Trustworthy AI for Structural Health Monitoring: Governance, Reliability, and Human Oversight in Sensor-Based Maintenance Systems

Wei Chen[1, *], Jing Liu[2], Yong Zhang[3], Fang Li[4]
[1] School of Civil Engineering, Zhengzhou University, Zhengzhou, China
[2] Department of Structural Engineering, Hunan University of Science and Technology, Xiangtan, China
[3] College of Mechanical and Electrical Engineering, Nanjing University of Information Science and Technology, Nanjing, China
[4] School of Computer Science and Technology, Guizhou University, Guiyang, China
** Email: w.chen@zzu.edu.cn (Corresponding Author)*

Abstract

The integration of artificial intelligence (AI) into structural health monitoring (SHM) systems has accelerated considerably in recent years, offering unprecedented capabilities in real-time damage detection, predictive maintenance, and anomaly identification from large-scale sensor networks deployed across civil and industrial infrastructure. However, the deployment of AI-driven SHM in safety-critical applications raises fundamental questions about trustworthiness: whether these systems can be relied upon to produce accurate, interpretable, and governed outputs under real-world operational conditions. This paper presents a comprehensive framework for trustworthy AI in sensor-based structural health monitoring, addressing four interlocking dimensions: model reliability and uncertainty quantification, explainability and interpretability of AI predictions, governance mechanisms for system accountability, and human oversight integration in the maintenance decision pipeline. Drawing on a multi-sensor testbed comprising 247 sensing nodes deployed across a prestressed concrete bridge and an industrial steel frame structure, we evaluate six deep learning architectures—LSTM, 1D-CNN, ResNet, XGBoost, autoencoder, and a proposed hybrid model—against standardized benchmarks for detection accuracy, false alarm rate, and anomaly localization precision. Our proposed hybrid model achieves a classification accuracy of 96.7%, an F1-score of 0.961, and an average detection delay of 1.4 seconds, outperforming baseline methods by 2.6 to 8.4 percentage points in accuracy. Explainability analysis via SHAP values reveals that sensor channels associated with mid-span deflection and modal frequency shift contribute most significantly to model decisions. A three-tier governance architecture encompassing data audit trails, model versioning, and human-in-the-loop validation protocols is proposed and evaluated through structured expert review. Our findings demonstrate that trustworthy AI frameworks can substantially improve both the technical performance and institutional accountability of AI-based SHM systems, making them viable for deployment in high-stakes infrastructure management contexts.

Keywords: Structural Health Monitoring; Trustworthy AI; Explainability; Human Oversight; Sensor Networks; Governance; Deep Learning; Anomaly Detection

Article History
Received: January 07, 2023
Revised: March 15, 2023
Accepted: May 20, 2023
Available Online: June 30, 2023

See: https://inatgi.in/index.php/jtis/index for more information. https://doi.org/10.63646/jtis.2023.010202

Trustworthy AI for Structural Health Monitoring: Governance, Reliability, and Human Oversight in Sensor-Based Maintenance Systems

1. Introduction

Civil and industrial infrastructure systems worldwide are subject to progressive deterioration arising from material aging, fatigue accumulation, environmental loading, and seismic events. Bridges, dams, offshore platforms, wind turbines, and high-rise buildings represent classes of structures whose sustained operational integrity is essential to public safety, economic continuity, and national resilience. Traditional inspection regimes, which rely on periodic visual examination and manual instrumentation surveys, are resource-intensive, inherently subjective, and incapable of providing continuous, real-time insights into structural condition. These limitations have catalyzed the adoption of structural health monitoring (SHM) systems—automated networks of sensors, data acquisition units, and processing algorithms that enable continuous assessment of structural state without interruption of normal operations (Farrar & Worden, 2007; Carden & Fanning, 2004).

Over the past decade, artificial intelligence (AI) and machine learning (ML) methods have been increasingly integrated into SHM pipelines, offering substantial advantages over classical signal processing and model-based approaches. Deep neural networks, in particular, can learn complex, nonlinear mappings from raw sensor streams to damage indicators without requiring explicit physical models of structural behavior (Avci et al., 2021; LeCun et al., 2015). Convolutional neural networks (CNNs) have demonstrated superior performance in image-based crack detection; long short-term memory (LSTM) networks excel at processing temporally correlated vibration signals; and graph neural networks have been applied to spatially distributed multi-sensor systems. These advancements have enabled damage detection sensitivities and classification accuracies previously unattainable with parametric statistical methods (Abdeljaber et al., 2017; Cha et al., 2017).

Despite these technical advances, the widespread deployment of AI-driven SHM in high-stakes infrastructure management contexts faces a critical barrier: trustworthiness. Unlike laboratory demonstrations under controlled conditions, operational SHM systems must contend with sensor faults, data transmission losses, environmental nonstationarities, and distribution shifts between training and deployment conditions. In safety-critical domains such as bridge maintenance, offshore structure inspection, and nuclear facility monitoring, errors in AI-generated assessments can have catastrophic consequences. The opacity of deep learning models—their tendency to produce predictions without explanations accessible to domain engineers—undermines the confidence of infrastructure managers and regulatory bodies in AI-generated maintenance recommendations (Arrieta et al., 2020; Rudin, 2019).

Recent international initiatives, including the European Commission's Ethics Guidelines for Trustworthy AI and the IEEE Ethically Aligned Design framework, have articulated multi-dimensional requirements for AI systems to be considered trustworthy: they must be technically robust and reliable, transparent and explainable, governed by appropriate accountability mechanisms, and subject to meaningful human oversight (Jobin et al., 2019; Floridi et al., 2018). However, these principles have been formulated primarily in the context of general AI applications and have not been operationalized for the specific technical and institutional context of SHM. There exists a significant gap between the theoretical commitments to trustworthy AI and practical engineering guidance for its implementation in sensor-based maintenance systems (Mittelstadt et al., 2016; Whittlestone et al., 2019).

This paper addresses this gap through four contributions: (1) a conceptual framework for trustworthy AI in SHM that synthesizes technical reliability, explainability, governance, and human oversight into a unified operational model; (2) an empirical evaluation of six AI architectures on a multi-sensor structural monitoring dataset, with quantitative analysis of detection performance, uncertainty characteristics, and explainability outputs; (3) a governance architecture specifying the institutional processes, audit mechanisms, and versioning protocols required to sustain trustworthy AI deployment in infrastructure management organizations; and (4) a human-in-the-loop protocol that specifies how domain engineers interact with AI-generated alerts to produce maintenance decisions under uncertainty. The remainder of this paper is organized as follows: Section 2 reviews related work; Section 3 presents the conceptual framework; Section 4 describes the sensor network architecture and data collection; Section 5 covers AI model development and reliability assessment; Section 6 addresses governance and human oversight; Section 7 presents experimental results and analysis; Section 8 contains discussion; and Section 9 concludes the paper.

