



SAPIENZA
UNIVERSITÀ DI ROMA

Advanced machine learning methods for damage detection in nonlinear dynamic systems

PhD in Structural Engineering
Department of Structural and Geotechnical Engineering (XXXVIII cycle)

Harrish Joseph

ID number 393462

Advisor

Prof. Walter Lacarbonara

Co-Advisor

Prof. Giuseppe Quaranta

Prof. Biagio Carboni

Academic Year 2022 - 2025

Thesis defended on 15 January 2026
in front of a Board of Examiners composed by:
Prof. Elena Benvenuti (chairman)
Prof. Filomena de Silva
Prof. Enrico Spacone

Advanced machine learning methods for damage detection in nonlinear dynamic systems

Sapienza University of Rome

© 2025 Harrish Joseph. All rights reserved

This thesis has been typeset by $\text{\LaTeX}$ and the Sapthesis class.

Author's email: harrish.joseph@uniroma1.it

Abstract

Structural Health Monitoring (SHM) is an interdisciplinary engineering field dedicated to the autonomous, real-time assessment of structural integrity, with critical applications ranging from civil infrastructure to aerospace, mechanical or more general engineering systems. A primary objective within SHM is the early detection of damage to ensure operational safety and optimize maintenance. In its modern standing, SHM leverages dense arrays of sensors that generate complex, multivariate data streams. These data are often captured under random excitations that constitute real-world operational conditions, and the structures however frequently exhibit nonlinear dynamic behavior. This combination of high-dimensional data and inherent nonlinearity poses a significant challenge, as traditional monitoring techniques based on linear system theory struggle to deduce meaningful insights from such complex responses. While deep learning has emerged as a powerful paradigm uniquely capable of extracting salient features from these datasets, its application has largely been confined to linear systems. This represents an important research gap, compounded by the practical scarcity of labeled data for damaged states, which necessitates the development of robust unsupervised learning methods.

This dissertation directly addresses this gap by presenting a systematic exploration of advanced deep learning architectures applied to a suite of challenging nonlinear systems. These benchmarks, subjected to random excitations, encompass numerical models with cubic stiffness, path-dependent hysteresis, and complex hybrid nonlinearities, and are further validated on physical experiments including a magneto-elastic beam and a large-scale lattice tower. To analyze these complex systems, the research follows a sought architectural progression. It starts with an investigation into 1D Convolutional Neural Network (CNN)-based models, where the application of a Convolutional Autoencoder and a Generative Adversarial Network successfully establishes the fundamental viability of using unsupervised feature learning to detect the onset of damage in nonlinear dynamics. This foundational study provides a critical contribution by validating these architectures in a domain where they have been seldom explored.

Building upon the insights from the convolutional models, which are powerful but offer limited diagnostic interpretability, the research culminates in the development of the SensorFusion Temporal Fusion Transformer (SensorFusionTFT). This novel and interpretable architecture is specifically adapted for multi-sensor based SHM. This progression from detection to diagnostics is enabled by a robust, multi-metric framework. Damage is identified through a synergistic suite of three distinct metrics: the foundational reconstruction error (RMSE), a highly sensitive Interval Score that quantifies the violation of learned uncertainty bounds, and a novel Attention

Anomaly Score derived from the model’s internal reasoning. The successful implementation of this framework is demonstrated across all benchmarks, showing a clear advancement in both sensitivity and diagnostic capability. This work thus establishes a complete pathway from foundational unsupervised models to a state-of-the-art, interpretable, and multi-metric Transformer-based solution, providing a powerful and practical framework for the next generation of SHM systems for real-world, nonlinear structures.

Related Publications

The research presented in this dissertation has contributed to the following peer-reviewed journal articles and manuscripts in preparation. The core content of the thesis chapters is derived from this body of work.

Journal Publications

1. Joseph, H., Quaranta, G., Carboni, B., & Lacarbonara, W. (2024). Deep learning architectures for data-driven damage detection in nonlinear dynamic systems under random vibrations. *Nonlinear Dynamics*, 112(23), 20611–20636.
Contribution: The work presented in this article forms the primary basis of Chapter 3 of this dissertation.

Manuscripts in Preparation

2. Joseph, H., Quaranta, G., Carboni, B., & Lacarbonara, W. (2025). SensorFusionTFT: A Dynamic Variable Selection Transformer for Unsupervised Damage Detection in Nonlinear Structural Dynamics. Manuscript under review at *Mechanical Systems and Signal Processing*.
Contribution: The work described in this manuscript forms the primary basis of Chapter 4 of this dissertation.

Acknowledgments

My profound gratitude goes to Prof. Walter Lacarbonara, whose trust and guidance gave me the opportunity to begin and carry forward this PhD journey. His leadership and approach to research have profoundly influenced the way I think, question, and solve problems.

My sincere thanks also go to Prof. Giuseppe Quaranta for his constant support and insightful guidance throughout the course of this research. Your steady motivation, encouragement, and willingness to discuss ideas and results, along with your habit of challenging me to refine and elevate my work were instrumental to the progress and quality of this research.

I would also like to thank Prof. Biagio Carboni for his involvement and support during my doctoral journey.

I truly appreciate Harshal for the friendship, companionship, encouragement, and constant motivation, especially during the difficult times we faced throughout this journey over the past eight years. Your support has meant a great deal to me. My heartfelt thanks to my colleagues and friends Pranath, Vinay and Sawan with whom I shared this doctoral journey over the past four years. Working alongside you made this experience not only productive but truly enjoyable.

To my girlfriend, Francesca, thank you for your love, trust, respect and unwavering support. Your belief in me has been a constant source of strength.

Finally, I extend my deepest gratitude to my family. To my father, for being a steady source of strength and stability guiding us through life's challenges; to my sister, for her constant encouragement; and to my mother, whose unconditional belief in me, even when no one else did, has shaped who I am today, your unwavering support at every step of my life is the foundation upon which I stand, and it is to you that I dedicate this thesis.

Contents

Abstract	iii
Related Publications	iv
Acknowledgments	v
Nomenclature	xiii
Acronyms	xvi
1 Introduction	1
1.1 The Domain of Structural Health Monitoring (SHM)	2
1.1.1 Aims and Scopes of SHM	2
1.1.2 Damage Assessment Hierarchy: Identification, Localization, Quantification, and Prognosis	3
1.2 Handling Nonlinearities in Structural Health Monitoring	6
1.2.1 Sources and Effects of Nonlinearity in Structural Systems . .	6
1.2.2 A Critical Review of Existing Damage Detection Methods . .	7
1.3 The Rise of Data-Driven and Machine Learning Methods for SHM .	10
1.4 Research Objectives and Scientific Contributions	12
2 Theoretical Foundations	15
2.1 Numerical Archetypes of Nonlinear Dynamic Systems	16
2.1.1 Duffing Oscillator: Cubic Stiffness Nonlinearity	16
2.1.2 Bouc-Wen Model: Hysteretic Nonlinearity	17
2.1.3 Seismic Nonlinear Isolator: Hybrid Nonlinearity	21
2.2 Experimental Case Studies of Nonlinear Systems	23
2.2.1 The Magneto-Elastic Beam System	23
2.2.2 The LUMO Lattice Tower Benchmark	24
2.2.3 Common Challenges in Unsupervised SHM Data and Simulation	26
2.3 Deep Learning Architectures for SHM	28
2.3.1 The Convolution Operator for Feature Extraction	29
2.3.2 Unsupervised Representation Learning with Convolutional Autoencoders	31
2.3.3 Unsupervised Anomaly Detection with Convolutional Genera- tive Adversarial Networks	32

2.3.4	Sequential Data Modeling with Attention Mechanisms and Transformer Architectures	34
2.4	Methodology for Unsupervised Damage Detection	40
2.4.1	The Unsupervised Learning Paradigm	41
2.4.2	Data Preparation and Pre-processing	41
2.4.3	Formulation of Training Objectives and Loss Functions	43
2.4.4	Inference and Formulation of Anomaly Scores	44
3	Damage Detection using Convolutional Architectures	48
3.1	Case Study Data Protocol	49
3.2	Application to a 1-DOF Duffing Oscillator	49
3.2.1	Autoencoder Performance	49
3.2.2	Generative Adversarial Network Performance	50
3.2.3	Comparative Discussion	51
3.2.4	On the Role of the Input Representation: 2D GAN with Wavelet Transforms	52
3.3	Application to a 2-DOF Nonlinear System	53
3.3.1	Autoencoder Performance	54
3.3.2	Generative Adversarial Network Performance	54
3.3.3	Comparative Discussion	55
3.4	Application to a Nonlinear Seismic Isolator	56
3.4.1	Autoencoder Performance	58
3.4.2	Generative Adversarial Network Performance	58
3.4.3	Comparative Discussion	59
3.5	Experimental Validation on a Magneto-Elastic Beam	60
3.5.1	Autoencoder Performance	61
3.5.2	Generative Adversarial Network Performance	61
3.5.3	Comparative Discussion	62
3.6	Operational Limitations and Critical Analysis	63
3.6.1	Black-Box Approaches vs. Traditional Statistical Methods	64
3.6.2	The Instability of Adversarial Detection	64
3.6.3	Noise Sensitivity and the Limit of Detection	65
3.6.4	Saturation and Non-Monotonicity	65
3.7	Summary and Findings	65
4	Damage Detection using an Interpretable Sensor Fusion Transformer	68
4.1	Case Study Protocol for Transformer-Based Analysis	69
4.2	Application to a Hysteretic Tuned Mass Damper	69
4.2.1	Performance as a Dynamic System Surrogate	70
4.2.2	Quantitative Damage Detection via Anomaly Metrics	70
4.2.3	Diagnostic Insights from Interpretability Mechanisms	71
4.3	Revisiting the Hysteretic Seismic Isolator	72
4.3.1	Performance as a Dynamic System Surrogate	73
4.3.2	Quantitative Damage Detection via Anomaly Metrics	74
4.3.3	Diagnostic Insights from Interpretability Mechanisms	75
4.4	Experimental Validation on the Magneto-Elastic Beam	77

4.4.1	Performance on Experimental Data	78
4.4.2	Quantitative Damage Detection via Anomaly Metrics	79
4.4.3	Interpretability and Damage Localization Insights	79
4.5	Large-Scale Experimental Validation: The LUMO Lattice Tower . .	81
4.5.1	Damage Detection with Reconstruction-Based Metrics	82
4.5.2	Diagnostic Insights from Internal Mechanisms and Localization	82
4.6	Summary and Findings	82
5	Conclusion and Future Research	86
5.1	Summary of Main Findings	87
5.1.1	Critical Comparative Evaluation of the Investigated Methods	88
5.1.2	General Guidelines for Applying Deep Learning to Nonlinear SHM	88
5.2	Limitations and Recommendations for Future Research	90
5.3	Concluding Remarks	91
	Bibliography	92

List of Figures

1.1	Hierarchical representation of damage assessment levels in SHM. Adapted from Rytter's framework [68], illustrating the progression from data-driven detection to physics-based prognosis.	4
1.2	The conceptual framework of the research. The diagram maps the evolution from traditional linear and nonlinear physics-based methods (Top) to the proposed data-driven framework (Bottom). It highlights the dead ends in traditional machine learning caused by manual feature engineering and data scarcity, motivating the focus on Un-supervised Deep Learning. The thesis progression is visualized in the lower section, advancing from black-box Convolutional Neural Networks (Phase I) to the interpretable, memory-aware SensorFusionTFT (Phase II), specifically engineered to handle the complexities of nonlinear structural dynamics.	14
2.1	Frequency response curves (FRCs) of the 1-DOF Duffing oscillator, illustrating the hardening nonlinearity and amplitude-dependent resonance for different damage levels.	18
2.2	Frequency response curves (FRCs) of the 2-DOF Duffing system, showing complex modal interactions and bent resonant peaks for different damage levels.	19
2.3	Layout of the magneto-elastic beam experiment, consisting of a cantilever beam, a permanent magnet to induce nonlinearity, and a bolted connection for simulating damage. Figure adapted from [85].	25
2.4	The LUMO experimental test setup, showing the 9-meter lattice tower, sensor measurement levels (MLs), and damage locations (DAMs). The complexity and environmental exposure of this benchmark provide a realistic test for SHM algorithms. Figure adapted from [91].	27
2.5	Architecture of the 1D Convolutional Autoencoder (AE) used for unsupervised damage detection. The encoder compresses the input signal into a latent space, and the decoder reconstructs it.	33
2.6	Architecture of the 1D Convolutional Generative Adversarial Network (GAN) used for damage detection. The Generator creates synthetic signals from noise, while the Discriminator learns to distinguish real (healthy) signals from synthetic/anomalous ones.	35

2.7	Core components of the SensorFusionTFT architecture: (a) Gated Residual Network (GRN), (b) Gated Linear Unit (GLU), (c) Positional Encoding, and (d) Interpretable Multi-Head Attention.	39
2.8	The proposed SensorFusionTFT architecture. The model dynamically fuses information from $\mathcal{N}$ sensors via a Time-Varying Sensor Fusion Network and processes the resulting unified time series through a Temporal Fusion Decoder to capture long-range dependencies.	40
3.1	Reconstruction of the 1-DOF Duffing system response via the Autoencoder for different damage levels. The mismatch between the actual and reconstructed response grows with increasing damage.	50
3.2	Fidelity of the trained GAN Generator. Comparison of a real time-history response from the undamaged 1-DOF system with a synthetic response from the Generator, shown in both the time and time-frequency domains.	51
3.3	Comparative damage sensitivity for the 1-DOF system. Top row: Scatter plots of the raw anomaly scores (AE reconstruction error and GAN discriminator output). Bottom row: Trend of the normalized average anomaly scores, showing a clear correlation with damage level.	52
3.4	Comparative evaluation of input representations for the 1-DOF Duffing system. Left: 1D-GAN operating on time-series. Right: 2D-GAN operating on CWT Scalograms. The 2D approach yields higher sensitivity (+76.5% vs +65.1%) by leveraging spectro-temporal features.	54
3.5	Reconstruction of the 2-DOF Duffing system response via the Autoencoder for different damage levels. The model is applied to each degree-of-freedom independently.	55
3.6	Fidelity of the trained GAN Generator for the 2-DOF system. Comparison of real vs. synthetic time-frequency representations for both degrees of freedom in the undamaged state.	56
3.7	Comparative damage sensitivity for the 2-DOF system. While both AE and GAN show a clear correlation with damage, the AE demonstrates higher sensitivity in this more complex case.	57
3.8	Reconstruction of the seismic isolator response via the Autoencoder for different damage levels. The reconstruction error grows as the damage severity increases.	58
3.9	Fidelity of the trained GAN Generator for the seismic isolator. The time-frequency comparison reveals that the Generator captures the main dynamics of the real signal.	59
3.10	Comparative damage sensitivity for the seismic isolator. The AE maintains satisfactory sensitivity across all damage levels, whereas the GAN's sensitivity is significantly reduced, only detecting damage at higher severity.	60
3.11	Reconstruction of the experimental magneto-elastic system response via the Autoencoder for different physical damage levels.	61
3.12	Fidelity of the trained GAN Generator on experimental data. The time-frequency comparison shows a high similarity between the actual undamaged response and the synthetic data generated by the GAN.	62

3.13	Comparative damage sensitivity on the magneto-elastic experimental data. Both architectures demonstrate high sensitivity to the incipient 2% damage level, confirming their utility for early damage detection.	63
4.1	Comparative predictions of the Total Restoring Force for the TMD. The divergence between the ground truth (solid) and the model's prediction (dashed) grows with damage severity, signaling an anomaly.	70
4.2	Normalized distribution of RMSE for the TMD's Total Restoring Force. Each box plot shows the distribution of scores across all prediction windows for a given damage case. The red line connects the mean RMSE value of each distribution, illustrating the clear and monotonic trend of increasing average reconstruction error with damage severity.	71
4.3	Normalized distribution of the Interval Score (IS) for the TMD's Total Restoring Force. Each box plot represents the distribution of scores for a given damage case. The red line, which connects the mean IS value of each distribution, illustrates the dramatic and monotonic increase in the mean score as damage severity increases.	72
4.4	Interpretability dashboard for the TMD application. This three-panel output shows (Top) which sensors the model focuses on, (Middle) how its temporal reasoning changes with damage, and (Bottom) a quantitative score of that change, all showing a clear correlation with damage.	73
4.5	Comparative prediction performance for the nonlinear isolator. The model's predictions for Displacement (left) and Total Restoring Force (right) are compared against the ground truth (solid black) across four damage states.	74
4.6	Normalized distribution of RMSE by damage case for the isolator. The trend lines show a clear, monotonic increase in mean reconstruction error distribution with damage severity for both target variables.	75
4.7	Normalized distribution of Interval Score (IS) by damage case for the isolator. The sharp, exponential growth of the mean IS with increasing damage highlights its high sensitivity as a damage indicator.	76
4.8	Interpretability dashboard for the seismic isolator. The VSN's focus on acceleration (top), the deviation in temporal attention (middle), and the rising Attention Anomaly Score (bottom) provide a multi-faceted and interpretable view of the damage detection process.	77
4.9	Comparative prediction performance for the magneto-elastic beam. The divergence between the ground truth (solid) and the model's prediction of the healthy response (dashed) serves as the primary damage indicator.	78
4.10	Normalized distribution of RMSE for the magneto-elastic beam. The trend of mean value for each damage distribution demonstrates a clear increase in reconstruction error with physical damage severity.	80

4.11	Normalized distribution of Interval Score (IS) for the magneto-elastic beam. The pronounced increase of mean values of IS highlights the model's sensitivity to anomalous behavior that violates its learned uncertainty bounds.	80
4.12	Interpretability dashboard for the magneto-elastic beam. The VSN's focus on the sensor nearest the damage (top), the deviation in temporal attention (middle), and the rising Attention Anomaly Score (bottom) demonstrate the model's diagnostic capabilities.	81
4.13	Distribution of normalized RMSE scores across all sensors and structural states for the LUMO tower. A clear increase in the magnitude and variance of the errors is observable for the damaged cases. . . .	83
4.14	Distribution of normalized Interval Scores (IS) for the LUMO tower. The damaged cases exhibit a significantly larger and more widespread distribution of scores, indicating a fundamental breakdown in the model's learned uncertainty.	83
4.15	Distribution of the Attention Anomaly Score by damage case for the LUMO tower. The statistically significant increase in the score for all damage scenarios confirms its utility as a robust, complementary damage indicator.	84

List of Tables

3.1	Configuration of the deep learning architectures for the 1-DOF Duffing system.	52
3.2	Configuration of the deep learning architectures for the 2-DOF Duffing system.	57
3.3	Configuration of the deep learning architectures for the magneto-elastic system.	63
5.1	Critical comparative evaluation of the investigated deep learning architectures.	89

Nomenclature

$\mathbf{M}$	Mass coefficient or matrix
$\mathbf{C}$	Viscous damping coefficient or matrix
$\mathbf{K}$	Linear stiffness coefficient or matrix
$\mathbf{x}(t)$	Displacement at time t
$\dot{\mathbf{x}}(t)$	Velocity at time t
$\ddot{\mathbf{x}}(t)$	Acceleration at time t
$\mathbf{F}(t)$	External excitation force at time t
d	Damage level (in percent)
Ω	Excitation frequency
t	Time variable
K_1, K_3	Linear and cubic stiffness coefficients for the Duffing oscillator
f_{res}	Total restoring force
f_i, f_n, f_s	Hysteretic, negative stiffness, and superelastic force components
$z, \tilde{z}$	Internal hysteretic state variable (Bouc-Wen model)
α, β, γ, n	Parameters of the Bouc-Wen model of hysteresis
K_i, K_n, K_s, K_m	Stiffness coefficients for the seismic isolator model
Y	Yielding force in the superelastic model
$\mathbf{x}$	Input data vector or time-series
$\tilde{\mathbf{x}}$	Reconstructed output vector
$\mathbf{h}$	Latent space representation (Autoencoder)
$\mathbf{z}$	Random noise vector (GAN)
$g(\cdot)$	Encoder function
$f(\cdot)$	Decoder function

$G(\cdot)$	Generator network function
$D(\cdot)$	Discriminator network function
θ	Vector of trainable model parameters (weights and biases)
ℓ	Layer index in a neural network
$\mathbf{K}_i^\ell$	Kernel (convolutional filter) of the i -th filter in layer ℓ
b_i^ℓ	Bias term of the i -th filter in layer ℓ
$y(j)$	Pre-activation output at position j
$a(j)$	Post-activation output at position j
$\mathbf{Q}, \mathbf{K}, \mathbf{V}$	Query, Key, and Value matrices in the attention mechanism
N_H	Number of attention heads
d_{model}	Hidden dimension of the Transformer model embeddings
$\mathbf{e}_j$	Embedding vector for sensor j
$\mathbf{w}_t$	Vector of VSN sensor importance weights at time t
$\mathbf{f}_t$	Fused embedding vector at time t
W	Window size for CNN pooling or TFT look-back
F	Forecast horizon length for TFT
N	Number of sensors
$\sigma(\cdot)$	Sigmoid activation function
$\odot$	Element-wise multiplication (Hadamard product)
$\mathcal{L}$	General loss function
λ	L_2 regularization coefficient
η	Learning rate for the optimizer
β_1, β_2	Exponential decay rates for the ADAM optimizer moments
$\mathbf{m}_\tau, \mathbf{v}_\tau$	First and second moment vectors in the ADAM optimizer
τ	Quantile level for probabilistic forecasting
$\hat{y}^{(\tau)}$	Model's prediction for the τ -th quantile
S_{AE}, S_{GAN}	Anomaly scores for the AE and GAN architectures
$S_{RMSE}, S_{IS}, S_{ROI}$	RMSE, Interval Score, and ROI-based anomaly scores for the TFT
$\mathcal{R}_{hot}, \mathcal{R}_{cold}$	Hotspot and Coldspot masks for the Attention Anomaly Score

Acronyms

SHM	Structural Health Monitoring
DOF	Degree-of-Freedom
FEM	Finite Element Model
FRF	Frequency Response Function
RUL	Remaining Useful Life
TMD	Tuned Mass Damper
LUMO	Leibniz University Test Structure for Monitoring
VRMS	Voltage Root Mean Square
LTI	Linear Time-Invariant
NNM	Nonlinear Normal Mode
HHT	Hilbert-Huang Transform
IMF	Intrinsic Mode Function
ML	Machine Learning
DL	Deep Learning
CNN	Convolutional Neural Network
RNN	Recurrent Neural Network
LSTM	Long Short-Term Memory
NN	Neural Network
GPU	Graphics Processing Unit
AE	Autoencoder
GAN	Generative Adversarial Network
TFT	Temporal Fusion Transformer
VSN	Variable Selection Network

GLU	Gated Linear Unit
GRN	Gated Residual Network
ReLU	Rectified Linear Unit
RMSE	Root Mean Square Error
MAE	Mean Absolute Error
IS	Interval Score
ROI	Region of Interest

Chapter 1

Introduction

This chapter establishes the foundational context and core motivation for the research presented in this dissertation. It begins by defining the broad domain of Structural Health Monitoring (SHM), outlining its primary objectives and the established hierarchical framework for damage assessment. The discussion then narrows to the central challenge that drives this work, the pervasive and complex issue of nonlinearity in structural dynamics, a phenomenon that fundamentally limits the efficacy of traditional monitoring techniques. Through a critical review of existing methods, a distinct research gap is identified. This gap motivates the subsequent turn to data-driven methods, leading to the formulation of this dissertation's specific research objectives and a clear statement of its scientific contributions to the field of SHM for nonlinear systems.

1.1 The Domain of Structural Health Monitoring (SHM)

SHM has emerged as a critical engineering discipline aimed at implementing a damage detection and characterization strategy for aerospace, civil, and mechanical engineering infrastructure [26, 77]. The fundamental objective of SHM is to manage risks and make appropriate decisions associated with an asset, including providing information for risk assessment and understanding how those risks might develop over time [4]. By instrumenting structures with sensors and analyzing the resulting data, SHM provides a data-driven paradigm for moving beyond traditional schedule-based maintenance towards condition-based interventions, thereby enhancing safety, reducing operational costs, and extending the service life of critical infrastructure [78].

1.1.1 Aims and Scopes of SHM

The primary goals of structural health monitoring can be categorized into three main aspects: design validation, assessment of structural performance, and the improvement of asset management [4].

Design Validation. SHM serves to verify whether a built asset behaves as planned during its design stages. This validation process is distinct across engineering disciplines but shares the common goal of model correlation.

- *Aerospace Engineering.* Design validation is critical for flight certification. This is typically achieved through Ground Vibration Testing (GVT) on full-scale aircraft to validate aeroelastic models and ensure flutter margins are safe before the first flight. Similarly, for space systems, rigorous acoustic and vibration testing is performed on satellites and launch vehicles to ensure survivability during the extreme dynamic loads of launch and deployment.
- *Mechanical Engineering.* In sectors such as automotive or power generation, validation involves experimental modal analysis of prototypes such as wind turbine blades or automotive chassis to ensure the dynamic properties (natural frequencies and damping) align with design specifications before mass production.

- *Civil Engineering.* Validation is often carried out immediately after the completion of a structure, sometimes during the construction state.. For instance, controlled load tests are performed on bridges to confirm their response characteristics match the Finite Element (FE) models.

Assessment of Structural Performance. This is the core function of SHM, conducted throughout the service life and operation of an asset. The development of SHM techniques is oriented toward the identification of features sensitive to small levels of damage such as fatigue cracks in aircraft fuselages, gear faults in rotating machinery, or degradation in reinforced concrete elements and the ability to distinguish these effects from those due to changing ambient and operational conditions. In civil structures, these conditions include temperature and traffic loading; in aerospace and space systems, they encompass varying flight envelopes, extreme thermal cycling in orbit, and potential hypervelocity impacts from micrometeoroids; while in mechanical systems, they involve fluctuating rotational speeds and torque [4, 46]. Consequently, the development of statistical models is essential to foster a better comprehension of damaged versus undamaged configurations under these varying states. A critical aspect of this is the optimization of sensor placement to accurately capture structural behavior and detect potential damage [4, 46].

Improving Asset Management. By providing continuous information about how an asset or its elements are behaving, SHM enables a deeper understanding of trends in physical state changes over time. This facilitates the optimization of interventions and maintenance strategies, allowing for a shift from rigid schedule-based maintenance to predictive or Condition-Based Maintenance (CBM). For aerospace and industrial mechanical systems, this evolution is crucial for reducing aircraft downtime and preventing costly unplanned outages in power generation assets, effectively avoiding future malfunctions and reducing costs associated with loss of structural performance [4].

The scope of SHM is vast, encompassing a wide range of structures from civil infrastructure such as bridges, tunnels, and buildings [65], to aerospace systems like aircraft and reusable launch vehicles, and mechanical systems such as wind turbines [27]. The technology has gained significant attention due to its capability of providing information about the global state of a structure, often without prior knowledge of the damage location and without needing physical access to the damaged portion [45].

1.1.2 Damage Assessment Hierarchy: Identification, Localization, Quantification, and Prognosis

The process of damage assessment in SHM is systematically organized into a hierarchy of increasing complexity and detail, a framework first formally proposed by Rytter [68], which answers a series of fundamental questions at increasing levels of sophistication [11, 66].

