SPECTRA-CQ: Incipient Anomaly Detection in Bridge Health Monitoring via Capacity-Constrained Quantum and Classical Autoencoders

Yesha Pandya[*†1], **S N Pandya**[†2], **and Dr. N K Solanki**[2]

[1] Department of Computer Science and Engineering, School of Technology, GSFC University, Vadodara, India
(Email: pandya.yesha.n@gmail.com)
[2] Department of Applied Mechanics, Faculty of Technology, The Maharaja Sayajirao University of Baroda, Vadodara, India
(Email: sarika.pandya-polyappmech@msubaroda.ac.in, nksolanki-appmech@msubaroda.ac.in)

[*] *Corresponding author.*
[†] *These authors contributed equally to this work.*

Abstract

Detecting incipient structural damage in civil infrastructure requires isolating transient acoustic emissions that are heavily obscured by macroscopic environmental noise. This paper introduces SPECTRA-CQ (Spectrogram Pattern Evaluation via Classical and Quantum Architectures), a hybrid anomaly detection framework evaluated on continuous vibration telemetry from the operational openLAB reference bridge in Bautzen, Germany. Raw 1-dimensional accelerometer signals are mapped into 2-dimensional 64 × 1000 energy-density spectrograms via Continuous Wavelet Transform (CWT) to preserve critical time-frequency localization. A synthetic acoustic emission (AE) proxy is then injected into a target node to establish strict physical detection limits. This AE is bounded at an amplitude of 0.35 to mathematically simulate early-stage micro-cracking rather than obvious macroscopic failure.

Initial unconstrained baseline ablations establish that both architectures are mathematically sound. Both Classical Convolutional Autoencoders (CAE) and Hybrid Quantum Autoencoders (HQAE) easily capture the structural variance at a latent space of 8 qubits on a single sensor, and yield highly accurate reconstructions with a mean squared error of approximately 0.008. A fundamental bottleneck emerges when enforcing Noisy Intermediate-Scale Quantum (NISQ) capacity constraints. Aggressively compressing the latent space to just 2 qubits forces isolated single-sensor models to fail. Neither the classical nor the quantum architecture can effectively separate low-amplitude anomalies from standard reconstruction noise without a broader spatial context. This limitation, called “Single-Sensor Spatial Blindness” prevents reliable incipient damage detection.

This barrier is overcome by expanding the architecture into a 6-sensor Siamese convolutional graph. It fuses spatial data from all six sensors while strictly maintaining the two-qubit bottleneck for each specific sensor. This setup allows the model to anchor itself against the geometric state of adjacent healthy nodes.

This multi-sensor fusion successfully recovers localized detection for the classical CAE architecture, and scales to an AUC-ROC of 0.888 for incipient damage evaluated at higher spectral mask widths. The quantum equivalent, however, hits a hard scaling wall. Ablation studies on circuit topology investigate the exact cause of this collapse.

Standard parameterized quantum circuits — ranging from completely uncoupled rotations to fully entangled arrays — are evaluated against two custom entanglement topologies. These custom designs include a rigidly structured global mapping and a Physics-Informed Neural Network (PINN) configuration designed to mathematically embed continuous wave attenuation. Despite these targeted configurations, forcing global entanglement across 12 qubits still reliably triggers gradient vanishing in the current NISQ-era hardware. Heavily entangled spatial quantum networks succumb to barren plateaus and degrade incipient detection back to baseline levels. Ultimately, the data highlights a clear divergence. Classical architectures readily exploit spatial coherence to bypass extreme latent compression. Forced quantum entanglement currently acts as a strict mathematical barrier for continuous multi-sensor structural monitoring.

Keywords: Bridge Health Monitoring, Quantum Machine Learning, Autoencoders, Anomaly Detection, Multi-Sensor Spatial Fusion

1 INTRODUCTION

Modern civil infrastructure is aging rapidly. Traditional maintenance relies on schedule-based visual inspections, which are fundamentally reactive and incapable of identifying internal structural degradation before macroscopic yielding occurs. Continuous Structural Health Monitoring (SHM) using vibration telemetry provides a data-driven alternative. Machine learning algorithms can theoretically detect structural anomalies in real-time by processing the dynamic response of a bridge under operational loading.

However, detecting incipient damage, specifically the high-frequency elastic stress waves generated by localized micro-cracking (Acoustic Emissions), remains a severe computational challenge. Incipient anomalies are characterized by low-amplitude transient energy bursts that are easily buried beneath the macroscopic environmental and operational noise floor of the structure. Classical deep learning architectures such as Convolutional Autoencoders (CAEs) excel at mapping this noise floor. But, they suffer from spatial blindness when evaluated on isolated single-sensor data. A solitary telemetry node mathematically lacks the geometric context required to differentiate a subtle acoustic emission from stochastic environmental variance. Overcoming this requires multi-sensor spatial fusion. Yet, naively scaling classical networks to process massive multi-node spatial graphs inevitably leads to parameter explosion and multi-channel overfitting.

To address these combined physical and computational bottlenecks, this paper evaluates classical spatial fusion against parameterized quantum topologies for early-stage damage detection. Both paradigms are benchmarked on real-world accelerometer telemetry recorded from a full-scale research bridge deck under operational loading [1]. The underlying codebase, network artifacts, and data processing pipeline are archived in an open-access repository [2].

1.1 Theoretical Rationale for using Hybrid Quantum Autoencoders

Unsupervised machine learning has become the standard paradigm for anomaly detection in SHM because physical structural failure events are extremely rare in operational infrastructure. Classical deep autoencoders (CAEs) and variational autoencoders (VAEs) excel at mapping baseline operational behavior without requiring labeled damage data. However, when CAEs process dense multi-sensor arrays to capture spatial cross-correlations, their fully connected latent layers scale quadratically ($O(d^2)$) with feature dimension d. These high parameter counts encourage over-parameterization since training datasets consist almost entirely of healthy operational telemetry. As a result, dense classical bottlenecks are prone to overfitting. They tend to memorize localized environmental noise and temperature-induced variance rather than learning generalizable structural dynamics.

Hybrid Classical-Quantum Autoencoders (HQAEs) circumvent this parameter explosion by substituting the dense classical bottleneck with a Variational Quantum Circuit (VQC). Recent comparative studies in anomaly detection across time-series and physical signal domains demonstrate that QAEs can achieve competitive or superior anomaly isolation compared to deep classical autoencoders while operating with order-of-magnitude parameter reductions [3-5]. Rather than scaling quadratically, parameter complexity in a VQC scales linearly ($O(N \cdot L)$) with N qubits and L circuit layers [6]. This occurs because trainable weights are stored merely as single qubit rotation angles rather than large matrix grids.

Quantum bottlenecks offer another geometric advantage beyond parameter reduction: Hilbert space projection [7]. Encoding compressed latent vectors onto N qubits maps the telemetry into a 2^N-dimensional Hilbert space:

$$\mathcal{H} \simeq \mathbb{C}^{2^N}$$

where $\mathbb{C}$ denotes the complex vector space. Under severe compression constraints, low-amplitude micro-cracking acoustic emissions overlap directly with macroscopic environmental vibrations in low-dimensional classical space. Lifting these features into a 2^N -dimensional Hilbert space increases feature separability. Subtle structural perturbations that were previously buried in operational noise become geometrically separable without inflating classical weight counts.

