A Novel Hybrid Wavelet-Deep Learning Framework for Advanced Structural Damage Detection

Oumayma Najem[1], Mohammed Benbrahim[2], Mohammed Nabil Kabbaj[3], and Jaouad Boumhidi[4]

[1,2,3] *Systems Engineering, Modeling and Analysis Laboratory, Faculty of Sciences, Sidi Mohamed Ben Abdellah University, Fez, Morocco*

oumayma.najem@usmba.ac.ma
mohammed.benbrahim@usmba.ac.ma
n.kabbaj@usmba.ac.ma

[4]*L3IA Laboratory, Faculty of Sciences, Sidi Mohamed Ben Abdellah University, Fez, Morocco*
jaouad.boumhidi@usmba.ac.ma

ABSTRACT

Infrastructures globally are nearing or surpassing their designated lifespans, with recent structural failures serving as a reminder. The aging of infrastructure is a natural eventuality that causes a decline in the structures' mechanical characteristics, consequently affecting their serviceability. The rapid aging of global infrastructure necessitates the development of precise and effective damage detection techniques to preserve public safety. Traditional inspection techniques cannot adequately address modern structural issues, which calls for more sophisticated methods. This study proposes a novel hybrid wavelet-CNN-Transformer framework for structural damage detection that simultaneously extracts localized damage signs and gradual changes in the overall behavior of structural vibrations. Our proposed framework uses the Ben wavelet transform to convert raw acceleration signals into time-frequency representations, which are then processed through a parallel CNN and Transformer branches to extract spatial and temporal features before fusion. We validated this approach on two datasets: the Z24 Bridge dataset and the Qatar University Grandstand Simulator (QUGS) dataset. Our proposed framework achieved 98.85% on the Z24 Bridge dataset and 97.9% on the QUGS dataset, representing a 1.35% improvement over state-of-the-art methods. The proposed framework identifies both the sharp structural discontinuities and the subtle shift in the global behavior of the structure. Furthermore, the model successfully performed multi-class damage classification with 91% accuracy.

https://doi.org/10.36001/IJPHM.2025.v16i2.4568

1. INTRODUCTION

Infrastructures globally are exposed to extreme stressors, whether from increased traffic loads or extreme environmental events, accelerating the natural and unavoidable aging of structures much earlier than expected. Recent studies of structural failures demonstrated that structures all over the world are failing faster than their designed life span, which underlines the need for sophisticated and targeted assessment methods (Garg et al., 2022). Visual inspections, as one of the first structural evaluation techniques, were proven to be insufficient since they rely on the personnel's expertise and can't detect internal defects.

Non-destructive testing (NDT) techniques such as thermography, magnetic waves, and ultrasound were presented as alternatives that go beyond surface-level damage in structures (Schabowicz, 2019). NDTs, although they present several advantages over visual inspections, were also proven to be sensitive to external factors such as environmental conditions. They also remain periodic inspections that lack real-time monitoring capabilities, therefore lack early warning signs.

Structural Health Monitoring (SHM) systems, containing sensor networks, were able to address this issue by enabling real-time, continuous monitoring of structures (Sun et al., 2020). Different types of sensors, depending on the monitored parameter, either mounted on or embedded within a structure, acquire data, which is then processed and interpreted for structural damage detection (Deng et al., 2023). The large amount of data gathered by sensors in different locations of the structure contributes to having accurate results and making knowledgeable decisions about the structural conditions, the additional tests to be conducted, or the corrective measures to be taken (He et al., 2022).

One of the most utilized techniques in SHM is vibration-based approaches, as they can identify internal damage before it visibly appears on the structure (Saidin et al., 2022). This technique builds on the structural definition of damage, stating that damage alters the structural characteristics; therefore, a behavior change indicates the existence of damage in the monitored structure (Rabi et al., 2024). Vibration-based SHM uses signal processing techniques to detect damage-sensitive features from acceleration data. Methods like Fast Fourier Transform (FFT), Short-Time Fourier Transform are traditional methods that are powerful for analyzing stationary signal. However, they present significant challenges when applied to nonlinear and non-stationary signals, which is problematic in real-world applications, given that structures are inherently exposed to dynamic loads (Xu et al., 2022).

Wavelet Transforms (WT) represent the time signals in the frequency domain. It has been one of the most used signal processing techniques in recent years for damage detection of civil structures. Demirlioglu and Erduran (Demirlioglu & Erduran, 2024) demonstrated that wavelet processing identifies damage without prior knowledge of the undamaged state using accelerometers mounted on a vehicle. Among wavelet types, the Mexican Hat wavelet is excellent at detecting localized abrupt changes in the signal. For instance, Tan et al. (Tan et al., 2022) identified damage by analyzing axle acceleration using the Mexican Hat Wavelet. Hester et González (Hester & González, 2017) used the Continuous Wavelet Transform (CWT) with the Mexican hat wavelet to detect localized stiffness loss for a drive-by bridge monitoring method. Another wavelet type is the Morlet wavelet, which is useful for global pattern identification of non-stationary signals. Qiu et al. (Qiu et al., 2024) employed the Morlet wavelet on the KW51 bridge vibration data, accurately localizing damage on the bridge.

The Ben Wavelet is a complex-valued Wavelet Transform that combines the intrinsic properties of both the Mexican Hat wavelet and the Morlet wavelet. Similar to the Morlet wavelet, the Ben wavelet captures both amplitude and phase data, while also remaining sensitive to the signal's continuities like the Mexican Hat wavelet (Benbrahim et al., 2005).

While WTs improve feature extraction, employing them as standalone techniques lacks automation. Hence, supervised deep learning algorithms are increasingly employed for damage identification, localization, and classification. CNNs are deep learning models that process time-frequency representations to automatically extract damage-sensitive features from the data. CNNs, with one or multiple layers, are capable of effectively extracting spatial hierarchies within images and time-frequency representations. The early layers detect simple damage while deeper layers detect complex patterns in the data. Wu et al. (Wu et al., 2022) used wavelet packet decomposition (WPD) for feature extraction in acceleration signals and CNN for damage detection. Nguyen et al. (Nguyen et al., 2024) integrated CWT with CNN to automate structural damage detection. Chen et al. (Chen et al., 2024) combined CWT and 2D-CNN to indirectly detect bridge damage from vehicle vibrations. Najdi et al. (Najdi et al., 2025) used Synchrosqueezing Wavelet Transform (SSWT) with ResNet-based CNN for an effective time-frequency feature extraction. Song et al. (Song et al., 2024) used the Morlet wavelet transform to convert acceleration data into time-frequency scalograms, which are then analyzed with pre-trained CNNs. However, one limitation is that CNN algorithms process data locally, which can affect their overall data interpretation, as they may miss the broader relationships between data points.

Transformers have emerged as an alternative to recurrent models by utilizing self-attention mechanisms to process time sequences of data. Additionally, Transformers process data simultaneously rather than sequentially, which allows them to capture relationships between data regardless of their location or timing (Wan et al., 2023). Fukuoka et Fujiu (Fukuoka & Fujiu, 2023) employ a transformer-based image processing model for bridge damage detection, specifically delamination and rebar exposure. However, their ability to process data as a whole comes at the cost of their inability to detect localized anomalies in data and consequently miss the subtle damage signs in structural data.

To address these issues, we propose in this paper a novel approach combining the Ben wavelet transform, CNN, and Transformers to improve feature extraction and state classification accuracy. The CNN extracts spatial local features from the vibration signals, while the Transformers extract global temporal dependencies, effectively detecting small damage as well as changes that manifest themselves in the structure's global behavior. The features extracted from both branches are fused through a concatenation layer, followed by a fully connected layer for final classification. Our experimental results, validated on two renowned datasets, the Z-24 bridge dataset and the Qatar University Grandstand Simulator (QUGS) dataset, demonstrate that this hybrid approach outperforms both traditional methods and state-of-the-art deep learning techniques for structural health monitoring.

Despite these advances, the existing vibration-based methods address either the detection of localized damage features, like sharp discontinuities, or the global behavioral change of the structure, like the gradual loss of stiffness, but not both simultaneously. CNN-based approaches capture spatial anomalies in the time-frequency representations, but miss the temporal dependencies. In contrast, the Transformer-based methods capture global behavioral shifts but lack the spatial resolution needed for damage detection. Real-world structural damage manifests simultaneously through both mechanisms. Existing methods address these separately or sequentially.

To adress this gap, we propose a novel hybrid CNN-Transformer framework that processes wavelet-transformed vibration signals through parallel branches, to extract both the local spatial anomalies and the global temporal patterns, providing a comprehensive dual-scale damage detection.

This research paper will contribute the following:

- **Advanced feature extraction:** using the Ben wavelet transform to extract features from vibration signals to address the limitations of standard wavelet transforms.
- **Novel damage detection:** Unlike traditional approaches that rely solely on one method, we introduce a parallel processing structure that combines the strengths of CNNs with Transformers. We use CNNs and Transformers to capture both localized disruptions in the time-frequency representation of vibration signals and global effects that damage has on the overall behavior of the structure through changes in the initial vibration signals.