Level I: Detection. This foundational level addresses the question: is damage present in the system? Traditional vibration-based methods often rely on the

assumption of a linear behavior and then on the changes in the structure's global modal properties, such as natural frequencies [22, 24]. Since damage typically causes a reduction in structural stiffness, this results in decreased natural frequencies [21]. However, natural frequencies have been shown to be relatively insensitive to localized, incipient damage, and multiple distinct damage scenarios can produce similar frequency shifts, making them insufficient for precise identification [21].

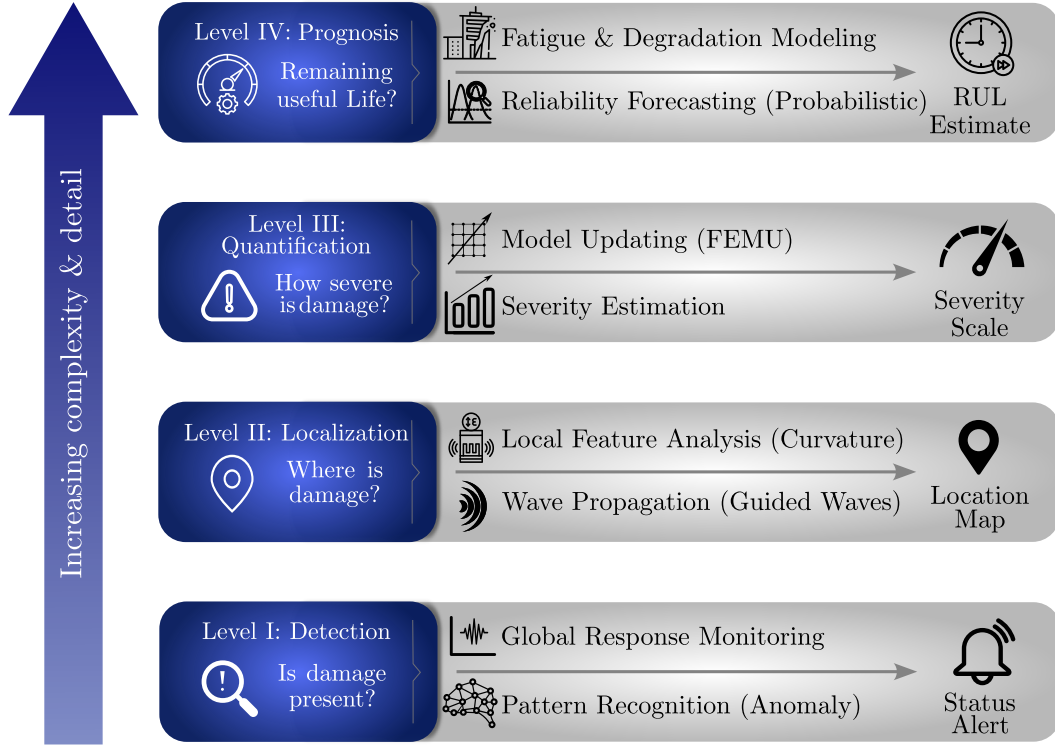


Figure 1.1. Hierarchical representation of damage assessment levels in SHM. Adapted from Rytter's framework [68], illustrating the progression from data-driven detection to physics-based prognosis.

Level II: Localization. Once damage has been detected, this level seeks to determine its physical location. Traditional tools for localization are diverse and include the following.

- *Modal curvature methods.* Curvature mode shapes are significantly more sensitive to local damage than displacement mode shapes. Changes in curvature are localized in the region of damage, providing a powerful tool for pinpointing its location [57].
- *Modal strain energy methods:* These methods utilize the change in modal strain energy distribution between the healthy and damaged states. Decomposing the strain energy into its constituent components (e.g., axial and transverse) can further enhance localization accuracy in complex structures [74, 42].
- *Frequency Response Function (FRF) methods.* FRFs are highly sensitive to local changes in mass, stiffness, and damping. Methods based on analyzing

changes in the FRF data have been successfully applied to locate damage in beam-like structures [40].

- *Ultrasonic guided wave methods.* These active-sensing techniques use high-frequency elastic waves that can travel long distances and are highly sensitive to small defects like cracks. They are particularly effective for inspecting pipelines, composite aerospace structures, and critical connections in civil infrastructure [79, 52]. Recent advances in this area include the use of deep learning to analyze the complex wave scattering patterns for improved localization [58].

Level III: Quantification. This level aims to assess the severity or extent of the damage. This is an exceptionally challenging task, with traditional methods including the following.

- *Finite Element Model Updating (FEMU).* This technique involves creating a high-fidelity numerical model of the structure and iteratively adjusting its parameters (e.g., elemental stiffness) to match the measured dynamic response of the damaged structure [3, 41]. The magnitude of the parameter adjustments quantifies the damage severity. Iterative methods based on modal strain energy have also been developed to quantify damage through successive updates [87].
- *Real-time quantification methods.* For laboratory settings, techniques such as Digital Image Correlation (DIC) can be combined with image processing to monitor the rate at which strain fields change during mechanical tests, providing a real-time measure of damage evolution [18].

Drawbacks of traditional quantification methods like FEMU include the high computational cost and the inevitable discrepancies between the analytical model and the actual structure, which can lead to non-unique or inaccurate solutions [41].

Level IV: Prognosis. The ultimate goal of SHM is prognosis, which addresses the estimation of the structure's Remaining Useful Life (RUL) [25]. This requires a prediction of future performance based on the current damage state and anticipated future loading. Traditional approaches include the following.

- *Physics-based models.* These methods compute RUL based on damage progression models, such as using fracture mechanics to predict fatigue crack growth. They often employ statistical techniques like Sequential Monte Carlo methods to manage and propagate uncertainty [8].
- *Probabilistic methods.* To avoid computationally expensive cycle-by-cycle calculations under random loading, probabilistic methodologies have been developed to transform random load histories into equivalent constant amplitude loads for efficient prognosis [93].
- *Adaptive procedures.* These approaches integrate online state estimation algorithms with offline predictive models to provide condition-based RUL estimates that adapt over time [54]. Recent advances include Bayesian methods that update RUL distributions as new data becomes available [17].

The successful development of prognosis capabilities remains a significant challenge, requiring the integration of robust hardware, predictive models, and a rigorous framework for quantifying uncertainty [25]. It is noteworthy that while the lower levels of the hierarchy can be addressed with data-driven pattern recognition, Level IV cannot be accomplished without a comprehensive understanding of the underlying physics of damage evolution [23].

1.2 Handling Nonlinearities in Structural Health Monitoring

Having established the general framework and hierarchical goals of SHM, we now turn to the central technical challenge that motivates this entire dissertation: the presence of nonlinearity in structural systems. While the hierarchical framework provides a clear roadmap, the efficacy of many traditional tools is predicated on the assumption that the structure behaves as a linear time-invariant system. Nonlinear characteristics degrade the performance of these methods [94]. This assumption often fails to capture the true dynamic behavior of real-world structures, where nonlinearity is frequently present and can be a primary indicator of damage itself.

1.2.1 Sources and Effects of Nonlinearity in Structural Systems

Nonlinearity in structural dynamics arises from multiple sources, and understanding them is crucial for developing effective damage detection strategies [12, 15].

- **Geometric Nonlinearity.** This occurs due to large deformations where the relationship between displacement and strain becomes nonlinear. For instance, large rotational displacements can induce nonlinear effects on vibration frequencies [88], and in plates and shells, large out-of-plane bending can couple with in-plane axial deformation [100].
- **Material Nonlinearity.** This is an intrinsic property where the stress-strain relationship is nonlinear, including plasticity, hyperelasticity, and damage-induced material degradation [12].
- **Boundary and Contact Nonlinearity.** This arises from bolted joints, interfaces between components, and breathing cracks. Contact nonlinearity is particularly important in structures with joints, clearances, and gaps, where components may impact or slide against each other [95, 99].
- **Damage-Induced Nonlinearity.** Damage itself is a major source of nonlinearity. A classic example is a breathing crack, which exhibits bilinear stiffness as it opens and closes during vibration cycles, creating a time-varying system stiffness [95, 14, 29].

Effects of Nonlinearity. Nonlinear systems exhibit characteristic dynamic phenomena that distinguish them from linear systems [12, 14].

- **Amplitude-dependent natural frequencies.** The resonant frequencies of nonlinear systems depend on the amplitude of excitation, a stark contrast to the constant natural frequencies of linear systems.
- **Harmonic generation.** A nonlinear system excited at a single frequency ω will generate higher-order harmonics ($2\omega, 3\omega$, etc.) in its response, a phenomenon known as superharmonic resonance [70, 47].
- **Subharmonic and combination resonances.** Resonances can occur at fractional multiples of the excitation frequency, or at sum-and-difference frequencies when subjected to multi-frequency excitation.
- **Jump phenomena and multiple solutions.** The frequency response curves of nonlinear systems can bend, leading to jump phenomena where the response amplitude changes abruptly as the excitation frequency is swept.
- **Internal resonances.** In multi-degree-of-freedom systems, energy can be transferred between modes when their natural frequencies satisfy certain integer ratio relationships.
- **Path dependence.** Many nonlinear systems, particularly those with hysteresis, exhibit path-dependent behavior, implying that the instantaneous response is a function of the entire deformation history.

Collectively, these phenomena constitute a fundamental violation of the principles of linear superposition and time-invariance, the very axioms upon which traditional modal analysis is predicated. In the presence of such nonlinearities, features that are assumed to be global system invariants - such as natural frequencies and mode shapes - strongly depend on the instantaneous energy state and deformation history. Consequently, the direct application of standard linear monitoring algorithms to these systems entails a significant risk of misdiagnosis, where benign operational variations may be conflated with structural damage. This fundamental incompatibility necessitates a critical evaluation of available damage detection strategies to delineate their capabilities and limitations within the nonlinear domain.

1.2.2 A Critical Review of Existing Damage Detection Methods

The challenge of detecting damage in the presence of nonlinearity has given rise to a broad spectrum of methodologies. Broadly, these can be dichotomized into *Model-Based* approaches, which rely on updating a high-fidelity physical model (e.g., Finite Element Model) to match experimental observations [28], and *Data-Driven* approaches, which utilize signal processing to identify damage-sensitive features directly from sensor data [27]. While model-based methods offer physical interpretability, they are computationally expensive and struggle to accurately model complex boundary conditions and damping. Consequently, data-driven methods have become the dominant paradigm for continuous monitoring.

To provide a comprehensive context, existing methods are critically reviewed below based on their underlying assumptions regarding system linearity.

Methods Assuming Linear Behavior Before and After Damage. The vast majority of traditional vibration-based SHM techniques fall into this category, as detailed in Section 1.1.2. These methods like *Modal Analysis* rely on the premise that damage manifests as a change in the global modal parameters of the system, such as shifts in natural frequencies, or alterations in mode shapes, modal curvatures, and flexibility matrices [5, 22]. Prior to deep learning, statistical time-series methods were standard which uses *Statistical Pattern Recognition* methods like Principal Component Analysis (PCA) or Auto-Regressive (AR) modeling fit linear statistical models to healthy data and define damage as a statistical outlier [78]. While effective for a specific class of problems, these methods are fundamentally inadequate when applied to systems that are inherently nonlinear from the outset, or when damage induces significant nonlinear behavior. They may misinterpret amplitude-dependent frequency changes as damage and are often insensitive to damage types, such as breathing cracks, that do not significantly alter the global linear modal properties.

Methods Assuming a Linear-to-Nonlinear Transition. A more refined class of methods originates from the assumption that the undamaged structure behaves linearly, while the occurrence of damage (e.g., a breathing crack, loosening of a joint) introduces nonlinearity into the system [92]. The objective then shifts from tracking linear parameter changes to detecting the very presence of nonlinearity itself. Key techniques include the following.

- **Harmonic Analysis.** The presence of damage-induced nonlinearity, when excited by a single-frequency sinusoid, will cause the generation of super-harmonics (multiples of the excitation frequency) in the response spectrum [14]. The detection of these harmonics serves as a sensitive qualitative indicator of damage. This method's reliability can be compromised by other sources of nonlinearity, such as those from actuators or sensors, which can lead to false positives. Furthermore, for complex broadband or random excitation, the harmonic signature becomes smeared and difficult to interpret.
- **Bilinear System Identification.** For damage like breathing cracks, the system's stiffness can be modeled as bilinear (one value for when the crack is open, another for when it is closed). Methods have been developed to identify these two distinct stiffness values from the response data [95]. This approach relies on a highly simplified model of damage. Many real-world damage mechanisms and other nonlinearities do not conform to a simple bilinear model, limiting its general applicability.
- **Time-Frequency Analysis.** Techniques like the Hilbert-Huang Transform (HHT) are used to extract the instantaneous frequency of the response. For a system with a softening nonlinearity like a breathing crack, the instantaneous frequency will show a characteristic dependence on the vibration amplitude. HHT is susceptible to a phenomenon known as mode mixing, which can corrupt the identification of instantaneous properties. The physical interpretation of the decomposed modes (in terms of intrinsic mode functions, IMFs) is not always straightforward, and the methods are highly sensitive to measurement noise.

Methods Assuming Nonlinear Behavior Before and After Damage. This is the most challenging and realistic scenario, addressing structures that are inherently nonlinear in both their healthy and damaged states. The goal is to distinguish subtle changes in the nonlinear characteristics due to damage from the system's baseline nonlinear behavior. This dissertation focuses primarily on this domain. Notable traditional methods include the following.

- **Analysis of Bifurcation Phenomena.** For some systems, damage can alter the parameters such that the system's dynamic behavior undergoes a qualitative change, or bifurcation. It has been shown that model reduction near these bifurcation points can reveal features that are sensitive to damage [1]. This method is only applicable if the system is operating in the vicinity of a bifurcation point, which is a restrictive condition and not representative of general operational states.
- **Nonlinear Normal Modes (NNMs).** The concept of modal analysis can be extended to nonlinear systems through NNMs, which describe amplitude-dependent vibration modes. It has been shown analytically that the effective nonlinearity coefficients governing the shape of the NNM backbone curves are more sensitive to damage than linear natural frequencies [39]. This theoretical evidence has been confirmed by experimental results [19, 13]. The experimental identification of NNMs is significantly more complex and challenging than linear modal analysis, particularly for multi-degree-of-freedom systems under operational (non-laboratory) conditions.
- **Describing Functions.** In the frequency domain, the describing function method can be used to approximate the response of a nonlinear system under sinusoidal excitation. Damage can then be detected by tracking changes in this function [62]. This is an approximate method that is typically only accurate for systems with a dominant first harmonic and under single-frequency excitation, limiting its utility for systems subjected to random or broadband loads.
- **Restoring Force Surface Method.** This non-parametric technique attempts to directly reconstruct the nonlinear restoring force as a function of the system's state (displacement and velocity). Changes in the shape of this reconstructed surface are direct indicators of a change in the system's internal mechanics due to damage [51]. This method is notoriously sensitive to measurement noise, requires accurate measurement of acceleration to estimate the inertial force, and suffers from the curse of dimensionality, making it practically infeasible for systems with more than a few degrees of freedom.
- **Stochastic System Identification.** Advanced methods like stochastic subspace identification have been employed to track the changing modal parameters of structures exhibiting nonlinear behavior under strong, random excitation such as earthquakes [83]. These methods provide a powerful framework for analyzing operational data but often rely on linearizing assumptions that may not fully capture the underlying physics. The identification of equivalent linear parameters can obscure the true nature of the nonlinearity, and distinguishing

parameter changes due to damage from those due to amplitude-dependent effects remains a significant challenge.

- **Volterra Series.** The Volterra series provides a rigorous mathematical framework for representing nonlinear systems with memory, effectively acting as a higher-order extension of the linear impulse response function. Several proposals use the Volterra kernels to describe the system response; damage is detected by observing changes in these kernels [75, 84, 85]. Output-only versions have also been developed [61]. The primary limitation of the Volterra series is its immense complexity and data requirement. The number of terms and the difficulty of identifying the kernels grow combinatorially with the order of the nonlinearity and the system’s memory, making the identification of higher-order kernels computationally prohibitive for many practical systems.

The collective shortcomings of these traditional methods - their reliance on specific signal types, susceptibility to noise, computational intensity, and difficulty in scaling to complex systems - underscore a significant research gap. There is a clear need for a new class of methods that can robustly and automatically learn the complex features of nonlinear dynamic behavior directly from raw sensor data without relying on restrictive prior models. This critical need serves as the primary motivation for this dissertation’s focus on advanced deep learning architectures for damage detection in nonlinear systems.

1.3 The Rise of Data-Driven and Machine Learning Methods for SHM

The inherent limitations of physics-based models for complex nonlinear systems, coupled with the increasing availability of rich data from sensor networks, have catalyzed a paradigm shift toward data-driven approaches in SHM [7, 6]. These methods, which range from innovative data acquisition strategies like drive-by monitoring using sensors mounted on trains or other vehicles [53] to advanced machine learning algorithms, offer the significant advantage of learning complex patterns and damage signatures directly from data. Deep learning techniques, in particular, have demonstrated remarkable promise, offering potential solutions to challenges that traditional methods struggle to address [6, 34].

Early deep learning architectures employed in SHM included Recurrent Neural Networks (RNNs) and Long Short-Term Memory (LSTM) networks, which are well-suited for capturing temporal patterns in time-series data [80, 86]. However, these recurrent architectures face inherent limitations when dealing with very long-term dependencies due to the vanishing gradient problem, which can impede their ability to learn relationships across extended temporal horizons [59, 73].

Building on these architectural foundations, recent research has explored hybrid frameworks that strategically combine the strengths of multiple components. For instance, architectures integrating Convolutional Neural Networks (CNNs) for local feature extraction with LSTMs for temporal modeling have shown effectiveness in capturing spatio-temporal dependencies in vibration signals [2]. Furthermore, to overcome the limitations of purely data-driven approaches, Physics-Informed Neural

Networks (PINNs) have gained attention. By integrating physical conservation laws into the loss function, PINNs can regularize the learning process, improving generalization for nonlinear systems where labeled data is scarce [97]. Additionally, probabilistic approaches such as Variational Autoencoders (VAEs) and Bayesian deep learning are emerging to address the critical need for uncertainty quantification in safety-critical infrastructure monitoring [38]. Despite these advances, the challenge of applying these sophisticated methods to unsupervised damage detection in highly nonlinear systems remains significant.

To address these challenges, significant attention has been focused on unsupervised deep learning architectures. Autoencoders (AEs) and Generative Adversarial Networks (GANs), in particular, appear to be among the most promising for damage detection [5]. Convolutional AEs have shown success by automatically extracting features from time-series signals to learn a compressed representation of the healthy state, with damage flagged through elevated reconstruction errors [63, 76]. This approach has gained significant traction, with recent studies demonstrating its effectiveness for the unsupervised, real-time health assessment of major civil infrastructure, such as bridges [30], and for detecting fatigue damage in full-scale experimental tests of wind turbine blades using frequency response functions as input [81]. However, the majority of these applications have been limited to linear dynamic systems or have not systematically explored performance on systems with strong, inherent nonlinearities. For instance, AE-based approaches have been proposed using linear modal properties as inputs [60], raw dynamic response from linear systems [50], cross-correlation functions from linear vibrations [71], and power cepstral coefficients to represent linear modal properties [43]. Similarly, applications of GANs for damage detection are even fewer, with the first examples for linear dynamic systems in civil structures appearing only recently [48, 49].

The application of machine learning to inherently nonlinear systems remains seldom explored. A seminal work in this context is due to [9], who used an autoregressive support vector machine for damage detection in a small-scale nonlinear system. More recently, the use of deep learning techniques has started to emerge. For example, [98] presented a numerical investigation using autoencoders for damage detection in seismic isolation bearings. While preliminary applications of GANs for nonlinear system identification have been proposed [96], no known proposals exist about their application to damage detection problems in this domain.

This leads to the most fundamental challenge: output-only, vibration-based damage detection in initially nonlinear systems is exceptionally difficult because the features of the response can change considerably with varying excitation conditions, even in the healthy state. Discerning the effects induced by damage from those caused by the baseline nonlinear response is a significant concern that deserves careful investigation [9].

To overcome the limitations of recurrent models in capturing long-term dependencies a critical requirement for characterizing hysteretic nonlinear systems with fading memory this research turns to the Transformer architecture [82]. Unlike RNNs, which process time-series sequentially, Transformers utilize a self-attention mechanism to compute relationships between all sequence positions in parallel [55]. In the context of nonlinear SHM, this allows the model to capture global temporal correlations that correspond to complex energy transfer mechanisms, analogous to

higher-order kernels in Volterra series, but learned directly from data. Building upon this, the Temporal Fusion Transformer (TFT) was introduced as a specialized architecture for interpretable time-series forecasting [44]. While TFTs have demonstrated exceptional performance in diverse forecasting domains [56, 72], their application to unsupervised damage detection in nonlinear structural dynamics remains largely unexplored. This is a critical research gap, as in real-world scenarios, labeled datasets of damaged structures are exceptionally rare. Consequently, the most viable strategy is unsupervised anomaly detection, where a model is trained exclusively on data from the structure’s healthy state.

1.4 Research Objectives and Scientific Contributions

The critical review of the existing literature reveals a distinct and compelling research gap: while deep learning is a promising avenue for SHM, there is a lack of systematic investigation into unsupervised methods specifically designed for and validated on systems where nonlinear dynamics are the dominant operational characteristic. Traditional methods for nonlinear systems have significant limitations, and existing deep learning applications have largely focused on linear systems.

The conceptual framework underpinning this research, mapping the evolution from physics-based paradigms to the proposed advanced unsupervised deep learning architectures, is illustrated in Figure 1.2.

1. **Benchmarking Unsupervised Architectures on Nonlinear Systems** to systematically evaluate the efficacy of 1D Convolutional Autoencoders and GANs for anomaly detection, establishing a performance baseline across a diverse suite of numerical and experimental nonlinear systems.
2. **Development of the SensorFusionTFT Architecture** to introduce a novel, **probabilistic forecasting architecture** specifically designed for high-dimensional nonlinear dynamics. Unlike standard multichannel Transformers, this model incorporates a **Variable Selection Network (VSN)** based on Gated Residual Networks (GRNs). This mechanism enables the model to learn time-varying **sensor importance weights**, dynamically suppressing noise from irrelevant channels while explicitly modeling the nonlinear coupling between sensors at each time step.
3. **Development of an Interpretable Diagnostic Framework** to formulate a robust decision-making framework that transcends simple reconstruction error by leveraging **Quantile Regression** to establish dynamic confidence intervals for anomaly detection. Furthermore, this objective seeks to utilize the model’s internal **temporal attention maps** and **spatial sensor weights** to not only detect but potentially localize damage by identifying deviations in the learned dependency structures.
4. **Comparative Validation** to provide a rigorous comparative analysis between convolutional (local feature) and attention-based (global dependency) models, delivering specific guidelines for architecture selection based on the type of structural nonlinearity (e.g., stiffness degradation vs. contact nonlinearity).

The principal scientific contribution of this work is a holistic and comparative evaluation that maps the progression of deep learning architectures—from foundational unsupervised models to a state-of-the-art interpretable Transformer—to the specific and challenging domain of nonlinear structural health monitoring. This research provides not only novel methodologies but also a practical and critical assessment of their applicability, thereby advancing the state-of-the-art in data-driven SHM for complex, real-world engineering systems.

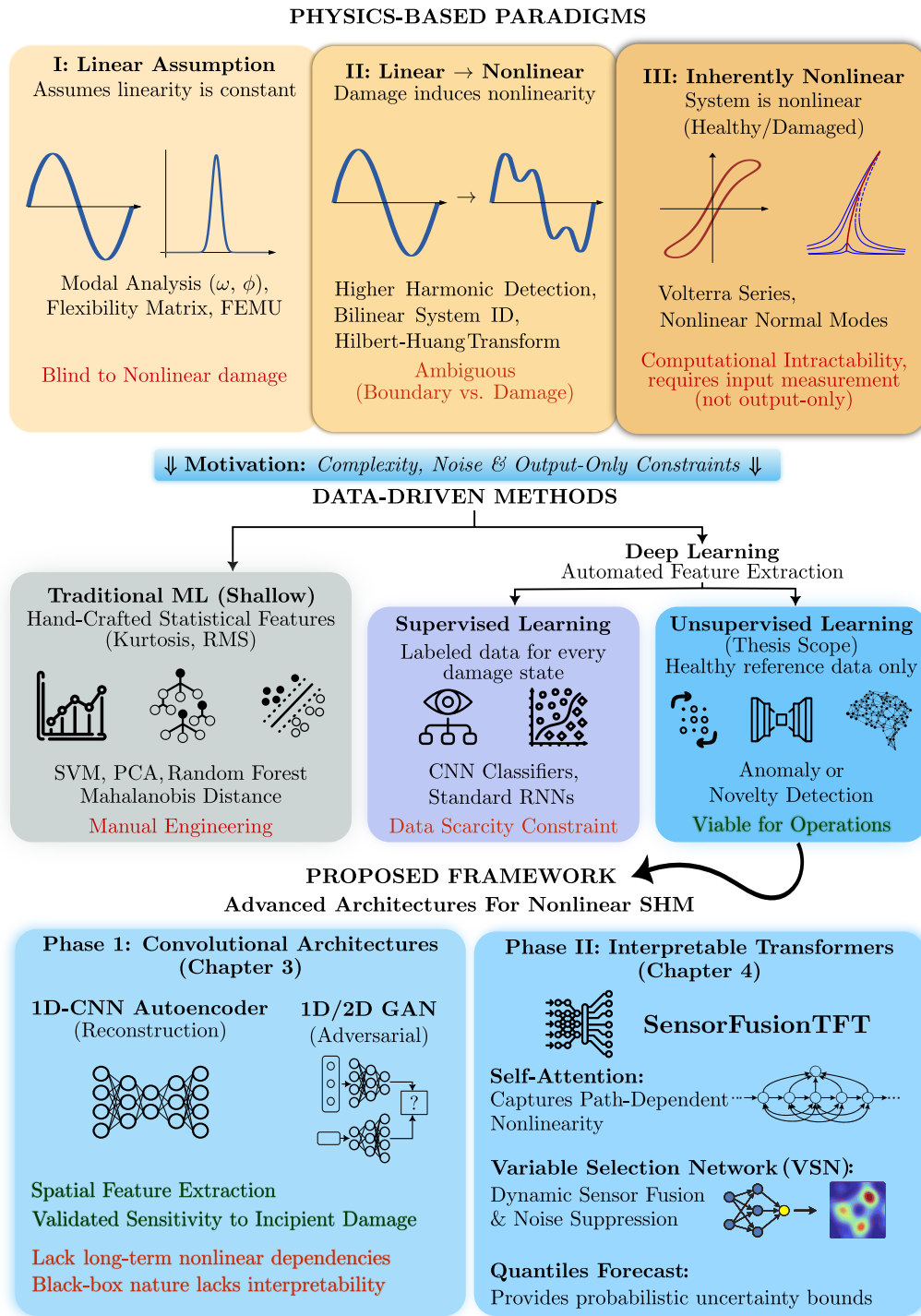


Figure 1.2. The conceptual framework of the research. The diagram maps the evolution from traditional linear and nonlinear physics-based methods (Top) to the proposed data-driven framework (Bottom). It highlights the dead ends in traditional machine learning caused by manual feature engineering and data scarcity, motivating the focus on Unsupervised Deep Learning. The thesis progression is visualized in the lower section, advancing from black-box Convolutional Neural Networks (Phase I) to the interpretable, memory-aware SensorFusionTFT (Phase II), specifically engineered to handle the complexities of nonlinear structural dynamics.