Despite these theoretical advantages, raw time-frequency telemetry cannot be loaded directly into quantum hardware due to the physical limits of Noisy Intermediate-Scale Quantum (NISQ) devices [8]. Mapping thousands of spectral energy pixels per sensor directly into quantum state amplitudes requires deep state-preparation gate sequences. The execution time for these deep gate sequences far exceeds physical qubit relaxation (T_1) and dephasing (T_2) coherence times on contemporary quantum processors. This causes the quantum state to lose its phase information and degrade into thermal noise long before the circuit completes execution (similar analogy to how a damped structural vibration signal decays into ambient noise). A classical 2D convolutional frontend is therefore necessary to filter raw telemetry and compress spatial-spectral matrices down to low-dimensional latent vectors prior to quantum encoding.

This research deploys this hybrid setup to address two core engineering questions. First, does entangling multi-sensor data in quantum space actually improve early damage detection under severe compression? Second, at what point do vanishing gradients — specifically barren plateaus $\mathrm{Var}[\partial_\theta \mathcal{L}] \to 0$ — cause model training to stall on real bridge telemetry [9, 10]?

2 RELATED WORK

The application of deep learning to structural dynamics has matured significantly over the past decade. Early foundational work by Abdeljaber et al. [11] demonstrated the efficacy of 1-D Convolutional Neural Networks (CNNs) for real-time vibration-based damage detection. Recognizing that transient mechanical events are poorly localized in the raw time domain, subsequent research transitioned to time-frequency representations. Zhang et al. [12] proved that mapping acceleration signals into 2D frequency-domain images dramatically improves the detection of concrete deck cracking by separating resonant frequencies from high-frequency noise.

More recently, the focus has shifted toward unsupervised learning to handle the extreme data imbalance in SHM, since physical bridges rarely experience structural failure. Abraham et al. [13] and Kim et al. [14] successfully deployed unsupervised deep learning frameworks to forecast dynamic responses and isolate deviations from normal structural behavior without requiring labeled damage data. While these classical models achieve high accuracy on macroscopic anomalies, they remain computationally expensive when scaled across dense multi-sensor arrays.

To circumvent classical capacity limits, QML is emerging as a novel paradigm in civil engineering. Alves et al. [15] recently established a new perspective on SHM by extracting features from raw acceleration signals and encoding them into unsupervised quantum classifiers to detect structural anomalies. Despite these advancements, a critical gap remains in the intersection of QML and structural physics.

Standard VQCs natively treat input parameters as independent mathematical features and apply rigid entanglement topologies (such as all-to-all CNOT layers) without inherent geometric context. If deployed naively for multi-sensor continuous telemetry, this mathematical formulation ignores physical reality: kinetic energy propagates through concrete and steel subject to strict spatial attenuation [16]. There is currently no framework in the literature that mathematically embeds continuous elastic wave mechanics into the entanglement architecture of a Variational Quantum Circuit. SPECTRA-CQ directly addresses this gap by comparing rigidly entangled quantum graphs against a physics-informed, distance-penalized quantum topology.

3 METHODOLOGY: DATA AND PROBLEM FORMULATION

Machine learning models are strictly bound by the physical integrity of their input data. In SHM, algorithms do not process abstract arrays; they map the real-time kinetic energy of massive concrete and steel structures. This section defines the empirical baseline, continuous wave mechanics, and the strict mathematical boundaries of the synthetic anomaly proxy.

3.1 The openLAB Reference Dataset

To establish a rigorous baseline, this study utilizes telemetry from the undamaged openLAB research bridge in Bautzen, Germany [1]. Data is acquired via piezoelectric accelerometers (PCB 393A03) routed through a Gantner Instruments Q-station.

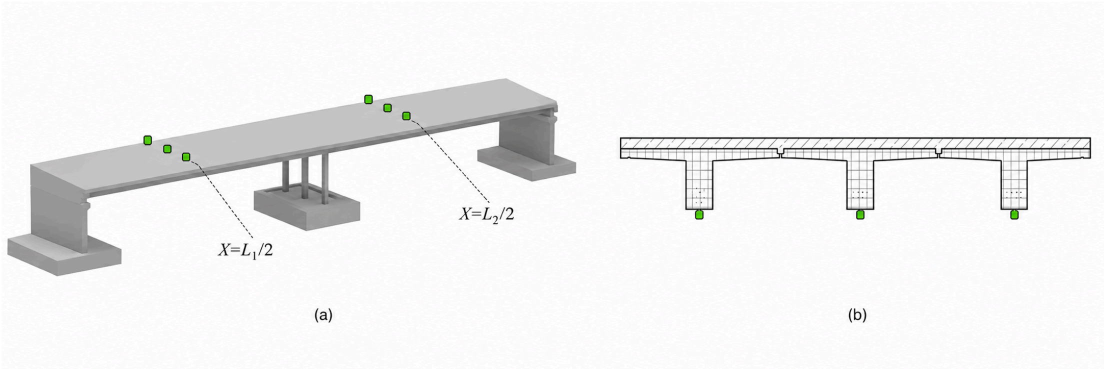


Figure 1: (a) Spatial topology and (b) cross-sectional piezoelectric accelerometer array of the openLAB reference bridge [1].

The hardware employs a threshold trigger. Measurements initiate only when localized vibration exceeds 2×10^{-4} m/s². Each triggered event captures a continuous 70-second burst sampled at 500 Hz. Upstream processing removes the DC offset and applies a 4th-order Butterworth filter to enforce a 0.5 to 100 Hz bandwidth.

In this study, node `PE11` (located on the bottom edge of precast element span 1, girder 1) was designated as the isolated target for all evaluations to guarantee a mathematically consistent baseline across both single-sensor ablations and multi-sensor spatial fusion.

3.2 Spatio Temporal Signal Processing

Standard Fast Fourier Transforms (FFT) isolate resonant frequencies but strip away temporal localization, rendering them insufficient for mapping transient acoustic emissions [17]. Therefore, a Continuous Wavelet Transform (CWT) was applied using the Complex Morlet Wavelet to capture both frequency shifts and exact temporal onset. The admissibility parameter is strictly set to $\omega_0 = 6.0$ to symmetrically balance time and frequency resolution according to the Heisenberg-Gabor limit [18].

The 70-second bursts were segmented into non-overlapping 2-second windows (1000 time steps at 500 Hz). The CWT maps these across 64 logarithmically spaced frequency scales. This converts 1D vibrations into 2D spatial matrices of shape $(64, 1000)$. All kinetic energy values are then strictly bounded to $[0, 1]$ range using Min-Max normalization to optimize the tensors for stable gradient descent.

