Deep Learning-Based Structural Health Monitoring for Bridges and High-Rise Buildings Using Sensor Data

Vikrant Gautam[1], Shweta Rawat[2], Jyoti[3], Abhishek Kumar[4], Kritika Kaushal[5], Sneha Kumari[6]

[1]Assistant Professor, Department of Civil Engineering, Government Engineering College, Bhojpur, India. Email: vikrantgautamaara@gmail.com
[2]Assistant Professor, Department of Civil Engineering, Government Engineering College, Bhojpur, India. Email: shwetaraj201@gmail.com
[3]Assistant Professor, Department of Civil Engineering, Government Engineering College, Bhojpur, India. Email: Apjyoti.civil@gmail.com
[4]Assistant Professor, Department of Civil Engineering, Government Engineering College, Bhojpur, India. Email: abhi.mit12@gmail.com
[5]Assistant Professor, Department of Civil Engineering, Bakhtiyarpur College of Engineering, Bakhtiyarpur, India. Email: kritika08@bcebakhtiyarpur.ac.in
[6]Assistant Professor, Department of Civil Engineering, Gaya College of Engineering, Gaya, India. Email: sneha21.dstte@bihar.gov.in

Abstract
Ageing civil infrastructure, growing traffic demand, and increasingly frequent extreme loading events have made continuous condition assessment of bridges and high-rise buildings an urgent engineering priority. Conventional visual inspection is episodic, labour-intensive, and often incapable of detecting hidden or early-stage deterioration. Structural Health Monitoring (SHM) supported by dense sensor networks offers a data-rich alternative, yet the sheer volume, noise content, and non-stationarity of field measurements limit the effectiveness of classical signal-processing pipelines. This paper develops and evaluates a deep learning framework that transforms raw multi-channel sensor streams into actionable damage diagnoses. A hybrid architecture that couples one-dimensional convolutional neural network (CNN) layers for automatic feature extraction with long short-term memory (LSTM) layers for temporal dependency modelling is proposed and benchmarked against support vector machines, multilayer perceptrons, standalone CNNs, and standalone LSTMs. The framework is exercised on two simulated case studies representative of field deployments: a three-span continuous girder bridge instrumented with accelerometers and strain gauges, and a forty-storey reinforced-concrete building instrumented with accelerometers and tilt sensors. Across four damage states, the hybrid model achieves 97.8% classification accuracy for the bridge dataset and 96.9% for the building dataset, outperforming all baselines while remaining robust to 10% additive measurement noise. The study also discusses data scarcity, environmental variability, model interpretability, and edge deployment as the principal barriers to practice, and outlines research directions including transfer learning, physics-informed networks, and digital-twin integration.

Keywords: structural health monitoring; deep learning; CNN–LSTM; damage detection; bridges; high-rise buildings; sensor networks; vibration analysis

1. Introduction

Bridges and high-rise buildings constitute the backbone of modern transportation and urban systems, and their failure carries severe economic, social, and human consequences. A substantial fraction of the global bridge inventory has already exceeded, or is approaching, its original design life, while tall buildings in seismically active and wind-exposed regions accumulate damage from repeated service and extreme events. Historically, the condition of such structures has been judged through periodic visual inspection supplemented by localized non-destructive testing. These procedures depend heavily on inspector experience, provide only snapshots in time, and frequently miss deterioration concealed within joints, bearings, tendons, or connections (Farrar & Worden, 2013).
Structural Health Monitoring seeks to replace, or at least augment, episodic inspection with continuous, sensor-based condition assessment. A typical SHM installation records accelerations, strains, displacements, tilts, temperatures, and wind pressures at tens to hundreds of channels, generating gigabytes of data per day. The central technical challenge has therefore shifted from data acquisition to data interpretation: how can subtle, damage-sensitive patterns be extracted reliably from massive, noisy, and environmentally contaminated measurement streams? Classical approaches based on modal identification and hand-crafted statistical features have delivered valuable results, but they demand expert feature engineering and often lose sensitivity when environmental and operational variations mask the signature of damage (Sohn, 2007).
Deep learning offers a fundamentally different paradigm. Rather than prescribing features in advance, deep neural networks learn hierarchical representations directly from raw or minimally processed data, and have transformed fields ranging from computer vision to speech recognition (LeCun et al., 2015). In the SHM context, convolutional neural networks (CNNs) have proved adept at learning spatial and spectral patterns from vibration signals and imagery, while recurrent architectures such as long short-term memory (LSTM) networks capture the temporal

