Chapter 13
Full-field Measurements for Anomaly Detection of Mechanical Systems using Convolutional Neural Networks and LSTM Networks

Carlos Quiterio Gómez Muñoz, Mariano Alberto García Vellisca, Celso T. do Cabo, Yujie Xi, and Zhu Mao

Abstract Structural health monitoring (SHM) of mechanical systems is of great importance for a variety of industries, and an early detection and localization of damage in such situations are often desired. One of the features that can be used for SHM is its dynamic characteristics, which are traditionally obtained by using contact and sparse sensing. In nowadays, technologies that can collect full-field information and perform noncontact sensing have become more advantageous. Among such equipment, laser Doppler vibrometers (LDV) can provide high-quality vibrational data with full-field information. However, damage detection based on vibrational data can also be difficult in complex systems. With the recent advancement of computational power, data-driven approaches have become increasingly popular for complex systems. One of the most known approaches is the convolutional neural network (CNN), which has great capabilities for image processing and classification. Since the LDV measurements can provide full-field information, these data can be converted into 2D matrices and employed in CNN as images. However, CNNs commonly do not keep any temporal dependency between the frames, making it difficult to be employed in vibration analysis where data are all time series. This work proposes combining CNNs with long short-term memory (LSTM) networks, to process video data instead of frames, allowing time-dependent image processing for damage detection. In addition, the results are expected to predict not only if there is damage to the system, but also to classify different types of damage.

Keywords Machine Learning · Convolutional Neural Network · Long Short-Term Memory · Laser Doppler Vibrometer

Introduction

The classification of normal and abnormal vibrations in structure has gained substantial attention due to its implications in Structural Health Monitoring (SHM) and fault detection. Vibrational analysis is critical for identifying potential issues in industrial machinery, infrastructures, and mechanical systems [1–3]. By detecting anomalies early, it is possible to prevent costly repairs and ensure the safety of operations. With the increase interest in machine learning techniques, many studies have been performed in the field of structural vibration for anomaly detection [4, 5]. Specifically, Convolutional Neural Networks (CNNs) have great capabilities for image processing and classification, been employed previously to dynamic systems [6–8]. However, CNN architectures only capture the dependance between pixels in an image. For time-series datasets representing the time dependency of the signal is challenging, leading to the usage of Long Short-Term Memory (LSTM) networks. Such architectures can be combined with CNN to obtain time dependency between frames for video-based sensing [9, 10], or applied into Remaining Useful Life (RUL) applications [11].

This study focuses on analyzing vibrations of a metallic plate subjected to different frequency excitations in both linear and non-linear states. The linear state is characterized by the absence of external disruptions, where the plate vibrates freely. In contrast, the non-linear state introduces rubber bumpers placed at various positions on the plate, altering its vibrational

Carlos Quiterio Gómez Muñoz · Mariano Alberto García Vellisca
HCTLab Research Group, Universidad Autónoma de Madrid, Madrid, Spain
e-mail: carlosq.gomez@uam.es; marianoa.garcia@uam.es

Celso T. do Cabo · Yujie Xi · Zhu Mao
Department of Mechanical and Materials Engineering, Worcester Polytechnic Institute, Worcester, MA, United States
e-mail: ctdocabo@wpi.edu; yxi@wpi.edu; zmao2@wpi.edu

Brian Damiano et al. (eds.), *Structural Health Monitoring & Machine Learning, Vol. 12*, Conference Proceedings of the Society for Experimental Mechanics Series, https://doi.org/10.13052/97887-438-0157-3_13

behavior. These bumpers simulate damage or interference, generating abnormal vibration behavior, which can mimic real-world structural irregularities.

By utilizing machine learning techniques, particularly CNNs and LSTM networks, this research aims to distinguish between these vibrational states. The primary objective is to classify the plate's vibrations into their respective states (linear and non-linear) based on the frequency data collected, as well as to compare its efficiency when performing multiclass classification, by differentiating between different locations and number of bumpers in the structure.