The content of this paper is organized as follows: Section 2 is dedicated to the theoretical background of the proposed hybrid method. Section 3 presents the real-world validation and its results through two bridge vibration datasets. Finally, section 4 concludes the article and presents the limitations.

2. METHODOLOGY

This section presents our novel local-global detection framework, which combines wavelet transforms with a hybrid deep learning architecture. Figure 1 illustrates the complete methodology, starting with raw vibration signals from civil structures and ending with damage classification.

Our proposed framework addresses a current limitation in the literature approaches of structural health monitoring, which is the simultaneous detection of localized damage signs and the changes in the global behavior of the structure's vibrations. This multi-scale detection of damage is not fully captured by traditional single-scale methods, which are only adept at either subtle or major damage detection.

2.1 Wavelet Transform

We chose to apply the Ben wavelet, originally developed by Benbrahim et al. (Benbrahim et al., 2005) for seismic signal classification, to civil structural health monitoring applications. Wavelet Transforms extract features from non-stationary signals that are missed when Fourier Transforms are used (Kim & Melhem, 2004). This Ben wavelet combines Mexican Hat localization properties with the Morlet phase preservation capabilities, both needed for sharp discontinuities like cracks and bolt loosening, and gradual changes in structural behavior as stiffness loss and energy redistribution. The Ben wavelet is an asymptotically admissible wavelet defined as (Benbrahim et al., 2005), where ω_0 is the central frequency of the wavelet.:

$$\psi_{ben}(t) = \frac{2}{\sqrt{3}}\,\pi^{-\frac{1}{4}}(1-t^2)e^{i\omega_0 t}e^{-t^2/2} \tag{1}$$

The Ben Wavelet combines properties from two classic wavelets, the Mexican Hat Wavelet (Mallat, 1989) and the Morlet Wavelet (Goupillaud et al., 1984). Each component of the wavelet has a specific functional purpose. The term $\frac{2}{\sqrt{3}}\,\pi^{-\frac{1}{4}}$ is a normalization constant that ensures the unit energy. The term $(1-t^2)$ offers localized sensitivity, similarly to the Mexican Hat wavelet, and $e^{i\omega_0 t}$, inherited from the Morlet Wavelet, provides amplitude and phase information preservation. Finally, the Gaussian envelope $e^{-t^2/2}$ ensures a minimal Time-Bandwidth Product as per the Heisenberg-Gabor limit to balance the time localization and frequency resolution (Papoulis, 1977).

The Ben wavelet satisfies the admissibility condition required for continuous wavelet transform analysis. The Ben wavelet possesses zero mean, $\int \psi_{ben}(t)dt \rightarrow 0$as $\omega_0 \rightarrow \infty$, and $\psi_{ben}(t)dt \in L^2(\mathbb{R}) \cap L^1(\mathbb{R})$, ensuring $C_\psi = \int \frac{|FT_{\psi(\omega)}|^2}{|\omega|d\omega} < +\infty$ (Benbrahim et al., 2005).

The Ben wavelet is not symmetric overall, and that is due to its complex term $e^{i\omega_0 t}$. This asymmetry is essential for phase analysis and detecting directional frequency shifts. However, the real part of the Ben Wavelet remains symmetric to maintain a good temporal localization. Additionally, for a balanced oscillatory behavior with time localization, the admissibility condition for the central frequency is $\omega_0 \geq 7$.

This hybrid formulation allows the detection of both abrupt and gradual changes. The Continuous Wavelet Transform with the Ben wavelet generates scalograms that present the energy redistribution signs indicative of structural damage:

$$SCAL(a,t) = \left|CWT_x^{\psi}(a,t)\right|^2 \tag{2}$$

Where damaged structures display energy distribution across a broad frequency band compared to the concentrated energy signatures of undamaged states.

2.2 Hybrid CNN-Transformer Architecture

2.2.1. Architecture

Our hybrid architecture addresses the limitations of CNNs and Transformers for structural damage detection. CNNs detect frequency discontinuities in scalograms using hierarchical feature extraction; however, they are unable to interpret these structural changes on the global scale. Transformers can extract long-range temporal correlations through self-attention mechanisms; however, they lack spatial resolution for sudden time-frequency changes. The purpose of our parallel processing design is to extract both feature types for a complete picture of the damage.

2.2.2. CNN Branch: Parallel local damage feature extraction

The CNN branch works in parallel to capture spatial features from the wavelet's resulting scalograms (Lecun et al., 1998). The CNN architecture uses four progressive convolutional layers (16, 32, 64, 128) and max pooling to reduce dimensionality while retaining key features.

To extract local patterns, each convolutional layer uses learned filters. We also employed ReLU activation to maintain training stability.

$$S'_{s',t',f} = \sum_{i=1}^{k_S} \sum_{j=1}^{k_T} W_{i,j,f} \cdot S_{s'+i-1,t'+j-1} \tag{3}$$

This branch is especially employed for spatial damage signatures extraction, such as shifts and spectral anomalies, while the Transformer branch processes structure-wide damage patterns simultaneously.

2.2.3. Transformer Branch: Parallel Global Feature Extraction

The Transformer branch processes reshaped scalograms to extract global temporal dependencies (Vaswani et al., 2017), while operating in parallel with the CNN Branch. We first reshaped the input scalograms from 2D time-frequency representations to sequential tokens to process them as time series.
The Transformer architecture uses a multi-head self-attention mechanism with 8 attention heads and a key dimension d_k=128, to capture long-range patterns in structural vibrations:

$$Attention(Q,K,V) = softmax\left(\frac{QK^T}{\sqrt{d_k}}\right)V \tag{4}$$

This configuration computes attention weights across the entire time-frequency representation to facilitate the detection of gradual energy shifts that happen over long periods. This parallel design captures long-range dependencies manifesting as gradual energy shifts throughout the entire time-frequency domain. We employed Layer Normalization (Ba et al., 2016) and dropout (Salehin & Kang, 2023) to maintain

training stability, while Global Average Pooling aggregates aggregates temporal features into vectors for fusion with spatial features extracted by the parallel CNN branch.

2.2.4. Hyperparameter ablation study

We optimized the number of layers by testing various configurations to identify the optimal performance configuration, as shown in table 1. For the CNN branch, we tested multiple combinations of CNN architectures, from 2 to 5 convolutional layers, paired with different Transformer configurations, 1 and 2 layers, with varying attention heads and key dimensions. The 4-layer configuration (16-32-64-128 filters) with 1 layer, 8 head, key dimension 128, reached the highest accuracy of 98.85%.

Table 1 Hyperparameter ablation study results

CNN Configuration	Transformer Configuration	Accuracy
2 layers [16,32]	1 layer, 4 heads, key=64	92.19%
3 layers [16,32,64]	1 layer, 4 heads, key=64	96.09%
4 layers [16,32,64,128]	1 layer, 4 heads, key=64	98.16%
4 layers [16,32,64,128]	1 layer, 4 heads, key=128	98.66%
4 layers [16,32,64,128]	1 layer, 8 heads, key=64	98.77%
4 layers [16,32,64,128]	1 layer, 8 heads, key=128	98.85%
4 layers [16,32,64,128]	2 layer, 8 heads, key=128	97.55%
5 layers [16,32,64,64,128]	2 layer, 8 heads, key=128	92.19%

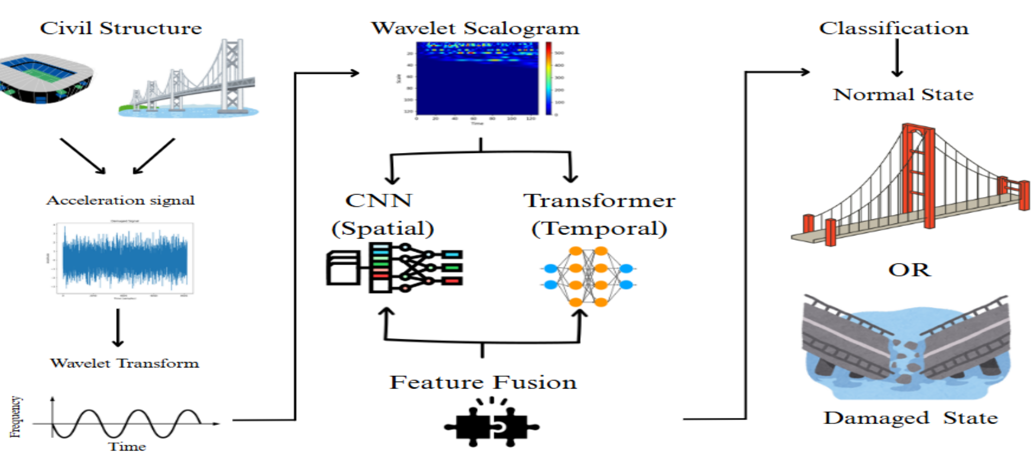


Figure 1 The proposed hybrid CNN-Transformer model

2.2.5. Feature Fusion

Feature fusion combines outputs from both the CNN and Transformer branches to create a complete and complementary representation. We employed concatenation fusion to preserve the features extracted from both branches while avoiding information loss. The fusion layer directly concatenated the feature vectors extracted from both branches:

$$Z_{concat} = \text{Concatenate}\,(Z_{CNN}, Z_{Trans}) \tag{5}$$

Where the Z_{CNN} are the spatial features extracted by the CNN branch, and the Z_{Trans} are the temporal features extracted by the Transformer branch. The CNN branch outputs feature vectors that encode local spatial patterns and frequency domain signatures, while the Transformer branch outputs temporal dependencies and long-range correlations, resulting in a fused feature vector that contains spatial and temporal damage signs.