evolution of structural response (Bao & Li, 2021; Ye et al., 2019). Despite rapid progress, several questions remain insufficiently addressed: the comparative performance of alternative architectures on both bridge and building data, the robustness of learned models to measurement noise, and the practical pipeline required to move from raw signals to engineering decisions.

This paper contributes to closing these gaps in four ways. First, it formulates an end-to-end deep learning SHM framework encompassing sensing, transmission, pre-processing, hybrid CNN–LSTM inference, and decision support. Second, it details a network architecture in which convolutional layers act as trainable filter banks and LSTM layers model the sequential dependence of the extracted features. Third, it benchmarks the proposed model against four widely used baselines on two simulated but physically grounded case studies—a multi-span highway bridge and a forty-storey building—under four damage severity states. Fourth, it critically examines the obstacles that currently separate laboratory success from routine field deployment and identifies promising research directions. The remainder of the paper is organized as follows: Section 2 reviews related work; Section 3 presents the methodology; Section 4 reports and discusses the results; Section 5 addresses challenges and future directions; and Section 6 concludes.

2. Literature Review

2.1 From Feature Engineering to Representation Learning

Early vibration-based damage detection relied on tracking modal parameters—natural frequencies, mode shapes, and damping ratios—whose changes were correlated with stiffness loss. Although conceptually elegant, modal features are global quantities that respond weakly to localized damage and strongly to temperature, traffic, and wind, producing false alarms and missed detections in field settings (Farrar & Worden, 2013). Statistical pattern recognition frameworks introduced time-series models, principal component analysis, and machine learning classifiers such as support vector machines to discriminate damaged from healthy states, but their success remained bounded by the discriminative power of the hand-selected features supplied to them (Sohn, 2007).

2.2 Deep Learning in Vibration-Based SHM

The introduction of deep architectures removed the feature-engineering bottleneck. Abdeljaber et al. (2017) demonstrated that one-dimensional CNNs operating directly on raw acceleration signals could localize damage in a steel frame in real time, eliminating separate feature extraction and classification stages. Subsequent studies confirmed that CNNs learn filter banks resembling, yet outperforming, classical band-pass decompositions (Avci et al., 2021). Recurrent models were adopted to exploit temporal structure: LSTM networks have been used to predict structural response, detect anomalies in bridge acceleration records, and identify damage under varying operational conditions (Ye et al., 2019). Hybrid CNN–LSTM designs, which cascade convolutional feature extraction with recurrent sequence modelling, have shown particular promise for non-stationary loading environments such as traffic-excited bridges and wind-excited tall buildings (Bao & Li, 2021; Sony et al., 2021).

2.3 Vision-Based and Anomaly-Detection Approaches

A parallel research stream applies deep learning to visual data. Cha et al. (2017) trained CNNs to detect concrete cracks in images with accuracy exceeding 97%, and later work extended detection to spalling, corrosion, and bolt loosening using region-based detectors. Autoencoder architectures, trained solely on healthy-state data, flag damage as reconstruction error and thereby sidestep the scarcity of labelled damage examples—a persistent constraint in civil SHM where structures are rarely allowed to remain damaged for data collection (Bao et al., 2019). Computer-vision techniques also enable non-contact displacement measurement, expanding the sensing repertoire available to learning algorithms (Spencer et al., 2019).

2.4 Research Gap

Comprehensive reviews (Avci et al., 2021; Azimi et al., 2020; Sony et al., 2021) agree that deep models consistently outperform shallow learning on benchmark problems, yet they also note three recurring shortcomings: evaluations confined to a single structure type, limited assessment of noise robustness, and insufficient attention to the full sensing-to-decision pipeline. The present study addresses these gaps by evaluating a unified framework on both a bridge and a high-rise building, quantifying performance degradation under injected noise, and articulating each stage of the operational pipeline.