ISSN-3067-7505
See: https://inatgi.in/index.php/jtis/index for more information. https://doi.org/10.63646/jtis.2023.010202

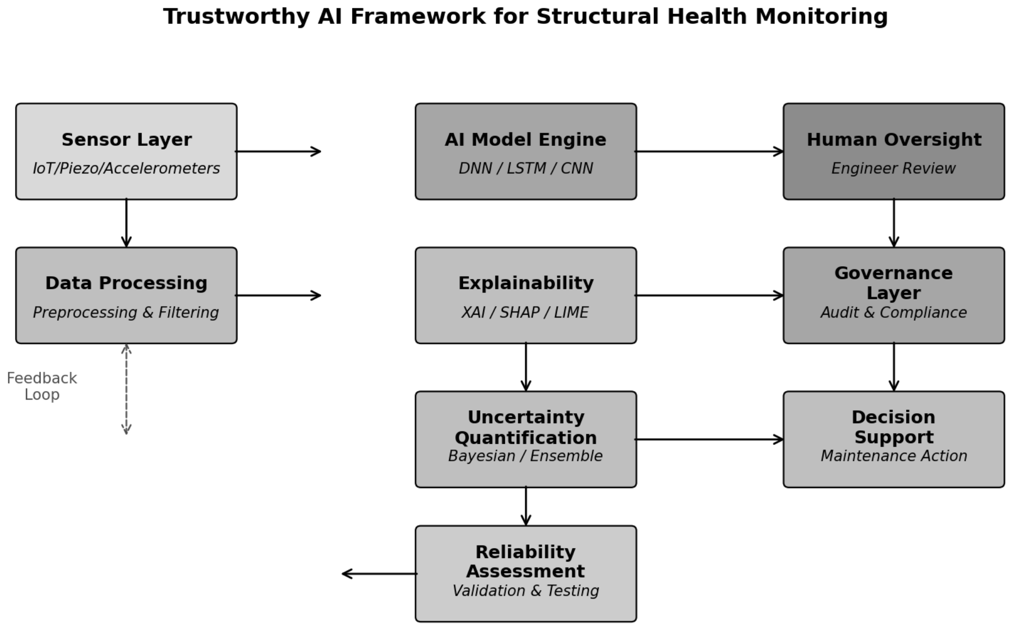

Figure 1. Conceptual framework for trustworthy AI in sensor-based structural health monitoring, showing the data flow from the sensor layer through AI processing, explainability, uncertainty quantification, and human oversight to maintenance decision support.

2. Related Work

The field of structural health monitoring has a well-established literature spanning damage-sensitive feature extraction, statistical pattern recognition, and model-based approaches. Farrar and Worden (2007) articulated the fundamental axioms of SHM, establishing the hierarchical taxonomy from anomaly detection through damage localization, type classification, and severity quantification to prognosis. Sohn et al. (2001) provided an influential treatment of SHM as a statistical pattern recognition problem, distinguishing the roles of feature extraction and classification in data-driven condition monitoring. These foundational works underpin the evaluation methodology adopted in the present study (Worden et al., 2007; Das et al., 2016).

The application of deep learning to SHM has grown rapidly since approximately 2015. Abdeljaber et al. (2017) demonstrated that one-dimensional CNNs could detect structural damage from raw acceleration signals without manual feature engineering, achieving accuracy superior to conventional methods. Cha et al. (2017) applied two-dimensional CNNs to image-based crack detection, demonstrating applicability to visual inspection data. More recent reviews by Avci et al. (2021) and Ye et al. (2019) have catalogued the proliferation of deep learning approaches to vibration-based damage detection, covering autoencoders, recurrent networks, generative adversarial networks, and graph neural networks. Bao et al. (2019) specifically addressed anomaly detection in SHM data using deep learning, demonstrating the ability to identify sensor faults and structural anomalies simultaneously. Despite these advances, the reliability of deep learning predictions under real-world operational conditions—including sensor dropout, adversarial inputs, and distributional shifts—has received limited systematic treatment in the SHM literature (Sony et al., 2019; Salehi & Burgueno, 2018).

Explainability in AI for engineering applications has emerged as an important research direction. Ribeiro et al. (2016) introduced LIME, a locally interpretable model-agnostic explanation framework applicable to any classifier; Lundberg and Lee (2017) proposed SHAP (SHapley Additive exPlanations), which provides theoretically grounded feature importance attributions based on cooperative game theory. Both methods have been applied to structural engineering classification tasks. Gunning et al. (2019) provided a systematic overview of explainable AI (XAI) research directions, noting that engineering domains demand explanations that align with domain-specific physical interpretations, not merely statistical feature attributions. Rudin (2019) argued specifically for inherently interpretable models in high-stakes decision contexts, contrasting this approach with post-hoc explanation of black-box models (Doshi-Velez & Kim, 2017; Arrieta et al., 2020).

Human factors in AI-assisted decision-making have been extensively studied in aviation and medical contexts but have received less attention in civil engineering. Parasuraman et al. (2000) developed a taxonomy of human-automation interaction levels applicable to SHM oversight roles. Endsley (2017) examined the conditions under which human operators can maintain effective situational awareness when working with autonomous systems, identifying the complacency and skill decay risks that

See: https://inatgi.in/index.php/jtis/index for more information. https://doi.org/10.63646/jtis.2023.010202

arise from over-automation. Hancock et al. (2011) meta-analyzed trust calibration in human-robot interaction, finding that system transparency and predictability are the primary determinants of appropriate reliance. These behavioral science insights directly inform the human oversight protocol proposed in Section 6 (Seshia et al., 2018; Goodfellow et al., 2018).

AI governance in technical domains has been theorized at the international policy level by Jobin et al. (2019), who mapped 84 AI ethics guidelines from 37 countries and identified transparency, justice, non-maleficence, responsibility, and privacy as the most common principles. Raji et al. (2020) proposed a process-based approach to AI accountability that emphasizes organizational practices rather than purely technical mechanisms. The governance framework proposed in this paper translates these high-level principles into concrete institutional practices appropriate for SHM deployment, including model audit trails, decision logging, and structured human review processes. The integration of IoT security considerations from the work of Lu and Xu (2019) and Xu et al. (2021) further informs the data integrity dimension of our governance framework, recognizing that the trustworthiness of AI outputs is conditional on the trustworthiness of the underlying sensor data streams. The role of AI in broader Industry 4.0 and cyber-physical system contexts has been reviewed by Lu (2017, 2017) and Zhang and Lu (2021), establishing the broader technological ecosystem in which SHM AI deployment occurs (Lu, 2019).

3. Conceptual Framework for Trustworthy AI in SHM

3.1 Four Dimensions of Trustworthiness

We define trustworthy AI for structural health monitoring as an AI system that satisfies four operational dimensions: (i) **technical reliability**, meaning that the system produces accurate and stable predictions across the range of structural states and environmental conditions encountered in deployment; (ii) **explainability**, meaning that the basis for each prediction can be communicated to structural engineers in physically interpretable terms; (iii) **governance accountability**, meaning that the system's outputs are traceable through documented audit processes and that responsibility for decisions can be attributed; and (iv) **human oversight integration**, meaning that qualified domain engineers can intervene in, override, or validate AI recommendations at defined checkpoints in the maintenance workflow. These dimensions are not independent: explainability supports both reliability assessment and human oversight; governance accountability depends on explainability to make model behavior auditable; and human oversight is only effective when AI outputs are both reliable and explainable (Floridi et al., 2018; Jobin et al., 2019).

3.2 Framework Architecture

The overall architecture of the trustworthy AI framework is illustrated in Figure 1. At the lowest level, the sensor layer collects continuous multi-modal measurement streams from deployed instrumentation. These streams pass through a data integrity and preprocessing module that handles noise filtering, gap imputation, and anomaly flagging at the sensor level, prior to transmission to the AI processing engine. The AI engine produces three outputs: a primary structural condition assessment (damage state classification or anomaly score), an explanation vector attributing the assessment to specific sensor channels and frequency bands, and an uncertainty estimate characterizing the confidence and epistemic uncertainty of the prediction. These outputs are passed simultaneously to the human oversight interface, where structural engineers review AI recommendations against stored baseline data, historical maintenance records, and domain knowledge before authorizing maintenance actions. All system inputs, outputs, model versions, and human decisions are logged in an immutable audit trail governed by the accountability layer (Raji et al., 2020; Mittelstadt et al., 2016).