We selected Concatenation fusion over alternative feature techniques, such as element-wise addition or weighted fusion, because it keeps the full information content of both branches. This ensures that the spatial features extracted by the CNN and the contextual changes extracted by the Transformers contribute equally to the classification decision. We processed the concatenation features through a fully connected layer with dropout regularization (Salehin & Kang, 2023) to allow binary classification between damaged and undamaged structural states. Figure 2 details the hybrid parallel CNN-Transformer process.

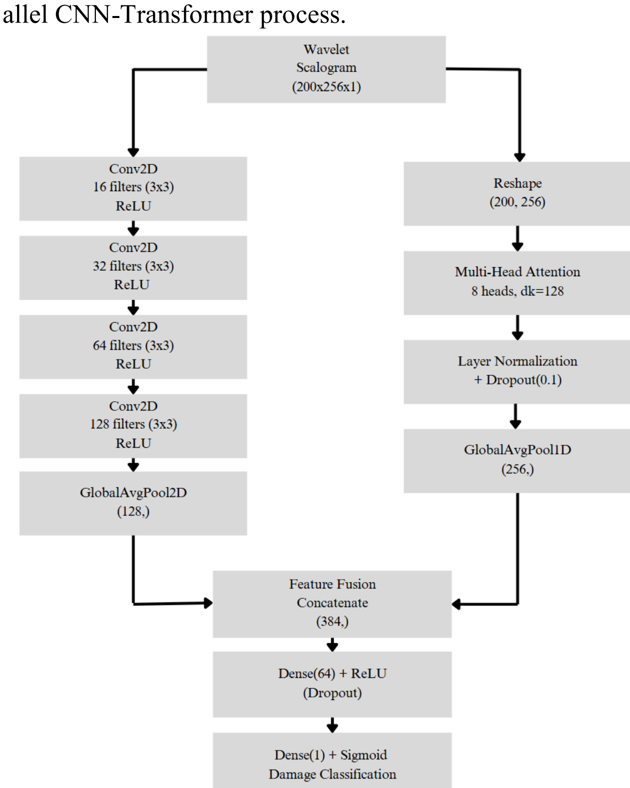


Figure 2 Hybrid CNN-Transformer architecture for structural damage detection.

2.3. Performance evaluation

2.3.1. Classification metrics

Hybrid CNN-Transformer architecture for structural damage detection.

Accuracy measures the correct predictions by the model:

$$\text{Accuracy} = \frac{TP+TN}{TP+TN+FP+FN} \tag{6}$$

Precision calculates the proportion of the correctly predicted damage cases among all predicted positive cases:

$$\text{Precision} = \frac{TP}{TP+FP} \tag{7}$$

Recall calculates how many true damage cases are identified:

$$\text{Recall} = \frac{TP}{TP+FN} \tag{8}$$

F1-score is a metric that accounts for both precision and recall for error quantification.

$$\text{F1-score} = 2 \times \frac{Precision \times Recall}{Precision+Recall} \tag{9}$$

Where, TP and TN are the true positives and true negatives, respectively. And, FP and FN are the false positives, and false negatives, respectively.

2.3.2. Feature Space Visualization

We employed t-distributed Stochastic Neighbor Embedding (t-SNE) (Kobak & Berens, 2019) to validate the discriminative capacities of the learned features extracted by our parallel processing architecture. t-SNE maps the high-dimensional learned features to a 2D space to visually validate the class separability between the damaged and undamaged states. This constitutes a qualitative validation of our parallel architecture's capability of creating separate clusters for different structural states.

3. RESULTS

We demonstrated the effectiveness and advantages of the proposed method for bridge damage detection on two benchmark datasets: the Z24 Bridge dataset and the Qatar University Grandstand Simulator (QUGS) dataset. We employed the Z24 dataset, consisting of progressive pier settlement scenarios, to evaluate performance against several established techniques. Additionally, the QUGS dataset, providing comprehensive data on steel frame joint loosening, was used to further validate the method's detection capabilities across different structural materials and damage mechanisms.

3.1. Dataset 1: Z24 Bridge Dataset

3.1.1. Experimental Setup

The Z24 bridge was a post-tensioned concrete bridge with a two-cell box-girder cross-section, located on the A1 highway between Bern and Zürich in Switzerland (Maeck & De Roeck, 2003). The monitoring was conducted as part of the European Brite EuRam research project BE-3157, titled 'System Identification to Monitor Civil Engineering Structures' (SIMCES). Figure 3 shows the Z24 Bridge structure.

16 accelerometers were mounted on the bridge to measure accelerations in multiple directions. Data was acquired at a sampling frequency of 100 Hz with an anti-aliasing filter set to a cutoff frequency of 30 Hz. One month before the bridge demolition, progressive damage was collected under multiple damage scenarios such as pier settlements, foundation tilt, and concrete spalling. Progressive damage was introduced through pier settlement scenarios, starting with 20 mm to 95 mm lowering. Table 1 summarizes the damage scenarios used in this study.

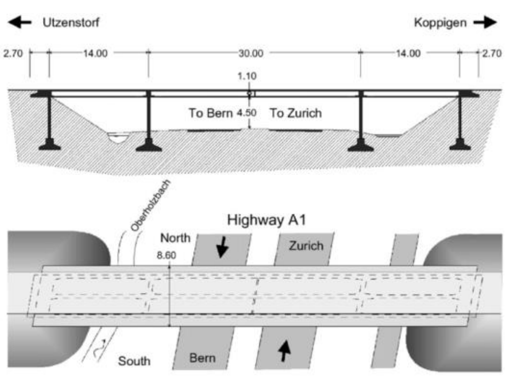


Figure 3 (Maeck & De Roeck, 2003)Z24 Bridge structure

Table 2 The Z24 bridge Pier settlement damage scenarios and their dates.

Date (1998)	Class label	Damage Scenario	Settlement (mm)
4 August	0	Undamaged bridge	0
10 August	1	Pier Lowering	20
12 August	2	Pier Lowering	40
17 August	3	Pier Lowering	80
18 August	4	Pier Lowering	95

3.1.2. Hybrid Model Architecture Configuration

The implemented Hybrid CNN-Transformer model follows a parallel-branch architecture that detects both local and global characteristics. Table 2 shows the hybrid model's implementation details.

Table 3 Hybrid model implementation details

Component	Configuration
Input Layer	Scalogram (S × T × 1)
CNN Branch (Local Feature Extraction)	
Conv2D-1	16 filters, (3,3), ReLU, same padding
MaxPool2D-1	(2,2) pooling
Conv2D-2	32 filters, (3,3), ReLU, same padding
MaxPool2D-2	(2,2) pooling
Conv2D-3	64 filters, (3,3), ReLU, same padding
MaxPool2D-3	(2,2) pooling
Conv2D-4	128 filters, (3,3), ReLU, same padding
MaxPool2D-4	(2,2) pooling
GlobalAvgPool2D	Average pooling
Transformer Branch (Global Feature Extraction)	
Reshape	(height, width × channels)
MultiHeadAttention	8 heads, key_dim=128
Dropout	rate=0.1
LayerNormalization	epsilon=1e-6
GlobalAvgPool1D	Average pooling
Local-Global Fusion & Classification	
Concatenate	CNN + Transformer features
Dense	64 neurons, ReLU, L2
Dropout	rate=0.3
Dense (Output)	1 neuron, sigmoid

We trained the model using binary cross-entropy as the loss function with Adam Optimizer at a learning rate of 0.001. Additionally, we used a batch size of 16 over 100 epochs and included early stopping with a patience of 3 epochs to track the validation loss to prevent overfitting. The dataset is split using stratified sampling, with 80% assigned for training and 20% for testing for balanced class representation across both sets.

3.1.3. Results

(i) Scalogram Visualization and Feature Analysis

The Ben wavelet transform efficiently extracts damage signatures in the time-frequency domain. Figure 4 shows a clear distinction between undamaged and damaged states. The pier settlement produces a redistribution of energy across frequency bands, with damaged conditions (Fig. 4d) showing a spread energy pattern when compared to the concentrated signatures of the undamaged state of the structures (Fig. 4c). The implementation generates scalograms of 200 scales × 256 time points, capturing both high-frequency transient events and low-frequency structural changes, creating **complementary** input for parallel CNN and Transformer processing.

(ii) Feature Space

We applied t-SNE dimensionality reduction to visualize the learned feature representations. Figure 5 shows a clear class separation in the 2D feature space between undamaged (Class 0, blue) and damaged (Class 1, green) structural states across all pier settlement scenarios (20mm, 40mm, 80mm, 95mm), forming distinct clusters. This clear separation demonstrated the hybrid CNN-Transformer architecture's ability to correctly learn discriminative features for the detection of structural damage.