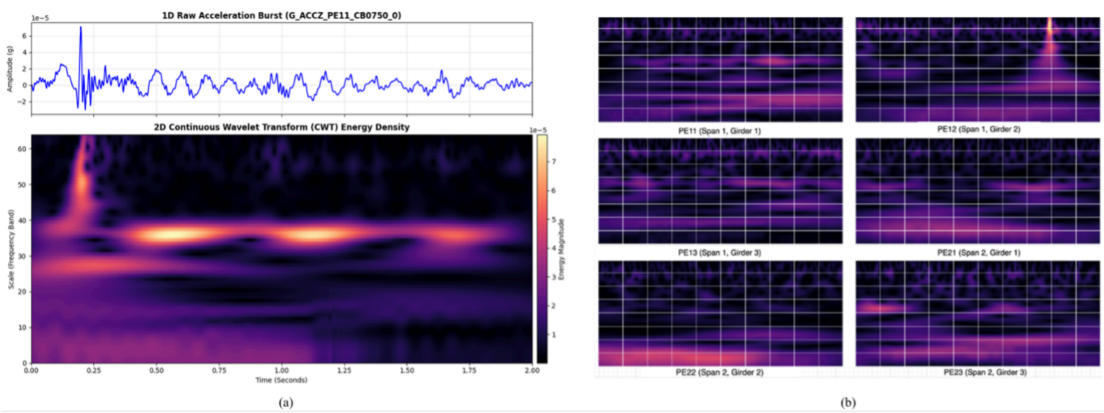


Figure 2: (a) 1D signal to CWT spectrogram mapping and (b) 6-sensor batch-channel stacking.

3.3 Physical Wave Mechanics and the Anomaly Proxy

The openLAB dataset provides a pristine baseline of a healthy structure; so, synthetic structural degradation must be introduced to evaluate anomaly detection. Bridges do not typically fail instantaneously. Degradation is a progressive phenomenon driven by cyclical fatigue. When a concrete or steel member experiences stress exceeding its localized tensile limit, molecular bonds fracture. This localized failure releases an Acoustic Emission (AE), i.e., a highly localized burst of kinetic energy that travels through the girder as an elastic stress wave [19].

To replicate this, a high-frequency AE proxy was injected directly into the normalized CWT matrices of the test data. This injection was isolated exclusively to the `PE11` channel to simulate a micro-crack occurring immediately adjacent to that specific accelerometer. Crucially, the energy amplitude of this synthetic wave was restricted to 0.35 on the $[0, 1]$ scale. The justification here is fundamental: Injecting a massive amplitude (e.g., 0.90) simulates catastrophic failure that triggers a trivial threshold alert; which renders advanced machine learning entirely unnecessary. Capping the amplitude at 0.35 buries the anomaly deep within the bridge's natural environmental noise floor. The network cannot rely on raw signal clipping to detect it, and is forced to rely on complex geometric reconstruction.

3.4 Spectral Masking and Damage Severities

If amplitude dictates a fracture's physical intensity, spectral width dictates its temporal footprint. A larger fracture generates a more sustained acoustic emission. To precisely quantify the detection limits of our architectures, we apply a sliding spectral mask that controls the duration of the injected high-frequency burst.

All evaluated masks represent incipient fatigue since the amplitude is capped at 0.35. In structural dynamics, a localized high-frequency perturbation without a corresponding low-frequency global modal shift inherently defines the early-warning damage regime. The degradation remains mathematically confined to the pre-macroscopic phase. We evaluate six distinct physical severities:

1. Mask Width 2 (3.1% Spectrum Loss): Isolated micro-cracking; a nearly instantaneous energy release.
2. Mask Width 3 (4.6% Loss): Developing micro-cracking yielding under cyclical load.
3. Mask Width 4 (6.2% Loss): Moderate micro-crack clustering with extended resonance.
4. Mask Width 5 (7.8% Loss): Intermediate incipient degradation; early localized yielding.
5. Mask Width 6 (9.3% Loss): Advanced micro-crack networking that is still physically undetectable to visual inspection.
6. Mask Width 8 (12.5% Loss): The absolute upper physical boundary of the incipient phase, immediately preceding macroscopic failure.

Receiver Operating Characteristic (ROC) curves were plotted across all spectral mask widths to evaluate anomaly isolation boundaries across progressive damage severities.

4 ARCHITECTURAL FRAMEWORK AND TOPOLOGIES

4.1 Classical Bottleneck and Spatial Blindness

The architecture directly ingests single-channel CWT energy density matrices of shape $(1, 64, 1000)$ from accelerometer node `PE11`. The deep classical frontend functions as a spatial-spectral low-pass filter to systematically strip stochastic operational noise. An initial 3×3 convolution expands the input to 16 channels, followed by an asymmetric 2×4 MaxPool layer. A secondary 3×3 convolution expands the representation to 32 feature maps before a 4×10 MaxPool layer downsamples the spatial grid. This sequence isolates dominant structural resonant frequencies and yields a compressed $(32, 8, 25)$ tensor that flattens into 6,400 classical features.

To establish a baseline, the CAE was evaluated under an unconstrained latent capacity of 8 continuous variables per node ($z_s \in \mathbb{R}^8$). This was matched to an unconstrained budget of $N_{\text{sensor}} = 8$ qubits per sensor node for HQAE models. Under this setting, the models achieved a best validation mean squared error of ~ 0.0089, confirming their foundational mathematical soundness. We then restrict $z_s \in \mathbb{R}^2$) to correspond to a Noisy Intermediate-Scale Quantum (NISQ) budget constraint of $N_{\text{sensor}} = 2$ qubits per node. Compressing $64,000$ input pixels down to 2 continuous latent variables severely starves the network across `CAE`, `HQAE[None]`, and `HQAE[Strong]` single-sensor ablations. Evaluated in isolation, the single node mathematically fails to distinguish a low-amplitude acoustic emission from its own blurred baseline reconstruction error. This empirically validates single-sensor spatial blindness under extreme latent compression.

4.2 Siamese Multi-Sensor Spatial Fusion

To overcome spatial blindness without violating the $N_{\text{sensor}} = 2$ latent constraint, this study transitions to a global 6-sensor spatial graph. Processing six independent $(64, 1000)$ matrices simultaneously increases parameter exposure and risks multi-channel overfitting; which is prevented using a weight-tied Siamese formulation. Prior to feature extraction, the input tensor is dynamically folded across the batch and channel dimensions from $(B, 6, 64, 1000)$ into $(B \times 6, 1, 64, 1000)$, forcing all six accelerometer streams through identical classical convolutional filters simultaneously.

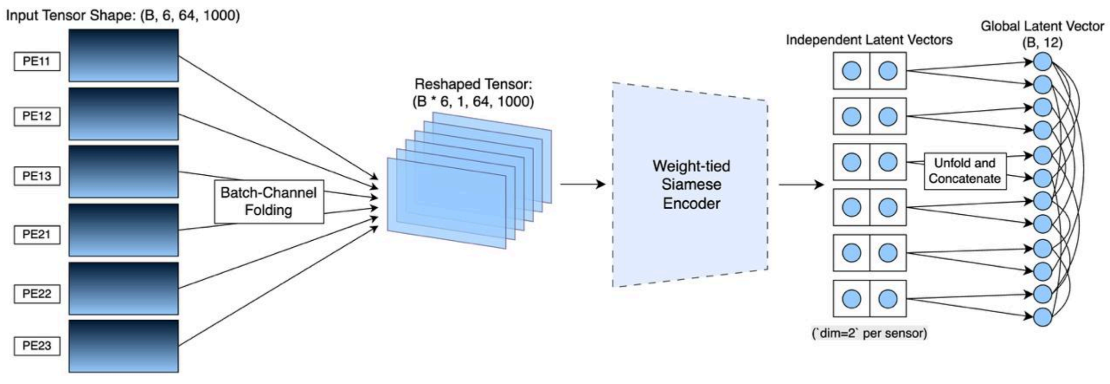


Figure 3: Siamese batch-channel folding of 6 spectrograms into the global latent graph

The Siamese frontend outputs six independent 2-dimensional latent feature vectors ($z_s \in \mathbb{R}^2$ for $s \in \{1, ..., 6\}$). These vectors are then unfolded and concatenated into a single 12-dimensional global latent tensor ($z_{\text{global}} \in \mathbb{R}^{12}$). This establishes a 1:1 mapping with the 12-qubit global register ($N = 12$) utilized in the quantum bottleneck. Passing this global representation to the decoder leverages spatial cross-correlation across the sensor array. By exploiting the spatial coherence of five healthy adjacent nodes, the network confidently isolates the out-of-phase transient anomaly at the target node and restores damage sensitivity without expanding individual node latent allocations.

