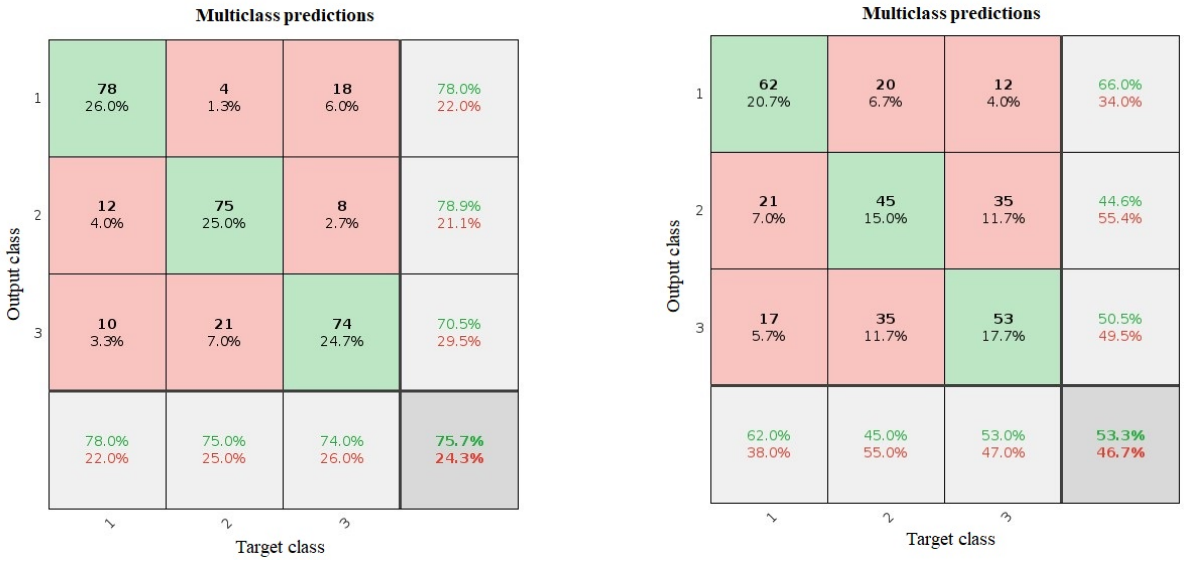


Figure 18. Confusion matrix from SVM. **(a)** Confusion matrix for the multiclass predictions using SVM trained by optimal FEM data; **(b)** confusion matrix for the multiclass predictions using SVM trained by nominal FEM data.

5. The development process of the RAG model

5.1. The construction of the knowledge graph

This research adopted a bottom-up approach to construct the knowledge graph (KG), illustrated in Figure 19. The initial data collection phase incorporates multiple sources, such as guidelines, case studies, and relevant online materials. Next, data cleaning by using a deep-learning language error correction model was conducted to enhance the quality of the KG. The study employed NLP-based named entity recognition (NER) to identify and extract entities along with their relationships. The NER method, which incorporates parts of speech (POS) tagging, identifies subjects and objects as entities, while the verb or main action in the sentence is classified as a relationship. Then, entity identification, relation extraction, and attribute definition were conducted. After that, knowledge integration is required. Some operations, including cloning, merging, insertion, or the removal of nodes on the graph, have been conducted. Next, the processed structured data was imported and stored in Neo4j to complete the task. Finally, the developed KG was used as an external database. Figure 20 presents the tools employed during the CKG development process.