(iii) Training performance

The hybrid model shows strong training characteristics, achieving convergence within 15-25 epochs while training accuracies reach 99% and validation accuracies are maintained above 98%. The early stopping mechanism activates at epoch 22±3 across damage scenarios which prevents overfitting and confirms effective parameter learning. Figure 6 shows the training and validation curves.

(iv) Detection performance results

Table 4 shows a stable performance of the proposed hybrid method across all settlement scenarios of the settlement magnitude. The Precision-recall balance (98.38% vs 98.32%) indicate low false positive/negative rates. The 80mm settlement case shows a slightly lower precision (96.63%) with high recall (98.85%), which suggests the models sensitivity to changes in intermediate damage scenarios.

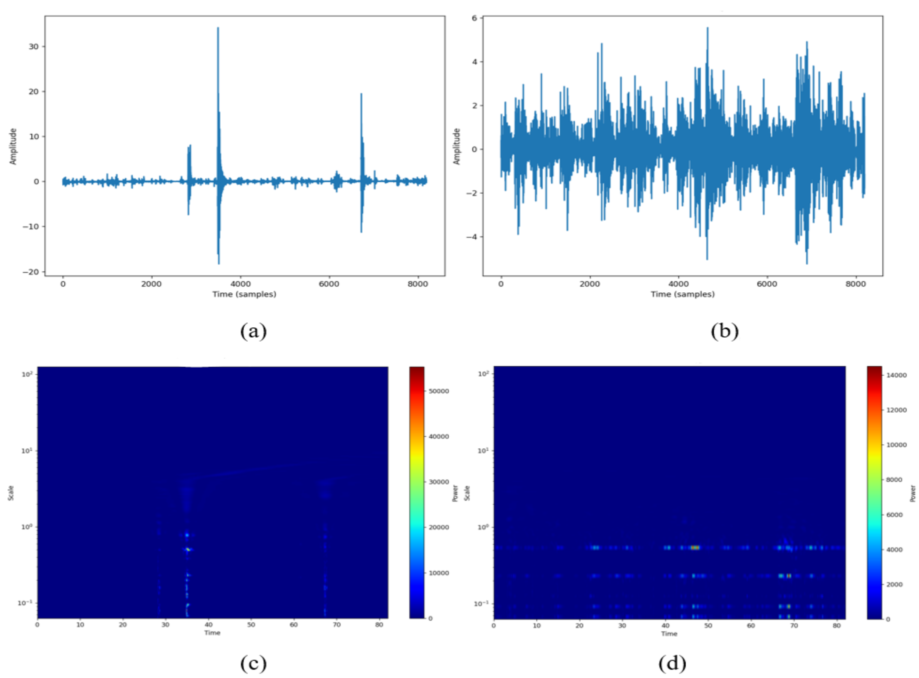


Figure 4 Ben wavelet scalogram comparison showing damage detection capability. (a) Undamaged case acceleration signal showing normal structural response, (b) Damaged signal (80mm pier settlement) showing increased amplitude and altered vibration characteristics, (c) Baseline scalogram with concentrated energy distribution, (d) Damaged scalogram revealing clear energy redistribution and additional frequency components. This representative comparison demonstrates the method's ability to distinguish between undamaged and damaged structural states.

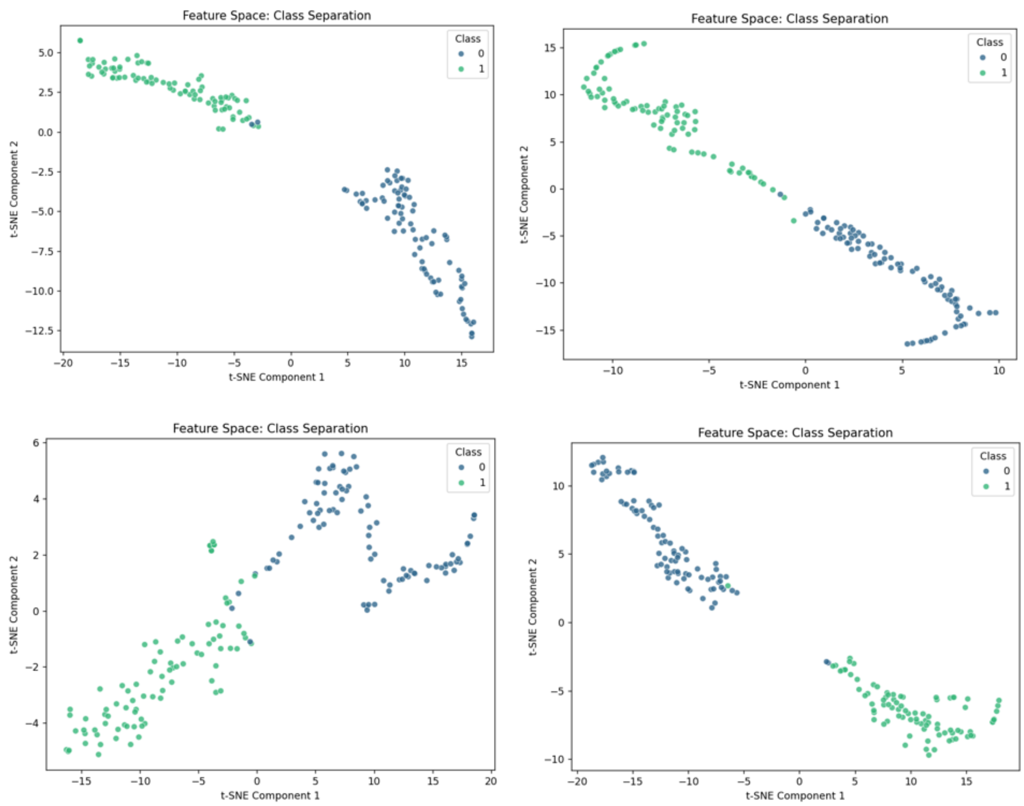
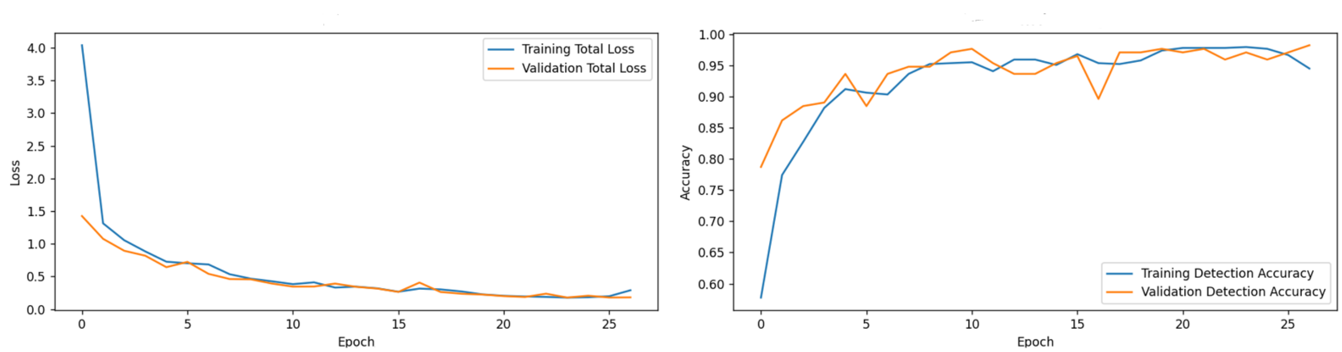


Figure 5 t-SNE visualization of learned feature representations from the hybrid CNN-Transformer model on the Z24 Bridge dataset.



(a) (b)

Figure 6 (a) Training and validation loss convergence over epochs, (b) Training and validation accuracy progression over epochs.

Table 4 Z24 Bridge Implementation Results.

Damage scenario	Accuracy	Precision	Recall	F1-Score
20 mm Settlement	98.85%	98.85%	98.85%	98.85%
40 mm Settlement	98.28%	98.84%	97.70%	98.27%
80 mm Settlement	97.70%	96.63%	98.85%	97.73%
95 mm Settlement	98.28%	97.73%	98.85%	98.29%
Mean ± Std	98.35 ± 0.36%	98.38 ± 0.44%	98.32 ± 0.42%	98.35 ± 0.36%

3.1.4. Comparison and analysis of detection methods.

To evaluate the performance of our proposed hybrid CNN-Transformer method within the existing SHM techniques. We reviewed state-of-the-art methods applied to the Z-24 bridge dataset to underline our proposed method's superiority. The experimental results in Table 4 summarize the performance of the nine detection methods on the Z-24 bridge dataset. We chose to compare our method to the following papers:

- Santaniello et Russo (Santaniello & Russo, 2023) convert 1D acceleration signals into time-frequency images with synchrosqueezing continuous wavelet transform (SCWT). The authors evaluate several CNN architectures including ResNet50, MobileNet v1, and DenseNet121, and propose two refinement techniques: image-splitting and signal-splitting.
- Sony et al. (Sony et al., 2022b) proposes a windowed Long Short-Term Memory (LSTM) network method for vibration-based multiclass damage detection and localization in civil structures.
- Sony et al. (Sony et al., 2022a) proposes an optimally-tuned windowed 1D CNN approach for multiclass damage identification using vibration responses

Table 4 and Figure 7 demonstrate our hybrid model's outperformance within the literature. Our hybrid model achieves an accuracy improvement of 1.35% over signal-splitting ResNet50, representing a 36% reduction in error rate (from 2.5% to 1.15%). Deep learning methods' accuracies range from 95-97%, while traditional approaches vary from 72% to 94%. This performance gap defines three method categories: our hybrid method at 98.85% accuracy, advanced deep learning from 95-97%, and convolutional techniques 72-94%.