4.3 HQAE Framework and Hardware Integration

The Hybrid Quantum Autoencoder (HQAE) retains the exact Siamese 2D-CNN encoder and deconvolutional decoder used in the classical baseline. The sole architectural divergence occurs at the bottleneck, where the classical dense layer is replaced by a 12-qubit Variational Quantum Circuit (VQC) partitioned into six 2-qubit local registers ($N_{\text{sensor}} = 2$ qubits per sensor node).

Executing hybrid neural networks introduces computational memory challenges: deep 2D-CNN feature maps are computed on GPU accelerators (`cuda`), whereas quantum state-vector simulators operate on host CPU memory (`default.qubit`). Custom autograd operators handle tensor transfers across the hardware bridge to manage this device handoff without breaking PyTorch's automatic differentiation graph.

Let $s \in \{1, ..., 6\}$ denote the physical sensor node index across the bridge deck, and $k \in \{1, 2\}$ denote the latent feature component index (corresponding to the two qubits allocated per sensor node). For each sensor s, the Siamese encoder outputs an unbounded 2-dimensional classical feature vector $\boldsymbol{z}_s$:

$$\boldsymbol{z}_s = \left[z_{s,1}, z_{s,2}\right]^T \in \mathbb{R}^2$$

Prior to quantum state preparation, each feature scalar $z_{s,k}$ is bounded to prevent Bloch sphere phase-wrapping (where rotation angles exceeding 2π radians over-rotate the quantum state and scramble spatial representation) via the transformation:

$$x_{s,k} = \pi \cdot \tanh(z_{s,k}), \quad \text{for} \quad s \in \{1, ..., 6\} \quad \text{and} \quad k \in \{1, 2\}$$

This operation clamps all classical feature values to bounded rotation angles $x_{s,k} \in [-\pi, \pi]$. Each bounded angle $x_{s,k}$ is then mapped onto the k-th qubit of sensor s's dedicated sub-register via single-qubit Pauli-Y rotation gates (R_y angle embedding) applied to the initial computational ground state $|0\rangle$:

$$|\psi_s\rangle = R_y(x_{s,1})\ |0\rangle \otimes R_y(x_{s,2})\ |0\rangle \in \mathcal{H}_s \simeq \mathbb{C}^{2^2}$$

where $|\psi_s\rangle$ is the local 2-qubit state vector for sensor s, $\otimes$ denotes the tensor product operator, and $\mathcal{H}_s \simeq \mathbb{C}^4$ represents the 4-dimensional (2^2) local complex Hilbert state space.

Prior to inter-sensor entanglement, the global 12-qubit state vector $|\Psi_0\rangle$ exists as a separable product state constructed via the multi-register tensor product ($\bigotimes$) across all six local sensor sub-registers:

$$|\Psi_0\rangle = \bigotimes_{s=1}^{6} |\psi_s\rangle \in \mathcal{H}_{\text{global}} \simeq \mathbb{C}^{4096}$$

where $\mathcal{H}_{\text{global}} \simeq \mathbb{C}^{2^{12}} = \mathbb{C}^{4096}$ represents the 4,096-dimensional global Hilbert state space across the 12-qubit register.

After applying parameterized unitary transformations $U(\theta)$, continuous scalar outputs are extracted by computing the Pauli-Z expectation value for each individual qubit k within each sensor s's sub-register:

$$z'_{s,k} = \langle \Psi_0 \mid U^\dagger(\theta)\left(\sigma_z^{(s,k)}\right)U(\theta) \mid \Psi_0 \rangle \in [-1.0, 1.0]$$

where $U(\theta)$ represents the parameterized quantum circuit driven by trainable parameters θ, $U^\dagger(\theta)$ is its conjugate transpose (adjoint), $\sigma_z^{(s,k)}$ is the Pauli-Z measurement operator for qubit k of sensor node s, and $\langle \Psi_0 \mid \ldots \mid \Psi_0 \rangle$ denotes the quantum mechanical expectation value (inner product).

These measured expectation values are concatenated across all sensors s and features k into a continuous 12-dimensional vector ($\boldsymbol{z}' \in \mathbb{R}^{12}$) and transferred across the hardware bridge back to the GPU. The vector is then un-concatenated back into six 2-variable sensor representations and passed to the classical deconvolutional decoder for spatial reconstruction. Each sensor's 2-qubit circuit functions as an independent non-linear quantum activation block when evaluated without entangling operations (`HQAE[None]`). However, modeling spatial cross-correlations across the bridge deck requires introducing multi-qubit entangling gates between these local sub-registers.

4.4 Quantum Entanglement Topologies and Optimization Instability

Mapping six physical sensor nodes ($s = 1, \ldots, 6$) onto dedicated 2-qubit sub-registers $\{q_{2s-1}, q_{2s}\}$ defines $C(6, 2) = 15$ potential spatial graph edges across the bridge deck. Two domain-specific entanglement topologies that couple these local 2-qubit registers were also engineered to evaluate multi-sensor spatial correlation learning:

1. **`bridge_hard`:** A rigid geometric topology applying unparameterized Controlled-NOT (CNOT) gates between sub-registers corresponding to physical sensor pairs separated by Euclidean distances $D_{ij} \leq 15.0m$.
2. **`bridge_soft`:** A Physics-Informed Neural Network (PINN) topology replacing rigid CNOTs with parameterized Controlled-Y (CRY) gates [12]. The rotation angle applied to each CRY gate is exponentially attenuated based on physical node separation:

$$\theta_{\text{applied}} = \theta_{\text{learned}} \cdot e^{-\gamma D_{ij}}$$

where θ_{learned} is the trainable parameter generated by the classical optimizer, D_{ij} is the physical Euclidean distance between bridge nodes i and j, and $\gamma = 0.1m^{-1}$ is the structural spatial attenuation coefficient.

To enforce physical wave propagation constraints across the spatial graph, a wave mechanics regularization penalty is appended to the global loss function [20]:

$$L_{\text{Total}} = L_{\text{Recon}} + \lambda \parallel \frac{\partial^2 u}{\partial t^2} - c^2 \nabla^2 u \parallel^2$$

where L_{Recon} is the reconstruction Mean Squared Error (MSE), λ is a penalty scaling factor, u denotes the spatio-temporal elastic displacement field reconstructed by the decoder, c is the acoustic wave propagation velocity in concrete, and ∇^2 is the spatial Laplacian operator.

Gradients through the quantum circuit are computed analytically using the parameter-shift rule [21]. Quantum circuit weights are initialized near zero ($[-0.01, 0.01]$) to mitigate early gradient vanishing associated with Barren Plateaus [9, 10]. This forces the circuit to act initially as an identity mapping [30]. Despite these mitigation measures, enforcing global multi-qubit entanglement across the six local sub-registers induces severe optimization instability.