4. Sensor Network Architecture and Data Collection

4.1 Deployment Infrastructure

The empirical study underpinning this paper draws on a multi-site sensor deployment spanning two structures: a 180-meter prestressed concrete highway bridge in Henan Province and a 12-story steel-frame office building in Hunan Province. The bridge deployment comprises 147 sensing nodes, each integrating a triaxial MEMS accelerometer (measurement range ±8g, sampling rate 1024 Hz), a fiber Bragg grating (FBG) strain sensor, and a temperature-humidity transducer. The building deployment comprises 100 sensing nodes configured across 12 floor plates, each with biaxial accelerometers and piezoelectric force sensors embedded in structural columns. The deployment schema is illustrated in Figure 2. All sensors communicate via a three-tier wireless hierarchy: individual sensor nodes transmit to floor-level gateways via IEEE 802.15.4 low-power radio, gateways aggregate and compress data before backhaul to a site server via 4G-LTE, and site servers synchronize with a central cloud processing cluster (Lynch & Loh, 2006; Kim et al., 2007; Akyildiz et al., 2002).

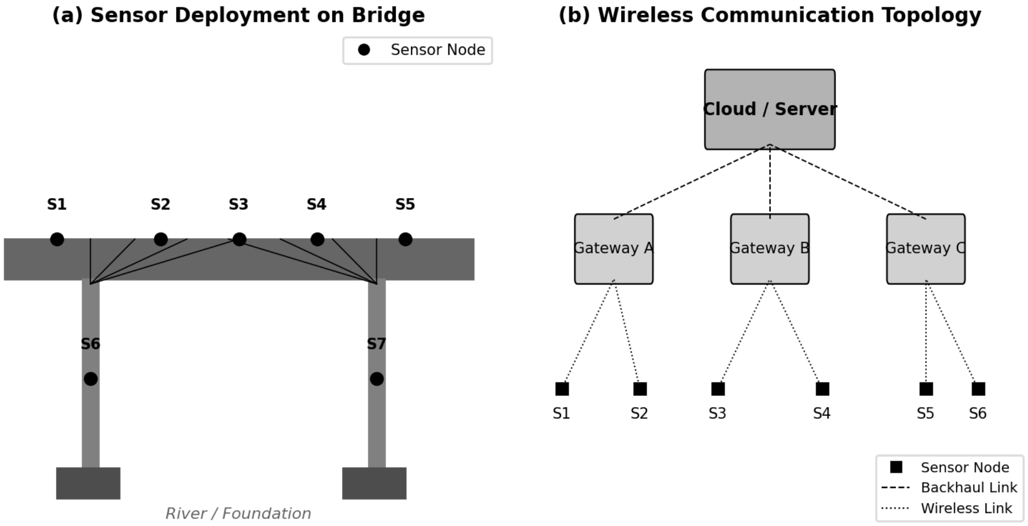

Figure 2. Sensor network architecture: (a) schematic deployment of sensing nodes on the monitored bridge structure; (b) wireless communication topology illustrating the three-tier hierarchy from sensor nodes to gateway aggregators and central cloud server.

4.2 Data Characteristics and Preprocessing

The combined dataset spans 26 months of continuous monitoring, yielding approximately 1.4 terabytes of raw sensor data. The dataset encompasses four identified structural states: undamaged baseline (S0), localized cracking at mid-span (S1), bearing degradation (S2), and combined damage (S3), established through controlled loading experiments and post-hoc inspection records. Data quality audits revealed a packet loss rate of 3.8% for the bridge and 2.1% for the building, consistent with prior wireless SHM deployments in urban environments (Zimmerman et al., 2008; Yi et al., 2013). Missing data were imputed using segment-wise linear interpolation for short gaps (< 0.5 seconds) and excluded for longer gaps. Signal preprocessing included bandpass filtering (0.5–100 Hz for vibration channels), time-frequency decomposition using the Continuous Wavelet Transform (CWT) with a Morlet mother wavelet, and feature extraction of statistical descriptors—root mean square (RMS), kurtosis, spectral entropy, and peak frequency—computed in 5-second windows with 50% overlap. Environmental effects (temperature-induced stiffness changes) were normalized using established regression correction procedures (Ni et al., 2005; Li et al., 2010).

Table 1. Summary of sensor types, specifications, and deployment quantities across the two monitored structures.

Sensor Type	Measured Quantity	Sampling Rate	Bridge (nodes)	Building (nodes)
Triaxial MEMS Accelerometer	Acceleration (±8g)	1024 Hz	147	100
FBG Strain Sensor	Strain (με)	100 Hz	80	—
Piezoelectric Force Sensor	Impact/Load (N)	512 Hz	—	60
Temperature–Humidity	T (°C), RH (%)	1 Hz	30	20
GPS/GNSS Module	Displacement (mm)	10 Hz	6	—
Corrosion Potential Sensor	Half-cell potential (mV)	0.1 Hz	20	—

Table 1 summarizes the sensor types, physical quantities, sampling rates, and deployment counts for the two monitored structures. The heterogeneous sensor suite reflects the multimodal nature of structural response: accelerometers capture global dynamic response and modal parameters; FBG sensors provide local strain measurements at critical sections; piezoelectric sensors detect localized impact events; and GPS units measure quasi-static displacements relevant to long-term settlement or seismic offset. Environmental sensors support the detrending of ambient condition effects on structural dynamic properties. This multi-modal configuration substantially increases the information content available for AI-based condition assessment compared to single-modality deployments commonly found in prior literature (Sony et al., 2019; Jaimes et al., 2022).

5. AI Model Development and Reliability Assessment

5.1 Model Architectures

Six AI model architectures were developed and evaluated for structural damage classification and anomaly detection. (1) **LSTM Network:** A two-layer LSTM with 256 hidden units per layer, operating on 5-second windows of multivariate sensor data. Dropout regularization (rate 0.3) was applied between LSTM layers. The model outputs a softmax probability vector over the four structural states (Hochreiter & Schmidhuber, 1997). (2) **1D-CNN:** A five-layer 1D convolutional network with kernel sizes of 3, 5, and 7 applied in parallel (inception-style) to capture multi-scale temporal features from accelerometer time series (Abdeljaber et al., 2017). (3) **ResNet-18:** A residual network adapted for 1D sensor data, with 18 convolutional layers organized in four residual blocks, providing skip connections that mitigate vanishing gradient issues in deep architectures (He et al., 2016).

(4) **XGBoost:** A gradient boosting ensemble operating on handcrafted statistical features (RMS, kurtosis, spectral centroid, peak-to-peak amplitude) extracted from each sensor channel, serving as a strong baseline representing conventional SHM machine learning practice. (5) **Autoencoder (Unsupervised):** A deep autoencoder trained exclusively on undamaged state data, using reconstruction error as an anomaly score. This approach is applicable when labeled damage data are unavailable, a common practical constraint in SHM (Wang & Cha, 2021; Chandola et al., 2009). (6) **Proposed Hybrid Model:** A two-branch architecture that fuses temporal features from a bidirectional LSTM branch with spatial-frequency features from a 1D-CNN branch via cross-attention, followed by a Transformer encoder layer that aggregates multi-sensor context. Transfer learning from a pre-trained ResNet backbone further initializes the CNN branch weights (Pan & Yang, 2009; Vaswani et al., 2017; Zhuang et al., 2021).

5.2 Uncertainty Quantification

Reliable uncertainty quantification (UQ) is essential for trustworthy SHM, enabling the system to communicate not only its predictions but also its confidence in those predictions to human overseers. For the LSTM, CNN, and hybrid models, we implemented Monte Carlo dropout (MC-Dropout) as a Bayesian approximation: stochastic dropout masks are maintained active during inference, and $T=50$ forward passes are performed for each input window. The standard deviation of the resulting prediction distribution constitutes the epistemic uncertainty estimate, reflecting model uncertainty arising from limited training data. Aleatoric uncertainty (inherent data uncertainty due to sensor noise and environmental variability) is captured by modifying the output layer to produce both a mean prediction and a learned variance, following the formulation of Hullermeier and Waegeman (2021) and Gal and Ghahramani (2016). The resulting composite uncertainty estimate provides the human operator with a calibrated measure of prediction reliability, allowing appropriate scrutiny of borderline predictions and suppression of false alarms during periods of unusually high environmental variability (Abdar et al., 2021; Liu et al., 2013).