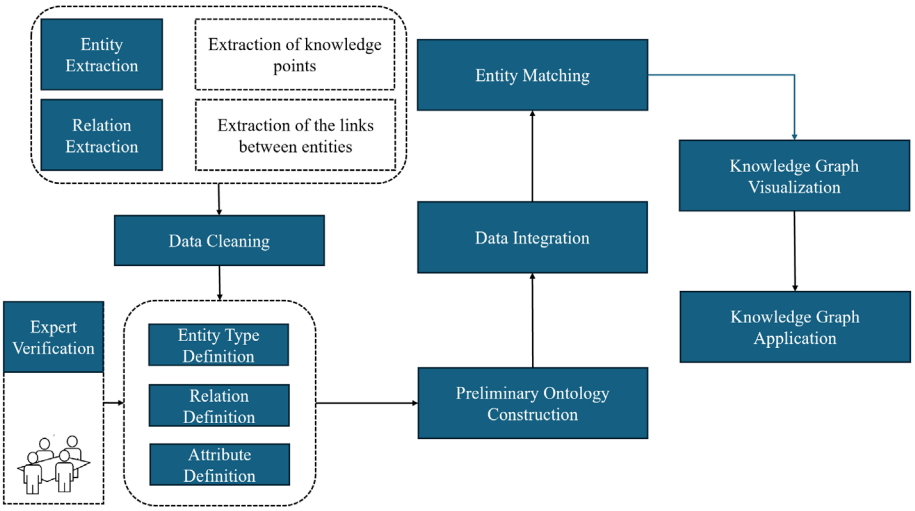


Figure 19. Flow chart of the course knowledge graph development.

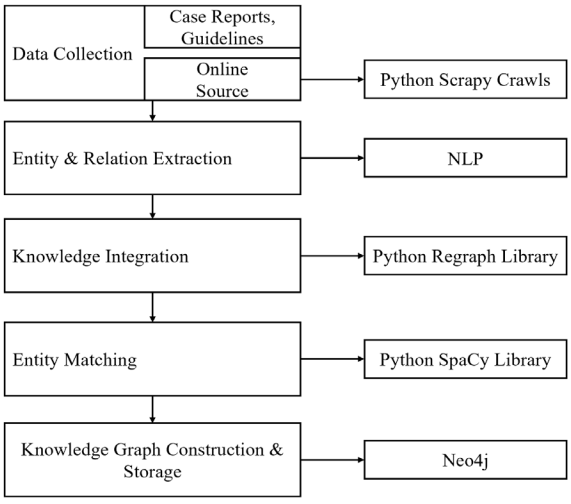


Figure 20. The development tools for every stage of the construction of the CKG.

5.2. The RAG model

The RAG (Retrieval-Augmented Generation) framework integrates three key components: knowledge databases, information retrieval systems, and generative AI models. The developed RAG model begins with knowledge graph queries before producing responses. A knowledge graph was constructed from structural health monitoring (SHM) guidelines and case studies. The knowledge graph serves as an interactive queryable database for the RAG model, enabling generation of more accurate and informed responses, especially in complex domains such as structural health monitoring.

The developed RAG framework integrates with GPT-4 through LangChain, utilizing the knowledge graph as an external knowledge base. When receiving graph-based query questions by using Langchain, users' queries are identified and then transferred into cypher queries. The Langchain library connects GPT with a Neo4j graph database. The cypher queries are employed to search the entities and relationships in the knowledge graph, which are then sent back to GPT to help generate accurate and reliable answers. The workflow is shown in Figure 21.

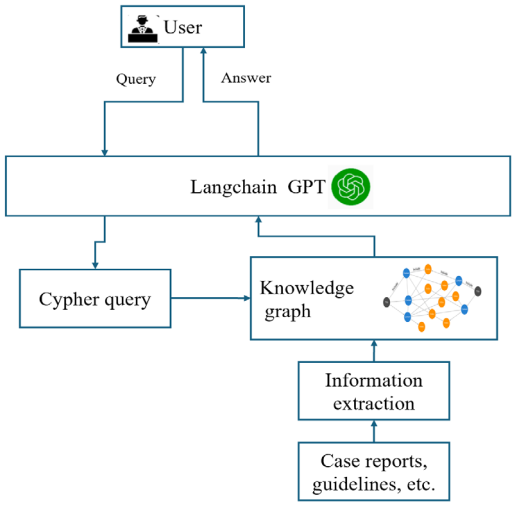


Figure 21. The development process of the RAG.