Crucially, while classical Siamese networks maintained tight optimization convergence across multiple deterministic seeds, the `bridge_soft` quantum PINN suffered an order-of-magnitude explosion in standard deviation across all tested degradation severities. This physically proves that under current NISQ limits, forced global entanglement across a multi-sensor graph acts as chaotic optimization noise. It actively degrades the anomaly detection boundaries established by classical spatial fusion.

5 EXPERIMENTAL SETUP AND EVALUATION PROTOCOL

All classical and hybrid architectures are constrained to an identical optimization framework to ensure strict mathematical comparability and isolate the representational behavior of the quantum bottleneck.

5.1 Hyperparameter Constraints and NISQ Regularization

Model training is executed using the Adam optimizer [23] with a fixed learning rate of $\eta = 1 \times 10^{-3}$. Batches are formatted to a constant size of 32 samples to provide sufficient gradient averaging to stabilize the stochastic evaluation of quantum expectation values. Training is bounded to a strict threshold of 15 epochs; empirical testing confirmed that extended epoch limits induce classical decoder overfitting without yielding further quantum convergence. All reported metrics represent a 3-seed deterministic ensemble (seeds 42, 100, and 2026) to account for optimization volatility inherent in NISQ simulation landscapes.

To stabilize backpropagation across the hybrid classical-quantum interface and delay the onset of gradient vanishing (Barren Plateaus), three complementary NISQ regularization protocols were implemented:

1. **Feature Space Bounding:** Unbounded classical encoder outputs $z_s \in \mathbb{R}^2$ are transformed via:

   $$x_{s,k} = \pi \cdot \tanh\left(z_{s,k}\right)$$

   This clamps all rotation inputs strictly to $x_{\{(s,k)\}} \in [-\pi, \pi]$; which prevents Bloch sphere phase-wrapping and ensures a deterministic 1-to-1 mapping onto single-qubit rotation gates (R_y).
2. **Near-Zero Identity Initialization:** Parameterized quantum weights θ are initialized from a uniform distribution over a restricted near-zero interval:

   $$\theta \approx U(-0.01, 0.01)$$

   This forces the parameterized quantum transformation matrix $U(\theta)$ (i.e., the overall unitary matrix representing the circuit) to act initially as an identity matrix ($U(\theta) \approx I$, where I is the identity matrix that leaves incoming quantum states unchanged). In structural signal processing terms, this creates a pass-through initialization at epoch 0: the circuit passes the state-prepared telemetry without adding initial phase distortion. This prevents the network from starting in a chaotic, maximally mixed state (i.e., white noise) and bypasses initialization-induced Barren Plateaus.
3. **Analytical Gradient Norm Clipping:** Backpropagation through the quantum circuit relies on the parameter-shift rule [13]. Small shifts in quantum expectation values can

trigger large backpropagation updates in the classical 2D-CNN encoder. So, we apply strict gradient norm clipping ($|g|_2 \leq 1.0$) prior to each optimization step to prevent high-variance quantum gradients from destabilizing classical feature extraction [24].

5.2 Target Node Isolation Protocol

Evaluating a multi-sensor spatial graph against a localized incipient structural anomaly introduces a significant metric hazard: spatial error dilution. Incipient damage (such as localized micro-cracking acoustic emissions) manifests as an isolated transient burst at target node `PE11`. The remaining five undamaged adjacent nodes (`PE12`, `PE13`, `PE21`, `PE22` `PE23`) continue to exhibit baseline healthy vibration signatures. Averaging the reconstruction loss across all six channels during evaluation would mathematically dilute the localized reconstruction error spike by a factor of $\frac{1}{6}$, masking subtle damage signatures and artificially inflating overall performance metrics.

Target Node Isolation protocol was enforced during inference to guarantee an equivalent performance comparison between the single-sensor baseline ($N_{\text{sensor}} = 2$) and the multi-sensor spatial graph ($N_{\text{total}} = 12$). While all six channels contribute to spatial feature extraction and global loss computation during training, evaluation metrics are calculated strictly on the target accelerometer node (`PE11`):

$$\mathrm{MSE}_{\text{target}} = \frac{1}{M} \sum_{m=1}^{M} \left\| \boldsymbol{X}_{\text{PE11}}^{m} - \hat{\boldsymbol{X}}_{\text{PE11}}^{m} \right\|_F^2$$

where $\boldsymbol{X}_{\text{PE11}}^{m}$ and $\hat{\boldsymbol{X}}_{\text{PE11}}^{m} \in \mathbb{R}^{64\times1000}$ denote the ground-truth and reconstructed CWT energy density matrices respectively for target node `PE11` on the m-th evaluation sample, $\|\cdot\|_F$ represents the Frobenius norm, and M is the evaluation sample count. Discarding the reconstruction error of undamaged neighboring nodes during evaluation isolates the model's true localized anomaly sensitivity.

6 RESULTS AND DISCUSSION

AUC-ROC scores are mapped across progressive spectral mask widths to quantify the boundaries of incipient detection. Table 1 establishes the single-sensor limits and Table 2 details the multi-sensor spatial fusion results. To guarantee a strict equivalent ablation, every configuration across both tables is constrained to a 2-variable latent bottleneck per sensor node ($z_s \in \mathbb{R}^2$). this is matched to a NISQ capacity limit of $N_{\text{sensor}} = 2$ qubits per node.

Table 1: *Single-Sensor Evaluation (Target Node: `PE11`, $N_{\text{sensor}} = 2$ qubits)*

Damage Severity	CAE (Baseline)	HQAE [None]	HQAE [Strong]
Mask 2 (3.1%)	0.620 ± 0.001	0.624 ± 0.010	0.634 ± 0.016
Mask 3 (4.6%)	0.673 ± 0.006	0.669 ± 0.006	0.678 ± 0.009
Mask 4 (6.2%)	0.720 ± 0.009	0.706 ± 0.009	0.720 ± 0.008
Mask 5 (7.8%)	0.752 ± 0.003	0.744 ± 0.007	0.747 ± 0.005
Mask 6 (9.3%)	0.781 ± 0.005	0.768 ± 0.010	0.777 ± 0.008
Mask 8 (12.5%)	0.819 ± 0.001	0.812 ± 0.011	0.813 ± 0.010

Table 2: *Multi-Sensor Spatial Fusion (6-Sensor Array, $N_{\text{sensor}} = 2$ qubits per node, $N = 12$ total qubits)*

Damage Severity	CAE (Baseline)	HQAE [None]	HQAE [Strong]	HQAE [Bridge Hard]	HQAE [Bridge Soft]
Mask 2	0.650 ± 0.015	0.629 ± 0.012	0.616 ± 0.010	0.617 ± 0.025	0.612 ± 0.048
Mask 3	0.705 ± 0.020	0.685 ± 0.012	0.664 ± 0.010	0.672 ± 0.035	0.647 ± 0.054
Mask 4	0.745 ± 0.016	0.729 ± 0.013	0.710 ± 0.010	0.709 ± 0.045	0.679 ± 0.064
Mask 5	0.788 ± 0.012	0.778 ± 0.017	0.749 ± 0.009	0.743 ± 0.054	0.712 ± 0.070
Mask 6	0.830 ± 0.013	0.819 ± 0.017	0.786 ± 0.010	0.770 ± 0.058	0.739 ± 0.078
Mask 8	0.888 ± 0.009	0.891 ± 0.018	0.839 ± 0.007	0.818 ± 0.063	0.801 ± 0.089

6.1 The Classical Spatial Regularization Advantage

The empirical data demonstrates a definitive performance hierarchy between the isolated and multi-sensor configurations. Transitioning from a solitary sensor to a Siamese spatial graph significantly enhances the classical architecture's sensitivity. At the upper boundary of incipient damage (Mask Width 8), the spatial CAE scales to an AUC-ROC of 0.888, outperforming the isolated baseline's 0.819.