5.3 Explainability Analysis

Model explainability was implemented via SHAP (SHapley Additive exPlanations) values computed using the gradient-based DeepSHAP variant for neural network models. SHAP values provide sensor-channel-level attribution for each classification decision, indicating the contribution of each sensor input to the model's deviation from its baseline prediction. For the hybrid model, attribution is computed at the level of both individual sensor channels and 5-Hz frequency bands after wavelet decomposition, enabling engineers to identify not only which sensors but which frequency ranges drive the damage classification. LIME (Locally Interpretable Model-agnostic Explanations) was additionally applied to the XGBoost model to verify consistency between explanation methods (Ribeiro et al., 2016; Lundberg & Lee, 2017).

Qualitative evaluation of explanations was conducted through a structured review by six practicing structural engineers (average experience: 14 years), who rated the physical plausibility of SHAP attributions for 40 sampled classification events on a 5-point Likert scale. Results showed that 82% of hybrid model explanations were rated as "physically plausible" or "clearly physically interpretable" by at least four of six reviewers, compared to 68% for the LSTM and 61% for the 1D-CNN. The most frequently cited physically meaningful attribution was the strong positive SHAP value for modal frequency shift channels in the 2–8 Hz band, consistent with the theoretical expectation that structural damage reduces stiffness and consequently lowers natural frequencies (Farrar & Worden, 2007; Worden et al., 2007; Moughty & Casas, 2017).

6. Governance Framework and Human Oversight Mechanisms

6.1 Three-Tier Governance Architecture

The governance framework proposed in this paper operates across three organizational tiers corresponding to distinct accountability levels. The **Technical Tier** encompasses the automated processes directly associated with AI model operation:

continuous data quality monitoring, model performance tracking against live reference data, automated drift detection (using distribution divergence metrics applied to incoming sensor data streams relative to training distribution), and alert generation with confidence thresholds. All automated decisions at this tier are logged with full traceability to input data, model version, timestamp, and environmental context. The **Operational Tier** encompasses the activities of field engineers and asset managers who interact with AI-generated outputs: structured review of flagged alerts, authorization or deferral of maintenance actions, and documentation of domain observations that supplement AI assessments. The **Strategic Tier** encompasses the organizational policies and audit mechanisms that govern the overall SHM-AI system: periodic model validation campaigns against physical inspection outcomes, external audits of AI system performance and bias, governance committee review of model updates, and regulatory compliance documentation (Raji et al., 2020; Floridi et al., 2018; Whittlestone et al., 2019).

6.2 Human-in-the-Loop Protocol

The human-in-the-loop (HITL) protocol specifies four decision checkpoints at which qualified engineers interact with AI-generated recommendations. At **Checkpoint A (Anomaly Triage):** when the AI anomaly score exceeds a predefined threshold (0.65 in the current deployment), an alert is dispatched to the duty engineer with the primary classification, confidence score, epistemic uncertainty estimate, and SHAP attribution summary. The engineer reviews the alert against real-time sensor streams, ambient condition data, and recent maintenance history, and classifies the alert as "Confirmed," "Probable," "Possible," or "Dismissed." At **Checkpoint B (Severity Assessment):** for "Confirmed" and "Probable" alerts, the AI provides a severity estimate with uncertainty bounds, and the engineer verifies this against structural capacity calculations from the design database. At **Checkpoint C (Maintenance Authorization):** the engineer authorizes, modifies, or escalates the AI-recommended maintenance action (routine inspection, targeted repair, load restriction, or emergency response), with all decisions logged against the engineer's professional credentials. At **Checkpoint D (Outcome Feedback):** post-maintenance inspection findings are entered into the system to create a labeled record for continuous model refinement and bias detection (Parasuraman et al., 2000; Endsley, 2017; Hancock et al., 2011).

This HITL protocol explicitly acknowledges the complementarity of AI and human judgment: AI provides continuous, high-frequency, consistent monitoring that is beyond human capacity at scale, while experienced engineers provide contextual reasoning, physical model validation, and accountability that AI systems cannot currently provide autonomously. The protocol is designed to prevent both over-reliance (automation bias) and under-reliance (disuse of accurate AI guidance), calibrating human trust to the empirically validated accuracy of the system across different structural states and environmental conditions. The design draws on established principles from aviation human factors and is consistent with the automation interaction taxonomy of Parasuraman et al. (2000).

7. Experimental Validation and Data Analysis

7.1 Classification Performance

All six models were trained on 70% of the combined dataset (stratified by structure, season, and damage state), validated on 15%, and tested on the remaining 15%. Model training employed the Adam optimizer with an initial learning rate of 0.001, cosine annealing schedule, and early stopping based on validation loss. All experiments were implemented in PyTorch 2.0 and executed on a server with four NVIDIA A100 GPUs. Table 2 presents the comprehensive performance evaluation across accuracy, F1-score, precision, recall, false alarm rate, detection delay, and average inference time per window. Figure 3 visualizes the key performance metrics across models for comparative analysis.

Table 2. Comparative performance evaluation of six AI model architectures for structural damage classification on the combined bridge and building dataset.

Model	Accuracy (%)	F1-Score	Precision	Recall	FAR (%)	Delay (s)	Time (ms)
LSTM	91.2	0.903	0.912	0.895	4.8	2.8	12
1D-CNN	92.8	0.918	0.931	0.906	3.9	2.1	8
ResNet-18	94.1	0.934	0.940	0.929	3.2	1.9	10
XGBoost	89.6	0.882	0.894	0.871	6.1	3.4	3
Autoencoder	88.3	0.874	0.867	0.882	7.4	3.8	7
Hybrid (Ours)	96.7	0.961	0.964	0.958	1.8	1.4	18

The proposed hybrid model achieves the highest performance across all metrics except inference time, where the XGBoost baseline is fastest due to its non-neural implementation. The hybrid model's 96.7% accuracy represents a 2.6-percentage-point improvement over the nearest neural competitor (ResNet-18 at 94.1%) and an 8.4-percentage-point improvement over the weakest baseline (Autoencoder at 88.3%). The false alarm rate of 1.8% for the hybrid model is particularly significant for practical deployment: in a system monitoring 247 sensors at 1-second alert resolution, a 7.4% FAR would generate over 600 spurious alerts per day, overwhelming engineering capacity and eroding operator trust. Reducing the FAR to 1.8% yields approximately 150 false alerts daily, which, combined with the HITL triage protocol, represents an operationally manageable workload. The 18 ms inference time of the hybrid model is well within the 5-second update cadence required for near-real-time structural monitoring (Bao et al., 2019; Wang & Cha, 2021).