6. Verification of the developed RAG model

The assessment was performed by comparing the SHM reports generated by the developed RAG model and GPT-4 model based on the project summary. Several sections should be included in the final SHM report to gain an understanding of the situation of the HFSS. The sections are 1-project profile, 2-monitoring objectives, 3-monitoring methods and equipment, 4-monitoring results, 5-data analysis and conclusion. This comparison is intended to evaluate how efficiently each model can generate relevant SHM reports for HFSS structures in accordance with the presented project scenarios and contrast their performance using several metrics.

The study utilized both cosine similarity and Euclidean distance metrics, computed through BERT [38] and OpenAI's embedding model. The cosine similarity metric ranges from 0 to 1, with values approaching 1 indicating greater similarity between compared items. Similarly, a smaller Euclidean distance value suggests greater similarity.

To evaluate text correspondence, the study employed both BLEU [39] and ROUGE [40] metrics to quantify alignment between reference reports and generated responses. The BLEU metric quantitatively

assesses the correspondence between generated responses and human productions, evaluating both linguistic accuracy and semantic coherence. On the other hand, the ROUGE metric evaluated the quality and reliability of the generated responses, focusing on how effectively the generated response captures the important content.

The comparative analysis incorporated 38 case studies to evaluate model efficacy, with detailed outcomes presented in Table 8. Quantitative metrics BLEU and ROUGE are shown in Table 8, reflecting the model's proficiency in identifying keywords and capturing domain-specific terminology from reference reports. The cosine similarity and Euclidean distance metrics quantitatively demonstrate the semantic similarity between model-generated responses and reference materials.

The computed cosine similarity and Euclidean distance metrics exhibit minimal differences among the evaluated models. This phenomenon reflects that the generated responses closely align with the referenced texts. However, both the BLEU and ROUGE scores exhibit significantly higher for the RAG model than those produced by GPT model. This indicates that the responses generated by RAG model are more closely aligned with the referenced reports. Compared to GPT model which usually generates more general information, the developed RAG model could deliver specialised knowledge using accurate and specific keywords and terminology related to SHM domain.

While the RAG model achieves superior performance in BLEU and ROUGE metrics, GPT exhibits greater cosine similarity scores. This stems from how large language models are trained: extensive datasets enable them to interpret queries and produce domain-agnostic responses. While such enhanced generalization performance is counterbalanced by a reduction in the production of domain-specific responses.

The verification phase reveals distinct response patterns: RAG generates domain-specific outputs through knowledge graph retrieval, whereas GPT produces more generic informational content. The generated response from RAG model surpasses that from GPT model. However, as the RAG model retrieves information based on the knowledge graph, it lacks diversity in responses.

Table 8. Quantitative metrics of the responses produced by the models.

		RAG	**GPT-4**
BLEU(Bi-gram)		0.054	0.039
BLEU(Uni-gram)		0.221	0.125
BERT	Cosine	0.873	0.846
	Euclidean	7.118	7.224
OpenAI	Cosine	0.623	0.635
	Euclidean	0.942	0.915
RougeL	Precision	0.052	0.033
	Recall	0.078	0.093
	F1	0.054	0.016
RougeLsum	Precision	0.067	0.042
	Recall	0.105	0.133
	F1	0.069	0.048
Rouge1	Precision	0.067	0.045
	Recall	0.101	0.124
	F1	0.071	0.059
Rouge2	Precision	0.006	0.002
	Recall	0.100	0.118
	F1	0.007	0.005

7. Conclusion

In conclusion, a DL model is provided where working statuses of a high-formwork support system are identified by a CNNs classifier. The proposed method can be divided into FE model updating, data production from the optimal FE model, and the CNNs classification steps. An experimental set-up of a real high-formwork support system was established to attain the measurements both for the optimization of FE model and the validation for the CNNs classifier. In the FE model updating process incorporating GBO, GA and GWO optimization algorithms, systematic evaluation showed the GA approach yielded the most reliable parameter estimations.