3.1.5. Architecture Ablation Study

To further validate our hybrid CNN-Transformer architecture, we performed ablation studies by comparing each component against the complete hybrid model. For the 20mm pier settlement damage scenario, we used the Ben wavelet transform to test three different configurations. The Transformer-only branch used Transformer components without the CNN components, whereas the CNN-only branch used CNN feature extraction without the Transformer components. The hybrid CNN-Transformer configuration represents our parallel processing architecture.
Figure 8 shows the performance comparison across the three architectural configurations. The ablation results show how

Table 5 Comparison with State-of-art Methods' Results

Method	Accuracy	Precision	Recall	F1-score	Reference
Proposed hybrid method	98.85%	98.85%	98.85%	98.85%	Our
Signal-splitting (SST with ResNet50)	97.5%	97.77%	97.34%	97.51%	(Santaniello & Russo, 2023)
Image-splitting (SST with ResNet50)	97.47%	97.39%	97.17%	97.27%	(Santaniello & Russo, 2023)
ResNet50 (with SST)	97.08%	97.22%	97.22%	97.22%	(Santaniello & Russo, 2023)
MobileNet v1 with SST	95.36%	-	-	-	(Santaniello & Russo, 2023)
LSTM	94%	95%	94%	94%	(Sony et al., 2022b)
DenseNet121 with SST	90.21%	-	-	-	(Santaniello & Russo, 2023)
1DCNN	83%	83%	83%	83%	(Sony et al., 2022a)
MLP	72.05%	73.17%	71.31%	71.42%	(Santaniello & Russo, 2023)

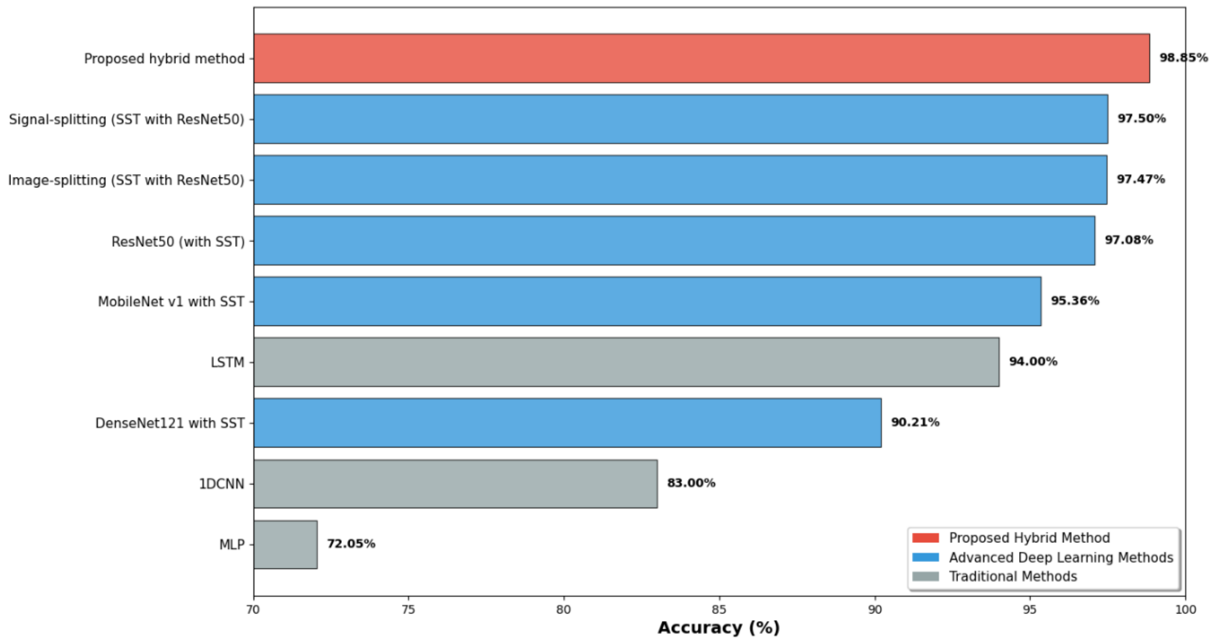


Figure 7 Performance comparison of the proposed hybrid method against traditional and advanced deep learning approaches.

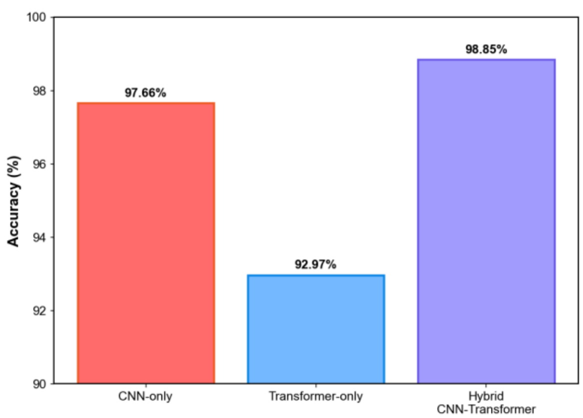


Figure 8 Architecture Ablation Study Results (20mm damage case)

the CNN and Transformer components are complementary. The Transformer-only method shows lower accuracy (92.97%) as it focuses on capturing global patterns, while the CNN-only method achieved good performance (97.66%) by capturing local damage features. The hybrid method performs better than both of its separate parts do, which demonstrates that the parallel processing of global and local features show in higher damage detection accuracy.

3.1.6. Wavelet Transform Comparison Study.

We compared three wavelets – the Morlet, Mexican Hat, and Cauchy – to justify the choice of the Ben Wavelet transform for time-frequency feature extraction. The Hybrid CNN-Transformer was applied to the 20mm settling case in all experiments.

Figure 9 presents the performance comparison across different wavelet transforms. The Ben wavelet achieves the highest accuracy value of 98.85%, surpassing the Cauchy (96.09%), Morlet (94.53%), and Mexican Hat (82.81%) wavelets. The Ben wavelet's effectiveness comes from its time-frequency localization features, which identify both localized and global damage features. The performance gaps between the different wavelets confirm that the wavelet choice influences the damage detection accuracy.

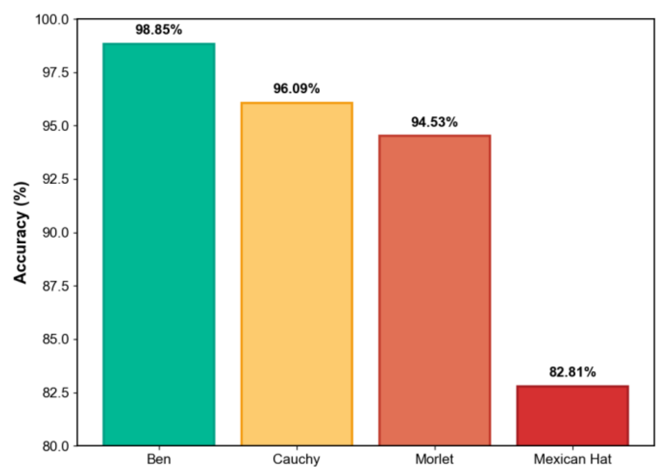


Figure 9 Wavelet Transform Comparison Results (20mm damage case)

3.1.7. Multi-class Damage Identification Results

Beyond the model's ability of binary classification of each damage level independently, we evaluated the hybrid model's capabilities for damage assessment through a multi-class classification task to identify damage severity levels. We tested the model on the task of simultaneously classifying all five damage scenarios of the Z24 bridge: Undamaged, 20 mm settlement, 40mm settlement, 80mm settlement, and 95mm settlement.

The model architecture remained unchanged except for the final output layer, which we modified from binary classification (sigmoid activation, 1 neuron) to multi-class classification (softmax activation, 5 neurons). We changed the function from binary cross-entropy to sparse categorical cross-entropy to accommodate integer class labels. We utilized training the same hyperparameters as binary classification: Adam optimizer (learning rate 0.001), batch size 16, early stopping with patience 10.

The evaluation of the multi-class performance achieved an overall accuracy of 91%, demonstrated the model's ability to not only detect damage presence but also assess the damage severity levels. As expected, the model shows systematic confusion between the adjacent damage levels. While per-scenario binary validation of the model achieved 98.85% as the best accuracy, this 5-class damage classification model shows the model's practical capability of identifying damage states without prior knowledge of the severity level. Table 6 reveals the detailed performance metrics.

Table 6 Multi-class Damage Identification Results

damage scenario	accuracy	precision	recall	F1-score
Undamaged	93.8%	95%	94%	94%
20mm Settlement	90.1%	85%	90%	87%
40mm Settlement	88.9%	87%	89%	88%
80mm Settlement	86.6%	96%	87%	91%
95mm Settlement	95.1%	93%	95%	94%

The confusion matrix, shown in figure 10, reveals expected patterns, most misclassifications occur between adjacent damage levels. The undamaged state and the 95mm settlement are the most distinct structural damage states with clear feature separation, therefore reaching the highest accuracies, 93.8% and 95.1% respectively.