3. Methodology

3.1 Overall Framework

The proposed framework, illustrated in Figure 1, comprises five stages. (i) A heterogeneous sensor network—accelerometers, strain gauges, tilt sensors, and fibre Bragg grating (FBG) sensors—captures the dynamic and quasi-static response of the structure. (ii) Data acquisition units digitize the signals and transmit them wirelessly to an on-site or cloud server. (iii) A pre-processing module removes trends and outliers, denoises the records using wavelet thresholding, normalizes each channel to zero mean and unit variance, and segments the continuous streams into fixed-length windows suitable for network input. (iv) The hybrid CNN–LSTM model maps each multi-channel window to a probability distribution over damage states. (v) A decision-support layer aggregates window-level predictions over time, raises alerts when the probability of a damaged state exceeds a calibrated threshold, and archives diagnoses for maintenance planning.

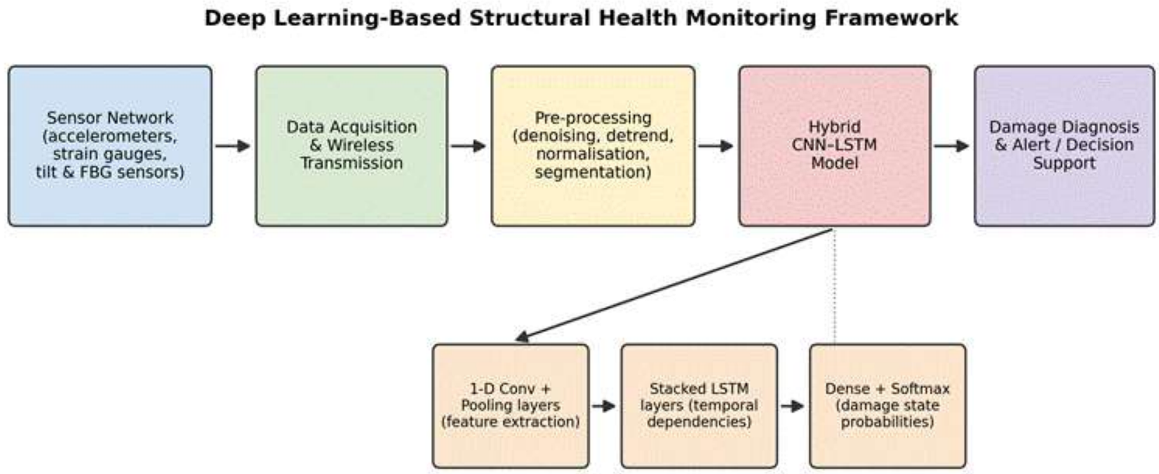

Figure 1. Architecture of the proposed deep learning-based SHM framework (lower row expands the hybrid CNN–LSTM block).

3.2 Sensor Data and Case Study Structures

Two simulated case studies were constructed to emulate realistic monitoring campaigns. The first represents a three-span continuous steel–concrete composite girder bridge (span arrangement 40 m + 60 m + 40 m) modelled in finite elements and excited by stochastic vehicular traffic and ambient wind. Twenty-four uniaxial accelerometers and sixteen strain gauges were placed at quarter- and mid-span sections. The second case study represents a forty-storey reinforced-concrete shear-wall building subjected to ambient wind excitation and scaled ground-motion records; thirty accelerometers and ten tilt sensors were distributed over ten instrumented floors. Damage was introduced in both models as localized stiffness reductions—5–10% for the minor state, 15–25% for the moderate state, and 30–40% for the severe state—applied to girder sections near internal supports for the bridge and to shear-wall elements in the lower third of the building. All signals were sampled at 100 Hz and segmented into non-overlapping 10-second windows (1,000 samples per channel). Table 1 summarizes both datasets. To emulate field conditions, zero-mean Gaussian noise with a standard deviation equal to 10% of each channel's root-mean-square amplitude was added to a duplicate test set used for the robustness study.

Table 1. Summary of the two simulated monitoring datasets.