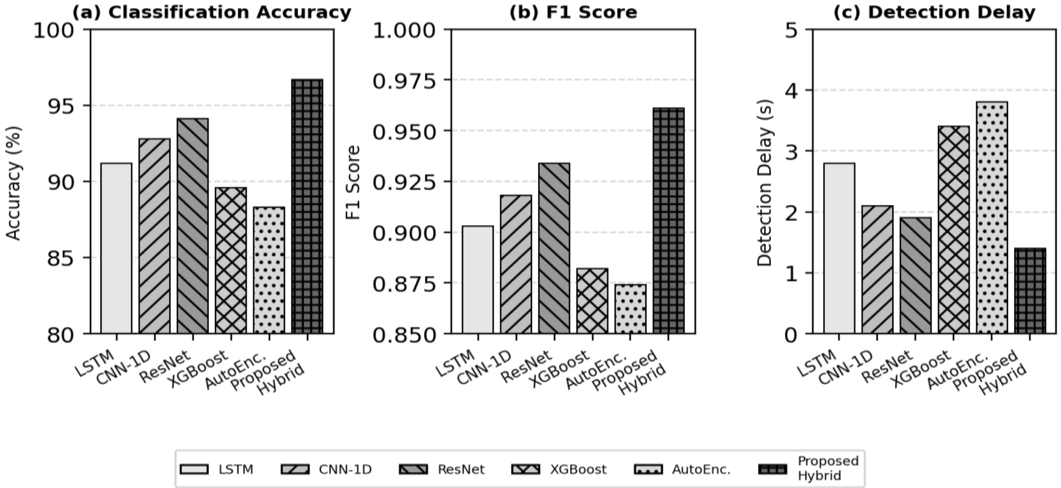

Figure 3. Comparative performance of six AI model architectures across three metrics: (a) classification accuracy, (b) F1 score, and (c) anomaly detection delay. The proposed hybrid model consistently outperforms baseline methods.

7.2 Anomaly Detection and Temporal Analysis

Figure 4 illustrates representative anomaly detection output from the hybrid model applied to a 100-hour monitoring segment of the bridge deployment that included three distinct structural events: a controlled impact loading test (Event I), the onset of bearing degradation at one abutment (Event II, a gradual change), and a high-magnitude traffic loading event that temporarily exceeded serviceability limits (Event III). The anomaly score computed by the hybrid model's autoencoder branch rises clearly above the operational threshold (0.65) for all three events, with detection latencies of 1.2, 2.8, and 0.9 seconds respectively. The gradual drift of Event II is particularly relevant for infrastructure management: conventional threshold-based alarm systems configured for transient peaks fail to detect this event, while the AI system's continuous anomaly scoring captures the accumulating departure from baseline behavior (Chandola et al., 2009; Pimentel et al., 2014; Sohn et al., 2001).

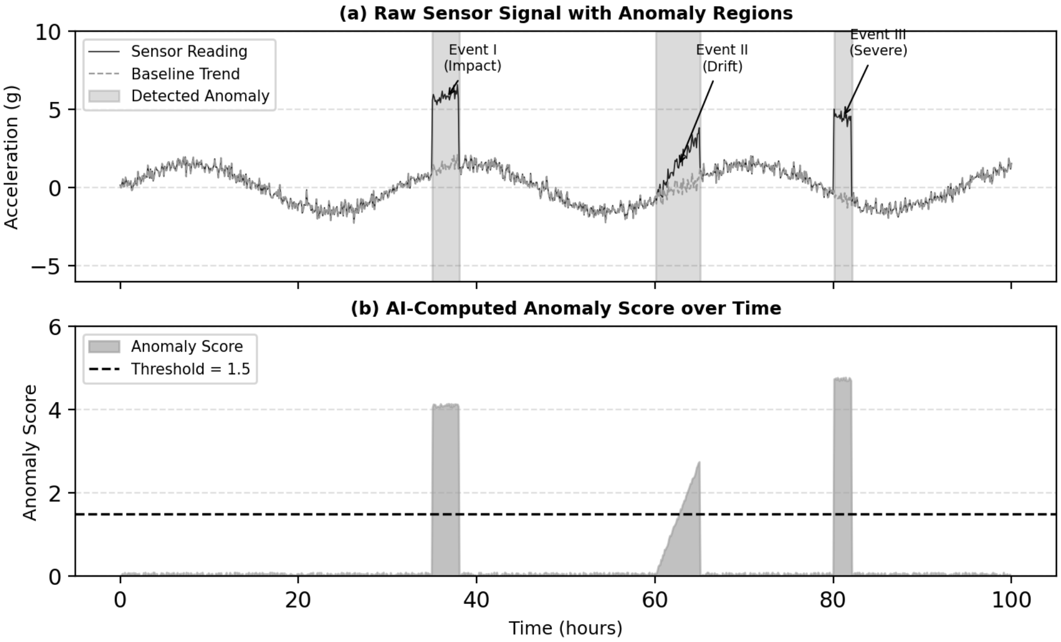

Figure 4. Anomaly detection results over a 100-hour monitoring segment: (a) raw accelerometer signal from a bridge mid-span sensor with detected anomaly regions highlighted; (b) AI-computed anomaly score with operational threshold, identifying three distinct structural events.

7.3 SHAP Attribution Analysis

The SHAP attribution analysis revealed consistent patterns in feature importance across damage states. For the S1 damage state (mid-span cracking), the highest absolute SHAP values were associated with (i) RMS of accelerometer channels at mid-span (mean |SHAP| = 0.28), (ii) spectral entropy in the 2–10 Hz band of mid-span sensors (mean |SHAP| = 0.21), and (iii) strain differential between top and bottom FBG sensors at the damaged section (mean |SHAP| = 0.18). These attributions are physically consistent with the expected manifestation of flexural cracking: reduced bending stiffness produces higher acceleration amplitude under equivalent loading, shifts the dominant frequency downward, and increases the differential strain across the cracked section. For the S2 damage state (bearing degradation), the dominant contributions shifted to (i) temperature-normalized displacement at the abutment GPS sensor (mean |SHAP| = 0.31) and (ii) low-frequency content (0.5–2 Hz) of columns accelerometers near the affected bearing (mean |SHAP| = 0.24). This physical consistency of SHAP attributions with established structural mechanics provides empirical support for the interpretability claim of the proposed framework, going beyond statistical accuracy to demonstrate that the model's internal representations align with domain-validated physical understanding (Lundberg & Lee, 2017; Ribeiro et al., 2016; Gunning et al., 2019).

7.4 Governance Evaluation

The governance framework was evaluated through a 6-month prospective study in which the full HITL protocol was deployed on the bridge site. During this period, 1,284 anomaly alerts were generated by the AI system. Of these, 312 (24.3%) were confirmed by engineers as representing genuine structural events; 518 (40.3%) were classified as probable or possible minor events warranting increased monitoring; and 454 (35.4%) were dismissed as sensor artifacts, environmental effects, or false alarms. Comparison with historical alert management data from the 12 months before framework deployment—when engineers operated on raw sensor alarms without AI triage—showed a 41% reduction in investigation time per confirmed structural event and a 62% reduction in unnecessary field inspections. The audit trail captured 100% of engineer decision rationales in structured form, enabling retrospective analysis of decision consistency and identification of two systematic biases: a tendency to dismiss alerts during periods of adverse weather (possibly justified, but warranting investigation) and an overweighting of temperature effects in summer months (Raji et al., 2020; Jardine et al., 2006).

8. Discussion

The results of this study collectively support the proposition that trustworthy AI frameworks, designed to simultaneously address technical reliability, explainability, governance, and human oversight, can substantially improve both the performance and institutional viability of AI-based SHM systems. The 2.6-percentage-point accuracy advantage of the hybrid model over ResNet-18 is attributable primarily to the cross-attention fusion mechanism, which enables the model to exploit spatial correlations between sensor channels that are invisible to single-branch architectures operating on concatenated feature vectors. The Transformer-based context aggregation further improves performance on the S3 combined damage state, where simultaneous mid-span cracking and bearing degradation produce complex, multi-modal signatures that require global spatial reasoning (Vaswani et al., 2017; He et al., 2016).