Chapter 2

Theoretical Foundations

Having established the broad challenges of SHM in nonlinear systems in chapter 1, this chapter provides the specific theoretical and mathematical foundations for the present work. The objective is to formally introduce the archetypal systems—both numerical and experimental—that will serve as the benchmarks for evaluating the deep learning architectures developed and analyzed in this dissertation. For each system, the governing equations of motion will be presented, the precise nature of its nonlinearity will be characterized, its dynamic response will be analyzed to visually demonstrate the fundamental difficulties that nonlinearity introduces to the damage detection problem, and all relevant system parameters will be defined. This chapter will thus serve as a self-contained reference for the subjects of my investigation, setting a clear and rigorous stage for the application and results chapters that follow.

2.1 Numerical Archetypes of Nonlinear Dynamic Systems

To systematically evaluate the performance of the proposed damage detection methods under controlled conditions, a series of numerical models are employed. These archetypes are chosen to represent distinct and fundamental classes of nonlinear behavior commonly encountered in structural and mechanical engineering.

2.1.1 Duffing Oscillator: Cubic Stiffness Nonlinearity

The first numerical archetype considered is the Duffing oscillator, a canonical model for systems exhibiting nonlinearity due to geometric effects. This type of behavior is common in structures like thin beams, plates, and cables undergoing large-amplitude vibrations, where membrane tension effects introduce a hardening cubic stiffness term. This system provides a fundamental benchmark for evaluating a model's ability to handle smooth, polynomial-type nonlinearities.

Governing Equations. The equation of motion for a single-degree-of-freedom (1-DOF) system is given by the Duffing equation:

$$M\ddot{x} + C\dot{x} + K_1x + K_3x^3 = F(t), \quad (2.1)$$

where x represents the displacement, M is the mass, C is the damping, K_1 and K_3 are the linear and cubic stiffness coefficients, and $F(t)$ is the external load. A two-degree-of-freedom (2-DOF) extension is also considered:

$$\begin{aligned} M_1\ddot{x}_1 + C_1\dot{x}_1 + K_{11}x_1 + K_{13}x_1^3 \\ - C_2(\dot{x}_2 - \dot{x}_1) - K_{21}(x_2 - x_1) - K_{23}(x_2 - x_1)^3 = F_1, \end{aligned} \quad (2.2a)$$

$$M_2\ddot{x}_2 + C_2(\dot{x}_2 - \dot{x}_1) + K_{21}(x_2 - x_1) + K_{23}(x_2 - x_1)^3 = F_2, \quad (2.2b)$$

System Parameters. For the numerical simulations, the system parameters are defined as follows:

- **1-DOF System:** $M = 412$ kg, $C = 405.95$ kg/s (corresponding to a damping ratio of 1%), $K_1 = 1$ kN/mm, and $K_3 = 0.001$ kN/mm³.
- **2-DOF System:** $M_1 = M_2 = 400$ kg, $C_1 = C_2 = 400$ kg/s (damping ratio of 1%), $K_{11} = K_{21} = 1$ kN/mm, and $K_{13} = K_{23} = 0.001$ kN/mm³.

Dynamic Characteristics and the Challenge for SHM. The fundamental difficulty of damage detection in nonlinear systems is vividly illustrated by their Frequency Response Curves (FRCs). Figure 2.1 shows the FRCs for the 1-DOF Duffing oscillator. Unlike a linear system, the Duffing oscillator exhibits a hardening behavior: as the excitation amplitude increases, the resonant frequency shifts to higher values and the response curve bends to the right, creating regions of bistability. This presents a profound challenge for SHM [9], as a measured frequency shift could be caused by a change in operational load rather than structural damage. Figure 2.1 also shows that damage (a stiffness reduction) shifts the curves to the left. However, the response of a damaged system at high excitation could be mistaken for the response of a healthy system at low excitation, leading to missed detections. These challenges are compounded in the 2-DOF system (Figure 2.2), which displays more complex modal interactions and multiple bent resonant peaks.

Damage Formulation. For both Duffing systems, damage is simulated as a parametric reduction of the stiffness coefficients. A damage level of $d\%$ corresponds to a reduction in both the linear and nonlinear stiffness terms by that percentage amount, without affecting mass or damping. This phenomenological approach is widely accepted for validating data-driven architectures as it provides a controllable and physically grounded proxy for structural degradation mechanisms, such as fatigue cracking or loosening of connections, which macroscopically manifest as a loss of restoring force. Crucially, for the unsupervised deep learning models evaluated in this work, this parametric reduction introduces a systematic drift in the underlying data distribution $P(x)$. In the context of nonlinear dynamics, this does not merely shift the resonant frequency (as in linear theory) but alters the shape of the nonlinear restoring force surface and the system’s amplitude-dependent response. This forces the Autoencoder and GAN to detect anomalies based on complex changes in the signal’s topology rather than simple spectral shifts, thereby rigorously testing their ability to learn the manifold of the healthy state.

2.1.2 Bouc-Wen Model: Hysteretic Nonlinearity

To investigate systems with memory and energy dissipation, this research employs a model governed by the Bouc-Wen formulation of hysteresis [10, 89]. This model represents a wide class of systems where the restoring force is path-dependent, a characteristic of material yielding or passive control devices like Tuned Mass Dampers (TMDs).

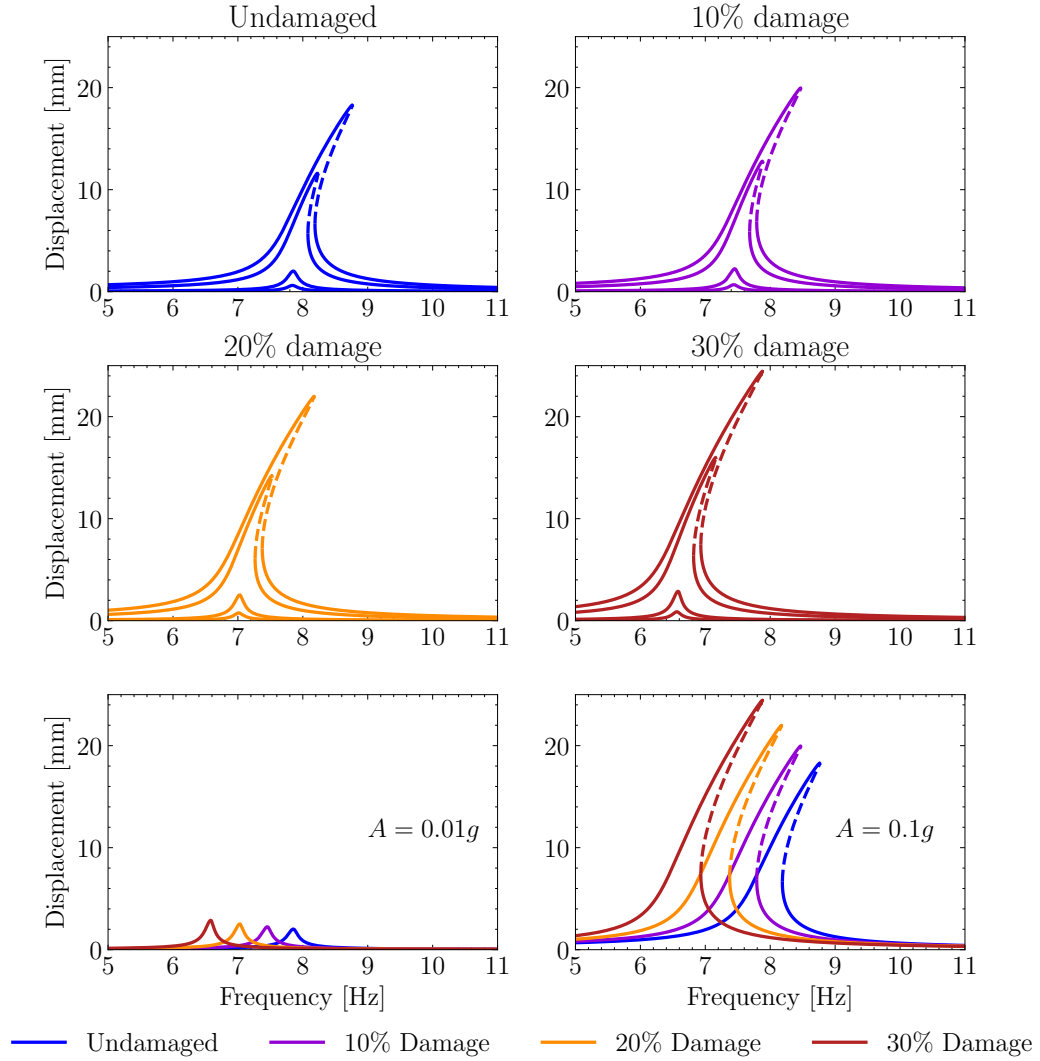


Figure 2.1. Frequency response curves (FRCs) of the 1-DOF Duffing oscillator, illustrating the hardening nonlinearity and amplitude-dependent resonance for different damage levels.

Type of Nonlinearity. The system exhibits a strong **material nonlinearity** that manifests as a **hysteretic restoring force**. Unlike elastic nonlinearity, where force is a function of displacement $f(x)$, hysteretic force is a functional of the motion history. The force at any given moment depends not only on the current displacement but on the specific trajectory taken to reach that state, introducing memory effects into the system dynamics.

Governing Equations and Parameter Definitions. The dynamics of the single-degree-of-freedom oscillator are fundamentally described by the equation of motion coupled with an evolutionary equation for the internal hysteretic variable z . To ensure numerical stability and generalizability, the system is analyzed in its

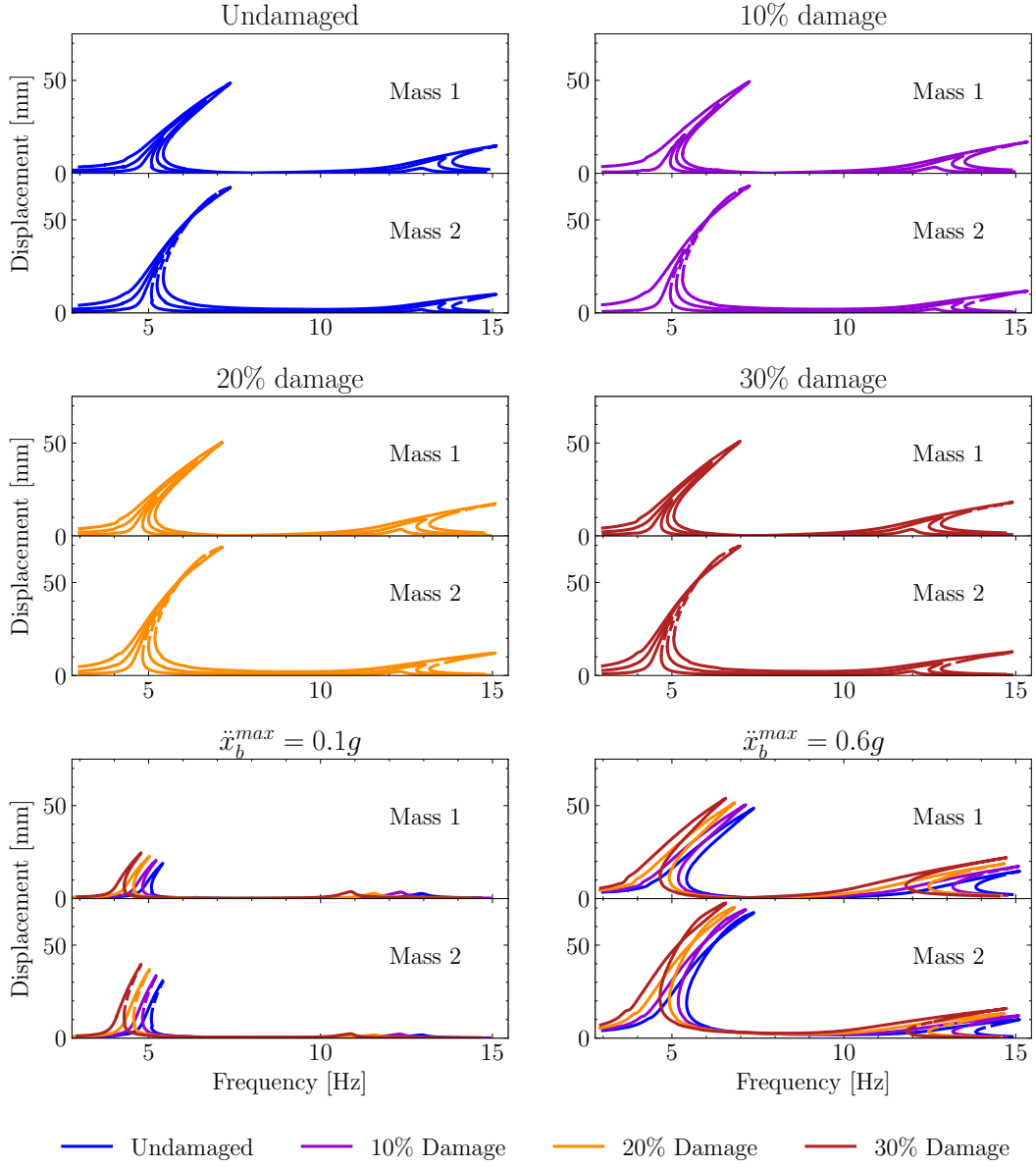


Figure 2.2. Frequency response curves (FRCs) of the 2-DOF Duffing system, showing complex modal interactions and bent resonant peaks for different damage levels.

non-dimensional form:

$$\ddot{\tilde{x}} + \delta \dot{\tilde{x}} + (1 - \delta)\tilde{z} + \tilde{k}_3 \tilde{x}^3 = \tilde{P} \sin(\tilde{\Omega} \tilde{t}) \quad (2.3)$$

$$\dot{\tilde{z}} = \left[\tilde{h}(\tilde{x}) - (\tilde{\gamma} + \tilde{\beta} \text{sgn}(\dot{\tilde{x}} \tilde{z})) |\tilde{z}|^n \right] \dot{\tilde{x}} \quad (2.4)$$

where $\tilde{x} = x/x_0$ is the non-dimensional displacement normalized by a characteristic length x_0 , and $\tilde{t} = \omega t$ is the non-dimensional time. The system includes a pinching function $\tilde{h}(\tilde{x}) = 1 - \xi e^{-\tilde{x}^2/\tilde{x}_c}$ which models the narrowing of the hysteresis loop. The non-dimensional parameters are derived from the physical system properties used in

the numerical simulations ($M = 400$ kg, $K_i = 1.1$ kN/mm). The conversion allows for the handling of parameters that differ by several orders of magnitude:

- $\delta = \frac{k_e}{k_e + k_d}$ represents the post-yield stiffness ratio ($\alpha = 0.2$), defined as the ratio of the linear elastic stiffness to the total initial stiffness.
- $\tilde{k}_3$ captures the cubic geometric nonlinearity normalized by the total linear stiffness.
- $\bar{\gamma}$ and $\bar{\beta}$ control the hysteresis loop size and energy dissipation. In the physical domain, these were set to $\beta = 0.0007$ and $\gamma = 0.000011$.
- n is the exponent controlling the sharpness of the yield transition, set to $n = 1$ for this study.

Manifestation of Damage and Nonlinear Dynamics. Damage to the hysteretic system is simulated by a parametric reduction of the stiffness parameters (k_e and k_d). In the context of the Bouc-Wen model, this reduction does not merely shift the natural frequency as in linear systems; it fundamentally alters the *topology* of the system’s response. As the stiffness decreases ($d\% \uparrow$), the system becomes more compliant, leading to larger displacement amplitudes that push the response further into the post-yield (inelastic) regime. Consequently, damage manifests as a severe distortion of the phase-space trajectory and an increase in the area of the hysteresis loop (increased damping). This amplitude-dependent behavior creates a non-unique relationship between damage and response, making it a challenging task for deep learning networks that rely on simple linear correlations.

Computational Challenges and Parameter Sensitivity

The simulation and data generation for systems governed by Eqs. (2.3) and (2.4) present specific challenges that must be addressed to ensure the reliability of the training dataset.

Parameter Sensitivity and Stiffness. A critical challenge in modeling hysteretic systems is that the governing parameters span several orders of magnitude. For instance, the shape parameters (β, γ) are often very small (10^{-4} to 10^{-3}), while stiffness values are large (10^3). This disparity makes the differential equations highly sensitive to inputs; small numerical perturbations can lead to significant deviations in the computed hysteresis loop, potentially resulting in bifurcations where the system switches between periodic and chaotic behaviors. This sensitivity implies that the healthy and damaged manifolds in the high-dimensional feature space may be close neighbors, complicating the decision boundary for the neural network.

Numerical Integration with Diffrax. The differential equation governing the hysteretic variable $\dot{\tilde{z}}$ is numerically stiff, particularly due to the sharp gradients introduced by the signum function $\text{sgn}(\hat{x}\tilde{z})$ and the pinching function. Standard explicit integration schemes may fail to capture these rapid transitions or require prohibitively small time steps. To address this, the dataset generation for this thesis utilized the Diffrax library, a high-performance differential equation solver in the JAX ecosystem. Specifically, the `Tsit5()` solver (Tsitouras 5/4 Runge-Kutta) was

employed. While Tsit5 is an explicit solver, its adaptive step-size controller allows it to robustly handle the non-smooth transitions inherent in the Bouc-Wen model while maintaining computational efficiency suitable for generating large-scale deep learning datasets.

2.1.3 Seismic Nonlinear Isolator: Hybrid Nonlinearity

The final and most complex numerical benchmark is a seismic isolator model proposed by [69], which combines multiple sources of nonlinearity to achieve advanced isolation performance. This system is chosen to serve as a formidable challenge for the deep learning architectures, testing their ability to learn and detect anomalies in the presence of highly complex, multi-faceted, and path-dependent dynamics¹.

Type of Nonlinearity. This system represents a hybrid of nonlinear behaviors, creating a rich dynamic response.

- **Material Nonlinearity** arising from superelastic components modeled on shape-memory alloys (SMAs). These exhibit a recoverable, non-Hookean response with distinct stiffnesses for different material phases (austenite and martensite)
- **Hysteretic Nonlinearity** derived from traditional elastomeric components, modeled using the Bouc-Wen formulation to capture path-dependent energy dissipation and memory effects
- **Stiffness Nonlinearity** introduced by a negative stiffness mechanism. This geometric nonlinearity is designed to create a quasi-zero stiffness effect, enhancing isolation by softening the force-displacement curve in a specific displacement range

Governing Equations and Parameter Definitions. The overall equation of motion describes a single degree of freedom system where the total restoring force is the superposition of three distinct nonlinear components:

$$M\ddot{x} + f_i + f_n + f_s = F(t), \quad (2.5)$$

where f_i , f_n , and f_s correspond to the hysteretic, negative stiffness, and superelastic forces, respectively.

Hysteretic Component (f_i) described by the Bouc-Wen model

$$f_i = C\dot{x} + \alpha K_i x + (1 - \alpha)K_i z, \quad (2.6a)$$

$$\dot{z} = \dot{x}[1 - (\gamma + \beta \text{sign}(z\dot{x}))|z|^n]. \quad (2.6b)$$

Here, $K_i = 1.1$ kN/mm is the initial stiffness, and $\alpha = 0.2$ is the post-yield to pre-yield stiffness ratio. The parameters $\beta = 0.0007$, $\gamma = 0.00001$, and $n = 1$ control the shape and smoothness of the hysteresis loop, determining the energy dissipated per cycle.

Negative Stiffness Component (f_n) which activates only within a specific displacement range:

$$f_n = \left(-K_n x + K_3 x^3\right) \frac{[1 + \text{sign}(x_f - |x|)]}{2}, \quad (2.7)$$

where $K_n = 0.5$ kN/mm is the negative linear stiffness which reduces the global system stiffness, and $K_3 = 0.00001$ kN/mm³ is a cubic hardening term ensuring stability at large displacements. The activation displacement is set to $x_f = 0.7x_u$, meaning this softening effect is conditional on the vibration amplitude.

Superelastic Component (f_s) which is modeled based on the formulation for SMAs, capturing phase transformations:

$$\dot{f}_s = (1 - s)K_s \left[\dot{x} - |\dot{x}| \text{sign}(f_s - \beta_s) \left(\frac{|f_s - \beta_s|}{Y} \right)^{n_s} \right] + sK_m \dot{x}, \quad (2.8a)$$

$$\beta_s = K_s \alpha_s \left\{ x - \frac{f_s}{K_s} + f_t \tanh(a_s x) \left[\frac{1 + \text{sign}(-x\dot{x})}{2} \right] \right\}, \quad (2.8b)$$

$$s = \frac{\tanh[c_s(|x| - x_m)] + 1}{2}, \quad (2.8c)$$

$$f_t = \frac{2Y - y_s Y}{\alpha_s K_s}, \quad (2.8d)$$

$$a_s = \frac{\tan^{-1}(\tilde{a}_s K_s)}{Y - y_s Y}. \quad (2.8e)$$

The system dynamics toggle between the austenitic stiffness K_s and martensitic stiffness K_m , governed by the phase fraction s . The parameters $x_u = 100$ mm (ultimate displacement) and $x_m = 0.7x_u$ define the phase transition points.

Manifestation of Damage and Nonlinear Dynamics. Damage is simulated by a parametric reduction of the stiffness components. Specifically, a damage level of $d\%$ corresponds to a simultaneous reduction in both the hysteretic stiffness K_i and the negative stiffness K_n . In this hybrid system, damage manifestation is highly non-trivial. Unlike linear systems where stiffness reduction simply lowers natural frequency, here it fundamentally reshapes the total restoring force surface. A reduction in K_i alters the yield point and energy dissipation capacity, while a reduction in K_n weakens the softening effect intended to isolate vibrations. Consequently, damage may lead to unexpected amplitude jumps or shifts in the activation of the super elastic regime. This creates a scenario where the healthy and damaged dynamic manifolds may intersect or diverge in complex ways depending on the excitation amplitude, posing a severe challenge for deep learning models that must disentangle these effects from noise and normal nonlinear operation.

Computational Challenges and Numerical Integration.

Simulating this hybrid system presents significant computational challenges due to the interactions between multiple nonlinear switching mechanisms.

Stiffness and Discontinuities. The governing equations involve multiple non-smooth functions, including $\text{sign}(\cdot)$ terms in the Bouc-Wen and negative stiffness models, and $\tanh(\cdot)$ functions in the superelastic model. These terms introduce sharp gradients and discontinuities in the vector field, making the system numerically stiff. Small changes in state can trigger rapid phase transformations or stiffness activations, requiring a solver capable of high temporal resolution during these transient events.

Adaptive Integration with Dopri5. To generate high-fidelity training data, the simulation utilized the Diffeq library. The dopri5 solver (Dormand-Prince 5/4) was selected for the numerical integration. While dopri5 is an explicit Runge-Kutta method, its embedded error estimation allows for robust adaptive step-size control. This feature proved critical for this application: the solver automatically refines the time step Δt in regions of rapid state change (e.g., during hysteresis reversals or phase transitions) to maintain stability and accuracy, while relaxing the step size during smooth trajectories to ensure computational efficiency. This approach ensured stable responses across different damage conditions, preventing numerical artifacts from contaminating the dataset used for deep learning validation.

2.2 Experimental Case Studies of Nonlinear Systems

While numerical simulations provide a controlled environment for initial validation, the ultimate test of any SHM methodology is its performance on real-world experimental data. Experimental benchmarks are indispensable for assessing the robustness and generalizability of the developed deep learning architectures. They introduce a host of challenges not present in idealized numerical models, including ambient measurement noise, unmodeled dynamic complexities, sensor non-idealities, and the subtle interplay of multiple physical phenomena. This section introduces the two experimental systems used in this dissertation to validate the proposed damage detection frameworks under these more realistic and demanding conditions.

2.2.1 The Magneto-Elastic Beam System

The first experimental benchmark is a laboratory-scale magneto-elastic system, utilizing the UNESP-MAGNOLIA dataset based on the setup reported by [75] and [85]. This system is specifically designed to create a strong, well-characterized source of nonlinearity with a hardening effect, serving as a physical realization of the Duffing-type behavior explored numerically in Section 2.1.1. It provides an essential bridge between pure simulation and complex, real-scale structures, offering a rich dataset for validating data-driven methodologies.

Physical Setup, Data Acquisition, and Experimental Protocol. The experimental apparatus, shown in Figure 2.3, consists of a 300 mm $\times$ 19 mm $\times$ 3.2 mm cantilever aluminum beam. The system's primary source of nonlinearity is a magnetic interaction mechanism: a small steel mass attached to the beam's free end interacts with a stationary neodymium magnet positioned at a fixed distance of 2 mm. The excitation is provided by an electrodynamic shaker attached to the beam at

a position 50 mm from the clamped end. To capture the system’s dynamic response with high spatial resolution, the setup employs a comprehensive sensor array:

- **Acceleration:** three piezoelectric accelerometers are distributed along the beam’s length to acquire the acceleration response at different mode shape nodes and anti-nodes.
- **Velocity:** a laser Doppler vibrometer is focused on the free end of the beam to measure the transverse velocity, ensuring high-fidelity capture of the tip kinetics where the nonlinear restoring force is most active.

To rigorously test the generalization capabilities of the deep learning models, the research employs training and testing on signals acquired from random excitations. The shaker drives the beam with band-limited Gaussian white noise (typically filtered between 10 Hz and 400 Hz). This broadband input serves two critical purposes: first, it forces the system to explore its full nonlinear geometric manifold rather than settling into a specific steady-state periodic orbit; second, the excitation amplitude is varied stochastically across different experimental runs (e.g., varying RMS voltage levels). This ensures that the training data encompasses multiple energy states, preventing the neural networks from overfitting to a specific amplitude-dependent frequency shift.

Type of Nonlinearity. A hardening-type *stiffness nonlinearity* originates from the magnetic interaction between the permanent magnet and the steel tip mass. As the beam vibrates, the magnetic attraction force changes as a non-linear function of the displacement, altering the system’s effective stiffness. This setup effectively creates a physical system whose dynamics are governed by a Duffing-like restoring force, exhibiting amplitude-dependent resonance and bifurcation phenomena depending on the excitation energy.

Damage Formulation. Damage is simulated in a controlled and physically realistic manner by altering a bolted connection positioned 150 mm from the free end. The baseline (healthy) state is defined with four nuts securing the connection. Damage is then progressively introduced by sequentially removing one nut at a time. This corresponds to a minor change in the local mass and stiffness of the beam, creating distinct damage classes approximately defined as 2%, 4%, and 6% of the total attached mass. This progressive degradation offers a granular challenge for the deep learning architectures, testing their sensitivity to detecting subtle, incipient damage buried within the complex nonlinear response.

2.2.2 The LUMO Lattice Tower Benchmark

To further assess the scalability of the proposed frameworks, a second, larger experimental benchmark is employed: the Leibniz University Test Structure for Monitoring (LUMO). This publicly available dataset, introduced by Wernitz et al. [91], provides a challenging and realistic testbed that moves beyond the controlled laboratory environment to an outdoor setting subject to uncontrolled environmental and operational variations (EOVs).