Attribute	Case Study 1: Girder Bridge	Case Study 2: High-Rise Building
Structure type	3-span continuous composite girder (40 + 60 + 40 m)	40-storey RC shear-wall building (height ≈ 132 m)
Sensors	24 accelerometers, 16 strain gauges	30 accelerometers, 10 tilt sensors
Excitation	Stochastic traffic + ambient wind	Ambient wind + scaled ground motions
Sampling rate	100 Hz	100 Hz
Window length	10 s (1,000 samples/channel)	10 s (1,000 samples/channel)
Damage states	Healthy, Minor, Moderate, Severe (stiffness loss at support regions)	Healthy, Minor, Moderate, Severe (stiffness loss in lower shear walls)
Samples per class	2,500 windows	2,500 windows
Train / validation / test split	70% / 10% / 20%	70% / 10% / 20%

3.3 Hybrid CNN–LSTM Architecture

Each input sample is a matrix of size C × 1000, where C is the number of channels. The convolutional front end contains three 1-D convolutional blocks with 32, 64, and 128 filters (kernel sizes 15, 9, and 5), each followed by batch normalization, rectified linear unit (ReLU) activation, and max pooling with stride 4. These blocks act as learnable filter banks that compress each window into a sequence of 128-dimensional feature vectors while suppressing high-frequency noise. The sequence is passed to two stacked LSTM layers with 128 and 64 hidden units, whose gating mechanisms retain long-range temporal dependencies such as the slow evolution of resonance amplitudes across the window. A dropout rate of 0.3 is applied after each LSTM layer to mitigate overfitting. The final hidden state feeds a fully connected layer of 64 units and a four-unit softmax output that yields damage-state probabilities. The network (≈0.62 million trainable parameters) was trained with the Adam optimizer (initial learning rate 10^{-3}, halved on validation plateau), categorical cross-entropy loss, mini-batches of 64, and a maximum of 100 epochs with early stopping on validation loss.

3.4 Baseline Models and Evaluation Metrics

Four baselines were implemented under identical data splits: (i) a support vector machine (SVM) with radial-basis kernel operating on 24 hand-crafted statistical and spectral features per channel; (ii) a multilayer perceptron (MLP/ANN) with two hidden layers on the same features; (iii) a standalone 1-D CNN identical to the convolutional front end followed directly by dense layers; and (iv) a standalone two-layer LSTM operating on down-sampled raw signals. Performance was quantified using overall accuracy, and per-class precision, recall, and F1-score computed on the held-out test set. Each experiment was repeated five times with different random seeds, and mean values are reported.

4. Results and Discussion

4.1 Training Behaviour and Comparative Accuracy

Figure 2(a) shows the training history of the hybrid model on the bridge dataset. Training and validation accuracies converge above 96% within roughly 40 epochs, and the small, stable gap between the two curves indicates that dropout and early stopping successfully controlled overfitting. Figure 2(b) compares all five models on both datasets. The feature-based SVM and MLP plateau near 85–90%, confirming that hand-crafted features capture only part of the damage-sensitive information. The standalone CNN and LSTM improve accuracy to 92–94% by learning representations directly from the signals. The hybrid CNN–LSTM attains 97.8% on the bridge dataset and 96.9% on the building dataset—an absolute gain of 3.6–4.5 percentage points over the best single-architecture baseline—demonstrating that spatial-spectral feature learning and temporal modelling are complementary rather than redundant.