The next sections of this papers present the background theory for the proposed approach, including an explanation on CNNs and LSTM networks, the experimental setup and preprocessing of the full-field measurements to obtain the dataset for the machine learning approach. Followed by the results and obtained for both simple nonlinearity detection and multiclass classification and a discussion of the results. Finally, a conclusion will be given with possible directions for future research.

Background

In recent years, machine learning, particularly deep learning techniques, has revolutionized various fields of data analysis and classification. One of the most critical applications is SHM, where machine learning models are employed to detect anomalies in vibrational patterns of structures. This growing field offers the potential to predict failures, optimize maintenance schedules, and improve the reliability of mechanical systems.

Machine learning, especially supervised learning, is widely used to classify vibration data. These models are trained on labeled datasets to predict the state of unseen vibrational data. Deep learning models like CNNs and LSTMs have been particularly effective in handling complex datasets. CNNs are suitable for analyzing time-frequency representations like spectrograms and scalograms, while LSTMs excel in processing temporal dependencies in time-series data.

CNNs process spatial data, making them ideal for analyzing the two-dimensional representations of vibrational signals. By using convolutional and pooling layers, CNNs capture essential patterns in frequency and time, classifying different vibrational states effectively.

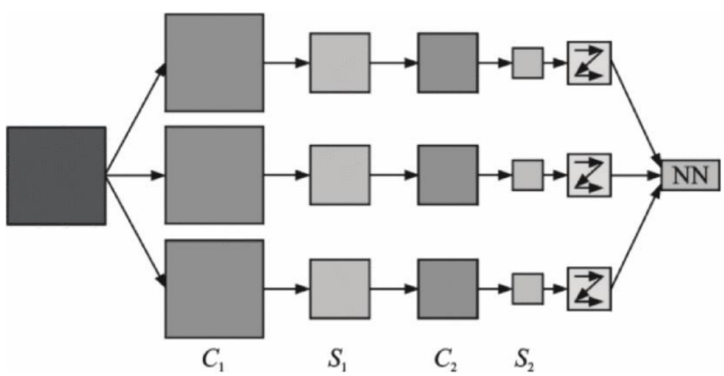

Fig. 1 Flowchart of CNN [12].

Convolution neural network is often designed to analyze two-dimensional data, and it has multi-layer inside the networks. CNNs uses convolution of kernels to learn pattern in 2D data. As shown in Figure 1, a simple convolution neural network structure has convolution layers (e.g. C1, C2) and sub-sampling layers (e.g. S1, S2). In this example, input image is convoluted by three convolution kernels, then three feature maps are generated in the C1 layer, for each feature map, their localized regions are weighted and averaged. Then three new feature maps are generated in S1 layer through nonlinear activation function. After S2, the feature maps will be flattened into a vector which will be used as input for a traditional neural networks (NN) for training [12].

While CNNs are capable of extracting spatial information. LSTMs are designed for sequential data, retaining relevant patterns over time through mechanisms called gates. This makes LSTMs ideal for detecting temporal anomalies in vibration sequences, which is essential in analyzing the vibrational behavior of the metallic plate. The overall layout of an LSTM layer can be seen in Figure 2.

The LSTM layer works by using cell states (C_t) which allows to retain or erase information from previous time-steps, depending on its relevance to predict future states. The states are calculated based on the previous state and with the current time-step information. The terms shown in Figure 2 are presented in Equations (1) to (6).

$$f_t = \sigma\left(W_f \cdot [h_{t-1}, x_t] + b_f\right) \tag{1}$$

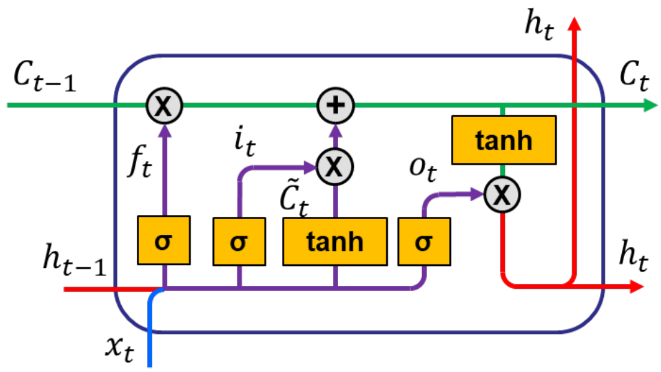

Fig. 2 Flowchart of LSTM cell [9].