The intermediate damage levels achieved lower accuracies; 80mm settlement achieved 86.6% accuracy as the model occasionally confused it with 40mm or 95mm. additionally, the high precision of 96% and low recall of 87% indicate that other classes are rarely misidentified as the 80mm settlement, preventing false alarms while maintaining an overall the detection capability.

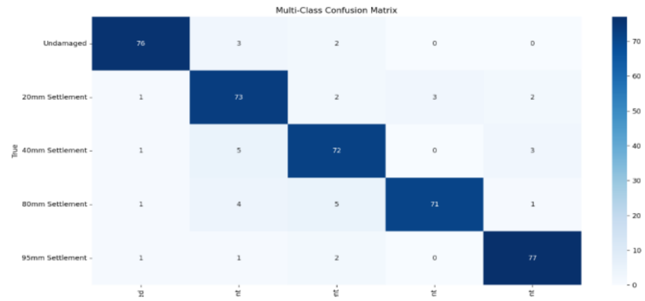


Figure 10 Multi-class confusion matrix

3.2. Dataset 2: Qatar University Grandstand Simulator (QUGS)

3.2.1. Experimental Setup and Data Preparation

The Qatar University Grandstand Simulator (QUGS) (Avci et al., 2022) is a laboratory-scale steel structure designed to simulate the behavior of modern stadium seating facilities (M.ASCE, 2018), (Abdeljaber et al., 2017). The dataset was developed to verify structural health monitoring techniques in a controlled environment before applying to real-life stadiums.

The structure has of a hot-rolled steel frame with footprint dimensions of 4.2m×4.2m, designed to carry 30 spectators. The steel frame has 8 girders, each 4.6m long, and 25 filler beams supported on 4 columns. The damage mechanism is based on loosening joints in the truss structure at specific locations, providing both undamaged and damaged state data. Figure 11 shows the QUGS Structure and Joint Locations.

30 accelerometers were place on the main girders at the 30 joints of the steel structure. Vibration signals from each joint location are processed similarly to the Z24 dataset. Signals are normalized, transformed using the spatial and temporal Ben Wavelet Transform into scalograms, and fed into the hybrid CNN-Transformer model for binary classification of joint damage states.

The QUGS dataset provides two independent datasets, collected in separate experimental runs to ensure that the model is evaluated on unseen data:

- Dataset A: For training and validation.
- Dataset B: for testing.

A random split is created with Dataset A to create a validation set for hyperparameter tuning and early stopping.

(i) Implementation results

Figure 12 shows high performance accuracies around 94-99% across the majority of joints, with joint 29 presenting a considerable performance outlier at 94.70%. The validation results show an average accuracy of 97.9% and a median of 98%, with most joints exhibiting high accuracy values, while a few others show lower values. Table 5 shows the details of this performance variation, revealing that the best performing near a perfect score above 98.9%, while challenging joints show balanced precision-recall values even with low accuracies.

3.2.2. Comparison analysis

To validate our approach, we compared our hybrid CNN-Transformer model with leading methods from the literature applied to the same QUGS dataset:

- Kuo and Lee, 2023 (Kuo & Lee, 2023) uses signals differences and a 1D fusion CNN (1D-FCNND) for structural damage detection.
- Truong et al., 2022 (Truong et al., 2022) uses a hybrid framework of 1D-CNN and GRU to learn spatial and temporal relationships from structural vibration signals.

Figure 13 shows validation results on joints 1-5. On this subset, our method achieves 98.2% of average accuracy compared to 96.7% for 1D-FCNND, which shows our method's superiority.

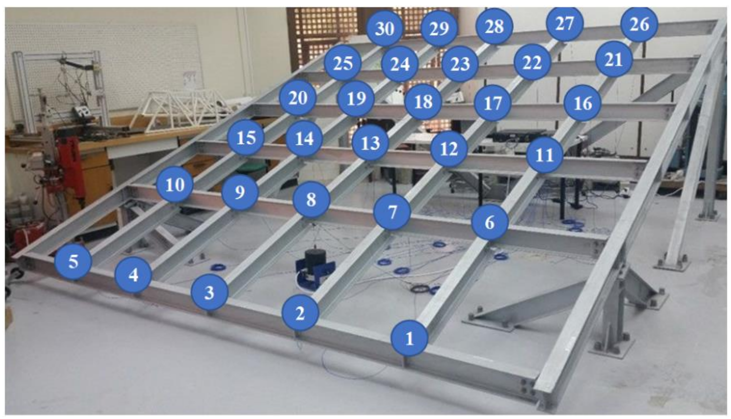


Figure 11 (Avci et al., 2022) QUGS Structure and Joint Locations

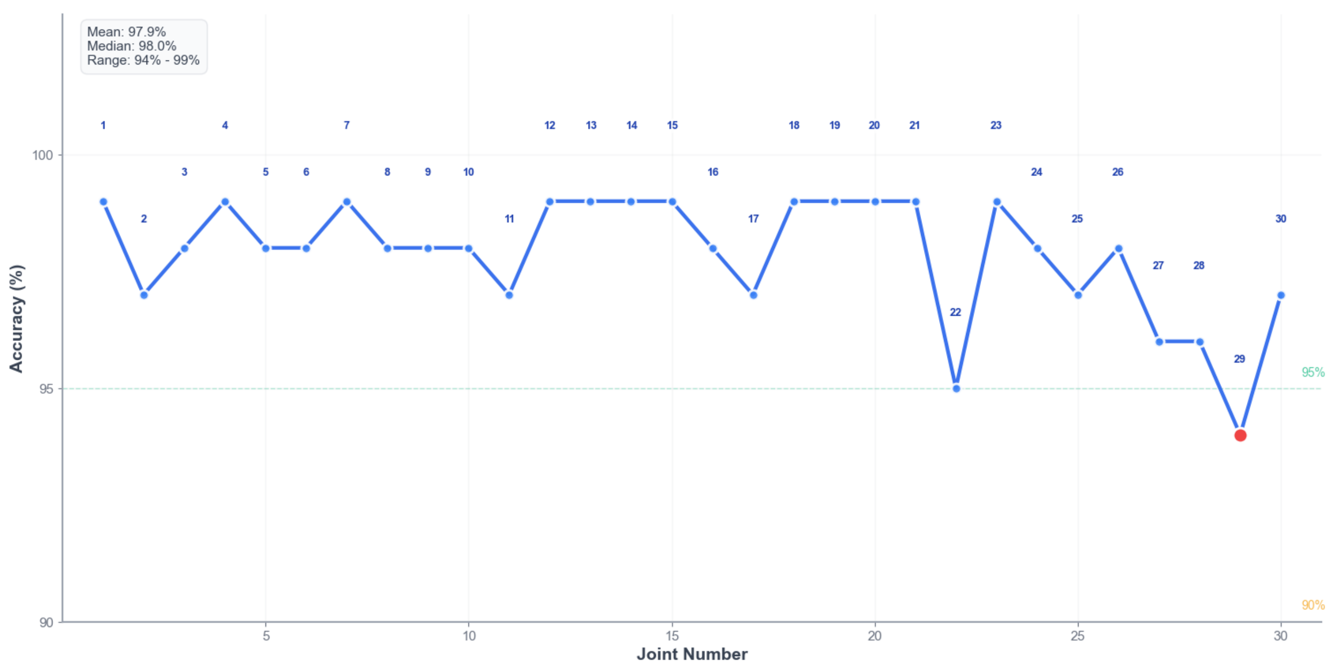


Figure 12 Detection accuracy across all 30 QUGS joint damage cases.

Table 7 Performance metrics for representative QUGS joints across different accuracy tiers.

Damage scenario	Accuracy	Precision	Recall	F1-Score
Joint 04	99.74%	100%	99.47%	99.73%
Joint 13	99.47%	99.47%	99.47%	99.47%
Joint 24	98.94%	99.47%	98.94%	98.94%
Joint 02	96.56%	96.32%	96.83%	96.57%
Joint 11	97.62%	97.37%	97.88%	97.63%
Joint 29	94.72%	95.00%	94.75%	94.87%