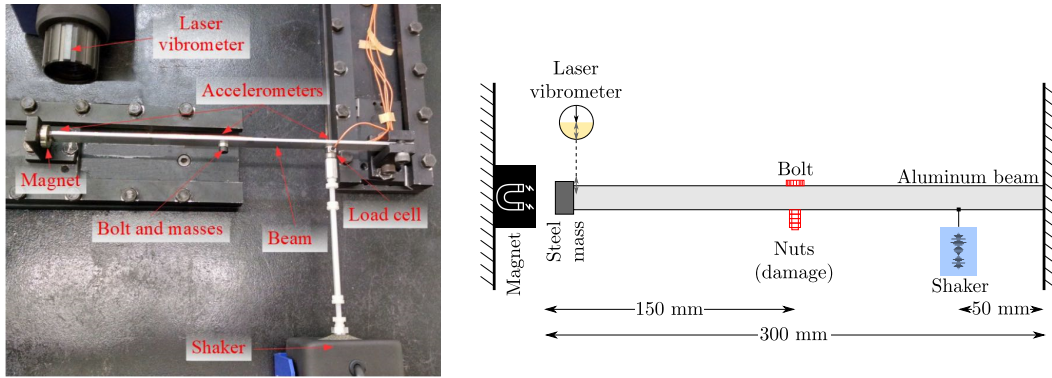


Figure 2.3. Layout of the magneto-elastic beam experiment, consisting of a cantilever beam, a permanent magnet to induce nonlinearity, and a bolted connection for simulating damage. Figure adapted from [85].

Physical Setup and Structural Design. The LUMO structure is a 9-meter-tall outdoor steel lattice mast, located in Hanover, Germany. It is composed of three identical 3-meter segments with an isosceles triangular cross-section formed by tubular legs¹. The segments are interconnected via elliptical flanges using M10 steel bolts. The entire structure is anchored to a reinforced concrete block foundation ($1.5 \times 1.5 \times 0.8$ m) embedded in cohesive soil, utilizing a 30-mm thick steel hinge plate that is bolted rigid during monitoring operations. Unlike laboratory setups, this structure is continuously exposed to real-world variability, including wind-induced aerodynamic loading and seasonal temperature cycles, making it an ideal candidate for evaluating the robustness of unsupervised SHM algorithms.

Data Acquisition and Sensor Network. The structure is instrumented with a high-density network of 22 sensors distributed across ten measurement levels (ML1 to ML10).

- **Structural Sensors.** The dynamic response is captured by 18 uniaxial piezoelectric accelerometers (MEAS 8811LF-01-005) mounted in orthogonal pairs to capture planar motion, along with two HBM strain gauges installed at the base of the legs to measure axial strain. A Pt100 thermocouple records the material temperature.
- **Acquisition System.** Data is digitized using a National Instruments cDAQ-9189 chassis equipped with NI-9234 and NI-9237 modules. The system operates continuously at a sampling rate of 1651.61 Hz, storing data in 10-minute blocks.
- **Environmental Monitoring.** Crucially, a separate meteorological mast located 20 meters from the structure acquires synchronized environmental data, including wind speed (average and maximum), wind direction, air temperature, and solar radiation.

Type of Nonlinearity. The LUMO structure exhibits nonlinearities that are more complex and subtle than those of the previous benchmarks. Wernitz et al.

characterize the system as having weak-to-moderate nonlinearity directly coupled to environmental effects. This manifests primarily as an amplitude-dependent dynamic response, where the natural frequencies of the structure show a clear decreasing trend with increasing wind speed. This phenomenon is attributed to a combination of aerodynamic damping and stick-slip friction behavior within the numerous bolted connections of the lattice structure. Furthermore, the structure exhibits a strong linear dependence on material temperature, which alters the global stiffness matrix and must be disentangled from damage-induced changes.

Damage Formulation. A distinguishing feature of the LUMO benchmark is its sophisticated system for introducing reversible damage. Specific bracing elements at predefined locations (DAM3, DAM4, and DAM6) are equipped with a specialized mechanism consisting of left- and right-hand screw sockets and a threaded connection rod. Damage is induced by loosening these sockets and removing the connection piece, effectively severing the bracing load path. As detailed in the dataset protocol, two distinct damage severities are simulated:

- **Moderate Damage.** The damage mechanism is activated on a single bracing strut at a given level. This simulates a localized loss of stiffness and a minor mass reduction (approx. 155 g).
- **Severe Damage.** The mechanisms are activated on all three bracing struts at a given level, creating a significant structural discontinuity and a larger reduction in both stiffness and mass.

This configuration provides a diverse set of damage scenarios, ideal for evaluating both the detection sensitivity and the localization capabilities of the proposed SHM frameworks under realistic noise conditions.

2.2.3 Common Challenges in Unsupervised SHM Data and Simulation

While data-driven frameworks offer significant promise, their practical implementation is often hindered by fundamental challenges inherent to both numerical modeling and experimental data acquisition. This section critically examines these hurdles, differentiating between the complexities of simulating nonlinear physics and the practical realities of handling imperfect experimental datasets.

Simulation and Modeling Complexities

Generating high-fidelity synthetic data for training deep learning models requires overcoming substantial computational barriers, particularly when nonlinearities are present.

Modeling Energy Dissipation. Accurately capturing damping remains one of the most persistent challenges in structural dynamics. In nonlinear regimes, energy dissipation is rarely viscous or linear; it often arises from friction, micro-slip, or hysteretic material behavior. Standard Rayleigh damping models frequently fail

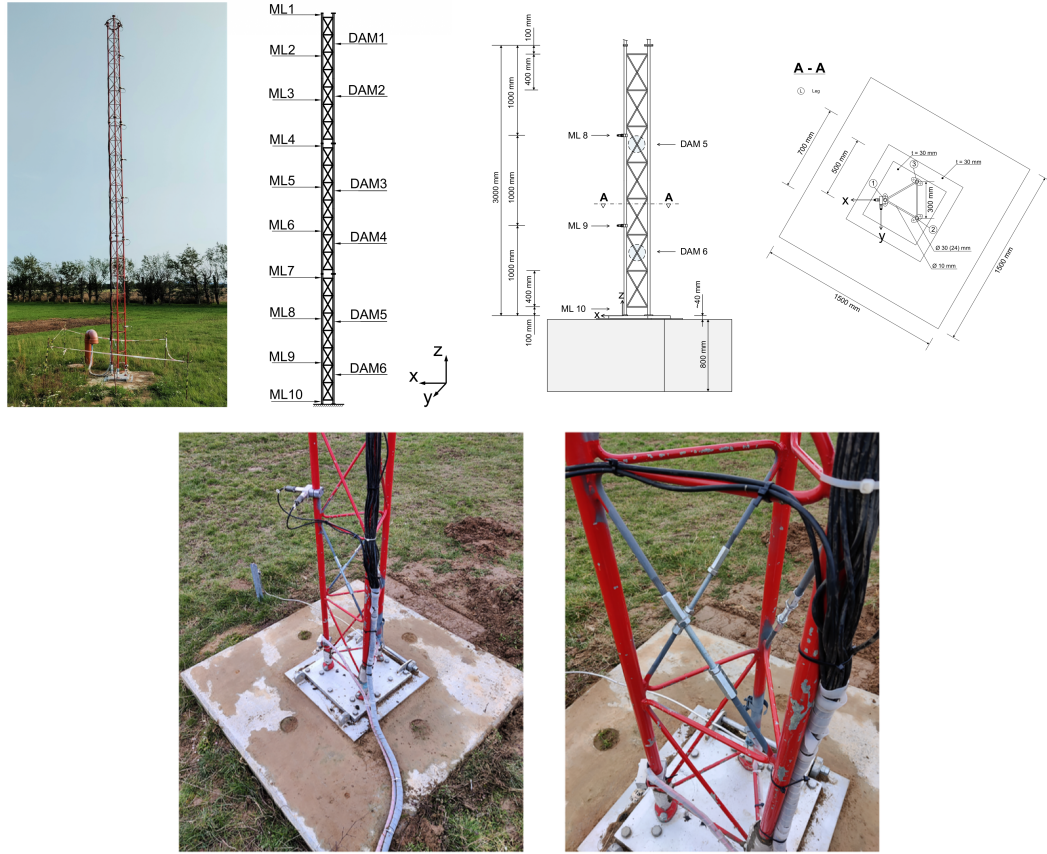


Figure 2.4. The LUMO experimental test setup, showing the 9-meter lattice tower, sensor measurement levels (MLs), and damage locations (DAMs). The complexity and environmental exposure of this benchmark provide a realistic test for SHM algorithms. Figure adapted from [91].

to capture amplitude-dependent dissipation, leading to discrepancies between the simulated and physical reality. Inaccurate damping models can result in synthetic datasets that underestimate the system's response amplitude or misrepresent the decay envelopes, potentially biasing the neural network's learning process.

Numerical Stiffness and Instability. Systems governed by rate-independent hysteresis (e.g., the Bouc-Wen model) or non-smooth transitions (e.g., contact/impact) often result in stiff differential equations. These equations are characterized by time scales that differ by orders of magnitude, where the solution can change rapidly in a short interval. Standard explicit integration schemes may become numerically unstable or require prohibitively small time steps to maintain accuracy. As discussed in Section 2.1.3, mitigating this requires the use of adaptive solvers (e.g., `dopri5` or `Tsit5`) capable of dynamically refining the integration step during transient events to prevent numerical artifacts from contaminating the training data.

Experimental Data Integrity and Pre-processing

Transitioning from simulation to experimental benchmarks introduces stochastic distortions that can severely degrade model performance if not rigorously addressed.

Sensor Noise and Signal Drift. Real-world sensors are subject to non-Gaussian noise, electromagnetic interference, and low-frequency thermal drift. Piezoelectric accelerometers, for instance, often exhibit a DC offset or slow baseline wander that does not reflect physical motion. If left uncorrected, deep learning models particularly Autoencoders may waste capacity learning to reconstruct this drift rather than the structural dynamics.

- *Mitigation Strategy.* In the processing of the Magneto-Elastic Beam (Magnolia) dataset, this issue was addressed by applying a digital high-pass filter to the raw acceleration signals. This pre-processing step effectively removed the low-frequency trend components and DC bias, isolating the dynamic content relevant to the damage detection task.

Multi-Rate Synchronization and Computational Constraints. Modern SHM systems are increasingly multimodal, combining high-frequency structural sensors (e.g., accelerometers at 100-1000 Hz) with low-frequency environmental sensors (e.g., temperature/wind at 1 Hz or lower). Fusing these disparate data streams requires careful temporal alignment.

- *Mitigation Strategy.* For the SensorFusionTFT application on the LUMO tower, the environmental data required upsampling or broadcasting to match the timestamps of the structural response. Furthermore, the quadratic computational complexity of Transformer attention mechanisms ($\mathcal{O}(L^2)$) presents a scalability bottleneck when dealing with long, high-frequency time series. To manage the computational cost without sacrificing determining dynamics, the structural data was downsampled and concatenated into a unified multivariate tensor. This trade-off balances the retention of spectral information with the memory constraints of GPU-accelerated training.

The Healthy Reference Assumption. Unsupervised learning rests on the critical axiom that the training data represents a truly healthy state. In widely deployed infrastructure, this assumption is often unverifiable; a structure may contain latent, pre-existing defects (e.g., micro-cracks) at the time of sensor installation. If a model is trained on data containing these features, it will learn to accept them as normal, effectively blinding the system to that specific damage type. Ensuring the integrity of the baseline dataset therefore remains a fundamental challenge that often requires hybrid validation using physics-based models or periodic inspections.

2.3 Deep Learning Architectures for SHM

Having established the potential of data-driven methods in chapter 1, this section provides a formal exposition of the deep learning architectures that form the methodological core of this dissertation. These architectures are presented in two main

categories, reflecting a progression in modeling capability and complexity, designed to address the specific challenges of analyzing nonlinear structural dynamics.

First, it is detailed the foundational principles of the Convolutional Neural Network (CNN), a paradigm that excels at learning hierarchical and spatially-invariant local features from sequential or multi-dimensional data. This capability to automatically discover salient patterns without manual feature engineering is fundamental to the data-driven approach of this thesis. It is then described how this convolutional operator serves as the primary building block for two powerful unsupervised models employed in this work: the Autoencoder (AE) for representation learning and the Generative Adversarial Network (GAN) for learning data distributions.

Second, to address the limitations of purely local feature extraction and to explicitly model complex temporal relationships over long horizons, the Transformer architecture is introduced. The Transformer’s self-attention mechanism provides a powerful framework for capturing global dependencies within time-series data. This forms the basis for the advanced diagnostic model developed in the later stages of this research, the SensorFusion Temporal Fusion Transformer (SensorFusionTFT).

This section will thus furnish a comprehensive theoretical toolkit, from the fundamental convolution operation to the sophisticated self-attention mechanism, providing the necessary background for the application chapters that follow.

2.3.1 The Convolution Operator for Feature Extraction

The foundational operation of the architectures used in chapter 3 is the convolution. A convolution is a mathematical operation that uses a small filter, known as a *kernel*, to slide across the input data and extract local features. The kernel itself is a matrix of learnable parameters (weights) that become specialized during training to detect specific patterns, such as spikes, oscillations, or trends in a time-series signal.

1D Convolution and Cross-Correlation. For a univariate 1D input signal, such as a time-series $\mathbf{x} \in \mathbb{R}^L$ of length L , and a 1D kernel $\mathbf{k} \in \mathbb{R}^K$ of size K , the generic output element $y[j]$ is computed as a weighted sum of adjacent input elements. Mathematically, this operation is defined as:

$$y[j] = (\mathbf{x} * \mathbf{k})[j] + b = \sum_{i=1}^K \mathbf{x}[j + i - 1] \cdot \mathbf{k}[i] + b, \quad (2.9)$$

where b is a learnable bias term. It is important to note a technical distinction: strictly speaking, Eq. (2.9) describes a *cross-correlation* operation. A true mathematical convolution would require the kernel weights to be flipped (reversed) relative to the input time axis. However, in the field of deep learning, this operation is universally referred to as convolution, as the learning process renders the distinction regarding weight orientation irrelevant.

Key Properties and Hyperparameters. The convolution operation is governed by several critical properties and hyperparameters that dictate how the network processes information:

- **Weight Sharing and Equivariance.** Unlike a fully connected layer where every output unit has its own unique weights, a Convolutional Neural Network (CNN) shares the same kernel $\mathbf{k}$ across all positions of the input. This drastic reduction in parameters enables the network to be *equivariant* with respect to translation: if the input signal is shifted in time, the resulting feature map is shifted by the same amount, ensuring that a damage signature is detected regardless of when it occurs in the window.
- **Padding:** The convolution operation requires decisions on how to handle the boundaries of the input signal where the kernel would otherwise exceed the valid range. Two common strategies are employed:
 - *Zero-padding.* artificial zero values are added to the signal boundaries. This allows the output size to match the input size, facilitating deep architectures, though it introduces artificial information (zeros) that does not strictly signify a null signal.
 - *Valid Convolution.* The operation is only performed where the kernel fully overlaps with the input. This avoids introducing artifacts but results in a reduction of the output dimension.
- **Stride and Dilation.** the *stride* parameter controls the step size of the kernel as it slides across the input. A stride greater than 1 results in a downsampling of the output feature map. Additionally, *dilation* introduces gaps between the kernel elements, expanding the kernel's field of view without increasing the number of parameters, which is particularly useful for capturing longer-range dependencies without increasing computational cost.

Nonlinearity and Channels. Since convolution is a linear operator, deep learning architectures introduce nonlinearities via activation functions to model complex data distributions. A common choice is the Rectified Linear Unit (ReLU). For a layer with kernel weights ω , bias b , and activation function ϕ , the hidden unit activation at time t is given by:

$$h^{(t)} = \phi \left[b + \sum_{j=1}^K \omega^{(j)} x^{(t+j-1)} \right]. \quad (2.10)$$

Furthermore, modern CNNs operate on multiple dimensions simultaneously. A layer typically consists of multiple kernels in parallel, each producing a distinct *feature map* or *channel*. This allows the network to learn multiple independent features (e.g., one channel detecting high-frequency noise, another detecting low-frequency trends) at the same hierarchical level.

Pooling and Receptive Fields. To reduce the spatial dimensionality and induce invariance to small local translations, pooling operations are employed. *Max-pooling*, for instance, extracts the maximum value within a defined window, summarizing the presence of a feature while discarding precise positional information. Crucially, as convolutional and pooling layers are stacked, the *receptive field* the region of the original input $\mathbf{x}$ that influences a specific hidden unit grows. Deeper neurons thus

capture increasingly global and abstract representations of the structural response, moving from simple local waveforms to complex damage signatures.

2D Convolution for Time-Frequency Representations. While 1D convolution excels at processing raw time-series, the principle extends naturally to 2D data structures, denoted as $\mathbf{X} \in \mathbb{R}^{H \times W}$. In the context of SHM, this input is typically a time-frequency representation, such as a spectrogram or a Continuous Wavelet Transform (CWT), where the two dimensions represent Time and Frequency rather than the spatial height and width found in image processing.

In this scenario, a 2D kernel $\mathbf{K} \in \mathbb{R}^{K_h \times K_w}$ slides across both dimensions of the input simultaneously. The output feature map $\mathbf{Y}$ at position (i, j) is computed via a double summation over the kernel dimensions:

$$y[i, j] = \sum_{m=1}^{K_h} \sum_{n=1}^{K_w} \mathbf{X}[i + m - 1, j + n - 1] \cdot \mathbf{K}[m, n] + b. \quad (2.11)$$

This operation allows the network to learn *spectro-temporal features*. For instance, a specific kernel might learn to identify a chirp signal—a frequency that increases over time which appears as a diagonal feature in a spectrogram, a pattern that a 1D convolution operating on raw time-series might struggle to isolate explicitly.

The hyperparameters governing 2D convolution extend those of the 1D case, but with independent controls for each dimension.

- **Kernel Size** (K_h, K_w). This defines the patch of the time-frequency plane the neuron looks at. A square kernel (e.g., 3×3) captures local dependencies in both time and frequency. However, rectangular kernels can be designed to capture specific structural behaviors: a kernel with shape $(1, K_w)$ acts as a pure frequency filter across a time window, while a kernel of shape $(K_h, 1)$ captures transient broadband impulses at a specific time instant.
- **2D Stride** (S_h, S_w). The stride parameter also becomes a tuple, controlling the downsampling rate independently for time and frequency. For example, setting $S_w > 1$ compresses the temporal resolution, while setting $S_h > 1$ reduces the frequency resolution, allowing the network to trade off fine-grained spectral details for a broader, more abstract view of the system’s energy distribution.
- **2D Padding**. Zero-padding is applied to both the top/bottom (frequency boundaries) and left/right (time boundaries) of the input matrix to preserve spatial dimensions and allow the kernel to process edge features, such as high-frequency content or onset events, without discarding them.

In this dissertation, 2D convolutional layers are central to the GAN architecture when processing Wavelet Transforms, enabling the generator to synthesize realistic joint time-frequency distributions of the structural response.

2.3.2 Unsupervised Representation Learning with Convolutional Autoencoders

The first architecture employed in this work is the Autoencoder (AE), an unsupervised neural network designed for representation learning and dimensionality reduction

[67]. The fundamental principle of an AE is to learn a compressed, low-dimensional representation of the input data, known as the *latent space*, and then reconstruct the original input from this representation with minimal distortion.

Architecture. The AE consists of two main components built from the 1D convolutional layers described in Section 2.3.1:

- **The Encoder (g).** This network takes a high-dimensional input signal $\mathbf{x} \in \mathbb{R}^n$ and maps it to a low-dimensional latent representation $\mathbf{h} \in \mathbb{R}^m$, where $m \ll n$. It is typically composed of a series of convolutional and pooling layers that progressively reduce the spatial dimensions while increasing the feature depth.

$$\mathbf{h} = g(\mathbf{x}) \quad (2.12)$$

- **The Decoder (f).** This network takes the latent representation $\mathbf{h}$ and attempts to reconstruct the original input, producing an output $\tilde{\mathbf{x}} \in \mathbb{R}^n$. It typically mirrors the encoder's architecture, using transposed convolutional layers (also known as deconvolutional layers) to upsample the feature maps back to the original input dimensions.

$$\tilde{\mathbf{x}} = f(\mathbf{h}) = f(g(\mathbf{x})) \quad (2.13)$$

Regularization and Latent Space Topology. A critical challenge in AE design is preventing the network from learning a trivial identity mapping, where the input is simply copied to the output without extracting meaningful features. To ensure robust representation learning, specific constraints are applied.

- **Undercomplete AE.** The dimension of the latent space (m) is required to be smaller than the input dimension (n). This bottleneck forces the network to discard noise and redundancy, retaining only the salient variations of the healthy dynamics.
- **Denoising and Sparse Strategies.** Robustness is further enhanced by introducing noise to the input while training the network to reconstruct the clean original signal (Denoising AE), or by imposing sparsity constraints (e.g., via dropout or L_1 regularization) on the latent activations.

The architecture is illustrated in Figure 2.5. The training process, which aims to minimize a reconstruction-based loss function using an optimizer like ADAM, is detailed in the overarching methodology in Section 2.4.

2.3.3 Unsupervised Anomaly Detection with Convolutional Generative Adversarial Networks

The second architecture is the Generative Adversarial Network (GAN), a powerful class of unsupervised models proposed by [32]. Unlike the reconstructive approach of the AE, the GAN is a generative model that learns to synthesize new data samples that statistically resemble the training distribution.

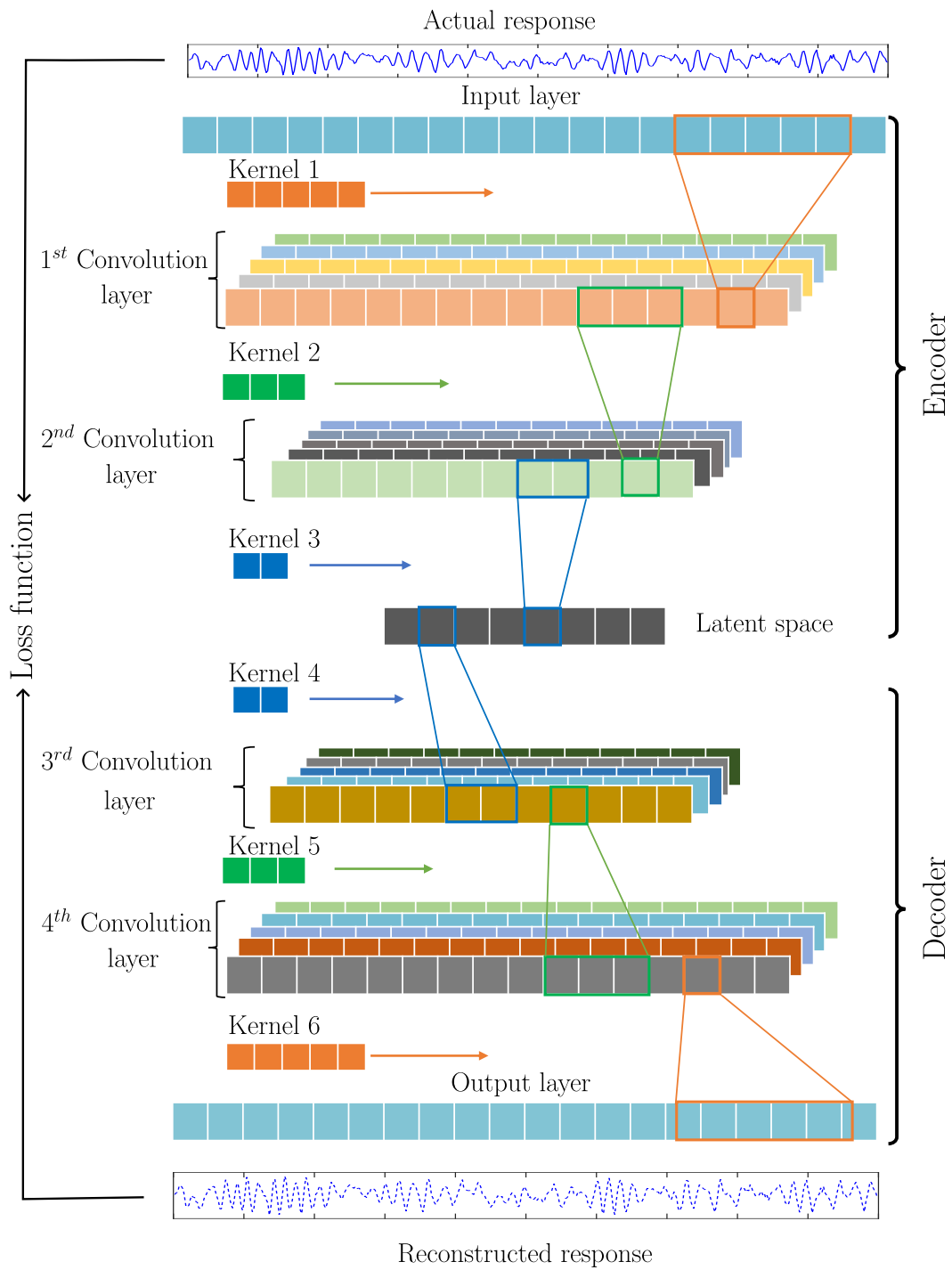


Figure 2.5. Architecture of the 1D Convolutional Autoencoder (AE) used for unsupervised damage detection. The encoder compresses the input signal into a latent space, and the decoder reconstructs it.

Architecture. The GAN framework consists of two competing networks, as depicted in Figure 2.6:

- **The Generator (G).** This network transforms a random noise vector $\mathbf{z}$ sampled from a prior distribution (e.g., Gaussian) into a synthetic time-series $G(\mathbf{z})$. To facilitate stable training of deep convolutional GANs (DCGANs), specific architectural guidelines are followed, such as the use of Batch Normalization layers to stabilize learning and specific activation functions (e.g., LeakyReLU) to prevent vanishing gradients.
- **The Discriminator (D).** This network acts as a binary classifier, receiving either a real data sample $\mathbf{x}$ from the training set or a fake sample $G(\mathbf{z})$ from the generator. Its objective is to correctly identify the source of the sample, outputting a probability close to 1 for real samples and 0 for fake ones. The discriminator is a deep CNN that uses standard convolutional layers to downsample the input time-series into a single probabilistic output.

This architectural framework can be implemented with either 1D CNNs for processing raw time-series directly, or with 2D CNNs for processing 2D time-frequency representations of the signals, allowing the network to learn from both temporal and spectral patterns.

Adversarial Training and Loss Formulation. The training is formulated as a minimax game where the Generator tries to minimize the probability that the Discriminator classifies its output as fake, while the Discriminator tries to maximize its classification accuracy. The objective function is derived from the binary cross-entropy loss:

$$\min_{\theta_G} \max_{\theta_D} V(D, G) = \mathbb{E}_{\mathbf{x} \sim P_{data}} [\log D(\mathbf{x})] + \mathbb{E}_{\mathbf{z} \sim P_z} [\log (1 - D(G(\mathbf{z})))] \quad (2.14)$$

Theoretically, this optimization minimizes the Jensen-Shannon (JS) divergence between the real data distribution P_{data} and the generated distribution P_g .