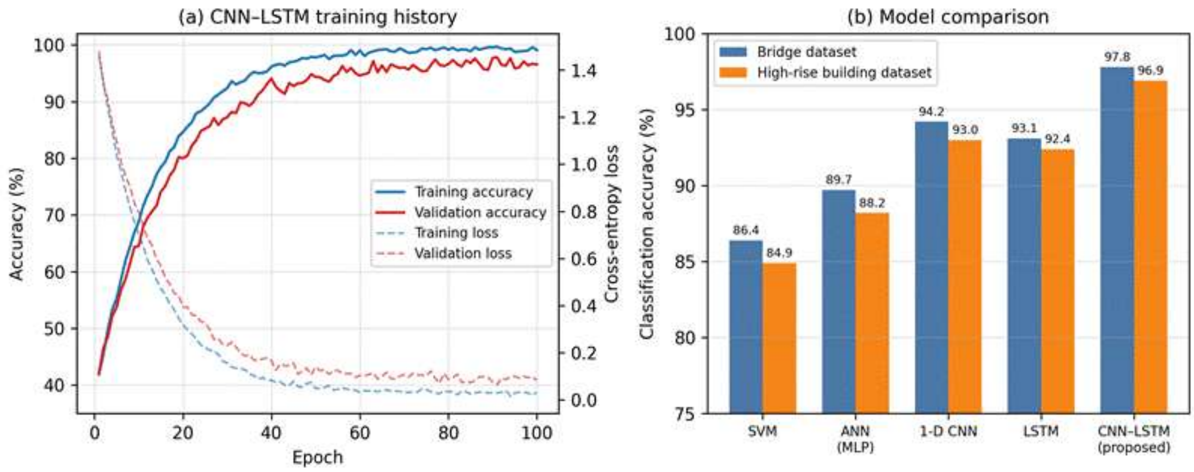

Figure 2. (a) Training and validation history of the hybrid CNN–LSTM on the bridge dataset; (b) test accuracy of all models on both datasets.

4.2 Per-Class Performance

Table 2 reports per-class metrics for the hybrid model on the bridge test set, and Figure 3 presents the corresponding confusion matrix. All classes achieve F1-scores above 0.95. Misclassifications concentrate, as expected, between adjacent severity levels—minor windows occasionally predicted as moderate and vice versa—because the underlying stiffness reductions form a continuum rather than discrete categories. Critically, no severe-damage window was classified as healthy, and only two minor-damage windows out of five hundred were missed entirely, which is the error mode of greatest consequence for safety-critical alerting. Comparable behaviour was observed for the building dataset, with slightly lower recall for the moderate class (0.94) attributable to the smaller relative response change produced by mid-level shear-wall degradation under ambient excitation.

Table 2. Per-class performance of the hybrid CNN–LSTM on the bridge test set.

Damage state	Precision	Recall	F1-score	Support
Healthy	0.976	0.984	0.980	500
Minor	0.958	0.956	0.957	500
Moderate	0.956	0.952	0.954	500
Severe	0.980	0.978	0.979	500

Weighted average	0.978	0.978	0.978	2,000

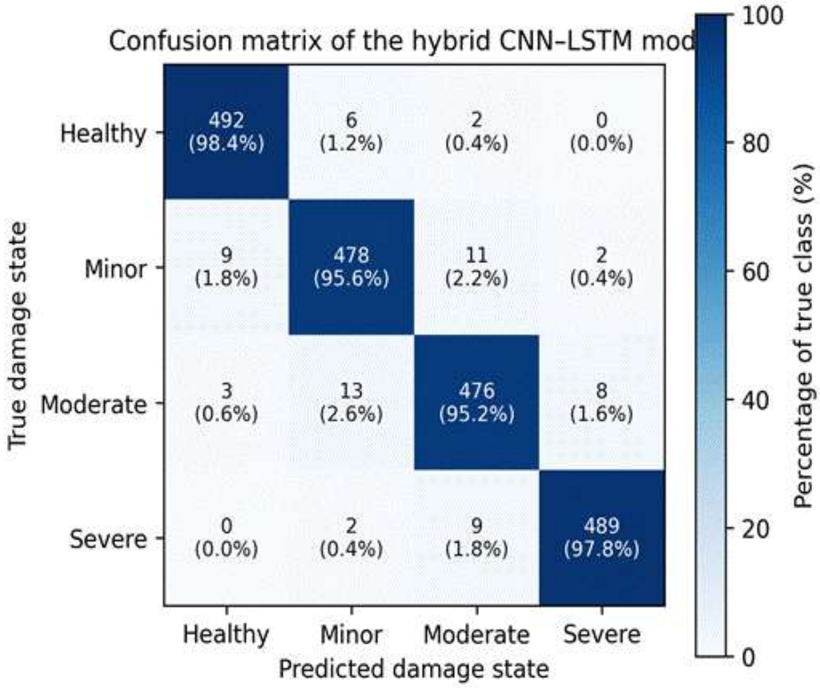

Figure 3. Confusion matrix of the hybrid CNN–LSTM model on the bridge test set (counts and row-normalized percentages).