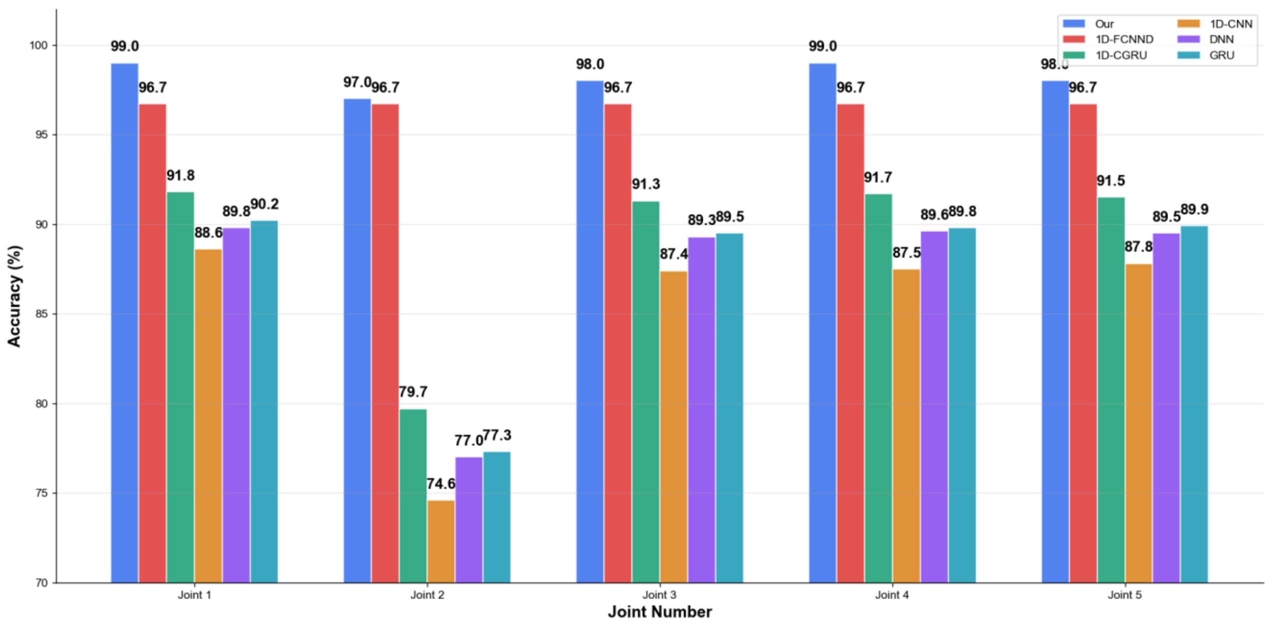


Figure. 13 Method comparison for Joint 1-5

4. DISCUSSION

Experimental results on two structural datasets demonstrate that the local-global hybrid framework effectively addresses the key challenge of detecting damage at multiple scales in Structural Health Monitoring (SHM). The wavelet scalograms, through time-frequency representations, reveal that damaged scenarios exhibit a broad frequency distribution, requiring both local analysis with CNN and global analysis with Transformer to accurately identify damage.

The ablation study shows that the CNN-only processing has an accuracy of 97.66% as it captures localized damage but neglects the global behavioral effects, whereas the Transformer-only processing has an accuracy of 92.97% as it lacks spatial resolution for abrupt changes, focusing on temporal patterns. Combining both capabilities through our hybrid model results in an accuracy of 98.85%.

The reduced precision of 96.63% of the 80mm pier settlement case reflects the challenges in structural state classification caused by the transitional energy patterns of the intermediate damage scenario. Similarly, QUGS joint 29's accuracy outlier (94.72%) shows variation across joint locations, highlighting the need for location-specific analysis in SHM to optimize detection strategies.

The proposed framework performs consistently across different damage mechanisms – pier settlement, joint loosening, and materials – concrete and steel, demonstrating its strong adaptability. The scalogram-based input representations also prove to be optimal for this dual analysis, as they offer multi-resolution time-frequency information, which allows the CNN and Transformer components to detect both localized changes and gradual shifts accurately.

From a computational perspective, the hybrid CNN-Transformer architecture is justified by both theoretical considerations and experimental results. All experiments were conducted on NVIDIA Tesla T4 GPU with 16GB VRAM. The parallel architecture processes both branches simultaneously on GPU hardware, resulting in training times slower than individual branches, however faster than the sequential training of the models. Additionally, it shows an improvement of 1.19% over the faster individual model, which is a 44% error reduction rate, making the time-accuracy trade-off advantageous.

More importantly, the proposed architecture's complexity addresses the critical limitation in structural damage detection, namely the necessity to detect both the localized discontinuities, such as cracks and bolt loosening, and gradual overall changes, such as stiffness loss. Our ablation study in Section 3.1.7, shows the hybrid model shows better performance than its individual branches (98.85%), while the CNN-only model reaches 97.66% by detecting only local shifts, while Transformer-only achieves 92.97% by capturing global behavioral shifts.

The complete computational pipeline consists of two distinct phases: the scalogram generation and the deep learning model. The scalogram generation step constitutes the dominant computational cost step. Despite this computational bottleneck, the Ben wavelet's feature extraction capability of combining localized sensitivity with phase preservation results in 4-16% accuracy improvement over alternative wavelets (Figure 9), justifying the additional cost.

Table 8 Computational performance analysis

method	dataset	Training time (in min)	Inference time per sample (average in ms)
CNN-only	Z24	1.5	40
Trans-only		2	60
hybrid		3	70
CNN-only	QUGS	0.2	8
Trans-only		0.1	10
hybrid		0.4	15

Despite the method's advantages, several limitations should be acknowledged. First, external factors such as environmental conditions, changing operational loads, and measurement noise were not addressed while validating this method, which would affect the performance in practical applications. Second, the method requires large training data, which limits its use on new or uncommon structures unless combined with transfer learning approaches. Third, the computational demands of the hybrid CNN-Transformer model are reasonable for offline analysis; however, they could present difficulties for real-time monitoring performance if not optimized further with strategies such as edge computing implementations.

5. CONCLUSION

We propose a novel multi-scale damage detection framework that enhances vibration-based structural monitoring by simultaneously detecting transient events and shifts in the overall structural behavior. The proposed hybrid CNN-Transformer achieved superior performance over many state-of-the-art methods, with an accuracy of 98.85% on the Z24 Bridge dataset and 97.9% on the QUGS dataset. Furthermore, the method achieved 91% accuracy for multi-class damage classification, proving its capability beyond binary damage detection. The results show that combining the spatial feature extraction with temporal pattern identification provides complementary insights into the structural damage state.

In this article, our key contributions include parallel processing for the simultaneous damage detection at the local and global scales. Additionally, the Ben wavelet combines time-frequency localization and phase preservation for optimal feature extraction from structural vibration signals. Lastly, hybrid architectures are used to learn damage characteristics, comprehensively capturing both small-scale and broad-scale damage features.

Future work should address environmental conditions, such as temperature variations, and optimize real-time processing for practical applications of the proposed method. This research contributes to the advancement of intelligent structural monitoring for a continuous assessment of structural health for safe and reliable structures.

ACKNOWLEDGMENT

This work is carried out with the support of the CNRST as part of the "PhD-Associate Scholarship- PASS" Program

REFERENCES

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.

Avci, O., Abdeljaber, O., Kiranyaz, S., Hussein, M., Gabbouj, M., & Inman, D. (2022). A New Benchmark Problem for Structural Damage Detection: Bolt Loosening Tests on a Large-Scale Laboratory Structure. In K. Grimmelsman (Ed.), *Dynamics of Civil Structures, Volume 2* (pp. 15–22). Springer International Publishing. https://doi.org/10.1007/978-3-030-77143-0_2

Ba, J. L., Kiros, J. R., & Hinton, G. E. (2016). *Layer Normalization* (arXiv:1607.06450). arXiv. https://doi.org/10.48550/arXiv.1607.06450

Benbrahim, M., Benjelloun, K., Ibenbrahim, A., & Daoudi, A. (2005). Classification of non stationary signals using Ben wavelet and artificial neural networks. *Proceedings of the 2nd World Enfomatika Congress*, February 25-27, 2005, Istanbul, Turkey.

Chen, D., Zhang, Y., Wan, R., Li, Z., Xu, S., & Yang, C. (2024). Indirect identification of bridge damage based on coupled vehicle–bridge vibration and 2D-CNN. *Measurement Science and Technology*, *35*(5), 055019. https://doi.org/10.1088/1361-6501/ad2ad5

Demirlioglu, K., & Erduran, E. (2024). Drive-by Bridge Damage Detection Using Continuous Wavelet Transform. *Applied Sciences*, *14*(7), Article 7. https://doi.org/10.3390/app14072969

Deng, Z., Huang, M., Wan, N., & Zhang, J. (2023). The Current Development of Structural Health Monitoring for Bridges: A Review. *Buildings*, *13*(6), Article 6. https://doi.org/10.3390/buildings13061360

Fukuoka, T., & Fujiu, M. (2023). Detection of Bridge Damages by Image Processing Using the Deep Learning Transformer Model. *Buildings*, *13*(3), Article 3. https://doi.org/10.3390/buildings13030788

Garg, R. K., Chandra, S., & Kumar, A. (2022). Analysis of bridge failures in India from 1977 to 2017. *Structure and Infrastructure Engineering*, *18*(3), 295–312. https://doi.org/10.1080/15732479.2020.1832539

Goupillaud, P., Grossmann, A., & Morlet, J. (1984). Cycle-octave and related transforms in seismic signal analysis. *Geoexploration*, *23*(1), 85–102. https://doi.org/10.1016/0016-7142(84)90025-5

He, Z., Li, W., Salehi, H., Zhang, H., Zhou, H., & Jiao, P. (2022). Integrated structural health monitoring in bridge engineering. *Automation in Construction*, *136*, 104168. https://doi.org/10.1016/j.autcon.2022.104168

Hester, D., & González, A. (2017). A discussion on the merits and limitations of using drive-by monitoring to detect localised damage in a bridge. *Mechanical Systems and Signal Processing*, *90*, 234–253. https://doi.org/10.1016/j.ymssp.2016.12.012

Kim, H., & Melhem, H. (2004). Damage detection of structures by wavelet analysis. *Engineering Structures*, *26*(3), 347–362. https://doi.org/10.1016/j.engstruct.2003.10.008

Kobak, D., & Berens, P. (2019). The art of using t-SNE for single-cell transcriptomics. *Nature Communications*, *10*(1), 5416. https://doi.org/10.1038/s41467-019-13056-x

Kuo, C.-C., & Lee, C.-H. (2023). An Enhanced Structural Damage Identification Approach Using Statistical Analysis and Optimization. *IEEE Sensors Journal*, *23*(24), 31082–31097.