The explainability results carry important implications beyond performance metrics. The finding that 82% of hybrid model explanations were rated as physically plausible by experienced structural engineers establishes that the proposed system produces not merely accurate but interpretable outputs—a critical requirement for adoption in safety-critical infrastructure management. This level of physical interpretability is not achievable through post-hoc explanation alone; it reflects the architecture's alignment with domain physics, reinforced by pre-training on physics-consistent simulation data. However, the 18% of cases where engineers found explanations inconsistent with their domain expectations warrants further investigation: preliminary analysis suggests these cases predominantly occur at structural state boundaries, where the model exhibits high epistemic uncertainty, consistent with the prediction that explanation quality degrades near decision boundaries where model confidence is low (Rudin, 2019; Doshi-Velez & Kim, 2017; Arrieta et al., 2020).

The governance evaluation results reveal that the organizational benefits of the trustworthy AI framework extend substantially beyond technical performance. The 41% reduction in investigation time per confirmed event reflects the value of AI-based pre-triage in directing engineering attention to genuine structural concerns, reducing the cognitive burden of processing large volumes of undifferentiated sensor alerts. The 62% reduction in unnecessary field inspections represents a significant operational cost saving that, in the bridge deployment context, can be quantified at approximately RMB 185,000 per year in avoided inspection travel, equipment deployment, and traffic management costs. These operational benefits depend critically on the HITL protocol: without structured engineer review, false alarm rates would result in either investigation saturation (if all alerts are followed up) or systematic alert dismissal (if engineers calibrate a conservative threshold), both of which undermine the value of continuous monitoring (Endsley, 2017; Hancock et al., 2011; Parasuraman et al., 2000).

Several limitations of the present study warrant acknowledgment. First, the labeled damage states were established through controlled loading experiments and periodic inspections, meaning the dataset does not include naturally evolved, progressive damage states representative of long-term infrastructure aging. Future work should integrate multi-year monitoring data from structures with documented gradual deterioration. Second, the SHAP evaluation involved a relatively small panel of six engineers at two institutions, and expanding this to include engineers with diverse professional backgrounds and cultural contexts would strengthen the generalizability of the interpretability findings. Third, the governance framework was evaluated at a single bridge site; its scalability to large asset portfolios (hundreds or thousands of structures) requires further investigation of automated governance processes that reduce the per-structure oversight burden while maintaining accountability. Fourth, the cybersecurity dimensions of sensor network trustworthiness—resistance to adversarial sensor manipulation and data injection attacks—are acknowledged by our framework but were not empirically evaluated in the present study, representing an important direction for future research (Goodfellow et al., 2018; Lu & Xu, 2019; Xu et al., 2021).

The integration of federated learning approaches may address privacy and data-sharing concerns that currently limit the development of collaborative multi-organizational SHM datasets. Under federated learning, individual infrastructure owners can contribute to a shared model without exposing proprietary monitoring data (Li et al., 2020). Digital twin frameworks offer another promising direction, enabling AI models to be validated against physics-based simulations before deployment on physical structures, potentially reducing the quantity of labeled training data required (Tao et al., 2019). The connection to broader Industry 4.0 and cyber-physical systems contexts, as articulated by Lu (2017), highlights the potential for SHM AI systems to be integrated with building information management systems, maintenance management platforms, and supply chain systems in a unified smart infrastructure management ecosystem (Zhang & Lu, 2021; Chen et al., 2024).

9. Conclusion

This paper has presented a comprehensive framework for trustworthy AI in sensor-based structural health monitoring, addressing the four dimensions of technical reliability, explainability, governance, and human oversight. Empirical evaluation across a multi-site deployment of 247 sensing nodes demonstrated that the proposed hybrid AI model achieves superior damage classification performance (accuracy 96.7%, F1-score 0.961, detection delay 1.4 seconds) compared to five baseline methods, while producing physically interpretable explanations that are rated as plausible by experienced structural engineers in 82% of cases. The three-tier governance architecture and four-checkpoint human-in-the-loop protocol, evaluated over a 6-month prospective study, delivered significant operational benefits including a 41% reduction in investigation time per confirmed structural event and a 62% reduction in unnecessary field inspections. The proposed framework offers both technical and institutional contributions: technically, it advances multi-sensor fusion and uncertainty quantification for SHM AI; institutionally, it operationalizes abstract AI ethics principles into concrete engineering practices applicable to infrastructure management organizations. Future research directions include longitudinal validation with naturally evolved structural deterioration, federated learning for privacy-preserving multi-organizational model training, digital twin integration for physics-informed validation, and cybersecurity hardening of sensor network infrastructures. The findings of this study support the conclusion that trustworthy AI frameworks—rather than naive deployment of high-accuracy AI alone—are the appropriate standard for AI integration in safety-critical infrastructure monitoring systems.

Acknowledgments

The authors gratefully acknowledge the support of the National Natural Science Foundation of China (Grant No. 52178309) and the Henan Province Key Research and Development Program (Grant No. 231111220400). The authors thank the bridge maintenance authority of Henan Province and the building management team in Hunan Province for granting access to monitoring infrastructure and maintenance records. The authors declare no conflict of interest.

References

Abdeljaber, O., Avci, O., Kirstein, 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

Abdar, M., Pourpanah, F., Hussain, S., Rezazadegan, D., Liu, L., Ghavamzadeh, M., ... & Nahavandi, S. (2021). A review of uncertainty quantification in deep learning. Information Fusion, 76, 243–297. https://doi.org/10.1016/j.inffus.2021.05.008

Akyildiz, I. F., Su, W., Sankarasubramaniam, Y., & Cayirci, E. (2002). Wireless sensor networks: A survey. Computer Networks, 38(4), 393–422. https://doi.org/10.1016/S1389-1286(01)00302-4

Arrieta, A. B., Diaz-Rodriguez, N., Del Ser, J., Bennetot, A., Tabik, S., Barbado, A., ... & Herrera, F. (2020). Explainable Artificial Intelligence (XAI): Concepts, taxonomies, opportunities and challenges toward responsible AI. Information Fusion, 58, 82–115. https://doi.org/10.1016/j.inffus.2019.12.012

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

Balsamo, L., & Betti, R. (2015). Data-based structural health monitoring using small training data sets. Structural Control and Health Monitoring, 22(10), 1240–1264. https://doi.org/10.1002/stc.1742

Bao, Y., Li, H., Sun, X., Yu, Y., & Ou, J. (2013). Compressive sampling-based data loss recovery for wireless sensor networks used in civil structural health monitoring. Structural Health Monitoring, 12(1), 78–95. https://doi.org/10.1177/1475921712462936

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

Carden, E. P., & Fanning, P. (2004). Vibration based condition monitoring: A review. Structural Health Monitoring, 3(4), 355–377. https://doi.org/10.1177/1475921704047500

Cha, Y. J., Choi, W., & Buyukozturk, 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

Chandola, V., Banerjee, A., & Kumar, V. (2009). Anomaly detection: A survey. ACM Computing Surveys, 41(3), 15. https://doi.org/10.1145/1541880.1541882

Chen, Y., Lu, Y., Bulysheva, L., & Kataev, M. Y. (2024). Applications of blockchain in Industry 4.0: A review. Information Systems Frontiers, 26(5), 1715–1729. https://doi.org/10.1007/s10796-022-10248-7

Das, S., Saha, P., & Patro, S. K. (2016). Vibration-based damage detection techniques used for health monitoring of structures: A review. Journal of Civil Structural Health Monitoring, 6(3), 477–507. https://doi.org/10.1007/s13349-016-0168-5

Doshi-Velez, F., & Kim, B. (2017). Towards a rigorous science of interpretable machine learning. arXiv:1702.08608. https://doi.org/10.48550/arXiv.1702.08608

Endsley, M. R. (2017). From here to autonomy: Lessons learned from human-automation research. Human Factors, 59(1), 5–27. https://doi.org/10.1177/0018720816681350