The training data for CNNs were obtained from FEM simulations. Results demonstrate that datasets produced by optimized FEM models deliver satisfactory forecasting accuracy, thereby minimizing the need for extensive experiments. In contrast, nominal FEM-generated data proves insufficient for addressing multi-class classification tasks. Therefore, implementing FEMU (Finite Element Model Updating) techniques at the preliminary stage can significantly enhance the predictive capability and robustness. Although a slight deviation was found between the experimental validation and the data generated from FEMU, it did not impact the classification of the working status prediction for HFSS. Finally, the CNNs and SVM were compared based on experimental measurements. The CNNs include multiple filters, which are able to learn important data properties and produce better results compared with SVM.

This research empirically validates the implementation potential of domain-specific LLMs for structural health monitoring applications. Through the deployment of a domain-specific LLM, this research demonstrated that the customized LLM can automatically generate accurate structural health monitoring reports for the HFSS. The developed RAG model streamlines the structural monitoring documentation process by reducing both time and labor costs.

The proposed method relies on simulation data from finite element models, whose accuracy is sensitive to parameter assumptions (e.g., linear material behavior). Unmodeled nonlinearities or localized defects in real-world structures may degrade classifier performance. The knowledge graph is domain-specific (HFSS-focused). When applied to other engineering scenarios (such as bridge monitoring), it requires the reconstruction of a domain-specific knowledge graph, thereby increasing deployment costs. The future study might focus more on the adoption of LLM in SHM to minimize domain-specific KG development efforts.

Abbreviation

AE: Autoencoder; ANFIS: Adaptive Fuzzy Neuro Inference System; ANNs: Artificial Neural Networks; AI: Artificial Intelligence; BNNs: Bayesian Neural Networks; BP: Back Propagation; CE: Cross Entropy; CNNs: Convolutional Neural Networks; DL: Deep Learning; FE: Finite Element; FEM: Finite Element Model; FEMU: Finite Element Model Updating; FIS: Fully Instability Status; GA: Genetic Algorithm; GANs: Generative Adversarial Networks; GBO: Gradient Based Optimizer; GPU: Graphics Processing Unit; GRNNs: General Regression Neural Networks; GSR: Gradient Search Rule; GWO: Grey Wolf Optimizer; HFSS: High-Formwork Support System; KGs: Knowledge Graphs; KL: Karhunen–Loève; LB: Lower Bound; LDA: Linear Discrimination Analysis; LEO: Local Escaping Operator; LIS: Local Instability Status; LLMs: Large Language Models; MAC: Modal Assurance

Criterion; ML: Machine Learning; MLP: Multi-Layer Perception; NLP: Natural Language Processing; NWS: Normal Working Status; OD: Outer Diameter; PDF: Probability Distribution Function; PSD: Power Spectral Density; PSO: Particle Swarm Optimization; RAG: Retrieval Augmented Generation; RAM: Random Access Memory; RF: Random Forest; SA: Simulated Annealing; SHM: Structural Health Monitoring; SVM: Support Vector Machine; UB: Upper Bound.

Acknowledgments

This research was funded by Beijing University of Technology, grant number 047000513102. The APC was funded by the two grants. The authors would like to thank Beijing University of Technology for its support through the research project. The authors would like to thank China Industry Associations for providing research data. In addition, I would like to thank all practitioners who contributed to this project.

Author's Contribution

Conceptualization, L.Z.; methodology, L.Z.; software, R.Z.; validation, L.Z., J.M.; resources, L.Z.; writing—original draft preparation, L.Z.; writing—review and editing, S.L.; visualization, R.Z.

Conflicts of interests

The authors declare no conflict of interest.

References

[1] Mozaffari S, Bruce M, Clune G, Xie R, McGee W, *et al*. Digital design and fabrication of clay formwork for concrete casting. *Autom. Constr.* 2023, 154:1–13.

[2] Reynolds J, Zhang H, Rasmussen KJR. Investigation of U-head rotational stiffness in formwork supporting scaffold systems. *Eng. Struct.* 2017, 136:1–11.

