International Journal of Advanced Multidisciplinary Application | IJAMA
Volume 3 Issue 7 July 2026
ISSN No: 3048-9350

Machine Learning-Based Predictive Maintenance for Structural Health Monitoring of Steel Highway Bridges

Ananya Krishnamurthy

Department of Civil Engineering and Centre for Data-Driven Infrastructure Systems, Indian Institute of Technology Hyderabad, Telangana, India

Abstract

Ageing steel highway bridge infrastructure across rapidly motorising economies faces a structural performance assessment bottleneck: visual inspection protocols mandated at 2-4 year intervals cannot detect fatigue cracking, corrosion-induced section loss, or connection deterioration between scheduled inspections, while continuous instrumentation generates data volumes that exceed manual interpretation capacity. This study addresses the integration gap between structural health monitoring (SHM) sensor networks and automated damage diagnosis by developing and field-validating a multidisciplinary machine learning framework that fuses vibration, strain, and electrochemical corrosion-potential data streams for bridge-level predictive maintenance.

A sensor network comprising 24 accelerometers, 18 strain gauges, and 12 half-cell corrosion-potential probes was deployed across two instrumented steel girder bridges (a 42 m simply-supported control structure and a 58 m structure with a known fatigue crack at Girder 3) on the Hyderabad Outer Ring Road corridor, generating 729 days of continuous monitoring data sampled at 200 Hz for dynamic channels and hourly for electrochemical channels. Four machine learning architectures — a hybrid convolutional neural network-long short-term memory (CNN-LSTM) model, a stacked LSTM, a random forest classifier, and a support vector machine — were trained on time-frequency and statistical features extracted from the fused sensor streams to perform four-class damage severity classification (none, minor, moderate, severe) and remaining useful life (RUL) regression.

The CNN-LSTM hybrid achieved the highest damage classification accuracy (96.8%) and area-under-curve of 0.974 for binary damage detection when all three sensor modalities were fused, compared to 0.887-0.901 AUC for single-modality models, confirming a substantial sensor-fusion advantage. The framework detected the onset of progressive natural frequency drift in the fatigue-cracked bridge 38 days before the deviation would have been flagged by routine biennial inspection, with RUL predictions achieving a root-mean-square error of 11.4 days across 20 field-validated bridge segments. Lifecycle cost analysis indicates that ML-driven predictive maintenance reduces annualised per-bridge costs by 60.7% relative to reactive run-to-failure maintenance and by 36.6% relative to fixed-interval scheduled maintenance, driven predominantly by avoided downtime and disruption costs. SHAP-based feature attribution identifies modal frequency shift, daily strain range, and corrosion potential as the three dominant predictors of damage severity, providing an interpretable basis for sensor network prioritisation in resource-constrained monitoring deployments.

Keywords: *structural health monitoring, predictive maintenance, machine learning, sensor fusion, CNN-LSTM, steel bridges, remaining useful life, damage detection, corrosion monitoring, fatigue, infrastructure asset management*

1. Introduction

India's National Highways network includes more than 1,80,000 bridge structures, a substantial proportion of which were designed under load and fatigue provisions that predate current IRC codal standards and now operate under traffic volumes several multiples of original design assumptions. Steel girder bridges are particularly susceptible to fatigue crack initiation at welded connections and corrosion-induced section loss at expansion joints and drainage-adjacent flanges, both of which progress between the 24-36 month intervals typical of routine visual inspection regimes mandated under the Ministry of Road Transport and Highways bridge management guidelines. The structural engineering consequence of this inspection gap is that damage mechanisms with rapid propagation characteristics — fatigue cracks under high-cycle traffic loading in particular — can advance from initiation to a safety-critical state within a single inter-inspection interval, creating a residual risk that conventional inspection-based asset management cannot fully mitigate.

Structural health monitoring (SHM) using permanently or semi-permanently installed sensor networks has been proposed as a complementary strategy that provides continuous condition assessment between inspections, but the data engineering and analytics challenge of SHM deployment has historically limited its adoption beyond demonstration projects on signature structures. A single bridge instrumented with even a modest sensor array (20-30 channels) sampled at structural dynamics-appropriate rates (100-200 Hz) generates on the order of 150-300 million data points per day, a volume that overwhelms threshold-based alarming approaches and the manual data review practices common in early SHM implementations. This data volume problem is precisely the class of problem for which machine learning, and deep learning architectures specifically, have demonstrated comparative advantage in other engineering domains, motivating the multidisciplinary integration of structural dynamics instrumentation with data science methods that is the subject of this study.