$$i_t = \sigma\left(W_i \cdot [h_{t-1}, x_t] + b_i\right) \tag{2}$$

$$\tilde{C}_t = \tanh\left(W_C \cdot [h_{t-1}, x_t] + b_C\right) \tag{3}$$

$$C_t = f_t \odot C_{t-1} + i_t \odot \tilde{C}_t \tag{4}$$

$$o_t = \sigma\left(W_o \cdot [h_{t-1}, x_t] + b_o\right) \tag{5}$$

$$h_t = o_t \odot \tanh\left(C_t\right) \tag{6}$$

Where W represents the weights and b are the biases of the neural network. σ is the sigmoid function and $tahn$ is the hyperbolic tangent used as activation functions for the gates. Meanwhile, $[h_{t-1}, x_t]$ represents a concatenation between the previous time-step output and the current input data.

This work proposes a combination of CNN and LSTM layers in order to extract both spatial and temporal information from the system for abnormal behavior classification, by identifying if there is a nonlinearity in the system and a multiclass classification to predict different locations of nonlinearities.

Experimental Setup

This study examines the classification of normal and abnormal vibrations in a metallic plate subjected to controlled vibrational tests. Both linear (normal) and non-linear (abnormal) states were induced, with data collected at various excitation frequencies.

The testing subject for the experiment consists of a metallic plate, with 6 inches of width, 17 inches height and 0.05 inches thickness, and it was tested using a mechanical shaker. A Laser Doppler Vibrometer (LDV) was used to record the vibrational data for 133 points on the plate's surface. Two states were analyzed:

1. **Linear State**: The plate vibrated freely, representing normal conditions.
2. **Non-Linear State**: Rubber bumpers were placed behind the plate, simulating structural damage and altering the vibrational patterns.

Different frequencies were employed for excitation including 298 Hz, 548 Hz, 1704 Hz, 2118 Hz, 2435 Hz, 2792 Hz, 3907 Hz, 4960 Hz, and 5300 Hz were used to test the plate's response, chosen based on modal analysis. In the non-linear state, bumpers were placed as show Figure 3.

Vibrational signals were recorded at each frequency, capturing the plate's oscillations. The data collected from all 133 points was divided into training and testing subsets, with linear and non-linear data used to train the machine learning models, and testing data to evaluate their classification performance. One sample of the data can be seen in Figure 4.

To optimize the training of the neural network and improve the efficiency of the simulations, exhaustive data preparation was carried out, which included:

- **Capturing Plate Movement:** Each of the 133 mesh points represents the movement at that specific point on the plate. An interpolation of these points was performed to recreate the continuous movement of the plate.