[3] Zhao L, Zhang H, Mbachu J. Multi-sensor data fusion for 3D reconstruction of complex structures: a case study on a real high formwork project. *Remote Sens.* 2023, 15(5):1–31.

[4] Yuan F, Zargar SA, Chen Q, Wang S. Machine learning for structural health monitoring: challenges and opportunities. In *Sensors and Smart Structures Technologies for Civil, Mechanical, and Aerospace Systems 2020*, Online, April 27–May 8, 2020, p. 1137903.

[5] Bono FM, Argentino A, Bernardini L, Benedetti L, Cazzulani G, *et al*. Automated operational modal analysis of a steel truss railway bridge employing free decay response. *J. Infrastruct. Intell. Resil.* 2025, 4:100145.

[6] Farrar CR, Worden K. *Structural health monitoring: a machine learning perspective*, 1st ed. London: Wiley, 2012.

[7] Fink O, Wang Q, Svensen M, Dersin P, Lee W, *et al*. Potential, challenges and future directions for deep learning in prognostics and health management applications. *Eng. Appl. Artif. Intell.* 2020, 92:1–15.

[8] Brozovsky J, Labonnote N, Vigren O. Digital technologies in architecture, engineering, and construction. *Autom. Constr.* 2024, 158:105212.

[9] Simoen E, De Roeck G, Lombaert G. Dealing with uncertainty in model updating for damage assessment: a review. *Mech. Syst. Sig. Process.* 2015, 56:123–149.
[10] Ye X, Chen B. Model updating and variability analysis of modal parameters for super high-rise structure. *Concurr. Comput. Pract. Exp.* 2019, 31:1–11.
[11] Farshadi M, Esfandiari A, Vahedi M. Structural model updating using incomplete transfer function and modal data. *Struct. Control Health Monit.* 2017, 24:1–13.
[12] He L, Reynders E, García-Palacios JH, Marano GC, Briseghella B, *et al*. Wireless-based identification and model updating of a skewed highway bridge for structural health monitoring. *Appl. Sci.* 2020, 10(7):1–13.
[13] Sun L, Li Y, Zhang W. Experimental study on continuous bridge-deflection estimation through inclination and strain. *J. Bridge Eng.* 2020, 25(5):1–14.
[14] Sun L, Li Y, Zhu W, Zhang W. Structural response reconstruction in physical coordinate from deficient measurements. *Eng. Struct.* 2020, 212:1–13.
[15] Tchemodanova SP, Sanayei M, Moaveni B, Tatsis K, Chatzi E. Strain predictions at unmeasured locations of a substructure using sparse response-only vibration measurements. *J. Civ. Struct. Health Monit.* 2021, 11(4):1113–1136.
[16] Razavi M, Hadidi A. Assessment of sensitivity-based FE model updating technique for damage detection in large space structures. *Struct. Monit. Maint.* 2020. 7(3):261–281.
[17] Yang X, Ouyang H, Guo X, Cao S. Modal strain energy-based model updating method for damage identification on beam-like structures. *J. Struct. Eng.* 2020, 146(11):1–13.
[18] Sun L, Xu Y. Modal parameter identification and finite element model updating of a long-span aqueduct structure based on ambient excitation. *J. Vibroeng.* 2020, 22(4):896–908.
[19] Zhou Y, Zhang J, Yi W, Jiang Y, Pan Q. Structural identification of a concrete filled steel tubular arch bridge via ambient vibration test data. *J. Bridge Eng.* 2017, 22(8):1–15.
[20] Marwala T. Finite-element model updating using computational intelligence techniques, 1st ed. London: Springer, 2010.
[21] Ghasemi S, Amiri GG, Dehcheshmeh MM. Structural damage assessment via model updating using augmented grey wolf optimization algorithm. *Int. J. Eng.* 2020, 33:1173–1182.
[22] Gao X, Govindasamy V, Xu H, Wang X, Zenger K. Harmony search method: theory and applications. *Comput. Intell. Neurosci.* 2015, 2015(1):258491.
[23] Rashedi E, Nezamabadi-Pour H, Saryazdi S. GSA: a gravitational search algorithm. *Inf. Sci.* 2009, 179(13):2232–2248.
[24] Kaveh A, Mahdavi VR. Damage identification of truss structures using CBO and ECBO algorithms. *Asian J. Civ Eng.* 2016, 17(1):75–89.
[25] Kirkpatrick S, Gelatt CD, Vecchi MP. Optimization by simulated annealing. *Science* 1983, 220(4598):671–680.
[26] Myers D, Mohawesh R, Chellaboina VI, Sathvik AL, Venkatesh P, *et al*. Foundation and large language models: fundamentals, challenges, opportunities, and social impacts. *Clust. Comput.* 2024, 27(1):1–26.
[27] Raiaan MAK, Mukta MSH, Fatema K, Fahad NM, Sakib S, *et al*. A review on large language models: architectures, applications, taxonomies, open issues and challenges. *IEEE Access* 2024, 12:26839–26874.