- The first term, $\mathbb{E}_{\mathbf{x}}[\log D(\mathbf{x})]$, encourages the discriminator to output 1 for real data.
- The second term, $\mathbb{E}_{\mathbf{z}}[\log (1 - D(G(\mathbf{z})))]$, encourages the discriminator to output 0 for generated data.

The solution to this game is a Nash equilibrium where the generator produces perfect samples ($P_g = P_{data}$) and the discriminator is reduced to random guessing ($D(\mathbf{x}) = 0.5$). In the context of SHM, I utilize the trained Discriminator as an anomaly detector: it outputs high probabilities for healthy data (which it recognizes as real) and low probabilities for damaged data (which it views as anomalous or fake), as detailed in the inference methodology.

2.3.4 Sequential Data Modeling with Attention Mechanisms and Transformer Architectures

While convolutional architectures excel at learning local, spatially-invariant features, they possess an inherent limitation in modeling structural dynamics: their receptive field is fixed. In nonlinear systems, particularly those exhibiting hysteresis or chaotic

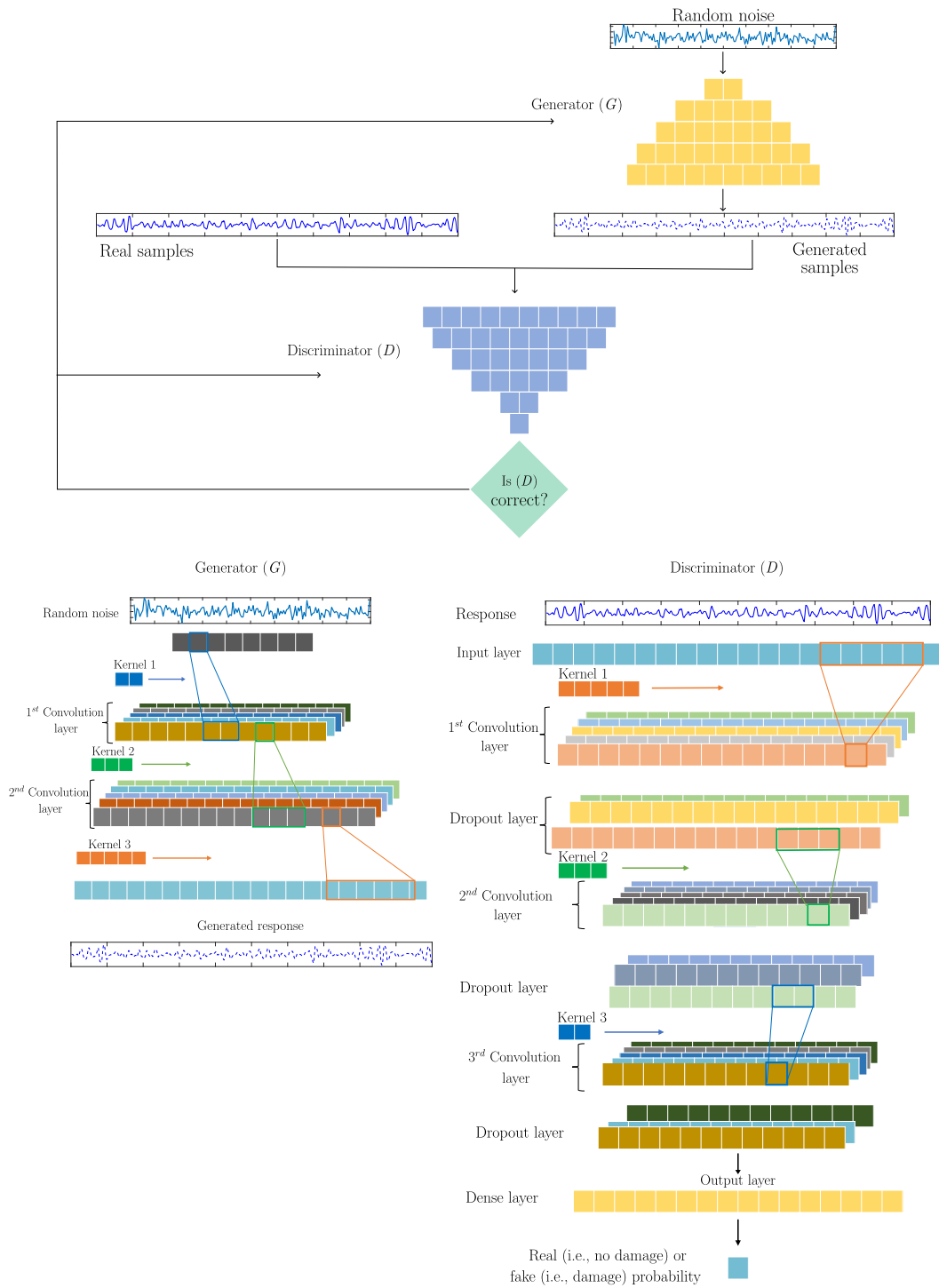


Figure 2.6. Architecture of the 1D Convolutional Generative Adversarial Network (GAN) used for damage detection. The Generator creates synthetic signals from noise, while the Discriminator learns to distinguish real (healthy) signals from synthetic/anomalous ones.

behavior, the current state of the system is often a function of its entire deformation history (fading memory). To address this challenge, this dissertation employs a more advanced architecture based on the Transformer, which utilizes a *self-attention* mechanism to capture global, long-range dependencies within a time-series regardless of the temporal distance between events.

A particularly sophisticated variant of this architecture is the Temporal Fusion Transformer (TFT), originally proposed by Lim et al. [44] for interpretable multi-horizon time-series forecasting. The original TFT was designed to handle a rich variety of inputs, including static covariates (e.g., sensor location) and known future inputs. In this work, this powerful framework is adapted to create the **Sensor-FusionTFT**, a specialized architecture for unsupervised damage detection. This section provides a rigorous mathematical exposition of the four fundamental building blocks employed in this architecture—Gating mechanisms, Residual Networks, Variable Selection, and Attention—explaining their specific utility in processing noisy structural response data.

Foundational Concepts: The Attention Mechanism

The Transformer architecture, first introduced by [82], represents a paradigm shift in sequence modeling. Unlike Recurrent Neural Networks (RNNs) that process data sequentially ($t \rightarrow t + 1$), the self-attention mechanism allows the model to process the entire sequence in parallel, weighing the importance of all past time steps when encoding the representation for the current step.

Positional Encoding. The self-attention mechanism is permutation-invariant; it treats the input as a bag of vectors without inherent order. For structural dynamics, where the order of events (causality) is critical, positional information must be explicitly injected. Following the canonical approach [82], a unique Positional Encoding (PE) vector is added to the input embeddings. This is calculated using fixed sinusoidal functions of varying frequencies:

$$\text{PE}_{(\text{pos}, 2i)} = \sin\left(\frac{\text{pos}}{10000^{2i/d_{\text{model}}}}\right), \quad \text{PE}_{(\text{pos}, 2i+1)} = \cos\left(\frac{\text{pos}}{10000^{2i/d_{\text{model}}}}\right) \quad (2.15)$$

Mathematically, this allows the model to learn to attend by relative positions, which is essential for capturing wave propagation phenomena where the time-lag between events is constant regardless of absolute time.

Interpretable Multi-Head Attention. The core of the Transformer is the calculation of attention weights. This operation can be conceptualized as a differentiable retrieval system involving three matrices: Queries (**Q**), Keys (**K**), and Values (**V**). The model computes a compatibility score between a query (representing the current state) and all keys (representing historical states), normalizes these scores via Softmax to obtain probabilities, and then uses these weights to create a weighted aggregate of the values.

Standard Transformers use multiple heads that learn different dependency patterns independently. However, this often results in disjoint attention maps that are difficult to interpret physically. To resolve this, an *Interpretable Multi-Head Attention*

mechanism is employed [44]. Here, the attention weights are shared across heads to produce a single, distinct attention matrix. Given input features transformed into $\mathbf{Q}, \mathbf{K}, \mathbf{V}$ projections, the mechanism averages the attention scores across N_H heads before aggregation:

$$\text{Attention}(\mathbf{Q}, \mathbf{K}, \mathbf{V}) = \mathbf{W}_H \left(\frac{1}{N_H} \sum_{h=1}^{N_H} \text{Softmax} \left(\frac{\mathbf{Q}_h \mathbf{K}_h^T}{\sqrt{d_k}} \right) \right) \mathbf{W}_V \mathbf{V} \quad (2.16)$$

In this formulation, the input tensor is projected into N_H separate Query ($\mathbf{Q}_h$) and Key ($\mathbf{K}_h$) matrices, one for each attention head, where N_H is a model hyperparameter. The term $\mathbf{Q}_h \mathbf{K}_h^T$ computes the dot-product similarity within each head, which is scaled by the dimension of the keys, d_k , to stabilize gradients. Crucially, the Value matrix ($\mathbf{V}$), which contains the actual information at each time step, is transformed by a single, shared projection matrix $\mathbf{W}_V$ to create a common value representation. The Softmax-normalized scores from all heads are then averaged, and this single, aggregated attention matrix is used to weight the shared value representation. Finally, a second learnable projection matrix, $\mathbf{W}_H$, transforms this result into the final output of the attention block. This modification is critical for SHM diagnostics as it forces the model to learn a single consensus over the temporal direction, allowing us to directly visualize which historical loading events the model considers significant for the current prediction.

Core Architectural Components

The SensorFusionTFT is constructed from a series of sophisticated and reusable building blocks designed to enable stable training, complex nonlinear processing and information flow.

Gated Linear Unit (GLU). In experimental SHM, sensor data is often contaminated by environmental noise or trivial fluctuations. To prevent this noise from propagating deep into the network, a gating mechanism is required. The Gated Linear Unit (GLU) [20] acts as a trainable switch. It takes an input $\mathbf{x}$ and splits it into two paths: a value path and a gate path.

$$\text{GLU}(\mathbf{x}) = \underbrace{(\mathbf{W}_4 \mathbf{x} + \mathbf{b}_4)}_{\text{Value}} \odot \underbrace{\sigma(\mathbf{W}_3 \mathbf{x} + \mathbf{b}_3)}_{\text{Gate} \in [0,1]}, \quad (2.17)$$

where σ is the Sigmoid activation function and $\odot$ is the element-wise Hadamard product. If the gate outputs near 0, the information is suppressed; if near 1, it passes through. This allows the network to learn to selectively ignore noisy or irrelevant time steps.

Gated Residual Network (GRN). Building upon the GLU, the Gated Residual Network (GRN) acts as the primary nonlinear processing unit (Figure 2.7a). Deep neural networks often suffer from the vanishing gradient problem, where learning signals degrade as they propagate back through many layers. The GRN addresses

this by combining a *Skip Connection* with nonlinear gating. For an input $\mathbf{a}$ and optional context vector $\mathbf{c}$:

$$\text{GRN}(\mathbf{a}, \mathbf{c}) = \text{LayerNorm}(\mathbf{a} + \text{GLU}(\mathbf{W}_2 \cdot \text{ELU}(\mathbf{W}_1 \mathbf{a} + \mathbf{W}_c \mathbf{c}))) \quad (2.18)$$

The architecture serves two mathematical purposes.

1. **Adaptive Complexity.** If the relationship in the data is simple (linear), the GLU can effectively suppress the nonlinear branch (setting weights to 0), and the skip connection ($\mathbf{a} + \dots$) allows the unmodified signal to pass through. This allows the model to default to a simpler representation if necessary.
2. **Gradient Flow.** The addition operation in the skip connection creates a gradient superhighway, ensuring stable training even for deep stacks of layers.

The SensorFusionTFT Architecture

The novel contribution of this dissertation is the architectural assembly of these blocks into the **SensorFusionTFT**. To tailor the original TFT for unsupervised damage detection, I introduce a specialized Time -Varying Sensor Fusion mechanism.

1. Parallel Input Embedding. Raw sensor data (scalars) typically lack sufficient capacity to represent complex states. Therefore, the first step is to project the scalar time-series $x_t^{(j)}$ of each sensor j into a high-dimensional vector space $\mathbf{e}_t^{(j)} \in \mathbb{R}^{d_{model}}$ using a linear transformation. This allows the model to learn rich feature representations for each sensor independently before mixing them.

2. Time -Varying Variable Selection Network (VSN). In a large sensor network, damage is often localized; at any given moment, only a subset of sensors may carry relevant information about the structural anomaly. Standard concatenation of all sensors would drown the useful signal in noise from unaffected sensors. To solve this, I re-engineered the VSN to perform dynamic, instance-wise feature selection.

Mathematically, the VSN computes a weight $w_{t,j}$ for each sensor j at each time step t . It first concatenates all sensor embeddings and passes them through a GRN to capture nonlinear correlations between sensors. The output is then normalized via a Softmax function across the sensor dimension:

$$\mathbf{v}_w = \text{GRN}_{\text{weight}}(\dots); \quad \mathbf{w}_t = \text{Softmax}(\mathbf{v}_w) \in \mathbb{R}^{N_{sensors}} \quad (2.19)$$

This ensures that $\sum_j w_{t,j} = 1$, effectively creating a probability distribution over the sensors that changes at every time step.

3. Weighted Sensor Fusion. The unified representation of the system state, $\mathbf{f}_t$, is then computed as a probability-weighted sum of the individual processed sensor embeddings:

$$\mathbf{f}_t = \sum_{j=1}^N w_{t,j} \cdot \text{GRN}_{\text{feat}}(\mathbf{e}_t^{(j)}) \quad (2.20)$$

This operation filters out irrelevant sensors (low $w_{t,j}$) and amplifies those showing significant activity, effectively performing automated, learned signal processing.

4. Temporal Fusion Decoder. Finally, the fused sequence $\mathbf{f}_{1:T}$ is processed by the interpretable multi-head attention block described in Eq. (2.16). This layer correlates the fused state at the current time t with significant events in the history $t - k$, enabling the detection of anomalies arising from long-term memory effects or hysteretic energy dissipation.

The specific training objectives and the formulation of damage metrics based on this architecture are detailed in Section 2.4.

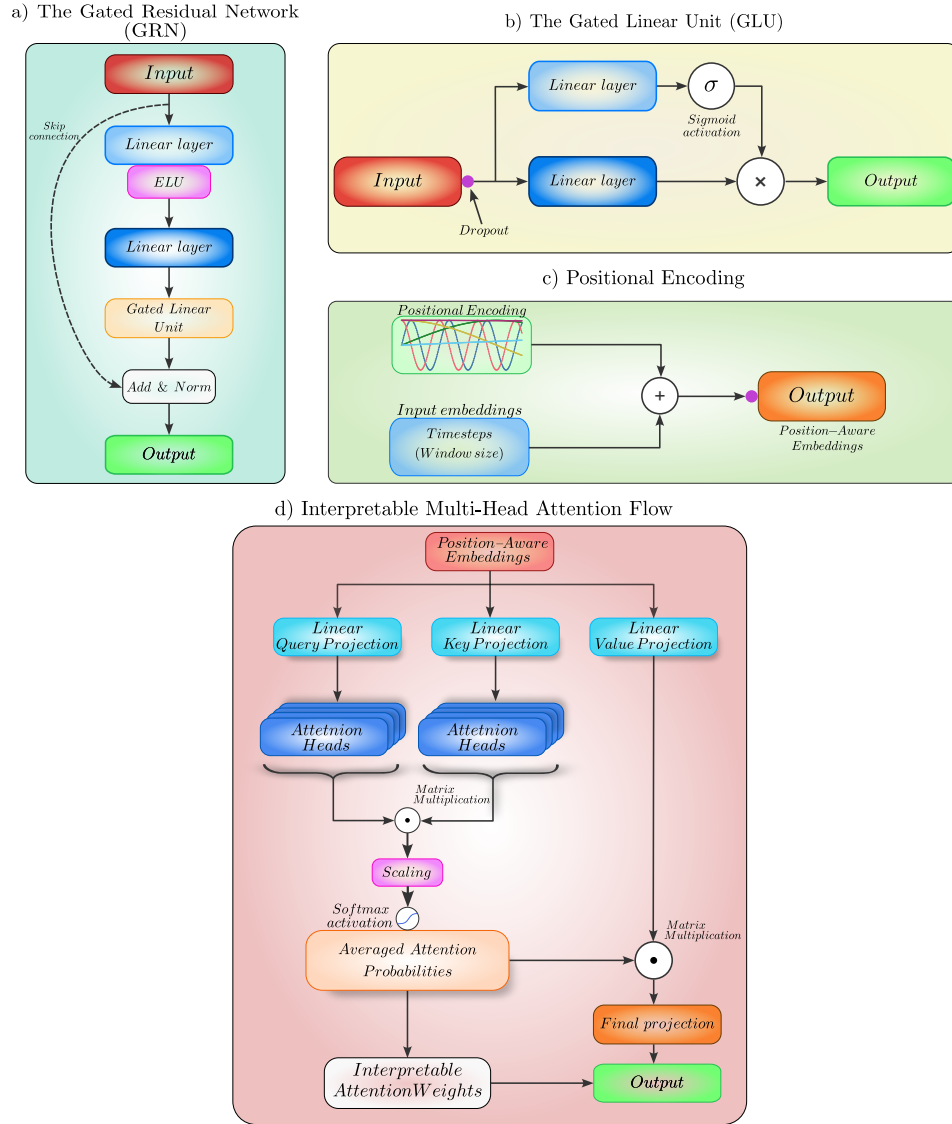


Figure 2.7. Core components of the SensorFusionTFT architecture: (a) Gated Residual Network (GRN), (b) Gated Linear Unit (GLU), (c) Positional Encoding, and (d) Interpretable Multi-Head Attention.

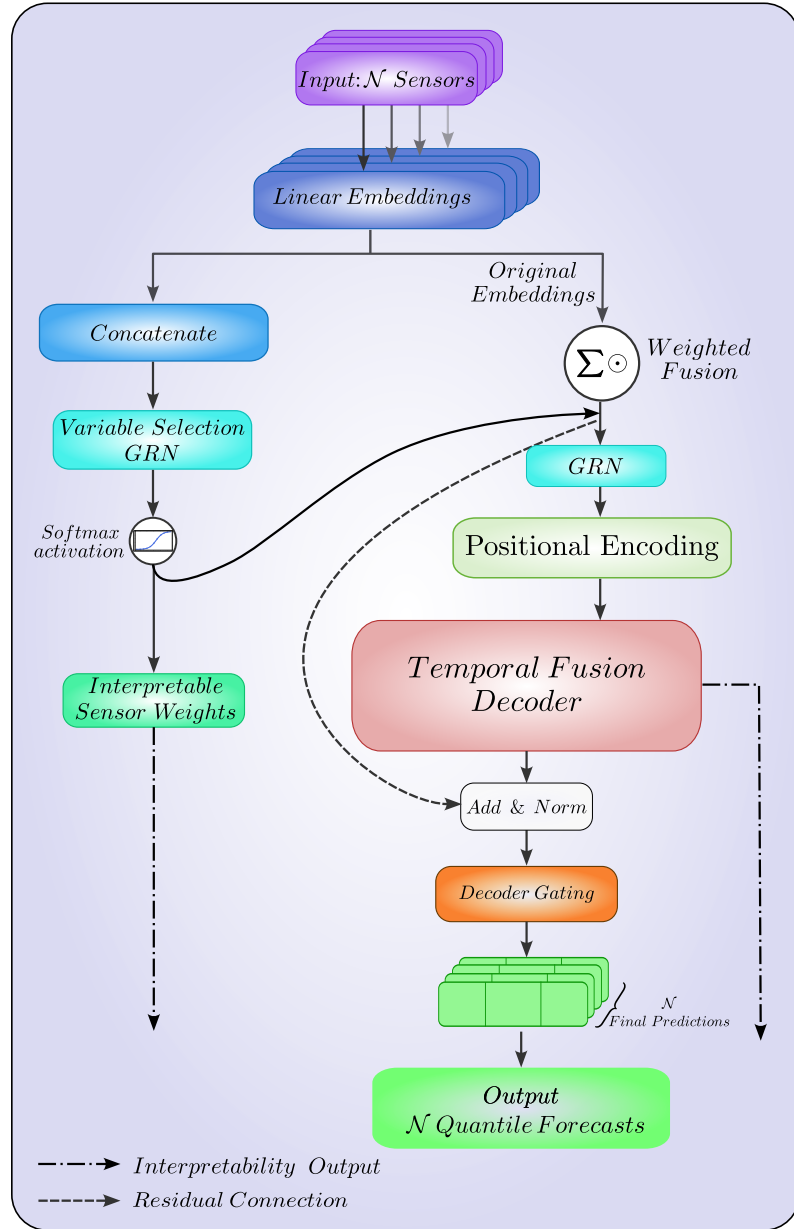


Figure 2.8. The proposed SensorFusionTFT architecture. The model dynamically fuses information from N sensors via a Time-Varying Sensor Fusion Network and processes the resulting unified time series through a Temporal Fusion Decoder to capture long-range dependencies.

2.4 Methodology for Unsupervised Damage Detection

This section details the unified computational framework employed throughout this dissertation to train, validate, and evaluate all deep learning architectures. The methodology is rooted in an unsupervised learning paradigm, where the models are tasked with learning a high-fidelity representation of a structure's normal, healthy behavior without exposure to any examples of damage. This approach is critically

important for real-world SHM, where comprehensive datasets of damaged states are rarely available. The following subsections will sequentially detail this framework, from the high-level paradigm and data preparation to the specific training objectives and the formulation of anomaly scores used for inference.

2.4.1 The Unsupervised Learning Paradigm

The core philosophy of this work is that of unsupervised anomaly detection, also known as novelty detection. The fundamental premise is that a deep learning model, when trained exclusively on data from a system in its healthy state, will become a specialized expert on that normal behavior. It learns to represent the underlying data distribution, $P(\mathbf{x}|\text{state}=\text{healthy})$.

During the subsequent inference phase, the trained model is presented with new, unseen data, $\mathbf{x}_{\text{test}}$. If this data originates from a healthy state, the model should be able to process it with high fidelity (e.g., reconstruct it with low error). However, if $\mathbf{x}_{\text{test}}$ originates from a damaged state, its statistical properties will deviate from the learned healthy distribution. The model, being an expert only on healthy dynamics, will consequently fail to process this anomalous data with the same fidelity. This failure is quantified by an *anomaly score*, $S(\mathbf{x}_{\text{test}})$. A score that exceeds a statistically defined threshold indicates the detection of damage. Furthermore, it is hypothesized that the magnitude of this anomaly score will be correlated with the extent of the damage. Ideally, a greater deviation from the learned healthy state, corresponding to a higher level of damage severity, should produce a monotonically increasing anomaly score.

In the context of the damage assessment hierarchy outlined in chapter 1, this framework is designed to address multiple levels. The detection of a statistically significant deviation in the anomaly score directly satisfies **Level I: Detection**. By establishing a correlation between the magnitude of this score and the severity of damage, this thesis makes a significant contribution towards **Level III: Quantification**. Finally, for advanced architectures such as the SensorFusionTFT, the analysis of internal model mechanisms-like the dynamic sensor importance weights-provides a pathway towards achieving **Level II: Localization**, by identifying which sensor channels are most affected by the anomalous structural behavior.

2.4.2 Data Preparation and Pre-processing

To prepare the raw sensor data for the deep learning models, a consistent pre-processing pipeline is applied. All time-series data, whether from numerical simulations or experimental measurements, are first normalized using min-max scaling to the range $[0, 1]$ specifically for the convolution based models. This step is crucial for stabilizing the training process, reducing the variability among data points from different scales, and ensuring that the models are not biased by the absolute amplitude of the response signals. It is acknowledged, however, that global min-max scaling assumes the training set covers the full dynamic range of the system. In scenarios where damage induces extreme amplitudes exceeding the healthy training bounds, this scaling may introduce saturation effects, a limitation addressed in the Transformer-based approach by using instance-wise standardization where appropriate. Following

normalization, the data is segmented into appropriate formats.

Data Formatting for Convolutional Networks (AE and GAN). For the 1D CNN-based architectures, the continuous time-series are segmented into fixed-length, 2-second windows sampled at 250 Hz. A total of 2,000 such time-series samples are generated from the healthy system response under various magnitudes of random excitation. This database is then partitioned into a training set (80%), used to learn the model parameters, and a validation set (20%), used to monitor for overfitting and to perform hyperparameter tuning. It is highlighted that neither the applied loading conditions nor the specific system characteristics are explicitly fed to the models, meaning both AE and GAN are designed to perform output-only damage detection on unknown dynamic systems.

Data Formatting for the SensorFusionTFT. For the Transformer-based architecture, a self-supervised, sliding window technique is employed. For a set of N sensor sequences $\{\mathbf{x}^{(j)}\}_{j=1}^N$, a look-back window of length W is used to predict a future horizon of length F , creating input-target pairs from the continuous healthy operational data:

$$\mathbf{X}_t = \{(x_{i,j})_{i=t+1}^{t+W}\}_{j \in \mathcal{C}_{in}}, \quad (2.21)$$

$$\mathbf{Y}_t = \{(y_{i,k})_{i=t+W+1}^{t+W+F}\}_{k \in \mathcal{C}_{out}}. \quad (2.22)$$

The full training dataset is generated by sliding this window across the time series of healthy operational data. This formulation is flexible and accommodates two distinct and powerful validation scenarios used in this dissertation:

Auto-Encoding for Reconstruction. In the primary framework, the input and target channel sets are identical ($\mathcal{C}_{in} = \mathcal{C}_{out}$). This tasks the model with a self-supervised objective: reconstructing the system's immediate future response based on its immediate past. Success in this task demonstrates that the model has learned the baseline temporal dynamics of the observed sensor signals.

Physics-Informed Surrogate Modeling. In a more advanced application of the framework, the input and target channels are deliberately different ($\mathcal{C}_{in} \neq \mathcal{C}_{out}$). Here, the model is tasked with learning a mapping from easily observable quantities to non-observable, physically meaningful internal variables. For instance, the input set may consist of measured kinematics (e.g., displacement, acceleration) and any known external excitation, while the target is the total internal restoring force of the system. This presents a more challenging objective: the model must learn to act as a high-fidelity, data-driven surrogate that emulates the underlying constitutive law of the system.

By training the model on these self-generated tasks using only healthy data, it becomes an expert on the baseline system dynamics. The overarching goal of unsupervised damage detection is then achieved during inference by identifying significant deviations between the model's predictions and new measurements.

2.4.3 Formulation of Training Objectives and Loss Functions

While all architectures follow the same unsupervised paradigm, each is trained to a unique objective defined by its specific loss function. A common element across all training procedures is the use of the Adaptive Moment Estimation (ADAM) optimizer, a gradient-based technique that computes adaptive learning rates for each parameter [36].

Reconstruction Objective for Autoencoders

The objective of the Autoencoder is to learn an efficient encoding and decoding of the data that minimizes the reconstruction error. This is accomplished through proper training, which consists of finding the optimal set of parameters θ that rule the encoder g and decoder f . Two main strategies are employed to ensure a meaningful latent representation and to avoid overfitting: creating a bottleneck by reducing the dimensionality m of the latent space compared to the data dimension n , and adding a regularization term. In this work, the L_2 regularization technique is implemented, leading to the following optimization problem:

$$\min_{\theta} \left\{ \mathbb{E}_{\mathbf{x} \sim P_{\text{healthy}}} [\Delta(\mathbf{x}, f(g(\mathbf{x}; \theta)))] + \lambda \|\theta\|_2^2 \right\}, \quad (2.23)$$

where the first term is a loss function (here, Mean Absolute Error, MAE) that quantifies the average reconstruction error, and the second term is the L_2 -norm of the parameters weighted by a coefficient λ . To further prevent overfitting, training is halted using an early stopping criterion.