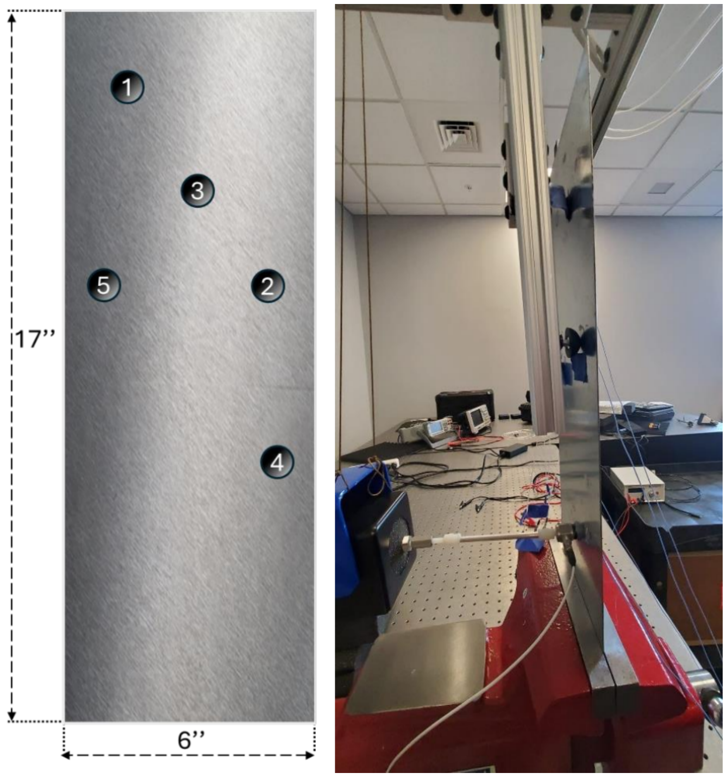

Fig. 3 Location of different bumpers on the plate to generate abnormal vibrations.

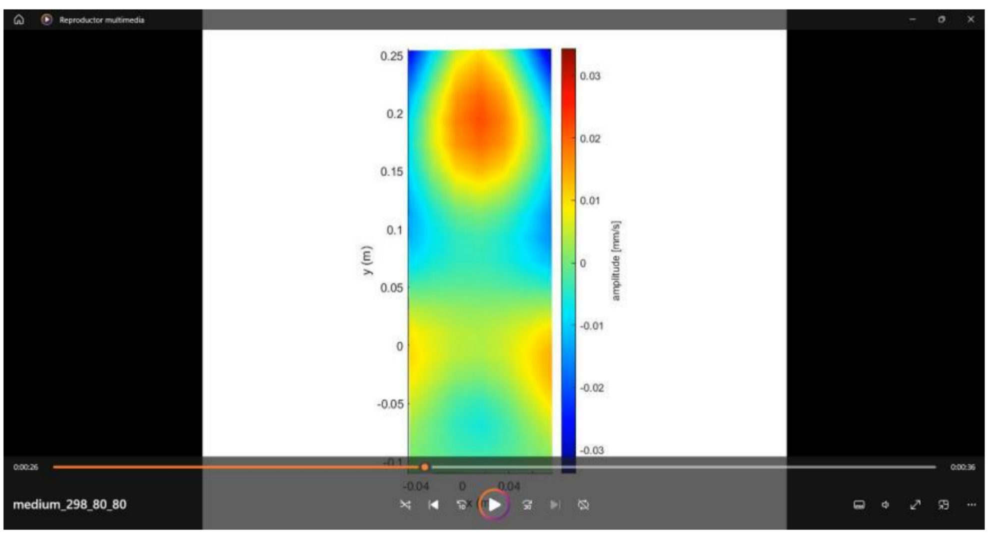

Fig. 4 Video image showing the movement of the place through pseud color representation.

- **Video Generation for Each Case:** Using the interpolated data, a video was generated for each case. This visual representation allowed for a clearer understanding of the plate's dynamics under different conditions.
- **Trimming video duration to 20 seconds**: It was determined that beyond this time, the videos did not provide additional relevant information. By removing segments that did not contribute significantly, processing was streamlined, and unnecessary computational resource consumption was avoided.
- **Balancing the number of videos per frequency and class**: The dataset was unified to have a homogeneous distribution, obtaining one video per frequency for each class (bumper). This allowed for better generalization during training and avoided biases towards classes with more data.

- **Generating copies of the videos**: Each original video was augmented with seven additional copies using data augmentation techniques, increasing the dataset to a total of **560 videos**. This helped improve the neural network's robustness against variations and prevent overfitting.
- **Incorporating and unifying classes**: New classes corresponding to different bumpers were added. After detailed analysis, it was decided to unify some classes to improve representation in the dataset.