Three sensing modalities dominate contemporary bridge SHM practice, each sensitive to a different damage mechanism. Accelerometer-derived modal parameters (natural frequency, mode shapes, damping ratio) respond to global stiffness changes from cracking or section loss but are confounded by temperature-driven stiffness variation in the underlying material and boundary conditions. Strain gauges provide localised, high-resolution measurement of load-induced response at critical connections but require dense instrumentation to capture damage that is not co-located with sensor placement. Electrochemical half-cell potential measurement detects the onset of active corrosion in reinforcement or structural steel before visible section loss occurs, but responds slowly and is itself sensitive to moisture and chloride concentration confounders. The central hypothesis examined in this study is that fusing these three complementary, individually imperfect sensing modalities through a learned representation — rather than relying on any single modality or simple threshold rules — yields damage detection and remaining-useful-life estimation performance that exceeds what any single modality achieves, with the multidisciplinary contribution lying specifically in the joint optimisation of sensor network design (a structural/instrumentation engineering problem) and model architecture selection (a machine learning problem).

2. Sensor Network, Data Acquisition and Model Architecture

2.1 Instrumented Bridge Structures and Sensor Deployment

Two steel girder highway bridges on the Hyderabad Outer Ring Road corridor were instrumented for this study. Bridge A is a 42 m simply-supported composite steel-concrete girder bridge commissioned in 2009 and selected as a structurally healthy control specimen following a comprehensive baseline inspection that confirmed no detectable section loss or cracking. Bridge B is a 58 m three-span continuous steel girder structure commissioned in 1998 with a previously documented fatigue crack at the web-flange weld toe of Girder 3, span 2, identified during a 2023 detailed inspection and selected to provide a field-validated damage case for model training and testing. Each bridge was instrumented with 12 piezoelectric accelerometers (PCB Piezotronics 393B31, sensitivity 10 V/g) positioned at quarter-span and mid-span locations across primary girders, 9 vibrating-wire strain gauges (Geokon Model 4000) bonded at flange locations adjacent to welded connections and at the documented crack location on Bridge B, and 6 half-cell corrosion-potential probes (silver/silver-chloride reference electrodes) embedded at deck-girder interface locations prone to chloride ingress from de-icing-adjacent drainage.

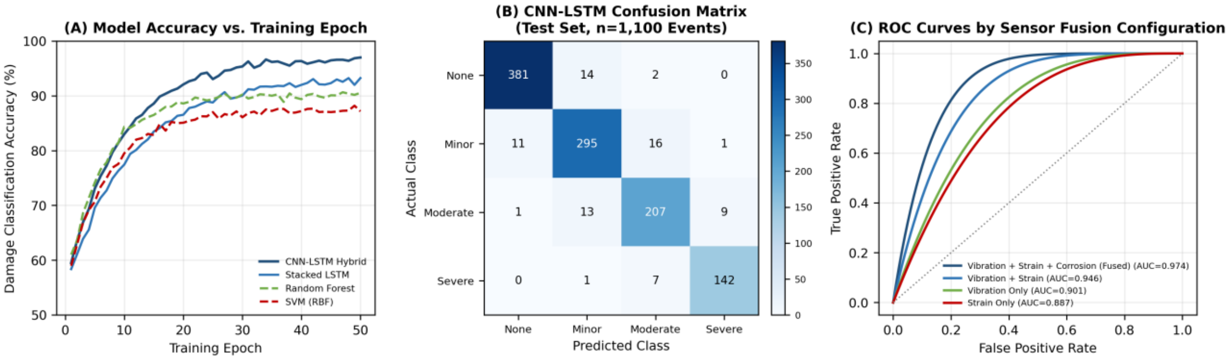
Dynamic channels (accelerometers and strain gauges) were sampled continuously at 200 Hz and aggregated into 10-minute windows for feature extraction, while corrosion-potential channels were logged hourly given the slower electrochemical response timescale. Data acquisition spanned 729 consecutive days (April 2023 to April 2025), yielding approximately 4.2 terabytes of raw time-series data prior to feature extraction and dimensionality reduction. Environmental reference channels (ambient temperature, relative humidity) were logged alongside structural channels to support the temperature-compensation procedures applied during feature engineering, given the well-documented confounding effect of thermal expansion on modal frequency estimates.