The network establishes a highly robust macroscopic baseline by forcing the shared convolutional filters to mathematically map six distinct geometric locations simultaneously [17]. This spatial regularization inherently suppresses the false-positive rate associated with localized environmental noise and allows the architecture to confidently isolate the target node anomaly despite the severe latent capacity constraint.

6.2 The Quantum Scaling Wall and Barren Plateaus

Initial single-sensor evaluations indicated that a highly entangled 2-qubit HQAE could marginally outperform the classical baseline (0.634 vs 0.620 at Mask Width 2). However, scaling this architecture to a 12-qubit spatial graph resulted in severe performance degradation. The all-to-all HQAE [`Strong`] topology dropped to an AUC-ROC of 0.616 at Mask Width 2, ultimately performing worse than its isolated single-sensor counterpart.

Notably, this drop persisted despite the implementation of standard barren plateau mitigations. Extending rigid CNOT topologies across 12 qubits inherently flattened the optimization landscape and rendered these regularizers ineffective, which terminally stalled gradient flow [18]. Consequently, the quantum model failed to learn the multi-node spatial graph. In contrast, the unentangled HQAE [`None`] maintained robust performance as the system scaled, achieving an AUC-ROC of 0.891 at Mask Width 8. This comparison strongly suggests that forced global entanglement, rather than total qubit volume or suboptimal initialization, serves as the primary computational bottleneck for multi-sensor structural fusion under current NISQ limitations.

6.3 The Physics-Informed Rejection (PINN Paradox)

The physics-informed topology (`bridge_soft`) yielded the highest optimization instability. This is proven by a tenfold increase in standard deviation (±0.089 at Mask Width 8). However, this represents a highly significant theoretical validation from a structural diagnostics perspective.

The synthetic acoustic emission was mathematically injected strictly into target node `PE11`. This left adjacent sensors unperturbed. In a physical continuum, elastic stress waves must propagate geometrically across the concrete medium [25].

The localized mathematical injection therefore created a physically impossible spatial discontinuity: $\nabla^2 u \to \infty$. The network's physics loss exploded because the PINN incorporates the continuous wave mechanics penalty. The PINN correctly classified the synthetic anomaly as a physical impossibility and rejected it. This strongly suggests that quantum PINNs possess substantial theoretical capability to filter non-physical sensor artifacts when applied to real-world, propagating structural degradation.

7 CONCLUSION AND FUTURE WORK

This study demonstrates that classical multi-sensor spatial fusion, leveraging weight-tied regularization, successfully isolates incipient acoustic emissions buried within macroscopic structural noise floors. Conversely, current NISQ-era quantum architectures hit a hard scaling wall. Forcing spatial entanglement across multi-sensor graphs reliably induces barren plateaus and catastrophic gradient vanishing. While the physics-informed `bridge_soft` quantum PINN struggled with synthetic localized damage, its mathematical rejection of spatial discontinuities ($\nabla^2 u \to \infty$) demonstrates immense theoretical promise for filtering non-physical sensor artifacts.

A critical limitation of this ablation study is its reliance on synthetic spectral masking, which fails to replicate the non-linear global modal shifts of true structural yielding. Future work will deploy this distance-penalized quantum topology against continuous field telemetry from destructive physical load tests. This would definitively corroborate its efficacy in mapping real-world, propagating macroscopic failure.

REFERENCES

[1] Jansen, A., Richter, B., Röder, R., Herbers, M., & Marx, S. (2024). *Monitoring data of the openLAB research bridge (2024-02-01 to 2024-10-31) and building information* [Data set]. OPARA Dataset Repository. https://doi.org/10.25532/OPARA-660

[2] Pandya, Y. (2026). *spectra-cq* [Source code]. GitHub. https://github.com/yeshapan/spectra-cq

[3] Sakhnenko, A., O'Meara, C., Neukart, F., & Bernabé-Moreno, J. (2021). Hybrid classical-quantum autoencoder for anomaly detection. *Quantum Machine Intelligence*, *3*(2), 27. https://doi.org/10.1007/s42484-021-00052-8

[4] Woźniak, S., Belis, V., Puljak, E., Grossi, M., Pierini, M., & Sofia, A. (2023). Anomaly detection in high-energy physics using a quantum autoencoder. *Physical Review D*, *108*(3), 036006. https://doi.org/10.1103/PhysRevD.108.036006

[5] Tscharke, K., Wendlinger, M., Ahouzi, A., Bhardwaj, P., Amoi-Taleghani, K., Schrödl-Baumann, M., & Debus, P. (2025). Quantum autoencoder for multivariate time series anomaly detection. *2025 IEEE International Conference on Quantum Computing and Engineering (QCE)*, 2470–2481. https://doi.org/10.1109/QCE65121.2025.00268

[6] Benedetti, M., Lloyd, S., Sack, S., & Fiorentini, M. (2019). Parameterized quantum circuits as machine learning models. *Quantum Science and Technology*, *4*(4), 043001. https://doi.org/10.1088/2058-9565/ab4eb5

[7] Schuld, M., & Killoran, N. (2019). Quantum machine learning in feature Hilbert spaces. *Physical Review Letters*, *122*(4), 040504. https://doi.org/10.1103/PhysRevLett.122.040504

[8] Preskill, J. (2018). Quantum computing in the NISQ era and beyond. *Quantum*, *2*, 79. https://doi.org/10.22331/q-2018-08-06-79

[9] McClean, J. R., Boixo, S., Smelyanskiy, V. N., Babbush, R., & Neven, H. (2018). Barren plateaus in quantum neural network training landscapes. *Nature Communications*, *9*(1), 4812. https://doi.org/10.1038/s41467-018-07090-4

[10] Wang, S., Fontana, E., Cerezo, M., Sharma, K., Sone, A., Cincio, L., & Coles, P. J. (2021). Noise-induced barren plateaus in variational quantum algorithms. *Nature Communications*, *12*(1), 6961. https://doi.org/10.1038/s41467-021-27045-6

[11] Abdeljaber, O., Avci, O., Kiranyaz, S., Gabbouj, M., & Inman, D. J. (2017). Real-time vibration-based structural damage detection using one-dimensional convolutional neural networks. *Journal of Sound and Vibration*, *388*, 154–170. https://doi.org/10.1016/j.jsv.2016.10.043

[12] Zhang, Q., Barri, K., Babanajad, S. K., & Alavi, A. H. (2021). Real-time detection of cracks on concrete bridge decks using deep learning in the frequency domain. *Engineering*, *7* (12), 1786–1796. https://doi.org/10.1016/j.eng.2020.07.026