The classes assigned to the bumpers were: **Bumper 1, Bumpers 1 and 2, Bumper 3, Bumper 4, Bumper 5 and Linear (no bumpers).**

In addition, the hyperparameters selected for the neural network are presented in Table 1.

Table 1 Hyperparameters of the proposed approach.

Hyperparameter	Value
Batch size	8
Learning rate	10^{-5}
Train/Test split	90% and 10%
Dropout	20%
# of LSTM neurons	2,000

Analysis and Results

The results presented in this section are divided into different cases, starting with Case 1, where the proposed method will do a simple classification of linear and nonlinear behavior. Case 2 will later present the results for the multiclass classification, with the linear response, and 5 different setups with bumpers as shown in Figure 3.

Case 1. Training with Linear and Non-Linear classes.
As a first step, a binary simulation was performed with only two classes: **Linear** and **Non-Linear** (which groups all bumpers). The accuracy was 97.14%.

Case 2. Training with Bumper Classes
Trainings were carried out using the classes corresponding to the different bumpers and with all the frequencies. The accuracy was 94.28%

The results indicate that:

- **Effectiveness of Preprocessing**: The reduction and balancing of the dataset, along with video preprocessing, contributed to improving the efficiency and precision of the simulations. Removing irrelevant segments allowed the network to focus on the most significant information.

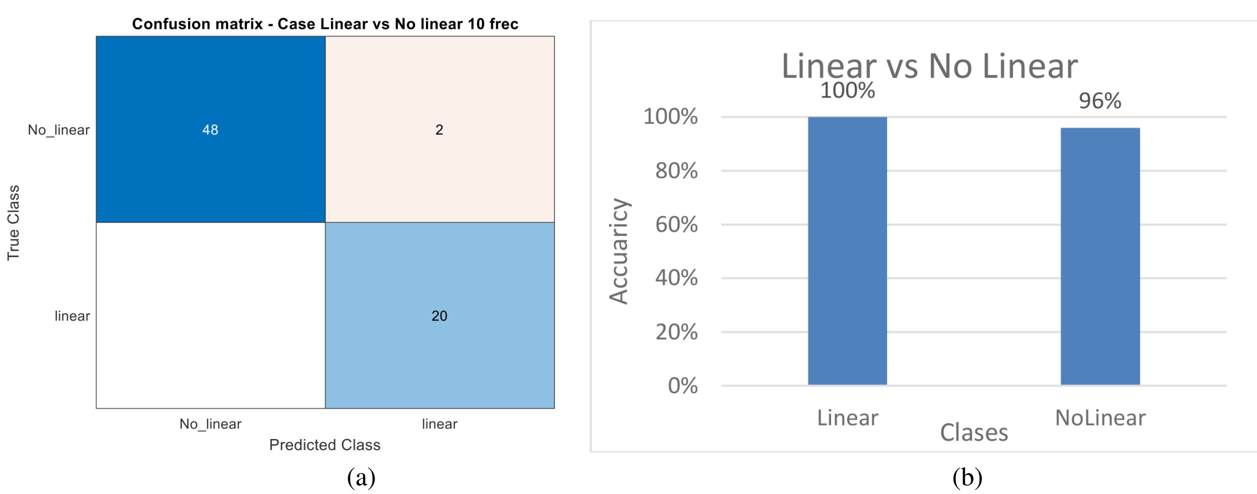

Fig. 5 Confusion matrix of linear and nonlinear cases (a) and accuracy of each class (b).