Adversarial Objective for Generative Adversarial Networks

The training of a GAN is a non-cooperative game, where optimization is attained at the Nash equilibrium point [64]. At this point, the Generator produces data so realistic that the Discriminator is no better than chance at distinguishing real from fake. This is achieved by solving the min-max optimization problem:

$$\min_{\theta_G} \max_{\theta_D} \{ \mathbb{E}_{\mathbf{x} \sim P_{\text{healthy}}} [\log D(\mathbf{x}; \theta_D)] + \mathbb{E}_{\mathbf{z} \sim P_z} [\log (1 - D(G(\mathbf{z}; \theta_G)))] \}, \quad (2.24)$$

where the training alternates between updating the Discriminator's parameters θ_D to maximize this objective (classify correctly) and updating the Generator's parameters θ_G to minimize it (fool the discriminator). Overfitting in the discriminator is controlled via dropout regularization.

Probabilistic Reconstruction Objective for the SensorFusionTFT

The SensorFusionTFT is trained not to predict a single point-estimate, but to forecast a full probabilistic distribution of the future response using quantile regression [37, 90]. The objective is to minimize the Quantile Loss (or Pinball Loss). For a specific quantile $\tau \in [0, 1]$, the loss is:

$$L_{\tau}(y, \hat{y}^{(\tau)}) = \begin{cases} \tau(y - \hat{y}^{(\tau)}) & \text{if } y \geq \hat{y}^{(\tau)} \\ (1 - \tau)(\hat{y}^{(\tau)} - y) & \text{if } y < \hat{y}^{(\tau)} \end{cases} \quad (2.25)$$

where y is the ground truth and $\hat{y}^{(\tau)}$ is the model's prediction for the τ -th quantile. The total loss is the sum of this function over all predicted quantiles (e.g., $\mathcal{T} = \{0.1, 0.5, 0.9\}$) and all target variables. This loss function incentivizes the model to produce predictions where the true value lies between the predicted quantiles the correct percentage of the time. Total model loss:

$$\mathcal{L} = \frac{1}{NTQ} \sum_{j=1}^N \sum_{t=1}^T \sum_{\tau \in \mathcal{T}} L_{\tau}(y_{t,j}, \hat{y}_{t,j}^{(\tau)}) \quad (2.26)$$

In Eq. (2.26), the total loss $\mathcal{L}$ is the average of the quantile loss computed over all dimensions of the prediction task. The index j iterates from 1 to N , the total number of target variables (i.e., sensors). The index t iterates from 1 to T , the number of time steps in the prediction horizon. Finally, the summation over $\tau \in \mathcal{T}$ accumulates the loss for all Q predicted quantiles in the set $\mathcal{T}$. The normalization factor $1/NTQ$ ensures that the total loss is a mean value, making it independent of the number of sensors, the forecast horizon, or the number of quantiles. Quantile regression [37, 90] provides robust uncertainty intervals essential for SHM.

Optimizer: Adaptive Moment Estimation (ADAM). All models are trained using the ADAM optimizer. It computes adaptive learning rates for each parameter based on estimates of the first and second moments of the gradients. The update rules are as follows:

$$\mathbf{m}_{\tau} = \beta_1 \mathbf{m}_{\tau-1} + (1 - \beta_1) \nabla_{\tau}, \quad (2.27a)$$

$$\mathbf{v}_{\tau} = \beta_2 \mathbf{v}_{\tau-1} + (1 - \beta_2) \nabla_{\tau}^2, \quad (2.27b)$$

$$\hat{\mathbf{m}}_{\tau} = \frac{\mathbf{m}_{\tau}}{1 - \beta_1^{\tau}}, \quad (2.27c)$$

$$\hat{\mathbf{v}}_{\tau} = \frac{\mathbf{v}_{\tau}}{1 - \beta_2^{\tau}}, \quad (2.27d)$$

$$\boldsymbol{\theta}_{\tau+1} = \boldsymbol{\theta}_{\tau} - \frac{\eta}{\sqrt{\hat{\mathbf{v}}_{\tau}} + \epsilon} \hat{\mathbf{m}}_{\tau}, \quad (2.27e)$$

where $\mathbf{m}_{\tau}$ and $\mathbf{v}_{\tau}$ are the first and second moment estimates, β_1 and β_2 are their exponential decay rates, ∇_{τ} is the gradient of the objective function, η is the learning rate, and ϵ is a small constant for numerical stability.

2.4.4 Inference and Formulation of Anomaly Scores

Once trained, the models are deployed for inference on new data. For the numerical nonlinear systems considered, the damage dataset is composed of responses generated from systems with stiffness reductions ranging from 5% (incipient damage) to 30% (severe damage) in 5% intervals.

Autoencoder: Reconstruction Error. The suitability of the AE for damage detection relies on the evidence that a model properly trained on undamaged data will yield a large loss function value when attempting to reconstruct the response of a damaged system. The anomaly score S_{AE} is therefore the reconstruction error (MAE) for a given test sample.

GAN: Discriminator Output as a Damage Index. The suitability of the GAN for damage detection stems from the role of its trained discriminator, D . Having been trained adversarially against the generator, the discriminator becomes an expert at identifying the subtle characteristics of healthy structural response data. During the inference phase, its function is repurposed for anomaly detection.

When the trained discriminator is presented with a new test sample, $\mathbf{x}_{\text{test}}$, its output probability, $D(\mathbf{x}_{\text{test}})$, can be interpreted as a health index. If the data originates from a healthy state, it will conform to the distribution the discriminator learned, and the output will be close to 1. Conversely, if the data comes from a damaged system, its statistical properties will differ from the learned norm. The discriminator will correctly identify this deviation, classifying the sample as anomalous (similar to fake data) by outputting a probability close to 0.

To create a more intuitive *damage index*, where the score increases with the severity of the damage, I simply define the anomaly score S_{GAN} as the complement of the discriminator’s output probability:

$$S_{GAN}(\mathbf{x}_{\text{test}}) = 1 - D(\mathbf{x}_{\text{test}}) \quad (2.28)$$

With this formulation, the damage index S_{GAN} is naturally bounded between 0 and 1. A score approaching 0 signifies a healthy state, while a score approaching 1 indicates a high probability of damage. This provides a clear, interpretable metric for quantifying the deviation from the learned baseline of normal behavior.

SensorFusionTFT: Multi-Metric Assessment. The probabilistic nature of the TFT allows for a suite of three distinct and complementary anomaly scores:

- **Metric 1: Reconstruction Error (RMSE).** The most direct indicator of an anomaly is the reconstruction error, which measures the difference between the model’s central prediction and the actual sensor measurement. I use the Root Mean Square Error (RMSE) calculated from the median quantile forecast ($\tau = 0.5$). For a given sensor j and an evaluation period containing T_{eval} time steps, the RMSE is calculated as:

$$\text{RMSE}_j = \sqrt{\frac{1}{T_{\text{eval}}} \sum_{t=1}^{T_{\text{eval}}} (y_{t,j} - \hat{y}_{t,j}^{(0.5)})^2} \quad (2.29)$$

where $y_{t,j}$ is the ground truth measurement and $\hat{y}_{t,j}^{(0.5)}$ is the model’s median prediction for sensor j at time step t . A significant increase in RMSE relative to the healthy baseline indicates that the system’s dynamics have changed in a way the model cannot explain, signaling potential damage.

- **Metric 2: Interval Score.** While RMSE evaluates the accuracy of the central (median) forecast, a robust anomaly detection framework must also assess the reliability of the model’s uncertainty estimates. To this end, I employ the Interval Score, a widely used proper scoring rule for evaluating the quality and calibration of prediction intervals [31]. The Interval Score is designed to penalize intervals that are too wide (lacking sharpness) as well as those that fail to contain the true observation (lacking coverage).

For the purpose of damage detection, we are primarily concerned with mis-calibration, where the dynamics of the damaged system become more volatile than the model’s learned uncertainty bounds for the healthy state. My implementation therefore focuses on the penalty component of the score, which quantifies the magnitude of the breach whenever the true measurement, $y_{t,j}$, falls outside the predicted lower ($\hat{y}_{t,j}^{(\tau_{\text{lower}})}$) and upper ($\hat{y}_{t,j}^{(\tau_{\text{upper}})}$) quantile bounds. For a prediction interval defined by these quantiles (e.g., 0.1 and 0.9), the Interval Score (IS) for sensor j is calculated as the average penalty over the evaluation horizon T_{eval} :

$$\text{IS}_j = \frac{1}{T_{\text{eval}}} \sum_{t=1}^{T_{\text{eval}}} \left[\max(0, y_{t,j} - \hat{y}_{t,j}^{(\tau_{\text{upper}})}) + \max(0, \hat{y}_{t,j}^{(\tau_{\text{lower}})} - y_{t,j}) \right] \quad (2.30)$$

As formulated in Eq. 2.30, if the true value $y_{t,j}$ lies within the prediction interval, the term inside the summation is zero, and no penalty is incurred. However, if a breach occurs, the score increases proportionally to the magnitude of that breach. A significant increase in the Interval Score is therefore a powerful damage indicator, signifying that the observed structural dynamics are no longer consistent with the model’s learned representation of healthy operational uncertainty.

- **Metric 3: Attention Anomaly Score.** The final metric provides a sophisticated measure of temporal dislocation by analyzing changes in the model’s self-attention patterns. While attention mechanisms are widely used for interpretability, I introduce a novel damage indicator, termed the Attention Anomaly Score, derived from the model’s internal attention patterns. This metric is calculated by first establishing robust Regions of Interest (ROIs) from a master attention difference matrix, which captures consistent changes in temporal focus across attention weights. The score, formulated as a ratio of attention energy between damage-activated ‘hotspots’ and suppressed ‘coldspots’, provides a physics-agnostic and highly sensitive measure of deviation from the learned healthy structural behavior. I hypothesize that damage alters the causal relationships between past and future states, forcing the model to re-allocate its attention. To quantify this, I define Regions of Interest (ROIs) within the attention matrix based on characteristic changes observed across a general set of damage scenarios. First, an average attention difference matrix, $\Delta\mathbf{A}$, is computed by subtracting the mean healthy attention pattern from the mean pattern of various damage cases. From this matrix, I define two masks: Hotspot Mask ($\mathcal{R}_{\text{hot}}$): A set of indices (i, k) corresponding to the top percentile of positive values in $\Delta\mathbf{A}$, indicating where attention consistently *increases* with damage. Coldspot Mask ($\mathcal{R}_{\text{cold}}$): A set of indices (i, k) corresponding to the bottom percentile of negative values in $\Delta\mathbf{A}$, indicating where attention consistently *decreases*.

For any new attention matrix $\mathbf{A} \in \mathbb{R}^{W \times W}$ (where W is the window size) generated during inference, the ROI Damage Score is calculated as the ratio

of the total attention energy within these two regions:

$$\text{Score}_{\text{ROI}}(\mathbf{A}) = \frac{\sum_{(i,k) \in \mathcal{R}_{\text{hot}}} a_{ik}}{\sum_{(i,k) \in \mathcal{R}_{\text{cold}}} a_{ik} + \epsilon} \quad (2.31)$$

where a_{ik} are the elements of the matrix $\mathbf{A}$, and ϵ is a small constant for numerical stability. A high score signifies that the model's temporal focus has shifted to align with patterns previously identified as characteristic of a damaged state, providing a powerful, interpretable signal for damage detection and localization in the temporal domain.

Chapter 3

Damage Detection using Convolutional Unsupervised Learning Architectures

Building upon the theoretical foundations established in chapter 2, this chapter presents the first major investigation of this dissertation: the application of 1D convolutional neural network-based architectures for unsupervised damage detection. This methodological approach, centered on training models like the Autoencoder (AE) to learn a baseline of normal behavior from healthy data, aligns with the state-of-the-art in data-driven SHM and has proven to be a promising strategy for practical engineering applications, including the real-time health assessment of large-scale bridges [30]. Here, I extend this paradigm by systematically evaluating the performance of both the Autoencoder and the Generative Adversarial Network (GAN) across the suite of specifically *nonlinear* systems previously introduced. The primary objective is to assess their sensitivity to incipient damage and to perform a critical comparative analysis of their performance and characteristics. The chapter is structured on a system-by-system basis to facilitate a direct and clear comparison of the two architectural approaches. The core methodologies and results presented in this chapter form the basis of a peer-reviewed article published in the journal *Nonlinear Dynamics* [35].

3.1 Case Study Data Protocol

To evaluate the effectiveness of both the AE and GAN architectures, a series of numerical simulations is performed. The goal is to test the models' ability to detect damage in nonlinear systems using only their time-domain response. To replicate a common and challenging real-world scenario, the numerical models are subjected to random vibrations of varying intensity, and the simulated response data is contaminated with white Gaussian noise. This protocol ensures that the models are tested not under a single, idealized condition, but across a range of operational states, forcing them to learn the underlying dynamics rather than memorizing a specific response type. All data is prepared and pre-processed according to the framework detailed in Section 2.4.2.

3.2 Application to a 1-DOF Duffing Oscillator

The first benchmark for evaluation is the single-degree-of-freedom (1-DOF) Duffing oscillator, whose theoretical properties and governing equation were detailed in Section 2.1.1. As established by the analysis of its Frequency Response Curves (Figure 2.1), this system's hardening nonlinearity and amplitude-dependent resonance present a significant challenge, creating ambiguity that can mislead traditional damage detection methods. The now apply the unsupervised anomaly detection framework and the model-specific anomaly scores are applied, as defined in Section 2.4, to assess the capability of the AE and GAN to overcome this challenge.

3.2.1 Autoencoder Performance

The performance of the trained AE is first assessed by its ability to reconstruct the system's time-history response. As shown in Figure 3.1, the AE, having been trained exclusively on healthy data, accurately reconstructs the displacement response of the

undamaged system. However, when presented with data from the damaged system, its reconstruction fidelity degrades significantly. This mismatch between the input and the AE's output highlights a clear deviation from the learned conditions of the training process and thus points to the occurrence of damage. As hypothesized, the magnitude of this reconstruction error, which serves as the anomaly score, visibly increases as the damage level becomes more severe.

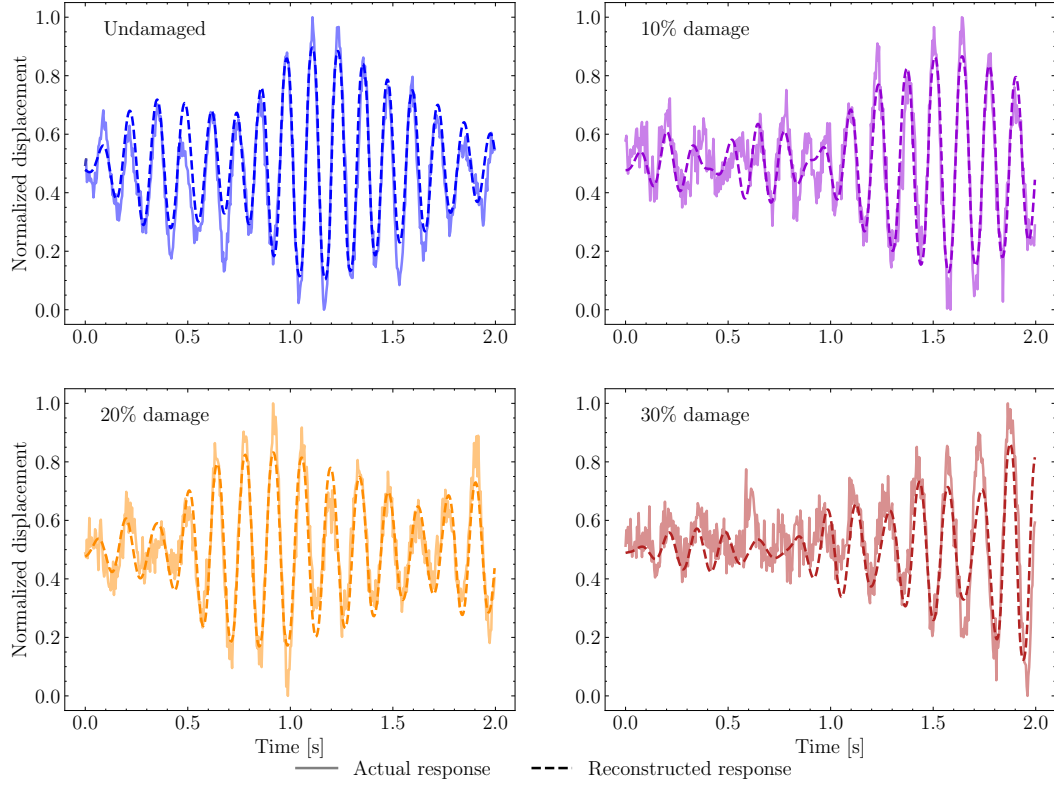


Figure 3.1. Reconstruction of the 1-DOF Duffing system response via the Autoencoder for different damage levels. The mismatch between the actual and reconstructed response grows with increasing damage.

3.2.2 Generative Adversarial Network Performance

In contrast to the AE's reconstructive approach, the GAN framework relies on the fidelity of its Generator to learn the healthy data distribution. The Generator must be able to mimic the undamaged system response with high accuracy to effectively train the Discriminator to be a sensitive detector of anomalies. Figure 3.2 demonstrates the success of this process. It shows a representative pair of actual and generated time-history displacement responses for the undamaged system, along with their corresponding time-frequency representations. The high degree of similarity confirms that the Generator has successfully captured the complex, nonlinear dynamics of the healthy state. This, in turn, implies that the Discriminator has been trained to be an expert on this specific dynamic behavior and will therefore be able to accurately recognize any deviation from it, such as that caused by the occurrence of damage.

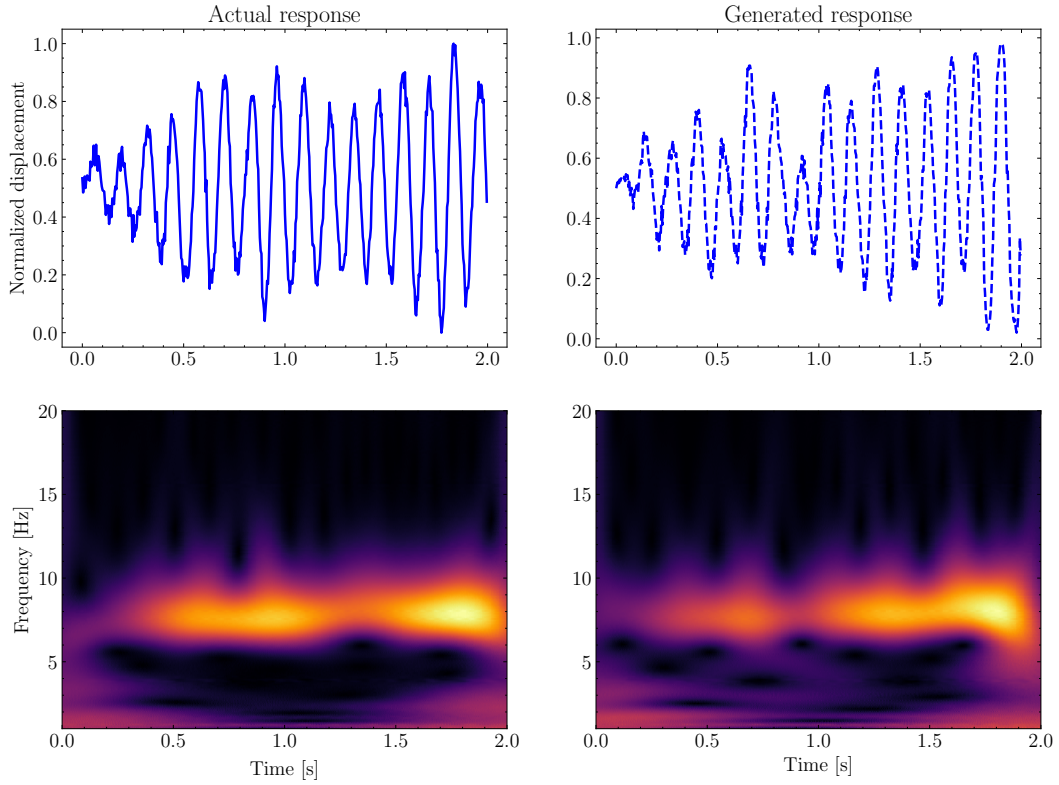


Figure 3.2. Fidelity of the trained GAN Generator. Comparison of a real time-history response from the undamaged 1-DOF system with a synthetic response from the Generator, shown in both the time and time-frequency domains.

3.2.3 Comparative Discussion

To quantitatively compare the damage detection sensitivity of both architectures, the anomaly scores for each are plotted against the damage level in Figure 3.3. The top-left plot shows the distribution of the AE's reconstruction error (MAE), while the top-right plot shows the raw output of the GAN's discriminator. As detailed in Section 2.4.4, the discriminator's output is transformed into an intuitive damage index, defined as its complement.

The results clearly demonstrate the efficacy of the unsupervised paradigm. For the AE, the reconstruction loss values consistently shift upwards as the damage severity increases. Similarly, for the GAN, the discriminator outputs are clustered near 1 for the healthy state, while a growing number of outcomes approach 0 as the damage level increases. The bottom plots in Figure 3.3 show the trend of the normalized anomaly scores. Both the AE's reconstruction loss and the GAN's damage index are properly correlated with the damage level. Notably, a significant deviation from the healthy baseline occurs even at the lowest 5% damage level, demonstrating a satisfactory sensitivity of both deep learning architectures to incipient damage.

For completeness, the specific hyperparameters of the trained AE and GAN architectures used for this case study are detailed in Table 3.1.

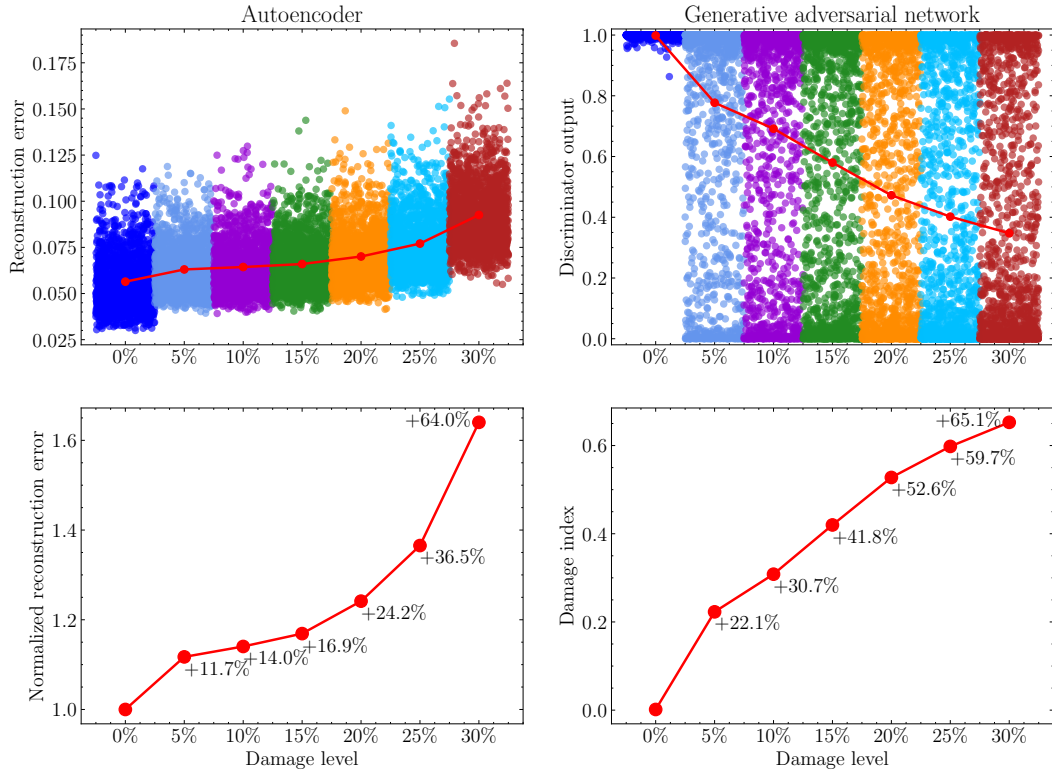


Figure 3.3. Comparative damage sensitivity for the 1-DOF system. Top row: Scatter plots of the raw anomaly scores (AE reconstruction error and GAN discriminator output). Bottom row: Trend of the normalized average anomaly scores, showing a clear correlation with damage level.

Table 3.1. Configuration of the deep learning architectures for the 1-DOF Duffing system.

Type	Component	Convolution layers	Trainable parameters	Kernel size	Number of filters	Strides
AE	Encoder	3	3,767	(100,50,25)	(10,5,2)	(1,1,1)
	Decoder	3	66,056	(25,50,100)	(5,10,1)	(2,2,1)
GAN	Generator	3	16,018	(2,3,6)	(32,32,1)	(2,1,2)
	Discriminator	3	8,117,729	(6,6,6)	(64,64,64)	(2,2,2)

3.2.4 On the Role of the Input Representation: 2D GAN with Wavelet Transforms

While the 1D convolutional architectures operating on raw time-series demonstrated satisfactory performance, they inherently treat the signal as a flat sequence. However, nonlinear structural responses often exhibit transient spectral characteristics such as instantaneous frequency shifts that are best analyzed in the time-frequency domain. To investigate whether explicit time-frequency features yield higher damage sensitivity, a supplementary study was conducted using a 2D Convolutional GAN (2D-GAN) applied to Continuous Wavelet Transforms (CWT) of the Duffing oscillator response.

Data Transformation and Architecture. Instead of raw acceleration vectors, the input data consists of Scalograms generated via the Continuous Wavelet Transform. The transformation was performed using the Complex Morlet wavelet (`cmor1.5-1.0`), chosen for its optimal balance between temporal and spectral resolution. Each time-series sample was converted into a 2D RGB image and resized to dimensions of 69×76 pixels to serve as the input for the 2D-GAN.

The architecture was adapted from the 1D GAN described in Section 2.3.3 to utilize 2D convolutional kernels (5×5) in both the Generator and Discriminator. This allows the network to exploit spatial correlations in the scalogram, effectively learning the joint time-frequency signature of the healthy system (e.g., the specific energy decay patterns and resonant bands).

Comparative Results: 1D vs. 2D Sensitivity. The performance of the 2D-GAN was evaluated using the same damage scenarios as the 1D study. Figure 3.4 presents the side-by-side comparison of the discriminator outputs and the normalized anomaly scores.

Qualitatively, the 2D-GAN (Right column) exhibits a similar clustering pattern to the 1D-GAN (Left column), with the healthy samples (blue) clustered near a discriminator output of 1.0. However, a quantitative analysis of the normalized anomaly scores reveals a distinct advantage in sensitivity for the 2D approach.