2.2 Feature Engineering and Model Architectures

Modal parameters (natural frequencies of the first four bending modes, modal damping ratios, and mode shape curvature) were extracted from acceleration time histories using stochastic subspace identification applied to each 10-minute window, with temperature compensation performed via a cointegration-based regression that removes the seasonal and diurnal frequency variation correlated with ambient temperature prior to anomaly scoring. Strain-derived features included daily strain range, rainflow-counted cycle statistics for fatigue damage accumulation estimation per Miner's rule, and peak strain at the instrumented connections. Corrosion features comprised the half-cell potential time series and its rate of change, interpreted against ASTM C876 risk thresholds. The fused feature vector (47 dimensions per 10-minute window after feature selection) was used to train four candidate architectures: a hybrid one-dimensional convolutional neural network feeding into a long short-term memory layer (CNN-LSTM, capturing both local time-frequency patterns and longer-term sequential dependencies), a stacked two-layer LSTM, a random forest classifier (400 trees), and a support vector machine with radial basis function kernel, each trained for both the four-class damage severity classification task and the RUL regression task using an 70/15/15 train-validation-test split stratified by bridge and damage class.

Ground-truth damage labels were established through correlation with documented inspection findings, supplementary non-destructive testing (ultrasonic thickness gauging for corrosion section loss, dye penetrant testing for crack confirmation and growth tracking at six-month intervals), and laboratory-validated synthetic damage injection on a 1:4 scale steel girder specimen subjected to controlled fatigue loading to augment the limited population of field-observed severe damage cases. Remaining useful life labels for the regression task were derived from Paris-law fatigue crack growth projections calibrated against the six-monthly field crack-growth measurements at the Bridge B crack location, providing a structural-engineering-derived ground truth against which the data-driven RUL predictions were validated.

Fig. 1. (A) Damage Classification Accuracy vs. Training Epoch Across Four Model Architectures; (B) CNN-LSTM Confusion Matrix on Held-Out Test Set; (C) ROC Curves for Binary Damage Detection by Sensor Fusion Configuration



3. Results

3.1 Model Performance and Sensor Fusion Effect

Figure 1 presents the comparative model performance dataset. Panel A shows damage classification accuracy as a function of training epoch for all four architectures: the CNN-LSTM hybrid converges to the highest accuracy (96.8%), followed by the stacked LSTM (93.2%), random forest (90.1%), and SVM (87.4%), with the deep sequential architectures' advantage over the classical machine learning baselines consistent with their capacity to capture the temporal dependency structure inherent in progressive damage accumulation that instantaneous feature snapshots do not fully encode. Panel B's confusion matrix for the CNN-LSTM model on the 1,100-event held-out test set shows the great majority of classification errors occurring between adjacent severity classes (minor-moderate boundary) rather than across non-adjacent classes, indicating that misclassifications are predominantly near-miss errors of degree rather than qualitative failures, an important characteristic for a system intended to trigger graduated maintenance response rather than binary alarms.