[13] Abraham, J. N., Tran, M. Q., Jayaraj, J. S., Matos, J. C., Valluzzi, M. R., & Dang, S. N. (2026). Unsupervised learning-based anomaly detection for bridge structural health monitoring: Identifying deviations from normal structural behaviour. *Sensors*, *26*(2), 561. https://doi.org/10.3390/s26020561

[14] Kim, S., Kim, H., & Hwang, Y. (2025). Data-driven dynamic response forecasting and anomaly detection in long-span bridges. *Journal of Civil Structural Health Monitoring*, *15*, 3045–3062. https://doi.org/10.1007/s13349-025-00952-8

[15] Alves, V. H. M., Gomes, R. F. I., & Cury, A. (2025). New perspectives on structural health monitoring using unsupervised quantum machine learning. *Mechanical Systems and Signal Processing*, *229*, 112489. https://doi.org/10.1016/j.ymssp.2025.112489

[16] Grosse, C. U., & Ohtsu, M. (2008). *Acoustic emission testing*. Springer Science & Business Media. https://doi.org/10.1007/978-3-540-69972-9

[17] Oppenheim, A. V., & Schafer, R. W. (2010). *Discrete-time signal processing* (3rd ed.). Prentice Hall. https://www.pearson.com/en-us/subject-catalog/p/discrete-time-signal-processing/P200000000305/9780134681171

[18] Mohanty, S., & Pani, S. K. (2026). Intelligent crack detection in building structures using coupled ultrasonic guided wave and acoustic emission sensing. *International Journal of Civil Engineering*, *13*(2), 226–246. https://doi.org/10.14445/23488352/IJCE-V13I2P116

[19] Abdul Kudus, S., Misnan, M. F., Muhammad, H., Sawahli, A. M. S., Ahmad, Z., & Shahid, K. A. (2026). Acoustic emission-based assessment of damage evolution in bridge systems: Methods, indicators, and advances in structural mechanics. *Journal of Nondestructive Evaluation*, *45*, 100. https://doi.org/10.1007/s10921-026-01388-w

[20] Raissi, M., Perdikaris, P., & Karniadakis, G. E. (2019). Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations. *Journal of Computational Physics*, *378*, 686–707. https://doi.org/10.1016/j.jcp.2018.10.045

[21] Schuld, M., Bergholm, V., Gogolin, C., Izaac, J., & Killoran, N. (2019). Evaluating analytic gradients on quantum hardware. *Physical Review A*, *99*(3), 032331. https://doi.org/10.1103/PhysRevA.99.032331

[22] Grant, E., Wossnig, L., Ostaszewski, M., & Severini, S. (2019). An initialization strategy for addressing barren plateaus in parametrized quantum circuits. *Quantum*, *3*, 214. https://doi.org/10.22331/q-2019-12-09-214

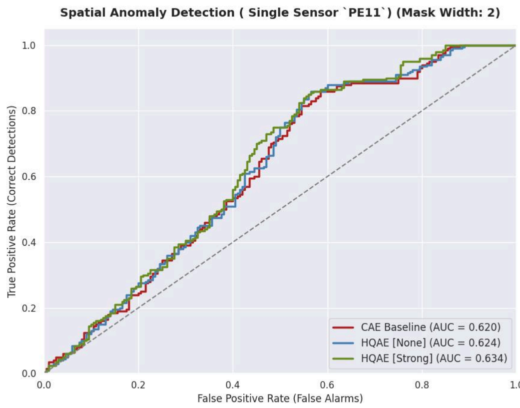
[23] Kingma, D. P., & Ba, J. (2015). Adam: A method for stochastic optimization. *3rd International Conference on Learning Representations (ICLR)*, San Diego, CA. https://arxiv.org/abs/1412.6980

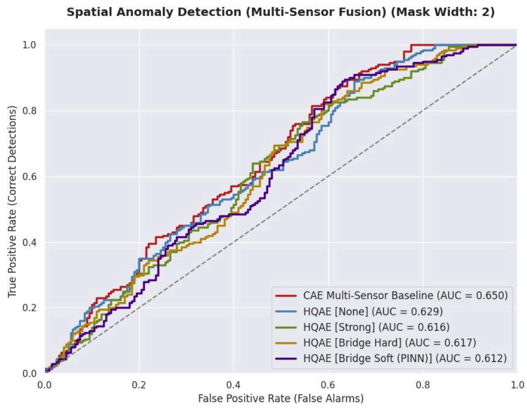
[24] Pascanu, R., Mikolov, T., & Bengio, Y. (2013). On the difficulty of training recurrent neural networks. *Proceedings of the 30th International Conference on Machine Learning (ICML)*, *28*, 1310–1318. https://proceedings.mlr.press/v28/pascanu13.html

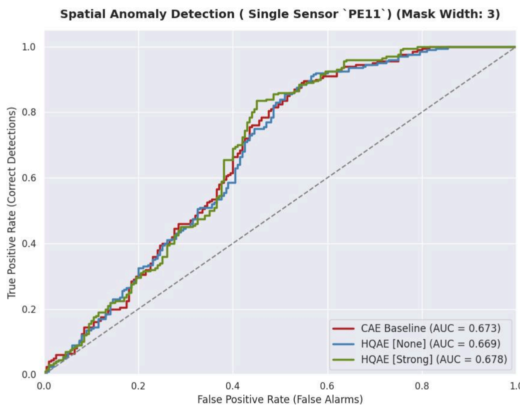
[25] Rose, J. L. (2014). *Ultrasonic waves in solid media*. Cambridge University Press. https://assets.cambridge.org/97811070/48959/frontmatter/9781107048959_frontmatter.pdf

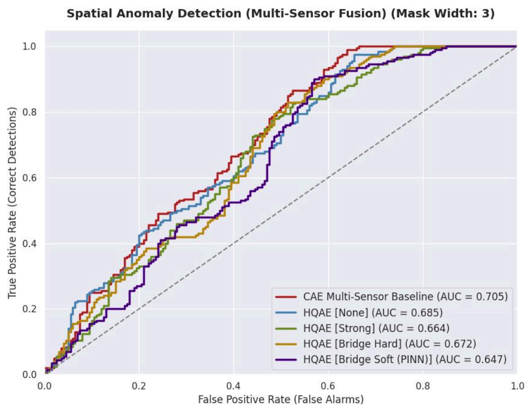
Appendix A: Extended ROC-AUC Sensitivity Curves

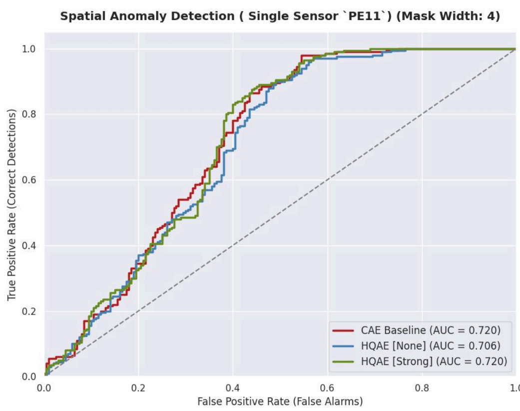
The following sensitivity curves detail the ROC-AUC performance across all evaluated damage severity levels (Mask 2 through Mask 8). The left column illustrates the single-sensor baseline ablations, while the right column demonstrates the enhanced stabilization achieved via multi-sensor spatial fusion.