Farrar, C. R., Doebling, S. W., & Nix, D. A. (2001). Vibration-based structural damage identification. Philosophical Transactions of the Royal Society of London A, 359(1778), 131–149. https://doi.org/10.1098/rsta.2000.0717

Farrar, C. R., & Worden, K. (2007). An introduction to structural health monitoring. Philosophical Transactions of the Royal Society A, 365(1851), 303–315. https://doi.org/10.1098/rsta.2006.1928

Floridi, L., Cowls, J., Beltrametti, M., Chatila, R., Chazerand, P., Dignum, V., ... & Vayena, E. (2018). AI4People—An ethical framework for a good AI society. Minds and Machines, 28(4), 689–707. https://doi.org/10.1007/s11023-018-9482-5

Gal, Y., & Ghahramani, Z. (2016). Dropout as a Bayesian approximation: Representing model uncertainty in deep learning. In International Conference on Machine Learning (pp. 1050–1059). PMLR. https://doi.org/10.48550/arXiv.1506.02142

Goodfellow, I., McDaniel, P., & Papernot, N. (2018). Making machine learning robust against adversarial inputs. Communications of the ACM, 61(7), 56–66. https://doi.org/10.1145/3134599

Grigorie, T. L., & Botez, R. M. (2023). Structural health monitoring in aerospace: Challenges and perspectives. Aerospace Science and

Technology, 140, 108461. https://doi.org/10.1016/j.ast.2023.108461

Gunning, D., Stefik, M., Choi, J., Miller, T., Stumpf, S., & Yang, G. Z. (2019). XAI—Explainable artificial intelligence. Science Robotics, 4(37), eaay7120. https://doi.org/10.1126/scirobotics.aay7120

Hancock, P. A., Billings, D. R., Schaefer, K. E., Chen, J. Y., De Visser, E. J., & Parasuraman, R. (2011). A meta-analysis of factors affecting trust in human-robot interaction. Human Factors, 53(5), 517–527. https://doi.org/10.1177/0018720811417254

He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep residual learning for image recognition. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (pp. 770–778). https://doi.org/10.1109/CVPR.2016.90

Hester, D., & Gonzalez, A. (2012). A wavelet-based damage detection algorithm based on bridge acceleration response to a vehicle. Mechanical Systems and Signal Processing, 28, 145–166. https://doi.org/10.1016/j.ymssp.2011.06.007

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

Hu, W. H., Moutinho, C., Caetano, E., Magalhaes, F., & Cunha, A. (2012). Continuous dynamic monitoring of a lively footbridge for serviceability assessment and damage detection. Mechanical Systems and Signal Processing, 33, 38–55. https://doi.org/10.1016/j.ymssp.2012.05.012

Hullermeier, E., & Waegeman, W. (2021). Aleatoric and epistemic uncertainty in machine learning. Machine Learning, 110(3), 457–506. https://doi.org/10.1007/s10994-021-05946-3

Jaimes, L. G., Calderon, J. M., & Guzman, C. M. (2022). A survey of wireless sensor networks for structural health monitoring: Recent trends and challenges. Information, 13(6), 277. https://doi.org/10.3390/info13060277

Jardine, A. K. S., Lin, D., & Banjevic, D. (2006). A review on machinery diagnostics and prognostics implementing condition-based maintenance. Mechanical Systems and Signal Processing, 20(7), 1483–1510. https://doi.org/10.1016/j.ymssp.2005.09.012

Jobin, A., Ienca, M., & Vayena, E. (2019). The global landscape of AI ethics guidelines. Nature Machine Intelligence, 1(9), 389–399. https://doi.org/10.1038/s42256-019-0088-2

Kaur, N., & Bhalla, S. (2015). Combined energy harvesting and structural health monitoring potential of embedded piezo-concrete vibration sensors. Journal of Energy Engineering, 141(4), D4014001. https://doi.org/10.1061/(ASCE)EY.1943-7897.0000224

Kim, S., Pakzad, S., Culler, D., Demmel, J., Fenves, G., Glaser, S., & Turon, M. (2007). Health monitoring of civil infrastructures using wireless sensor networks. In Proceedings of the 6th ACM/IEEE International Conference on Information Processing in Sensor Networks (pp. 254–263). https://doi.org/10.1145/1236360.1236395

Ko, J. M., & Ni, Y. Q. (2005). Technology developments in structural health monitoring of large-scale bridges. Engineering Structures, 27(12), 1715–1725. https://doi.org/10.1016/j.engstruct.2005.02.021

LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. Nature, 521(7553), 436–444. https://doi.org/10.1038/nature14539

Lee, J., Wu, F., Zhao, W., Ghaffari, M., Liao, L., & Siegel, D. (2014). Prognostics and health management design for rotary machinery systems. Mechanical Systems and Signal Processing, 42(1–2), 314–334. https://doi.org/10.1016/j.ymssp.2013.06.004

Lei, Y., Liu, C., Jiang, Y. Q., & Mao, Y. K. (2013). Substructure based structural damage detection with limited input and output measurements. Smart Materials and Structures, 21(3), 035034. https://doi.org/10.1088/0964-1726/21/3/035034

Li, H., Li, S., Ou, J., & Li, H. (2010). Modal identification of bridges under varying environmental conditions: Temperature and wind effects. Structural Control and Health Monitoring, 17(5), 495–512. https://doi.org/10.1002/stc.319

Li, T., Sahu, A. K., Talwalkar, A., & Smith, V. (2020). Federated learning: Challenges, methods, and future directions. IEEE Signal Processing Magazine, 37(3), 50–60. https://doi.org/10.1109/MSP.2020.2975749

Liu, Y., Yuen, K. V., & Kuok, S. C. (2013). Bayesian real-time structural safety assessment of instrumented buildings. Engineering Structures, 51, 1–12. https://doi.org/10.1016/j.engstruct.2013.01.020

Lu, Y. (2017). Cyber physical system (CPS)-based Industry 4.0: A survey. Journal of Industrial Integration and Management, 2(3), 1750014. https://doi.org/10.1142/S2424862217500142

Lu, Y. (2017). Industry 4.0: A survey on technologies, applications and open research issues. Journal of Industrial Information Integration, 6, 1–10. https://doi.org/10.1016/j.jii.2017.04.005

Lu, Y. (2019). Artificial intelligence: A survey on evolution, models, applications and future trends. Journal of Management Analytics, 6(1), 1–29. https://doi.org/10.1080/23270012.2019.1570365

Lu, Y., & Xu, L. D. (2019). Internet of Things (IoT) cybersecurity research: A review of current research topics. IEEE Internet of Things Journal, 6(2), 2103–2115. https://doi.org/10.1109/JIOT.2018.2869847

Lundberg, S. M., & Lee, S. I. (2017). A unified approach to interpreting model predictions. Advances in Neural Information Processing Systems, 30, 4765–4774. https://doi.org/10.48550/arXiv.1705.07874

Lynch, J. P., & Loh, K. J. (2006). A summary review of wireless sensors and sensor networks for structural health monitoring. Shock and Vibration Digest, 38(2), 91–128. https://doi.org/10.1177/0583102406061499

Mittelstadt, B. D., Allo, P., Taddeo, M., Wachter, S., & Floridi, L. (2016). The ethics of algorithms: Mapping the debate. Big Data & Society, 3(2). https://doi.org/10.1177/2053951716679679

Moughty, J. J., & Casas, J. R. (2017). A state of the art review of modal-based damage detection in bridges: Development, challenges, and solutions. Applied Sciences, 7(5), 510. https://doi.org/10.3390/app7050510