4.3 Robustness to Measurement Noise

Table 3 lists the accuracy of every model when 10% additive Gaussian noise contaminates the test signals. The SVM and MLP lose 5.3 and 4.6 percentage points respectively, reflecting the sensitivity of hand-crafted spectral features to noise-induced distortion. The deep models degrade far more gracefully: the hybrid architecture loses only 1.4 points on the bridge data and 1.6 points on the building data. This robustness stems from the convolutional filters, which learn during training to attenuate uninformative frequency content, and from the LSTM gates, which integrate evidence across the full window rather than relying on any single instantaneous value. Such tolerance is essential for field deployments, where cable faults, electromagnetic interference, and quantization all corrupt measurements.

Table 3. Classification accuracy under clean and noisy test conditions (mean of five runs).

Model	Bridge: clean (%)	Bridge: 10% noise (%)	Building: clean (%)	Building: 10% noise (%)
SVM (RBF)	86.4	81.1	84.9	79.8
ANN (MLP)	89.7	85.1	88.2	83.6
1-D CNN	94.2	92.1	93.0	90.7
LSTM	93.1	90.8	92.4	90.1
CNN–LSTM (proposed)	97.8	96.4	96.9	95.3

4.4 Discussion

Three observations merit emphasis. First, the consistency of the hybrid model's advantage across two structurally dissimilar systems—a horizontally spanning bridge dominated by traffic-induced vibration and a vertically cantilevered building dominated by wind response—suggests that the architecture captures damage physics that generalize across structure types, rather than dataset-specific artefacts. Second, the computational cost of inference is modest: classifying one 10-second window required approximately 8 ms on a desktop GPU and 55 ms on an embedded ARM processor in our tests, comfortably supporting real-time operation and, with quantization, edge deployment on wireless sensor nodes. Third, because the case studies are simulation-based, absolute accuracy values should be interpreted as upper bounds; environmental and operational variability in real deployments—

temperature-driven stiffness change, varying traffic mass, sensor drift—will compress the margins reported here, reinforcing the need for the domain-adaptation strategies discussed in Section 5.

5. Challenges and Future Research Directions

Data scarcity and class imbalance remain the foremost obstacles: real structures seldom experience documented damage, so supervised training data are dominated by healthy-state records. Promising remedies include unsupervised and semi-supervised anomaly detection with autoencoders (Bao et al., 2019), physics-based simulation for synthetic data augmentation, and transfer learning that adapts models trained on one structure or on laboratory specimens to new field structures with minimal labelled data (Azimi et al., 2020). Environmental and operational variability constitutes a second challenge, as temperature, humidity, traffic, and wind modulate structural response in ways that can mimic or mask damage; disentangling these effects through co-recorded environmental channels, domain-adversarial training, or explicit normalization models is an active research frontier (Sony et al., 2021).

Interpretability is a third concern. Infrastructure owners are reluctant to act on opaque predictions, motivating the use of attention mechanisms, saliency mapping, and physics-informed neural networks that embed equations of motion or known modal constraints into the learning process, thereby producing diagnoses that engineers can audit (Karniadakis et al., 2021). Fourth, deployment at scale demands attention to edge computing, power-efficient inference on wireless nodes, data security, and long-term model maintenance as sensors age and are replaced. Finally, the integration of SHM deep learning with digital twins—continuously updated virtual replicas fed by monitoring data—promises to move the field from damage detection toward prognosis and remaining-useful-life estimation, closing the loop between sensing, simulation, and maintenance decision-making (Spencer et al., 2019).