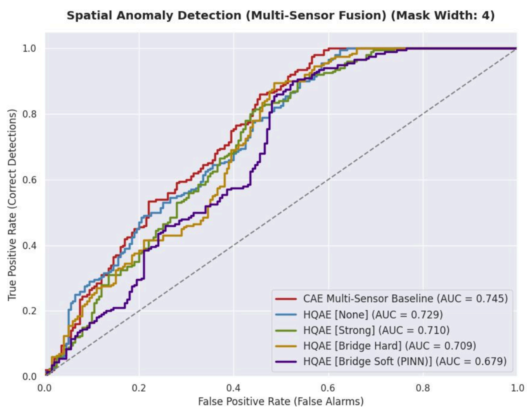


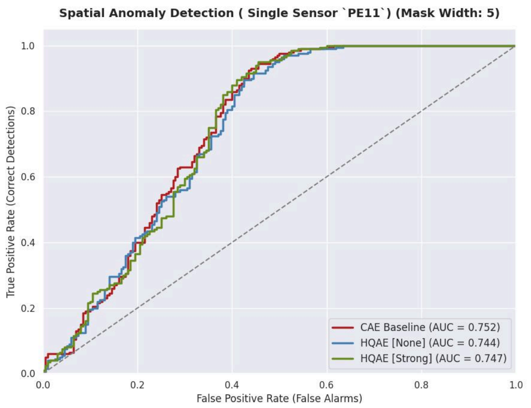


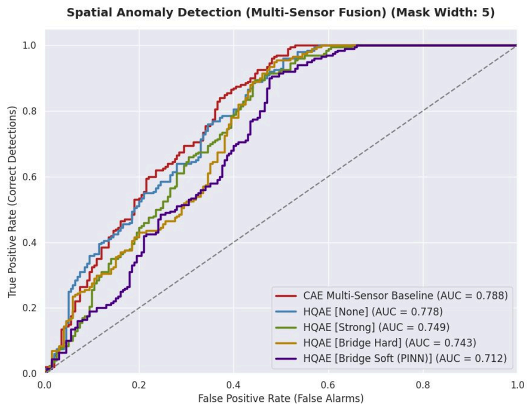


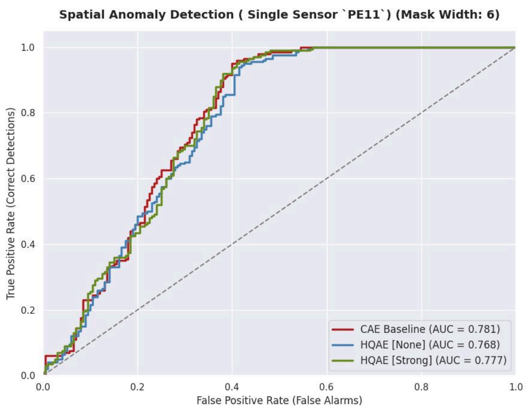


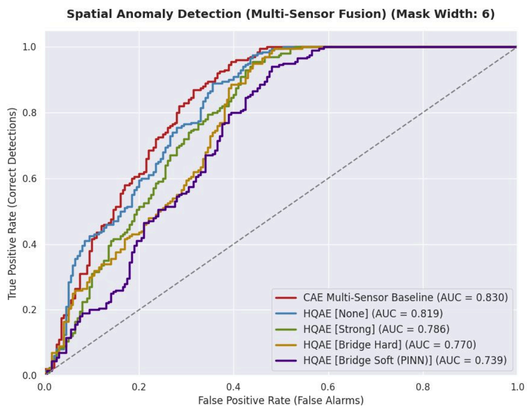


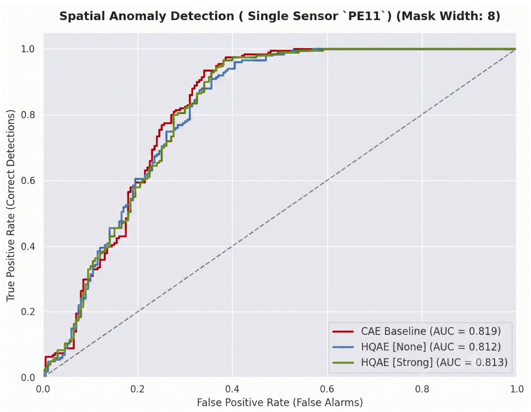


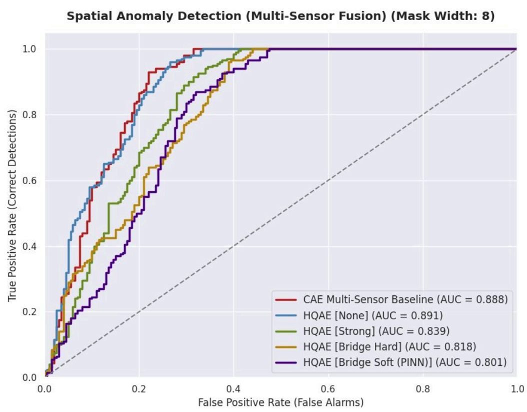

Spatial Anomaly Detection (Single Sensor `PE11`) (Mask Width: 5)
True Positive Rate (Correct Detections)
False Positive Rate (False Alarms)
CAE Baseline (AUC = 0.752)
HQAE [None] (AUC = 0.744)
HQAE [Strong] (AUC = 0.747)


Spatial Anomaly Detection (Multi-Sensor Fusion) (Mask Width: 5)
True Positive Rate (Correct Detections)
False Positive Rate (False Alarms)
CAE Multi-Sensor Baseline (AUC = 0.788)
HQAE [None] (AUC = 0.778)
HQAE [Strong] (AUC = 0.749)
HQAE [Bridge Hard] (AUC = 0.743)
HQAE [Bridge Soft (PINN)] (AUC = 0.712)


Spatial Anomaly Detection (Single Sensor `PE11`) (Mask Width: 6)
True Positive Rate (Correct Detections)
False Positive Rate (False Alarms)
CAE Baseline (AUC = 0.781)
HQAE [None] (AUC = 0.768)
HQAE [Strong] (AUC = 0.777)


Spatial Anomaly Detection (Multi-Sensor Fusion) (Mask Width: 6)
True Positive Rate (Correct Detections)
False Positive Rate (False Alarms)
CAE Multi-Sensor Baseline (AUC = 0.830)
HQAE [None] (AUC = 0.819)
HQAE [Strong] (AUC = 0.786)
HQAE [Bridge Hard] (AUC = 0.770)
HQAE [Bridge Soft (PINN)] (AUC = 0.739)


Spatial Anomaly Detection (Single Sensor `PE11`) (Mask Width: 8)
True Positive Rate (Correct Detections)
False Positive Rate (False Alarms)
CAE Baseline (AUC = 0.819)
HQAE [None] (AUC = 0.812)
HQAE [Strong] (AUC = 0.813)


Spatial Anomaly Detection (Multi-Sensor Fusion) (Mask Width: 8)
True Positive Rate (Correct Detections)
False Positive Rate (False Alarms)
CAE Multi-Sensor Baseline (AUC = 0.888)
HQAE [None] (AUC = 0.891)
HQAE [Strong] (AUC = 0.839)
HQAE [Bridge Hard] (AUC = 0.818)
HQAE [Bridge Soft (PINN)] (AUC = 0.801)