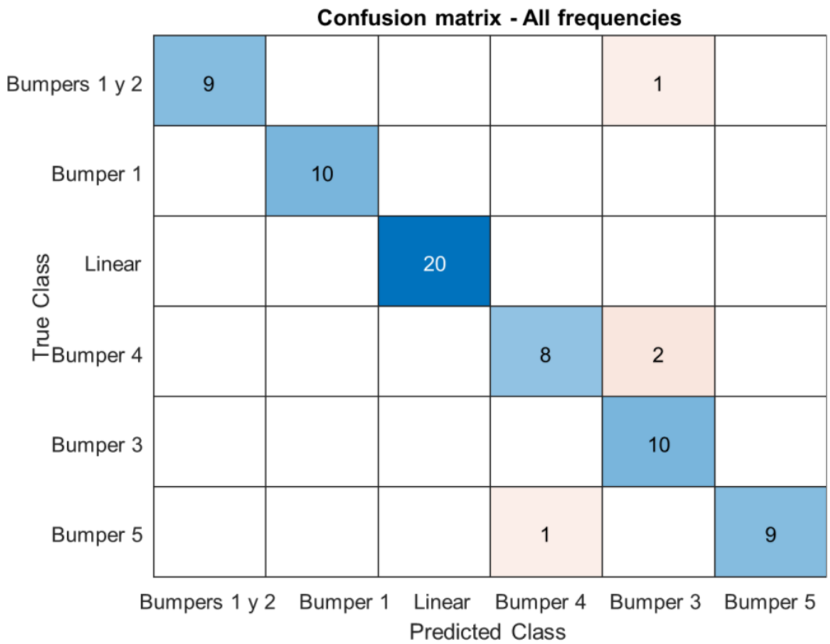

Fig. 6 Confusion matrix of the different bumpers located on the plate.

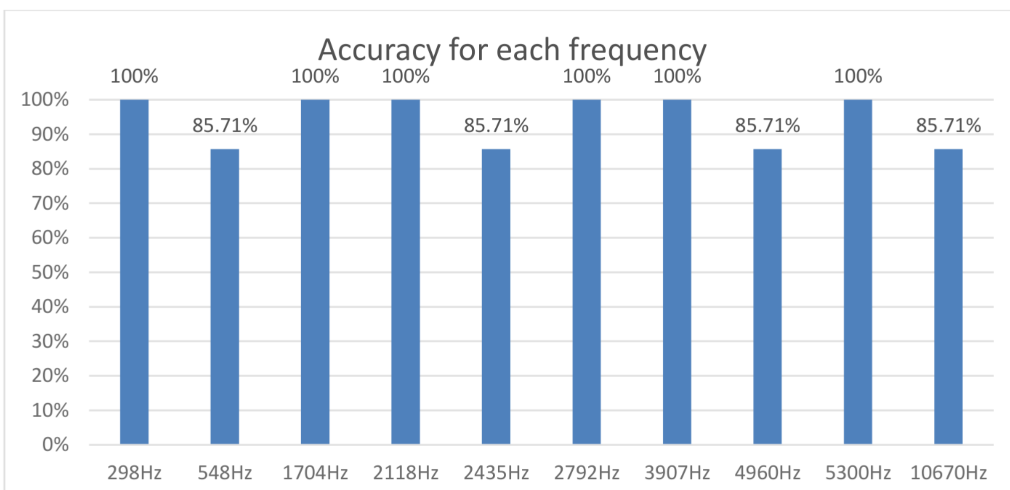

Fig. 7 Accuracy of the classification of all bumpers for each frequency.

- **Importance of Class Balancing**: By unifying the number of videos per class and frequency, bias towards classes with a larger amount of data was avoided, allowing better generalization of the model and a fairer evaluation of its performance.
- **Behavior by Classes (Bumpers)**: It was observed that certain classes presented more errors at specific frequencies. For example, the **Bumper 1** class showed more errors in the simulation with 8 frequencies, while **Bumpers 1 and 2** presented errors at specific frequencies in other simulations.

Conclusion

This study successfully demonstrated the use of LSTM networks for classifying normal and abnormal vibrational states in a metallic plate. By combining these models, it was possible to accurately detect linear and non-linear vibrational behaviors, critical for Structural Health Monitoring.