[28] Chang Y, Wang X, Wang J, Wu Y, Yang L, *et al*. A survey on evaluation of large language models. *ACM Trans. Intell. Syst. Technol.* 2024, 15(3):1–45.

[29] Afzal M, Li RYM, Shoaib M, Ayyub MF, Tagliabue LC, *et al*. Delving into the digital twin developments and applications in the construction industry: a PRISMA approach. *Sustainability* 2023, 15(23):16436.

[30] Siriwardhana S, Weerasekera R, Wen E, Kaluarachchi T, Rajib R, *et al*. Improving the domain adaptation of Retrieval Augmented Generation (RAG) models for open domain question answering. *Trans. Assoc. Comput. Linguist.* 2023, 11:1–17.

[31] Taiwo R, Bello IT, Abdulai SF, Yussif AM, Salami BA, *et al*. Generative AI in the construction industry: a state-of-the-art analysis. *arXiv* 2024, arXiv:2402.09939.

[32] Prieto SA, Mengiste ET, García de Soto B. Investigating the use of ChatGPT for the scheduling of construction projects. *Buildings* 2023, 13(4):857.

[33] Uddin SMJ, Albert A, Ovid A, Alsharef A. Leveraging ChatGPT to aid construction hazard recognition and support safety education and training. *Sustainability* 2023, 15(9):7121.

[34] Zhao L, Mbachu J, Wang B, Liu Z, Zhang H. Installation quality inspection for high formwork using terrestrial laser scanning technology. *Symmetry* 2022, 14(2):1–31.

[35] Behrendt M, Kitahara M, Kitahara T, Comerford L, Beer M. Data-driven reliability assessment of dynamic structures based on power spectrum classification. *Eng. Struct.* 2022, 268:1–13.

[36] Yan W, Ren W. An enhanced power spectral density transmissibility (EPSDT) approach for operational modal analysis: theoretical and experimental investigation. *Eng. Struct.* 2015, 102(1):108–119.

[37] Bono FM, Radicioni L, Cinquemani S, Conese C, Tarabini M. Development of soft sensors based on neural networks for detection of anomaly working condition in automated machinery. In *NDE 4.0, Predictive Maintenance, and Communication and Energy Systems in a Globally Networked World*, Long Beach, USA, March 25–28, 2024, pp. 56–70.

[38] Devlin J, Chang M, Lee K, Toutanova K. BERT: pre-training of deep bidirectional transformers for language understanding. In *2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, Minneapolis, USA, June 2–7, 2019, pp. 4171–4186.

[39] Papineni K, Roukos S, Ward T, Zhu W. Bleu: a method for automatic evaluation of machine translation. In *40th Annual Meeting of the Association for Computational Linguistics (ACL 2002)*, Philadelphia, USA, July 7–12, 2002, pp.1–8.

[40] Lin C. ROUGE: a package for automatic evaluation of summaries. In *ACL 2004 Workshop on Text Summarization Branches Out*, Barcelona, Spain, July 25–26, 2004, pp. 1–8.