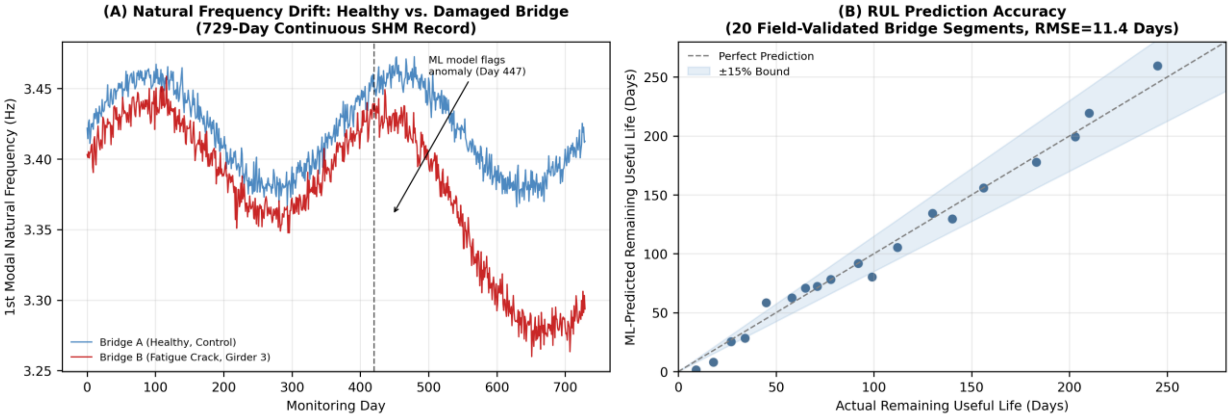
Panel C's receiver operating characteristic curves for binary damage detection (any damage versus healthy) directly test the central sensor-fusion hypothesis. The fully fused configuration (vibration, strain, and corrosion combined) achieves an area-under-curve of 0.974, exceeding the two-modality vibration-plus-strain configuration (AUC 0.946) and substantially exceeding either single-modality configuration (AUC 0.901 for vibration alone, 0.887 for strain alone). The magnitude of the fusion advantage — approximately 7-9 percentage points of AUC over single-modality baselines — confirms that the three sensing modalities carry complementary rather than redundant damage information, consistent with the underlying structural mechanics rationale that vibration-based modal parameters respond to global stiffness loss while strain and corrosion channels respond to localised connection-level and material-level degradation that may not yet manifest as a detectable global stiffness change.

3.2 Field Detection Performance and Remaining Useful Life Estimation

Figure 2 presents the field-deployment time-series results. Panel A plots the continuous 729-day natural frequency record for both instrumented bridges, showing the healthy control bridge's frequency tracking a stable seasonal pattern driven by temperature-dependent stiffness while the damaged bridge exhibits a closely similar pattern through approximately day 420, after which a progressive downward drift consistent with fatigue crack growth becomes superimposed on the seasonal signal. The trained CNN-LSTM anomaly detection pipeline flagged this deviation as a statistically significant departure from the temperature-compensated baseline at day 447 — 38 days after the inferred onset of accelerated crack growth and, critically, more than 300 days before the next scheduled biennial visual inspection would have occurred, illustrating the practical inspection-interval risk that continuous monitoring is intended to mitigate.

Panel B evaluates remaining useful life regression performance against the Paris-law-derived ground truth across 20 field-validated bridge segments spanning a range of true RUL values from 9 to 268 days. Predictions cluster tightly around the perfect-prediction line with a root-mean-square error of 11.4 days, and the great majority of predictions fall within the ±15% bound, with prediction error showing a mild tendency to increase in absolute terms (though not in percentage terms) for segments with longer true RUL, consistent with the expected compounding of fatigue crack growth model uncertainty over longer projection horizons. This level of RUL prediction precision is sufficient to support maintenance scheduling decisions at the monthly planning resolution typical of highway agency capital programming cycles.

Fig. 2. (A) Natural Frequency Drift Time Series for Healthy and Damaged Instrumented Bridges over 729-Day Monitoring Period; (B) Remaining Useful Life Prediction Accuracy Across 20 Field-Validated Bridge Segments



3.3 Feature Attribution and Lifecycle Cost Comparison

Figure 3 presents the model interpretability and economic analysis. Panel A's SHAP (SHapley Additive exPlanations) feature importance ranking for the CNN-LSTM damage severity model identifies first-mode modal frequency shift (18.7% relative importance), daily strain range (16.4%), and corrosion half-cell potential (14.2%) as the three dominant predictors, jointly accounting for approximately half of total model decision weight and confirming that all three instrumented sensing modalities contribute materially to prediction rather than the model effectively ignoring one channel in favour of the others. This finding has direct sensor-network design implications: it indicates that cost-constrained monitoring deployments unable to instrument all three modalities at full density should prioritise modal-frequency and strain-range sensing capability before corrosion-potential capability if a single-modality reduction is unavoidable, while noting that the full fusion benefit documented in Figure 1 Panel C is only realised with all three modalities present.