Key findings include:

1. **Effectiveness of Machine Learning Models**: LSTM networks proved to be highly effective, achieving excellent performance with a peak accuracy of 94.28% in the model. LSTMs excelled at capturing temporal dependencies in time-series data, making them particularly well-suited for tasks involving sequential patterns.
2. **Challenges in Complex Non-Linear States**: The models performed well overall but struggled with more complex non-linear states involving multiple bumpers, which introduced intricate vibrational patterns.
3. **Frequency-Based Performance Variations**: The ability to differentiate between linear and non-linear states depended on the frequency, with the models performing best at natural vibrational frequencies like 2118 Hz and 2792 Hz. However, accuracy declined at higher frequencies, particularly at 4960 Hz, where the distinction between the states became less clear.

The LSTM approach applied to videos generated with laser Doppler vibrometers showed great potential for real-time SHM applications, particularly in complex scenarios and high-frequency performance. This research contributes to the development of machine learning methods for real-time anomaly detection in structural systems.

Acknowledgments This work was supported by the GHAIA (Geometric and Harmonic Analysis with Interdisciplinary Applications) project, under the H2020-MSCA-RISE program of the European Commission. This project has received funding from the European Union's Horizon 2020 research and innovation program under the Marie Skłodowska-Curie grant agreement No 777822.

References

1. M. Arul and A. Kareem, "Data anomaly detection for structural health monitoring of bridges using shapelet transform," arXiv preprint arXiv:2009.00470, 2020.
2. N. D. Boffa, M. Arena, E. Monaco, M. Viscardi, F. Ricci, and T. Kundu, "About the combination of high and low frequency methods for impact detection on aerospace components," Progress in Aerospace Sciences, vol. 129, p. 100789, 2022.
3. D. Caballol, Á. P. Raposo, and F. Gil-Carrillo, "Non-destructive testing of concrete layer adhesion by means of vibration measurement," Construction and Building Materials, vol. 411, p. 134548, 2024.
4. N. M. Kalimullah, A. Shelke, and A. Habib, "A deep learning approach for anomaly identification in PZT sensors using point contact method," Smart Materials and Structures, vol. 32, no. 9, p. 095027, 2023.
5. X. Ye, P. Wu, A. Liu, X. Zhan, Z. Wang, and Y. Zhao, "A deep learning-based method for automatic abnormal data detection: Case study for bridge structural health monitoring," International Journal of Structural Stability and Dynamics, vol. 23, no. 11, p. 2350131, 2023.
6. C. T. do Cabo and Z. Mao, "An Optical Mode Shape-Based Damage Detection Using Convolutional Neural Networks," in Rotating Machinery, Optical Methods & Scanning LDV Methods, Volume 6: Springer, 2022, pp. 157–162.
7. Z. Tang, Z. Chen, Y. Bao, and H. Li, "Convolutional neural network-based data anomaly detection method using multiple information for structural health monitoring," Structural Control and Health Monitoring, vol. 26, no. 1, p. e2296, 2019.
8. S.-Y. Kim and M. Mukhiddinov, "Data Anomaly Detection for Structural Health Monitoring Based on a Convolutional Neural Network," Sensors, vol. 23, no. 20, p. 8525, 2023.
9. C. T. do Cabo and Z. Mao, "An Optical Temporal and Spatial Vibration-Based Damage Detection Using Convolutional Neural Networks and Long Short-Term Memory," in Rotating Machinery, Optical Methods & Scanning LDV Methods, Volume 6: Springer, 2022, pp. 159–165.
10. R. Yang et al., "CNN-LSTM deep learning architecture for computer vision-based modal frequency detection," Mechanical Systems and signal processing, vol. 144, p. 106885, 2020.
11. M. Ma and Z. Mao, "Deep-convolution-based LSTM network for remaining useful life prediction," IEEE Transactions on Industrial Informatics, vol. 17, no. 3, pp. 1658–1667, 2020.
12. A. A. M. Al-Saffar, H. Tao, and M. A. Talab, "Review of deep convolution neural network in image classification," in 2017 International conference on radar, antenna, microwave, electronics, and telecommunications (ICRAMET), 2017: IEEE, pp. 26–31.