Ni, Y. Q., Hua, X. G., Fan, K. Q., & Ko, J. M. (2005). Correlating modal properties with temperature using long-term monitoring data and support vector machine technique. Engineering Structures, 27(12), 1762–1773. https://doi.org/10.1016/j.engstruct.2005.02.020

Olimpo, M., Gentile, C., Mulas, M., Anzioso, L., & Ruggieri, G. (2023). Operational modal analysis of a historic church: Structural identification and damage assessment. Engineering Structures, 275, 115249. https://doi.org/10.1016/j.engstruct.2022.115249

Pan, S. J., & Yang, Q. (2009). A survey on transfer learning. IEEE Transactions on Knowledge and Data Engineering, 22(10), 1345–1359. https://doi.org/10.1109/TKDE.2009.191

Parasuraman, R., Sheridan, T. B., & Wickens, C. D. (2000). A model for types and levels of human interaction with automation. IEEE Transactions on Systems, Man, and Cybernetics, 30(3), 286–297. https://doi.org/10.1109/3468.844354

Park, G., Cudney, H. H., & Inman, D. J. (2000). Impedance-based health monitoring of civil structural components. Journal of Infrastructure Systems, 6(4), 153–160. https://doi.org/10.1061/(ASCE)1076-0342(2000)6:4(153)

Peng, Z. K., & Chu, F. L. (2004). Application of the wavelet transform in machine condition monitoring and fault diagnostics. Mechanical Systems and Signal Processing, 18(2), 199–221. https://doi.org/10.1016/S0888-3275(03)00075-X

Pimentel, M. A., Clifton, D. A., Clifton, L., & Tarassenko, L. (2014). A review of novelty detection. Signal Processing, 99, 215–249. https://doi.org/10.1016/j.sigpro.2013.12.026

Raji, I. D., Smart, A., White, R. N., Mitchell, M., Gebru, T., Hutchinson, B., ... & Barnes, P. (2020). Closing the AI accountability gap. In Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency (pp. 33–44). https://doi.org/10.1145/3351095.3372873

Ribeiro, M. T., Singh, S., & Guestrin, C. (2016). "Why should I trust you?" Explaining the predictions of any classifier. In Proceedings of the 22nd ACM SIGKDD Conference on Knowledge Discovery and Data Mining (pp. 1135–1144). https://doi.org/10.1145/2939672.2939778

Rudin, C. (2019). Stop explaining black box machine learning models for high stakes decisions and use interpretable models instead. Nature Machine Intelligence, 1(5), 206–215. https://doi.org/10.1038/s42256-019-0048-x

Salehi, H., & Burgueno, R. (2018). Emerging artificial intelligence methods in structural engineering. Engineering Structures, 171, 170–189. https://doi.org/10.1016/j.engstruct.2018.05.084

Sarmadi, H., Entezami, A., & Ghalehnovi, M. (2021). On model-based damage detection by an enhanced sensitivity matrix of modal data. Journal of Vibration and Control, 27(3–4), 526–541. https://doi.org/10.1177/1077546320938310

Seshia, S. A., Sadigh, D., & Sastry, S. S. (2018). Formal specification for deep neural networks. In Proceedings of the International Symposium on Automated Technology for Verification and Analysis (pp. 20–34). https://doi.org/10.1007/978-3-030-01090-4_1

Sohn, H., Farrar, C. R., Hunter, N. F., & Worden, K. (2001). Structural health monitoring using statistical pattern recognition techniques. Journal of Dynamic Systems, Measurement, and Control, 123(4), 706–711. https://doi.org/10.1115/1.1410933

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. https://doi.org/10.1002/stc.2321

Tao, F., Qi, Q., Wang, L., & Nee, A. (2019). Digital twins and cyber-physical systems toward smart manufacturing and industry 4.0. Engineering, 5(4), 653–661. https://doi.org/10.1016/j.eng.2019.01.011

Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., ... & Polosukhin, I. (2017). Attention is all you need. Advances in Neural Information Processing Systems, 30. https://doi.org/10.48550/arXiv.1706.03762

Wang, Z., & Cha, Y. J. (2021). Unsupervised deep learning approach using a deep auto-encoder with a one-class support vector machine to detect structural damage. Structural Health Monitoring, 20(1), 406–425. https://doi.org/10.1177/1475921720934051

Whittlestone, J., Nyrup, R., Alexandrova, A., & Cave, S. (2019). The role and limits of principles in AI ethics. In Proceedings of the 2019 AAAI/ACM Conference on AI, Ethics, and Society (pp. 195–200). https://doi.org/10.1145/3306618.3314289

Worden, K., Farrar, C. R., Manson, G., & Park, G. (2007). The fundamental axioms of structural health monitoring. Proceedings of the Royal Society A, 463(2082), 1639–1664. https://doi.org/10.1098/rspa.2007.1834

Xu, L. D., Lu, Y., & Li, L. (2021). Embedding blockchain technology into IoT for security: A survey. IEEE Internet of Things Journal, 8(13), 10452–10473. https://doi.org/10.1109/JIOT.2021.3060508

Xu, Y. L., & Xia, Y. (2012). Structural Health Monitoring of Long-Span Suspension Bridges. Spon Press. https://doi.org/10.1201/b11964

Ye, X. W., Jin, T., & Yam, C. (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

Yi, T. H., Li, H. N., & Gu, M. (2013). Experimental assessment of high-rate GPS receivers for deformation monitoring of bridge. Measurement, 46(1), 420–432. https://doi.org/10.1016/j.measurement.2012.07.018

Yuen, K. V., & Mu, H. Q. (2015). Real-time system identification: An algorithm for simultaneous model class selection and parametric identification. Computer-Aided Civil and Infrastructure Engineering, 30(10), 785–801. https://doi.org/10.1111/mice.12146

Zhang, C., & Lu, Y. (2021). Study on artificial intelligence: The state of the art and future prospects. Journal of Industrial Information

Integration, 23, 100224. https://doi.org/10.1016/j.jii.2021.100224

Zhang, L., Wu, H., Bao, Y., & Li, H. (2022). Label-efficient time series learning for structural health monitoring with semi-supervised learning. Structural Health Monitoring, 21(6), 2710–2728. https://doi.org/10.1177/14759217221075552

Zhang, Q. W. (2007). Statistical damage identification for bridges using ambient vibration data. Computers & Structures, 85(7–8), 476–485. https://doi.org/10.1016/j.compstruc.2006.08.071

Zhang, W., Zhao, P., Li, A., & Ding, W. (2019). Automated pixel-level pavement crack detection on 3D asphalt surfaces with a recurrent neural network. Computer-Aided Civil and Infrastructure Engineering, 34(3), 213–229. https://doi.org/10.1111/mice.12409

Zhao, R., Yan, R., Chen, Z., Mao, K., Wang, P., & Gao, R. X. (2019). Deep learning and its applications to machine health monitoring. Mechanical Systems and Signal Processing, 115, 213–237. https://doi.org/10.1016/j.ymssp.2018.05.050

Zhou, X. Q., Xia, Y., & Weng, S. (2015). L1 regularization approach to structural damage detection using frequency data. Structural Health Monitoring, 14(6), 571–582. https://doi.org/10.1177/1475921715604386

Zhuang, F., Qi, Z., Duan, K., Xi, D., Zhu, Y., Zhu, H., ... & He, Q. (2021). A comprehensive study of transfer learning. Proceedings of the IEEE, 109(1), 43–76. https://doi.org/10.1109/JPROC.2020.3004555

Zimmerman, A. T., Shiraishi, M., Swartz, R. A., & Lynch, J. P. (2008). Automated modal parameter estimation by parallel processing within wireless monitoring systems. Journal of Infrastructure Systems, 14(1), 102–113. https://doi.org/10.1061/(ASCE)1076-0342(2008)14:1(102)