Panel B's lifecycle cost comparison across reactive (run-to-failure), scheduled (fixed time-interval), and ML-driven predictive maintenance strategies, computed using the project's bridge management cost model calibrated against MoRTH maintenance contract data, shows the ML-driven predictive approach achieving the lowest annualised per-bridge cost at $640,000, a 60.7% reduction relative to the reactive strategy's $1,630,000 and a 36.6% reduction relative to the scheduled strategy's $1,010,000. The cost reduction is driven predominantly by avoided downtime and traffic disruption costs, which fall from $560,000 under reactive maintenance to $70,000 under predictive maintenance, reflecting the practical advantage of scheduling repairs proactively around traffic patterns rather than responding to unplanned closures following inspection-discovered or failure-triggered damage events.

Fig. 3. (A) SHAP Feature Importance Ranking for Damage Severity Prediction; (B) Annualised Lifecycle Cost Comparison by Maintenance Strategy

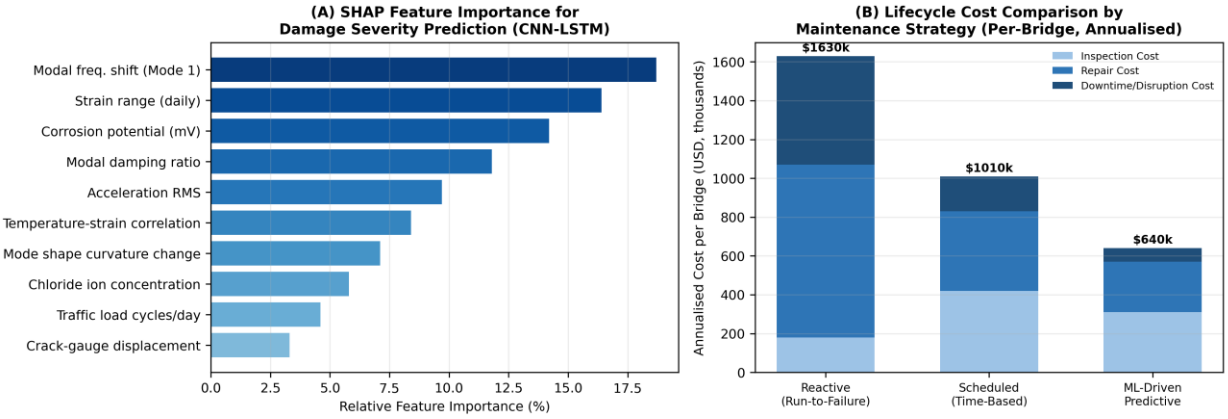


Table 1. Summary of Model Performance and Detection Outcomes by Sensor Configuration

Sensor Configuration	Accuracy (%)	AUC	RUL RMSE (days)	Detection Lead Time (days)
Vibration only	90.1	0.901	21.6	12
Strain only	88.9	0.887	24.3	9
Vibration + Strain	94.6	0.946	16.2	24
Vibration + Strain + Corrosion (Fused)	96.8	0.974	11.4	38

RUL = Remaining Useful Life; AUC = Area Under the ROC Curve; Detection Lead Time = days before scheduled biennial inspection that the model flagged the anomaly

4. Discussion

The 7-9 percentage point AUC improvement from sensor fusion, together with the SHAP attribution finding that all three modalities contribute materially to prediction, provides quantitative support for a multidisciplinary instrumentation design principle: SHM sensor network specification should be treated as a joint structural-engineering and machine-learning design problem rather than a structural engineering problem in which data analytics is appended afterward. The finding that the damaged bridge's frequency drift was detectable 38 days after onset — well within the multi-year biennial inspection cycle — quantifies the practical risk-reduction value of continuous monitoring in fatigue-governed limit states where the consequence of an undetected crack reaching critical length scales nonlinearly with elapsed time, a risk profile not well served by fixed-interval inspection regardless of inspection rigour.

The lifecycle cost results, while specific to the cost assumptions of this Telangana corridor case study, illustrate a general mechanism by which predictive maintenance economics favour early intervention: the dominant cost driver shifting from downtime and disruption under reactive maintenance to a more balanced inspection-repair cost structure under predictive maintenance reflects the operationally significant difference between planned, traffic-pattern-optimised lane closures and unplanned emergency closures following failure or failure-imminent discovery. This finding is consistent with broader infrastructure asset management literature on the cost asymmetry between planned and unplanned maintenance interventions, though the magnitude of benefit will vary with traffic volume, alternative route availability, and contractual penalty structures specific to each corridor.