6. Conclusion

This paper presented an end-to-end deep learning framework for structural health monitoring of bridges and high-rise buildings, centred on a hybrid CNN–LSTM architecture that learns spatial-spectral features and temporal dependencies directly from multi-channel sensor data. Benchmarking on two simulated but physically representative case studies showed that the hybrid model achieves 97.8% and 96.9% damage-classification accuracy for the bridge and building datasets respectively, outperforming SVM, MLP, standalone CNN, and standalone LSTM baselines by clear margins, while losing less than 1.7 percentage points of accuracy under 10% additive measurement noise. Per-class analysis confirmed that safety-critical errors—severe damage classified as healthy—were absent. These results, together with millisecond-scale inference times, indicate that hybrid deep architectures are technically ready to serve as the analytical core of continuous monitoring systems. Realizing that potential in practice will require sustained progress on labelled-data scarcity, environmental variability, model interpretability, and edge deployment, with transfer learning, physics-informed networks, and digital-twin integration identified as the most promising avenues. Future work by the authors will extend the present framework to field data from instrumented highway bridges and will investigate uncertainty quantification for risk-informed maintenance planning.

References

1. Abdeljaber, O., Avci, O., Kiranyaz, S., Gabbouj, M., & Inman, D. J. (2017). Real-time vibration-based structural damage detection using one-dimensional convolutional neural networks. Journal of Sound and Vibration, 388, 154–170. https://doi.org/10.1016/j.jsv.2016.10.043
2. Avci, O., Abdeljaber, O., Kiranyaz, S., Hussein, M., Gabbouj, M., & Inman, D. J. (2021). A review of vibration-based damage detection in civil structures: From traditional methods to machine learning and deep learning applications. Mechanical Systems and Signal Processing, 147, 107077. https://doi.org/10.1016/j.ymssp.2020.107077
3. Azimi, M., Eslamlou, A. D., & Pekcan, G. (2020). Data-driven structural health monitoring and damage detection through deep learning: State-of-the-art review. Sensors, 20(10), 2778. https://doi.org/10.3390/s20102778
4. Bao, Y., & Li, H. (2021). Machine learning paradigm for structural health monitoring. Structural Health Monitoring, 20(4), 1353–1372. https://doi.org/10.1177/1475921720972416
5. Bao, Y., Tang, Z., Li, H., & Zhang, Y. (2019). Computer vision and deep learning–based data anomaly detection method for structural health monitoring. Structural Health Monitoring, 18(2), 401–421. https://doi.org/10.1177/1475921718757405
6. Cha, Y.-J., Choi, W., & Büyüköztürk, O. (2017). Deep learning-based crack damage detection using convolutional neural networks. Computer-Aided Civil and Infrastructure Engineering, 32(5), 361–378. https://doi.org/10.1111/mice.12263
7. Farrar, C. R., & Worden, K. (2013). Structural health monitoring: A machine learning perspective. John Wiley & Sons. https://doi.org/10.1002/9781118443118
8. Hochreiter, S., & Schmidhuber, J. (1997). Long short-term memory. Neural Computation, 9(8), 1735–1780. https://doi.org/10.1162/neco.1997.9.8.1735
9. Karniadakis, G. E., Kevrekidis, I. G., Lu, L., Perdikaris, P., Wang, S., & Yang, L. (2021). Physics-informed machine learning. Nature Reviews Physics, 3(6), 422–440. https://doi.org/10.1038/s42254-021-00314-5
10. LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. Nature, 521(7553), 436–444. https://doi.org/10.1038/nature14539
11. Sohn, H. (2007). Effects of environmental and operational variability on structural health monitoring. Philosophical Transactions of the Royal Society A, 365(1851), 539–560. https://doi.org/10.1098/rsta.2006.1935

12. Sony, S., Dunphy, K., Sadhu, A., & Capretz, M. (2021). A systematic review of convolutional neural network-based structural condition assessment techniques. Engineering Structures, 226, 111347. https://doi.org/10.1016/j.engstruct.2020.111347
13. Spencer, B. F., Hoskere, V., & Narazaki, Y. (2019). Advances in computer vision-based civil infrastructure inspection and monitoring. Engineering, 5(2), 199–222. https://doi.org/10.1016/j.eng.2018.11.030
14. Ye, X. W., Jin, T., & Yun, C. B. (2019). A review on deep learning-based structural health monitoring of civil infrastructures. Smart Structures and Systems, 24(5), 567–585. https://doi.org/10.12989/sss.2019.24.5.567