- **1D-GAN Sensitivity:** The normalized anomaly score increases by approximately **+65.1%** at the maximum damage level (30% stiffness reduction).
- **2D-GAN Sensitivity:** The 2D architecture achieves a higher relative increase of **+76.5%** for the same damage level.

Critical Discussion on Computational Trade-offs. While the 2D-GAN demonstrates superior sensitivity, this improvement comes at a significant computational cost. The pre-processing step (CWT generation) increases the data storage requirements by two orders of magnitude compared to the raw time-series. Furthermore, the training of the 2D-GAN is considerably more resource-intensive due to the larger number of parameters in the 2D kernels and the processing of high-dimensional image tensors.

Therefore, the choice between 1D and 2D architectures represents a trade-off: the 1D-GAN offers a highly efficient, lightweight solution suitable for edge computing where moderate sensitivity is acceptable, whereas the 2D-GAN offers maximum diagnostic resolution for critical assets where the computational overhead of time-frequency analysis is justified.

3.3 Application to a 2-DOF Nonlinear System

Having demonstrated the efficacy of the convolutional architectures on a single-degree-of-freedom system, two-degree-of-freedom (2-DOF) Duffing oscillator is now addressed, as defined theoretically in Section 2.1.1. This system introduces modal interactions and a more complex dynamic behavior, as illustrated by its FRCs in

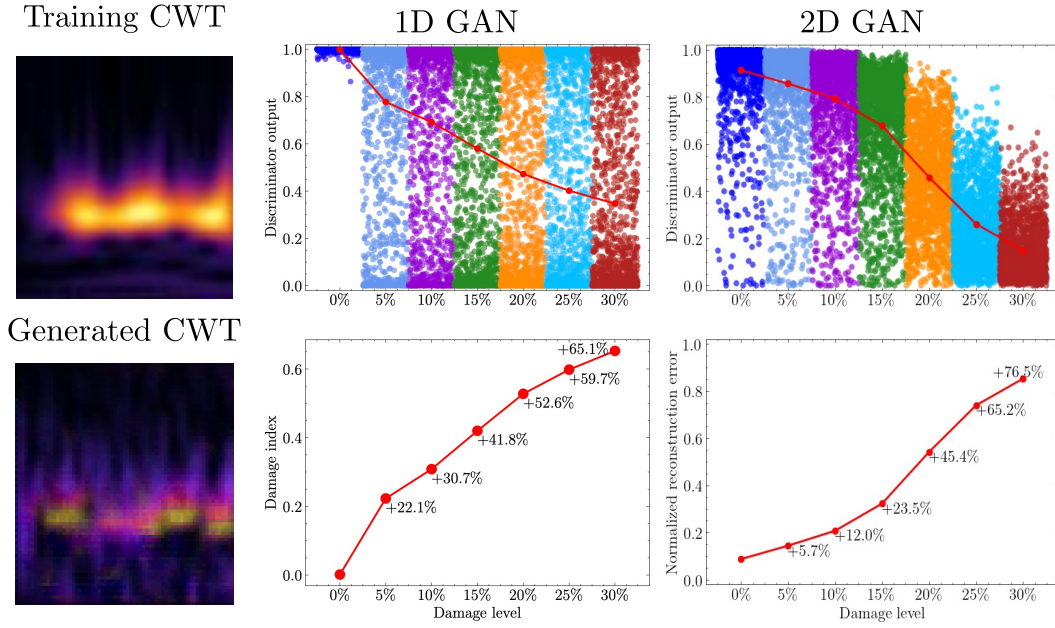


Figure 3.4. Comparative evaluation of input representations for the 1-DOF Duffing system. Left: 1D-GAN operating on time-series. Right: 2D-GAN operating on CWT Scalograms. The 2D approach yields higher sensitivity (+76.5% vs +65.1%) by leveraging spectro-temporal features.

Figure 2.2. This application serves to evaluate the robustness and scalability of the AE and GAN architectures when faced with a more intricate, multi-modal nonlinear response.

3.3.1 Autoencoder Performance

Despite the added complexity from the interaction of the two nonlinear oscillators, the AE architecture continues to perform reliably. As shown in Figure 3.5, the model successfully reconstructs the undamaged response of both oscillators after the training process. As expected, when the system is damaged, the reconstruction accuracy decreases noticeably. The AE, trained only on the dynamics of the healthy system, is unable to accurately reproduce the altered dynamics of the damaged state, resulting in a significant reconstruction error that serves as the damage indicator.

3.3.2 Generative Adversarial Network Performance

Similarly, the GAN architecture demonstrates a satisfactory performance in modeling the more complex 2-DOF system. Figure 3.6 confirms the fidelity of the trained Generator by comparing the real and generated responses of the undamaged oscillators in the time-frequency domain. The strong similarity indicates that the Generator has learned a robust representation of the healthy, multi-modal dynamics. This, in turn, implies that the Discriminator has been successfully trained to recognize this healthy response as genuine and will therefore be prone to classifying the response from a damaged system as anomalous or fake.

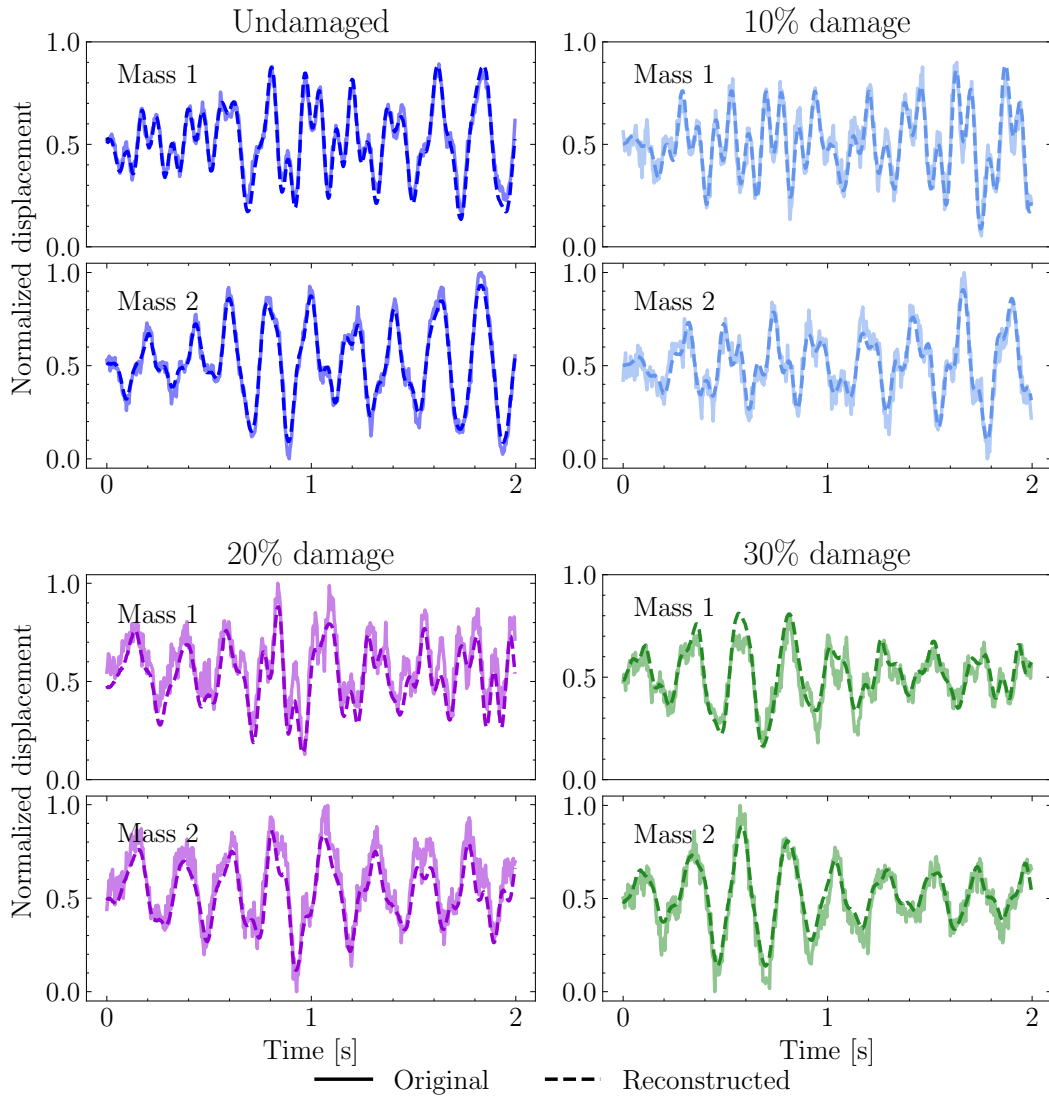


Figure 3.5. Reconstruction of the 2-DOF Duffing system response via the Autoencoder for different damage levels. The model is applied to each degree-of-freedom independently.

3.3.3 Comparative Discussion

A direct comparison of the damage detection performance for both architectures is provided in Figure 3.7. Since the models are applied individually to the response of each mass, the scatter plots illustrate the distribution of anomaly scores for each degree-of-freedom. The trend lines represent the average score across both masses to enable damage detection via a single, aggregated parameter.

A key observation from this more complex case is a divergence in performance stability. While the AE maintains a tight clustering of healthy scores, the GAN displays a concerning variance in the undamaged state (Figure 3.7, top right). The wide spread of discriminator outputs for the healthy condition overlaps significantly with the 5% damage distribution. In an operational setting, this overlap represents

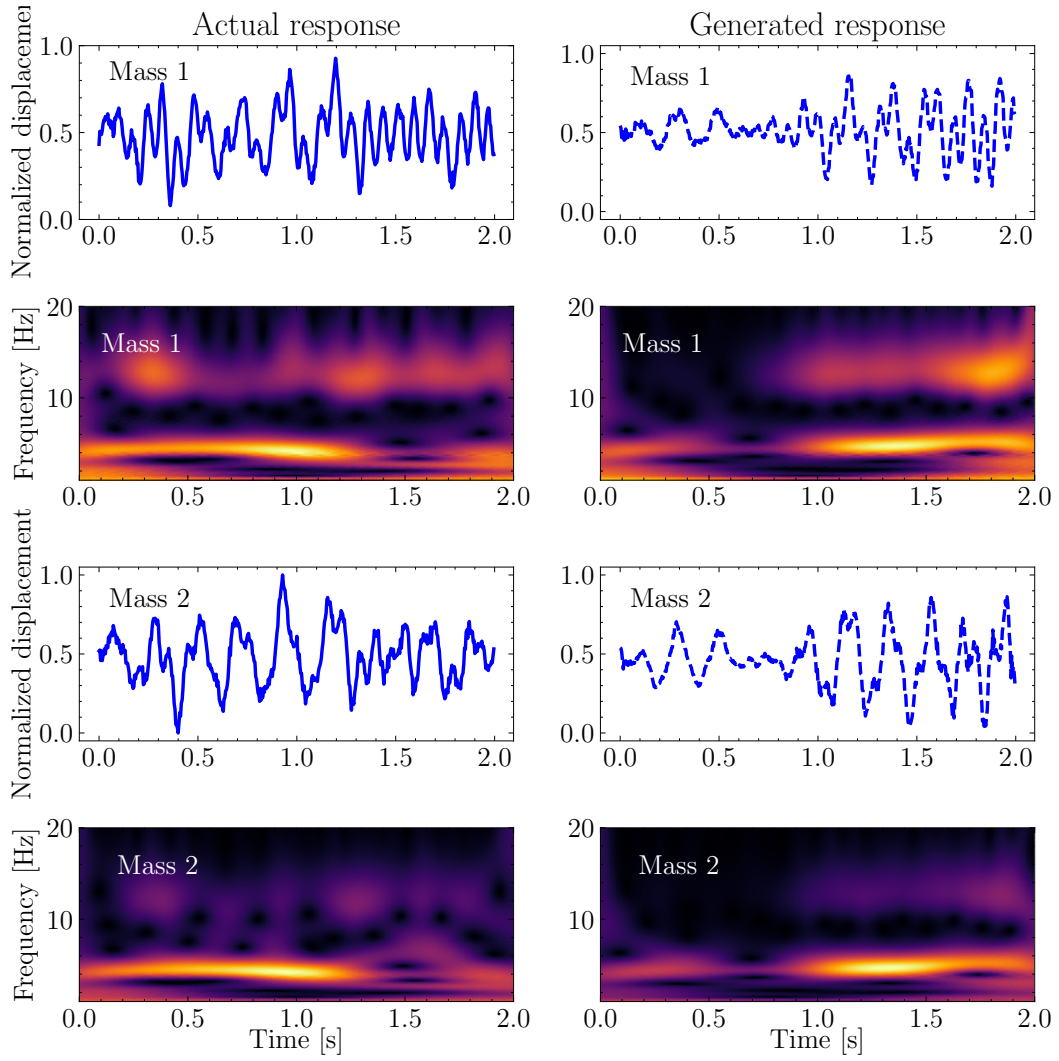


Figure 3.6. Fidelity of the trained GAN Generator for the 2-DOF system. Comparison of real vs. synthetic time-frequency representations for both degrees of freedom in the undamaged state.

a non-trivial probability of false positives, where a healthy system could be flagged as damaged due to the discriminator’s hypersensitivity to benign modal interactions. This highlights a fundamental instability in the adversarial training process for multi-modal nonlinear systems compared to the convex optimization landscape of the Autoencoder.

The specific hyperparameters of the trained AE and GAN architectures used for the 2-DOF system case study are detailed in Table 3.2.

3.4 Application to a Nonlinear Seismic Isolator

The final and most demanding numerical benchmark in this chapter is the hysteretic seismic isolator, whose complex hybrid nonlinearity was detailed in Section 2.1.3. This

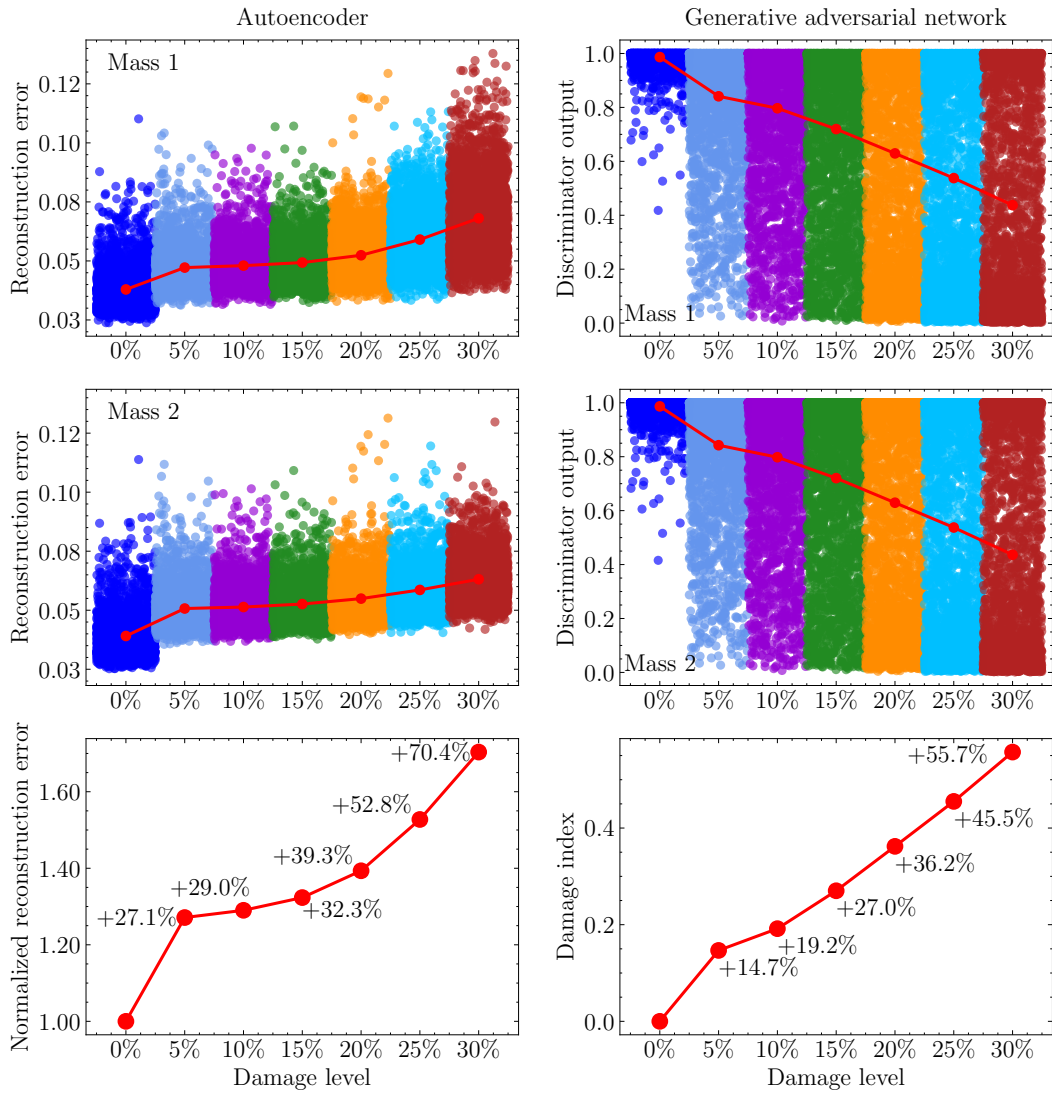


Figure 3.7. Comparative damage sensitivity for the 2-DOF system. While both AE and GAN show a clear correlation with damage, the AE demonstrates higher sensitivity in this more complex case.

Table 3.2. Configuration of the deep learning architectures for the 2-DOF Duffing system.

Type	Component	Convolution layers	Trainable parameters	Kernel size	Number of filters	Strides
AE	Encoder	3	150,175	(50,25,12)	(100,50,25)	(1,1,1)
	Decoder	4	2,142,902	(12,25,25,50)	(50,100,100,2)	(2,2,1,1)
GAN	Generator	3	38,151	(6,6,4)	(64,32,2)	(2,1,2)
	Discriminator	3	4,050,770	(6,6,6)	(32,32,32)	(2,2,2)

system, governed by the intricate dynamics of Eqs. (2.5)-(2.8), combines hysteretic, negative stiffness, and superelastic behaviors, presenting a formidable challenge for any data-driven model. This application serves as a crucial stress test to evaluate

the robustness and performance limitations of the AE and GAN architectures when confronted with highly complex, path-dependent dynamics and a higher level of simulated measurement noise.

3.4.1 Autoencoder Performance

Despite the increased complexity, the Autoencoder continues to demonstrate a satisfactory and predictable performance. As shown in the representative results in Figure 3.8, the AE successfully reconstructs the response of the undamaged isolator. As with the previous case studies, its reconstruction performance degrades as expected when the damage level increases, with the mismatch between the input and output signals serving as a clear indicator of the anomalous system behavior.

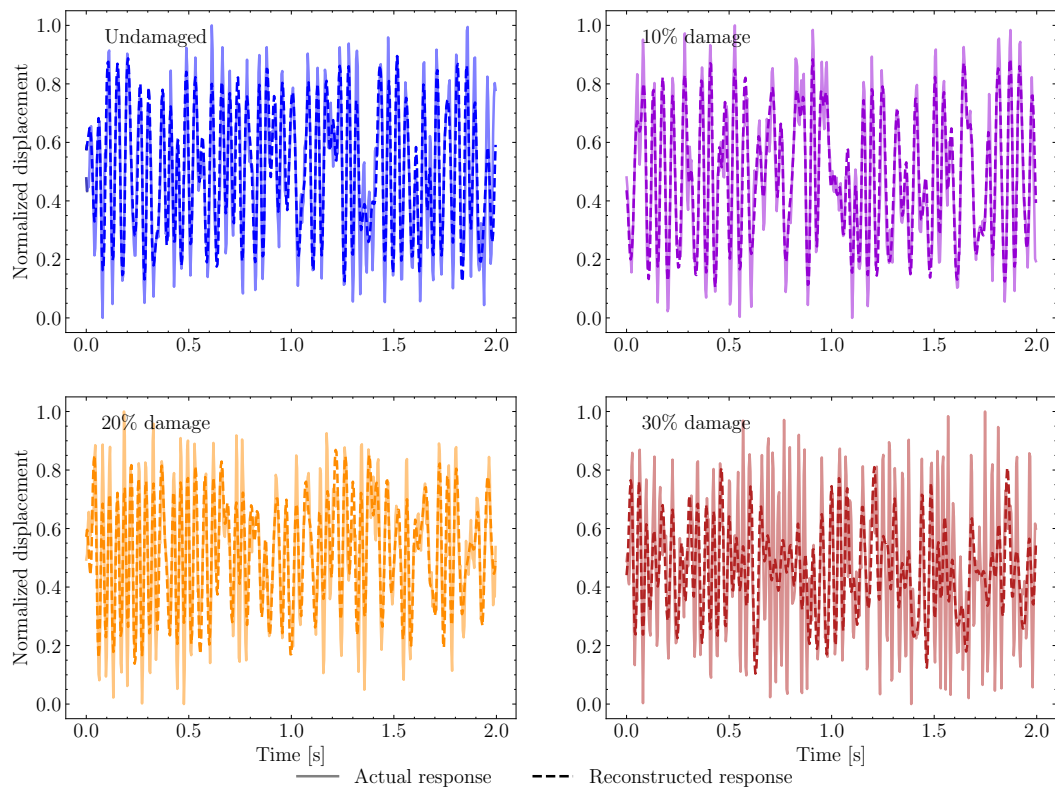


Figure 3.8. Reconstruction of the seismic isolator response via the Autoencoder for different damage levels. The reconstruction error grows as the damage severity increases.

3.4.2 Generative Adversarial Network Performance

In this highly challenging application, the Generative Adversarial Network shows a good capture of essential system dynamics avoiding high frequency noise. As can be inferred from the time-frequency comparison in Figure 3.9, the Generator perfectly replicates the rich dynamics of the healthy isolator; some frequency components present in the actual response are missing from the output produced by the generator.

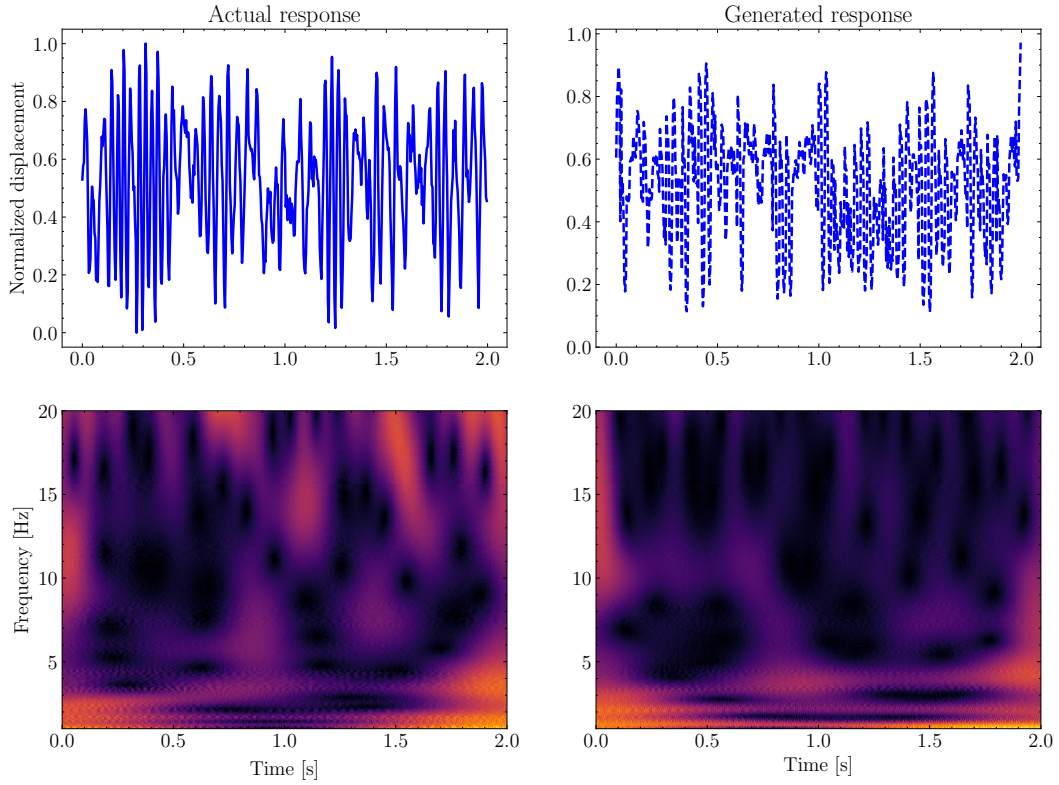


Figure 3.9. Fidelity of the trained GAN Generator for the seismic isolator. The time-frequency comparison reveals that the Generator captures the main dynamics of the real signal.

3.4.3 Comparative Discussion

The quantitative results presented in Figure 3.10 agree with and reinforce the qualitative evidence from the reconstruction and generation plots. The performance of both architectures on this complex benchmark reveals important insights into their respective robustness.

The AE properly detects the presence of damage across all severity levels. Its sensitivity, as indicated by the slope of the normalized reconstruction error, remains satisfactory, although it is generally lower compared to its performance on the simpler Duffing systems (Figures 3.3 and 3.7). This reduced sensitivity is likely attributable to the higher level of measurement noise (10%) considered in this numerical application, which makes it more difficult to distinguish small parametric changes.

Conversely, the GAN's damage detection capability is noticeably compromised. While damage is eventually recognized, it requires a significant level of severity (around 15%) before the damage index shows a meaningful increase (around 10%). At lower damage levels, the response is almost negligible. This is mainly because the discriminator tends to produce more false negatives in this application; that is, it incorrectly classifies some samples from the damaged isolator as true or healthy. This case study effectively highlights the AE's superior robustness in handling

highly complex dynamics and increased noise levels compared to the vanilla GAN architecture employed in this work.

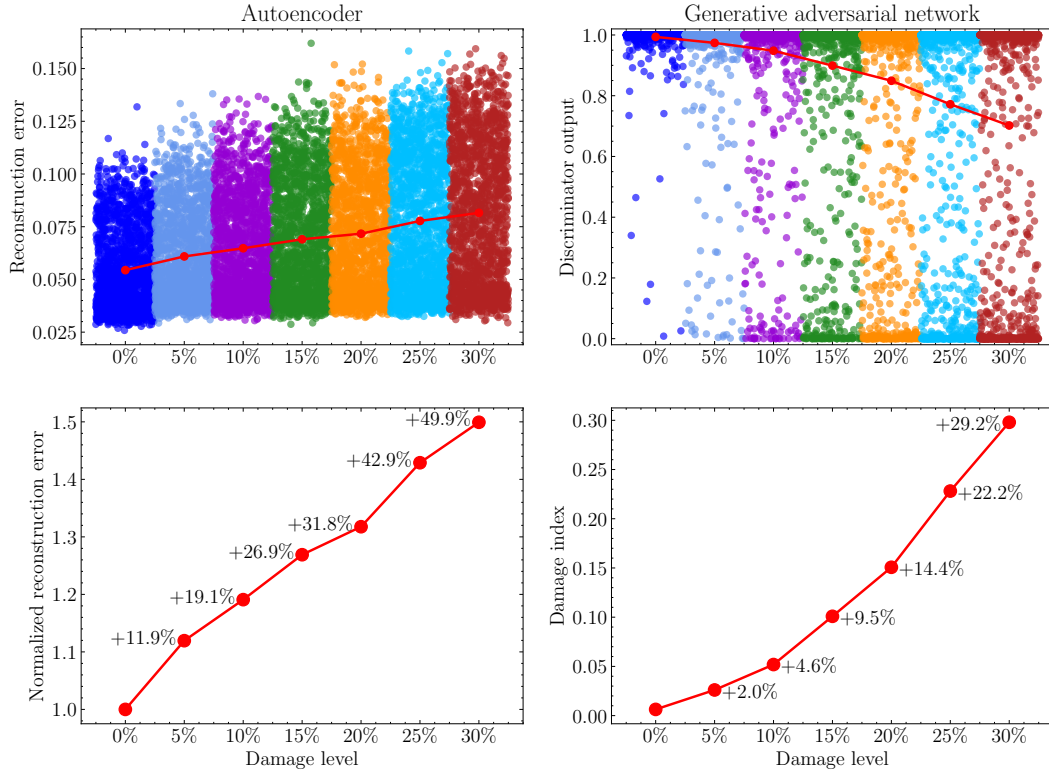


Figure 3.10. Comparative damage sensitivity for the seismic isolator. The AE maintains satisfactory sensitivity across all damage levels, whereas the GAN’s sensitivity is significantly reduced, only detecting damage at higher severity.