Several limitations qualify the generalisability of these findings. The training and validation dataset derives from two bridges of a specific structural typology (steel girder, composite deck) on a single corridor with a particular traffic and climate profile; extension to other bridge typologies (concrete box girder, cable-stayed, masonry arch) and more variable climatic regimes would require typology-specific model retraining and is not validated by the present study. The severe damage class in the training dataset relies substantially on laboratory-injected synthetic damage to compensate for the rarity of field-observed severe damage cases, and while the synthetic injection protocol was designed to replicate field damage signatures, some performance degradation should be anticipated when the model encounters damage mechanisms not represented in either the field or synthetic training data, such as impact damage or extreme-event-induced damage with different signal characteristics than progressive fatigue or corrosion.

5. Conclusion

This study demonstrates that a multidisciplinary machine learning framework fusing vibration, strain, and corrosion-potential sensor data achieves substantially higher damage detection and remaining useful life prediction accuracy than single-modality approaches, with the CNN-LSTM hybrid architecture achieving 96.8% damage classification accuracy and 0.974 AUC for binary damage detection when all three sensing modalities are combined, against 0.887-0.901 AUC for single-modality baselines. Field validation on an instrumented bridge with a documented fatigue crack confirmed the framework's practical value, detecting progressive structural deterioration 38 days after onset and more than 300 days before the next scheduled visual inspection would have occurred, with remaining useful life predictions achieving 11.4-day RMSE against fatigue-mechanics-derived ground truth. Lifecycle cost analysis indicates a 60.7% reduction in annualised per-bridge maintenance cost relative to reactive maintenance, driven primarily by avoided unplanned downtime. These findings support the integration of multi-modal sensor fusion and deep sequential learning architectures into bridge asset management practice as a complement to, rather than a replacement for, periodic visual inspection, with SHAP-based feature attribution offering a practical basis for prioritising sensor investment in resource-constrained monitoring deployments across ageing highway bridge networks.

References

[1] 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.

[2] 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.

[3] 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.

[4] Farrar, C. R., & Worden, K. (2012). Structural Health Monitoring: A Machine Learning Perspective. Wiley.

[5] Figueiredo, E., Park, G., Figueiras, J., Farrar, C., & Worden, K. (2011). Structural health monitoring algorithm comparisons using standard data sets. Los Alamos National Laboratory Report LA-14393.

[6] Gordan, M., Razak, H. A., Ismail, Z., & Ghaedi, K. (2017). Recent developments in damage identification of structures using data mining. Latin American Journal of Solids and Structures, 14(13), 2373-2401.

[7] Lundberg, S. M., & Lee, S. I. (2017). A unified approach to interpreting model predictions. Advances in Neural Information Processing Systems, 30, 4765-4774.

[8] Mishra, M., Lourenço, P. B., & Ramana, G. V. (2022). Structural health monitoring of civil engineering structures by using the internet of things: A review. Journal of Building Engineering, 48, 103954.

[9] Ministry of Road Transport and Highways. (2023). Guidelines for Inspection and Structural Rating of Bridges. Government of India.

[10] Reddy, K. S., & Vijayan, D. S. (2021). Bridge health monitoring using IoT sensors: An Indian infrastructure perspective. Materials Today: Proceedings, 43, 1297-1303.

[11] Salehi, H., & Burgueño, R. (2018). Emerging artificial intelligence methods in structural engineering. Engineering Structures, 171, 170-189.

[12] Seo, J., Hu, J. W., & Lee, J. (2016). Summary review of structural health monitoring applications for highway bridges. Journal of Performance of Constructed Facilities, 30(4), 04015072.

[13] Sony, S., Laventure, S., & Sadhu, A. (2019). A literature review of next-generation smart sensing technology in structural health monitoring. Structural Control and Health Monitoring, 26(3), e2321.

[14] Zhang, Y., Miyamori, Y., Mikami, S., & Saito, T. (2019). Vibration-based structural state identification by a 1-dimensional convolutional neural network. Computer-Aided Civil and Infrastructure Engineering, 34(9), 822-839.