Lecun, Y., Bottou, L., Bengio, Y., & Haffner, P. (1998). Gradient-based learning applied to document recognition. *Proceedings of the IEEE*, *86*(11), 2278–2324. https://doi.org/10.1109/5.726791

Maeck, J., & De Roeck, G. (2003). DESCRIPTION OF Z24 BENCHMARK. *Mechanical Systems and Signal Processing*, *17*(1), 127–131. https://doi.org/10.1006/mssp.2002.1548

Mallat, S. G. (1989). A theory for multiresolution signal decomposition: The wavelet representation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *11*(7), 674–693. https://doi.org/10.1109/34.192463

M.ASCE, O. A., Ph D. ,. P. E. (2018, December 14). *Qatar University Grandstand Simulator (QUGS)*. Onur Avci. http://onur-avci.com/benchmark/qugs/

Najdi, B., Benbrahim, M., & Kabbaj, M. N. (2025). Adaptive Res-LSTM Attention-based Remaining Useful Lifetime Prognosis of Rolling Bearings. *International Journal of Prognostics and Health Management*, *16*(1), Article 1. https://doi.org/10.36001/ijphm.2025.v16i1.4171

Nguyen, D. C., Salamak, M., Katunin, A., Poprawa, G., & Gerges, M. (2024). Vibration-based SHM of railway steel arch bridge with orbit-shaped image and wavelet-integrated CNN classification. *Engineering Structures*, *315*, 118431. https://doi.org/10.1016/j.engstruct.2024.118431

Papoulis, A. (1977). *Signal analysis* (16. [print]). McGraw-Hill.

Qiu, Y., Ahmed, B., Abueidda, D. W., El-Sekelly, W., de Soto, B. G., Abdoun, T., Ji, H., Qiu, J., & Mobasher, M. E. (2024). *Damage identification for bridges using machine learning: Development and application to KW51 bridge* (arXiv:2408.03002). arXiv. http://arxiv.org/abs/2408.03002

Rabi, R. R., Vailati, M., & Monti, G. (2024). Effectiveness of Vibration-Based Techniques for Damage Localization and Lifetime Prediction in Structural Health Monitoring of Bridges: A Comprehensive Review. *Buildings*, *14*(4), Article 4. https://doi.org/10.3390/buildings14041183

Saidin, S. S., Jamadin, A., Abdul Kudus, S., Mohd Amin, N., & Anuar, M. A. (2022). An Overview: The Application of Vibration-Based Techniques in Bridge Structural Health Monitoring. *International Journal of Concrete Structures and Materials*, *16*(1), 69. https://doi.org/10.1186/s40069-022-00557-1

Salehin, I., & Kang, D.-K. (2023). A Review on Dropout Regularization Approaches for Deep Neural Networks within the Scholarly Domain. *Electronics*, *12*(14), Article 14. https://doi.org/10.3390/electronics12143106

Santaniello, P., & Russo, P. (2023). Bridge Damage Identification Using Deep Neural Networks on Time–Frequency Signals Representation. *Sensors*, *23*(13), Article 13. https://doi.org/10.3390/s23136152

Schabowicz, K. (2019). Non-Destructive Testing of Materials in Civil Engineering. *Materials*, *12*(19), Article 19. https://doi.org/10.3390/ma12193237

Song, X., Li, D., & Cho, C. (2024). Image-based machine learning approach for structural damage detection through wavelet transforms. *Urban Lifeline*, *2*(1), 4. https://doi.org/10.1007/s44285-023-00010-z

Sony, S., Gamage, S., Sadhu, A., & Samarabandu, J. (2022a). Multiclass Damage Identification in a Full-Scale Bridge Using Optimally Tuned One-Dimensional Convolutional Neural Network. *Journal of Computing in Civil Engineering*, *36*(2), 04021035. https://doi.org/10.1061/(ASCE)CP.1943-5487.0001003

Sony, S., Gamage, S., Sadhu, A., & Samarabandu, J. (2022b). Vibration-based multiclass damage detection and localization using long short-term memory networks. *Structures*, *35*, 436–451. https://doi.org/10.1016/j.istruc.2021.10.088

Sun, L., Shang, Z., Xia, Y., Bhowmick, S., & Nagarajaiah, S. (2020). Review of Bridge Structural Health Monitoring Aided by Big Data and Artificial Intelligence: From Condition Assessment to Damage Detection. *Journal of Structural Engineering*, *146*(5), 04020073. https://doi.org/10.1061/(ASCE)ST.1943-541X.0002535

Tan, C., Zhao, H., Uddin, N., & Yan, B. (2022). A Fast Wavelet-Based Bridge Condition Assessment Approach Using Only Moving Vehicle Measurements. *Applied Sciences*, *12*(21), Article 21. https://doi.org/10.3390/app122111277

Truong, T. T., Lee, J., & Nguyen-Thoi, T. (2022). An effective framework for real-time structural damage detection using one-dimensional convolutional gated recurrent unit neural network and high performance computing. *Ocean Engineering*, *253*, 111202. https://doi.org/10.1016/j.oceaneng.2022.111202

Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, Ł. ukasz, & Polosukhin, I. (2017). Attention is All you Need. *Advances in Neural Information Processing Systems*, *30*. https://proceedings.neurips.cc/paper/2017/hash/3f5ee243547dee91fbd053c1c4a845aa-Abstract.html

Wan, H., Gao, L., Yuan, Z., Qu, H., Sun, Q., Cheng, H., & Wang, R. (2023). A novel transformer model for surface damage detection and cognition of concrete bridges. *Expert Systems with Applications*, *213*, 119019. https://doi.org/10.1016/j.eswa.2022.119019

Wu, C.-S., Peng, Y.-X., Zhuo, D.-B., Zhang, J.-Q., Ren, W., & Feng, Z.-Y. (2022). Energy Ratio Variation-Based Structural Damage Detection Using Convolutional Neural Network. *Applied Sciences*, *12*(20), Article 20. https://doi.org/10.3390/app122010220

Xu, C., Ni, Y.-Q., & Wang, Y.-W. (2022). A novel Bayesian blind source separation approach for extracting non-stationary and discontinuous components from structural health monitoring data. *Engineering Structures*, *269*, 114837. https://doi.org/10.1016/j.engstruct.2022.114837

Z24 Bridge benchmark. (n.d.). [Page]. Departement Burgerlijke Bouwkunde KU Leuven. Retrieved February 17, 2025, from https://bwk.kuleuven.be/bwm/z24

BIOGRAPHIES

Oumayma Najem received her Engineering degree in Civil engineering from the EuroMed University of Fez in 2023 and is currently pursuing a Ph.D in Physics at Sidi Mohamed Ben Abdellah University, Fez. Her research focuses on the integration of artificial intelligence with reinforced concrete structures maintenance for predictive monitoring and enhanced durability in Structural Health Monitoring (SHM). She can be contacted at: oumayma.najem@usmba.ac.ma

Mohammed Benbrahim received the B.Eng. degree in electromechanical engineering from the Higher National School of Mines in 1997 and the M.Sc. and Ph.D. degrees in automatic and industrial informatics from the Mohammadia School of Engineers in 2000 and 2007, respectively. Currently, he is a Full Professor at the Department of Physics, the director of the LIMAS laboratory and the program coordinator for the Master's program in Smart Industry at the Faculty of Sciences, Sidi Mohamed Ben Abdellah University. His research interests include robotics, automatic control, intelligent systems, predictive maintenance, modeling, and optimization. He can be contacted at email: mohammed.benbrahim@usmba.ac.ma.

Mohammed Nabil Kabbaj is a Full Professor at the Faculty of Sciences, University of Fez, where he is the Vice Dean in charge of Research and the Program Coordinator of the Mechatronics and Embedded Systems Bachelor's program. His research interests include control engineering, fault detection, and diagnosis of complex systems. Before joining the University of Fez, he received a Ph.D. degree from the University of Perpignan in 2004 and was a postdoctoral researcher at LAAS-CNRS in Toulouse. He can be contacted at email: n.kabbaj@usmba.ac.ma.

Jaouad Boumhidi is a Professor of Computer Science at the Faculty of Sciences in Fez, Morocco. He received his PhD in Computer Science from the University of Sidi Mohamed ben Abdellah in 2005. He was a Board of Governors member of the International Neural Network Society (2020-2022). His research interests include machine learning, deep learning, computational intelligence, and intelligent transportation systems. He can be contacted at email: jaouad.boumhidi@usmba.ac.ma