3.5 Experimental Validation on a Magneto-Elastic Beam

Having evaluated the performance of the convolutional architectures on a series of increasingly complex numerical models, the final and most critical validation of this chapter involves their application to real experimental data. This step is essential to assess the robustness and practical applicability of the proposed methods when faced with the challenges inherent in physical measurements, including ambient noise and unmodeled dynamics. For this purpose, data from the magneto-elastic beam experiment are utilized, a system whose physical setup and nonlinear characteristics were detailed in Section 2.2.1.

This experimental application deals with an output-only damage detection scenario under random vibrations. The measured system response under band-pass filtered white noise (10-420 Hz) is considered, with input excitation levels varying across a range of values (0.01 to 0.15 VRMS). This protocol ensures that the models are tested on their ability to detect damage irrespective of the specific degree of nonlinearity being experienced, without any prior information about the system or the applied excitation.

3.5.1 Autoencoder Performance

The performance of the trained Autoencoder on the experimental data is shown in Figure 3.11. The results confirm the effectiveness of the reconstruction-based anomaly detection strategy. The AE accurately reconstructs the response signal from the healthy experimental setup. As soon as a small, physical damage of 2% is introduced, the reconstruction error displays an abrupt and noticeable increment, signaling a clear deviation from the learned healthy baseline.

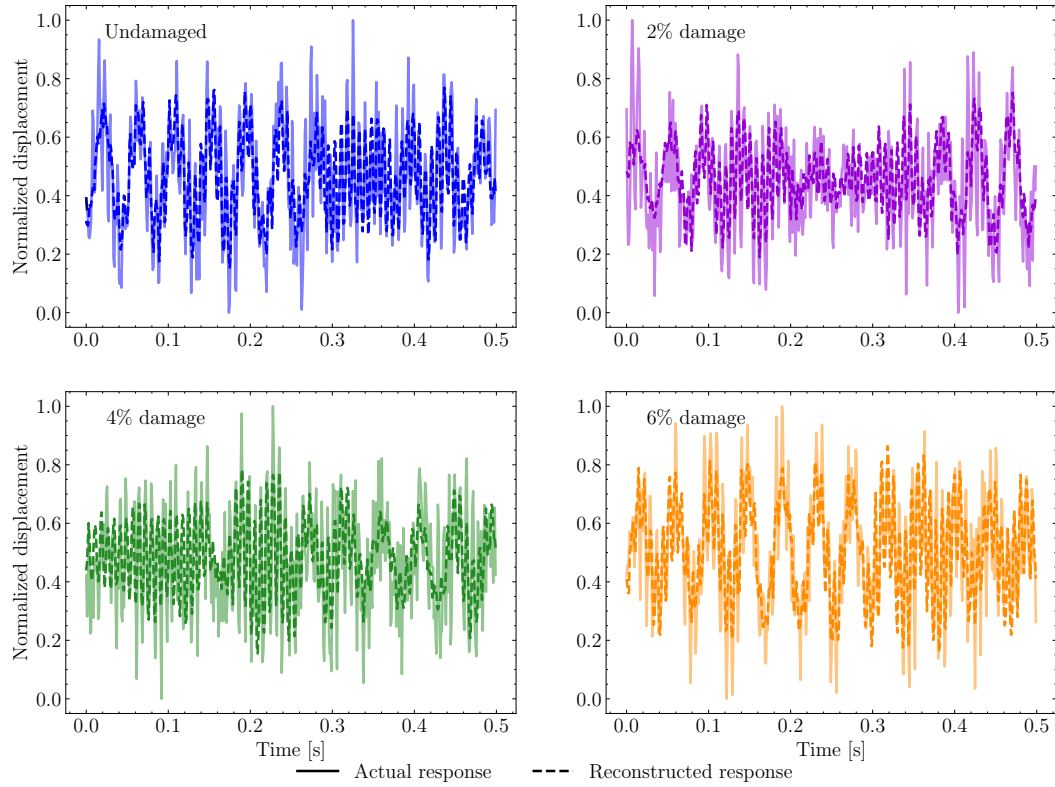


Figure 3.11. Reconstruction of the experimental magneto-elastic system response via the Autoencoder for different physical damage levels.

3.5.2 Generative Adversarial Network Performance

The GAN's Generator also demonstrates its ability to capture the dynamics of the physical system, as shown in Figure 3.12. The strong similarity between the actual and generated time-frequency representations of the undamaged system's response indicates that the Generator has learned a faithful model of the healthy experimental dynamics. Consequently, when the system is damaged, the resulting response no longer represents the dynamics of the healthy system, leading to a growing number of time-series samples being classified as fake by the trained Discriminator.

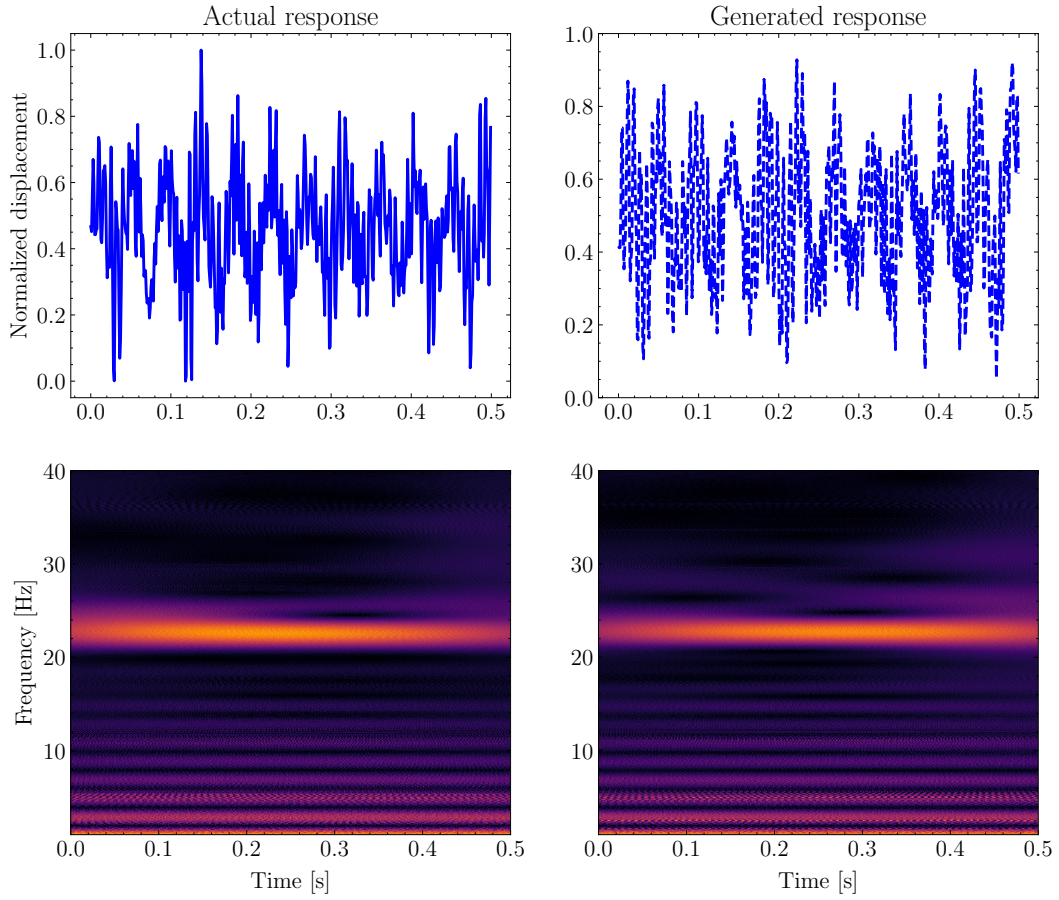


Figure 3.12. Fidelity of the trained GAN Generator on experimental data. The time-frequency comparison shows a high similarity between the actual undamaged response and the synthetic data generated by the GAN.

3.5.3 Comparative Discussion

The final comparative assessment in Figure 3.13 demonstrates the effectiveness of both architectures in this challenging experimental application. The results confirm the very good sensitivity of both models to the lowest damage level of 2%, validating their capability for early-stage damage detection in a physical system.

In particular, the damage index derived from the discriminator loss highlights both the sensitivity and the limitations of the GAN. While it successfully flags the onset of damage, the metric exhibits saturation at higher severity levels. As seen in Figure 3.13, the anomaly scores for the 4% and 6% damage levels are statistically indistinguishable. This plateauing effect suggests that while the GAN is an exceptional binary detector (healthy vs. damaged), its utility for damage quantification is limited in this experimental context, as the relationship between the physical damage magnitude and the discriminator's probability output becomes highly nonlinear and saturates once the signal deviation exceeds the generator's learned manifold.

Ultimately, this experimental validation provides the strongest confirmation of the

viability of the unsupervised convolutional framework. It proves that both AE and GAN architectures are able to detect physical damage without any prior information about the system's parameters or the applied dynamic loads, and irrespective of the degree of nonlinearity in the response. The specific hyperparameters for the models used in this experimental study are provided in Table 3.3.

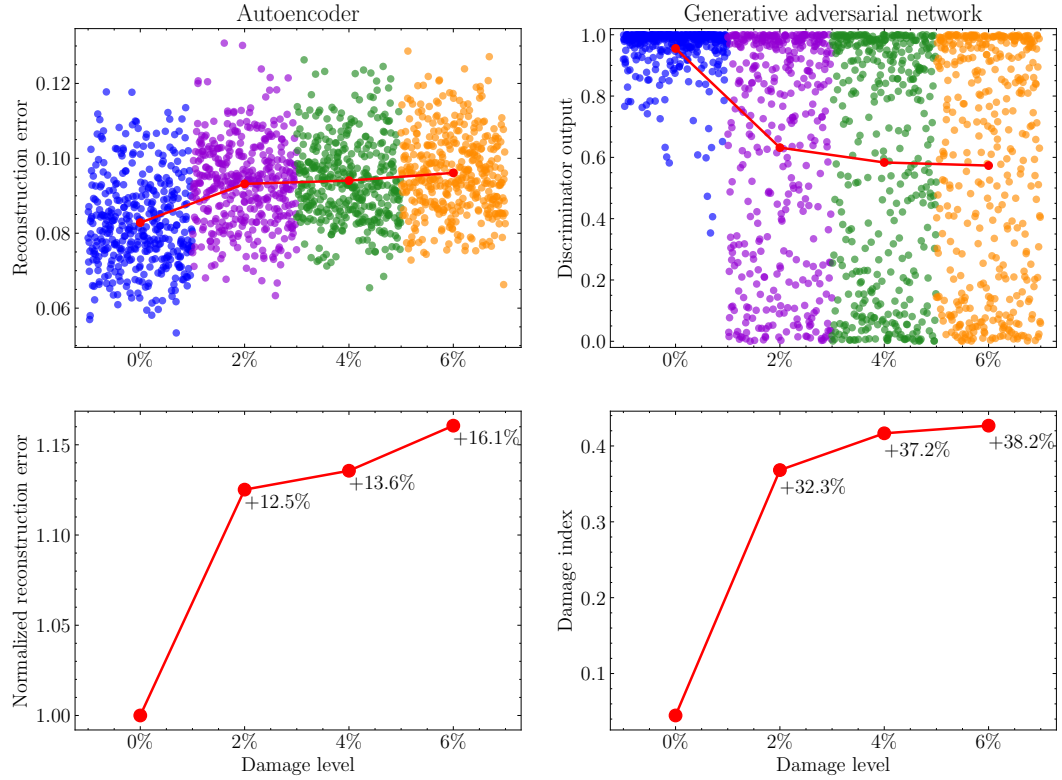


Figure 3.13. Comparative damage sensitivity on the magneto-elastic experimental data. Both architectures demonstrate high sensitivity to the incipient 2% damage level, confirming their utility for early damage detection.

Table 3.3. Configuration of the deep learning architectures for the magneto-elastic system.

Type	Component	Convolution layers	Trainable parameters	Kernel size	Number of filters	Strides
AE	Encoder	4	1,283,525	(20,100,50,25)	(100,100,50,25)	(1,1,1,1)
	Decoder	3	527,679	(50,50,20)	(50,100,1)	(1,2,2)
GAN	Generator	3	414,629	(10,20,2)	(100,200,1)	(1,2,2)
	Discriminator	4	2,357,201	(3,50,15,10)	(200,200,100,50)	(1,1,1,1)

3.6 Operational Limitations and Critical Analysis

The results presented in the preceding sections demonstrate the potential of deep convolutional architectures for nonlinear SHM. However, to provide a balanced

scientific assessment, it is necessary to critically examine the failures, stability issues, and limitations observed during this investigation. As noted by the results on the seismic isolator and the 2-DOF system, successful detection in a controlled study does not immediately translate to operational robustness. This section outlines the specific boundaries of applicability for these methods.

3.6.1 Black-Box Approaches vs. Traditional Statistical Methods

A fundamental question arises: what advantage do these computationally expensive deep learning models offer over traditional statistical baseline methods, such as Principal Component Analysis (PCA) or Mahalanobis Squared Distance (MSD)?

Traditional statistical approaches rely on the assumption that the healthy data follows a multivariate Gaussian distribution and that the system dynamics can be decomposed into linear orthogonal components. For the highly nonlinear systems investigated here (e.g., the hysteretic isolator), these assumptions are fundamentally violated. The data manifold of a hysteretic system is curved and folded; a linear projection like PCA would overestimate the variance, leading to a high floor of reconstruction error even in the healthy state.

However, deep learning models are not without their own disadvantages compared to these traditional baselines.

- **Convexity and Stability.** PCA is a convex optimization problem with a unique global solution. In contrast, training the GAN proved to be highly unstable, often requiring multiple restarts to avoid mode collapse, where the generator produces a limited variety of samples. This stochasticity in training is a significant barrier to automated deployment compared to deterministic statistical methods.
- **Interpretability.** when an MSD metric flags an anomaly, it is often possible to trace the contribution of specific variables. The Convolutional AE and GAN, however, act as black boxes. A high reconstruction error signals a global mismatch but offers no immediate insight into whether the error stems from a localized stiffness change or a benign environmental shift.

3.6.2 The Instability of Adversarial Detection

While the GAN demonstrated superior sensitivity in the experimental beam case study, this sensitivity comes at the cost of robustness. In the 2-DOF Duffing benchmark (Section 3.3), the Discriminator exhibited a hair-trigger response, producing a wide variance of outputs for the healthy state. This suggests that the decision boundary learned by the Discriminator is not smooth. While this study utilizes the standard minimax loss formulation based on the Jensen-Shannon divergence (Eq. 2.14), it is acknowledged that this objective can sometimes lead to vanishing gradients, particularly when the generated and real distributions have negligible overlap. In the literature, the Wasserstein GAN (WGAN) is often cited as a robust alternative. By leveraging the Earth Mover's Distance rather than the JS divergence, WGANs provide smoother gradients and a loss metric that correlates better with sample quality, thereby improving training stability. However, for the scope of this

work, the standard formulation proved sufficient. In a real-world scenario with fluctuating environmental conditions, this lack of smoothness would likely result in an unacceptable rate of false positives. The Autoencoder, by minimizing a continuous error metric (MAE), demonstrated a smoother manifold representation, making it a more conservative but reliable choice for safety-critical infrastructure.

3.6.3 Noise Sensitivity and the Limit of Detection

The performance degradation observed for the seismic isolator case study (Section 3.4) under 10% noise provides a practical bound on the Limit of Detection (LOD).

- **LOD for Stiffness Reduction.** across the numerical benchmarks, the effective LOD for the Autoencoder was approximately a 5% stiffness reduction. Below this threshold, the reconstruction error distribution of the damaged state overlapped significantly with the noise floor of the healthy state.
- **Impact of Signal-to-Noise Ratio (SNR).** for the GAN, high noise levels prevented the Discriminator from learning the fine-grained features of the hysteresis loops, effectively blinding it to damage levels below 15% in the Isolator case. This indicates that while Generative models excel at capturing complex distributions, they are less noise-robust than reconstructive models (AE) which effectively act as de-noising filters.

3.6.4 Saturation and Non-Monotonicity

Finally, the experimental results on the magneto-elastic beam revealed a saturation of the damage metrics. The distinction between moderate (4%) and severe (6%) damage was minimal. This non-monotonic behavior implies that while these 1D convolutional architectures are powerful *detectors* (Level I), their output cannot be linearly mapped to damage *severity* (Level III) without extensive calibration. This limitation stems from the use of the Tanh or Sigmoid activation functions in the final layers, which squash unbounded damage features into a bounded range, losing resolution at the extremes.

3.7 Summary and Findings

This chapter has presented a comprehensive and systematic investigation into the use of 1D convolutional deep learning architectures for unsupervised, data-driven damage detection in nonlinear dynamic systems. The performance of two distinct models, the Autoencoder (AE) and the Generative Adversarial Network (GAN), was thoroughly examined across a suite of three numerical benchmarks and one experimental application, each exhibiting different and increasingly complex nonlinear behaviors. The investigations were conducted under the realistic assumption that neither the specific system parameters nor the excitation characteristics were known, forcing the models to perform output-only damage detection based on learning the dynamics of the healthy state. From this extensive study, several key conclusions can be drawn.

- **Validation of the Unsupervised Paradigm.** The primary finding is that both the AE and GAN architectures are fundamentally capable of detecting the onset of damage in a variety of nonlinear dynamic systems. Across all case studies, the models demonstrated the ability to distinguish between healthy and damaged states, and the calculated anomaly scores showed a consistent and monotonic correlation with increasing damage levels, confirming the viability of the overall methodological framework.
- **Comparative Performance and Robustness.** A clear distinction in performance characteristics was observed. The Autoencoder consistently exhibited a more stable and robust performance, providing reliable damage detection across all benchmarks, including the highly complex seismic isolator and the noisy experimental data. The Generative Adversarial Network, on the other hand, exhibited significant operational instability. While capable of high sensitivity, it suffered from metric saturation at high damage levels and a reduced noise tolerance in complex hysteretic systems, suggesting it is best suited for binary detection tasks in low-noise environments rather than robust quantification.
- **Practical Implementation and Interpretability Trade-offs.** From a practical standpoint, establishing a well-performing AE architecture proved to be a more straightforward task than configuring and training the GAN, which required significantly more time and careful hyperparameter tuning. Conversely, a principal advantage of the GAN lies in its output. The Discriminator's raw output, bounded between 0 and 1, provides a naturally interpretable health index, which can be easily transformed into an intuitive damage index, a feature not inherently present in the AE's unbounded reconstruction error.

Contribution to the SHM Hierarchy. In the context of the damage assessment hierarchy established in Chapter 1, the work presented in this chapter makes clear and significant contributions. The demonstrated ability of both architectures to reliably detect the onset of incipient damage (as low as 2%-5%) in all case studies firmly satisfies the requirements of level I for damage detection. Furthermore, the consistent, monotonic relationship observed between the magnitude of the anomaly scores and the severity of the damage provides a strong foundation and a significant step towards achieving level III for damage quantification. While this work does not propose a full quantification model, it proves that the output of these unsupervised models carries quantitative information about the extent of damage, not just its presence.

Limitations and Transition to the Next Chapter. While the convolutional architectures have proven effective, this investigation has also highlighted their inherent limitations and opened avenues for further research. The methods presented here focused solely on damage detection and preliminary quantification; the challenge of level III related to damage localization was not addressed. Additionally, the potential influence of changing environmental and operational conditions was not considered.

More fundamentally, the 1D convolutional architectures process each sensor channel independently and are primarily designed to capture local features in time. This limits their ability to model long-range temporal dependencies or to dynamically fuse information from multiple sensors to understand the global state of a complex structure. To address these architectural limitations and to explore a more powerful and interpretable approach to unsupervised diagnostics, the subsequent chapter introduces an advanced methodology based on the Transformer framework.

Chapter 4

Damage Detection using an Interpretable Sensor Fusion Transformer

The previous chapter established the viability of unsupervised convolutional architectures for damage detection, while also revealing their limitations in interpretability and in modeling complex temporal dependencies. This chapter addresses these limitations by introducing and evaluating the novel SensorFusion Temporal Fusion Transformer (SensorFusionTFT). The central challenge motivating this shift is the nature of nonlinear structural dynamics itself: phenomena such as hysteresis introduce path-dependent memory effects where the current state depends on the entire history of motion. Capturing these non-local, history-dependent relationships is a fundamental problem that local feature extractors like CNNs struggle with. This necessitates an architecture with a mechanism, such as self-attention, that can learn global dependencies across time. The novel architecture and findings detailed in this chapter have been compiled into a manuscript for submission to a peer-reviewed journal.

This chapter will demonstrate how the SensorFusionTFT, adapted from the state-of-the-art forecasting model, provides a superior framework for this task. The convolutional autoencoder from chapter 3 will serve as a conceptual performance baseline, allowing for a clear assessment of the advancements offered by this more sophisticated architecture.

4.1 Case Study Protocol for Transformer-Based Analysis

To assess the capabilities of the SensorFusionTFT, its performance is evaluated on challenging numerical benchmarks chosen to represent distinct and complex forms of nonlinear behavior discussed in Chapter 2. To emulate realistic operational conditions, the numerical models were subjected to band-limited random noise excitation. This process yielded a comprehensive dataset for each system, comprising 340 distinct time-history simulations, which were partitioned into training (200), validation (40), and test (100) sets. Following the unsupervised training phase on healthy data, the model's diagnostic capabilities were tested against simulated damage scenarios with stiffness reductions ranging from 5% to 30%.

4.2 Application to a Hysteretic Tuned Mass Damper

The first numerical benchmark for the SensorFusionTFT is an hysteretic Tuned Mass Damper (TMD), whose governing equations were defined in Section 2.1.2. This system was specifically chosen to test the model's ability to learn path-dependent, hysteretic memory effects—a key challenge for many data-driven methods.

For this application, we employ the physics-informed surrogate modeling framework, a specialized use-case of the broader auto-encoding paradigm where the model learns to map observable kinematics to a non-observable physical quantity. The objective is to predict the *Total Restoring Force*, $\tilde{F}_{res}$, a quantity that holistically represents the system's internal constitutive law:

$$\tilde{F}_{res}(\tilde{t}) = \delta \dot{\tilde{x}}(\tilde{t}) + (1 - \delta) \tilde{z}(\tilde{t}) + \tilde{k}_3 \tilde{x}(\tilde{t})^3. \quad (4.1)$$

By training the model to predict $\tilde{F}_{res}$ using only observable inputs (displacement, acceleration, and external excitation), it is tested its ability to infer the underlying physics of the system.

4.2.1 Performance as a Dynamic System Surrogate

Figure 4.1 demonstrates the core capability of the trained SensorFusionTFT: its ability to act as a high-fidelity dynamic surrogate. For the healthy test case, the model exhibits exceptional fidelity. The median prediction precisely tracks the ground truth restoring force, indicating high point-forecast accuracy. Furthermore, the 80% confidence interval (CI) is shown to be both narrow and well-calibrated; its tightness signifies that the model is highly confident in its predictions, and the fact that the true signal is consistently enveloped by these tight bounds confirms that this confidence is justified. This dual success, an accurate median prediction coupled with a narrow, and calibrated confidence interval, validates that the model has successfully learned the complex, path-dependent nature of the healthy hysteretic behavior, establishing its credibility. When presented with data from damaged states, the model's predictions, which are based on its learned understanding of the healthy state, begin to diverge from the actual response. This divergence is the primary diagnostic signal, and it becomes visually apparent as damage increases.

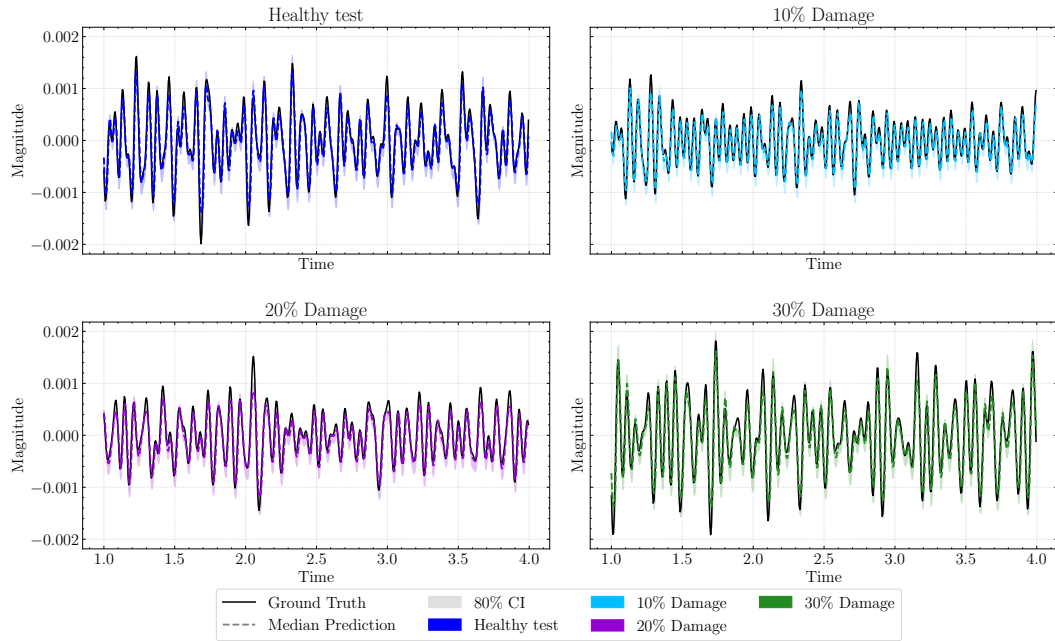


Figure 4.1. Comparative predictions of the Total Restoring Force for the TMD. The divergence between the ground truth (solid) and the model's prediction (dashed) grows with damage severity, signaling an anomaly.

4.2.2 Quantitative Damage Detection via Anomaly Metrics

To systematically quantify this divergence, I employ the multi-metric framework defined in Section 2.4.4. Figure 4.2 illustrates the normalized distribution of the

Root Mean Square Error (RMSE). A clear and monotonic trend is observable: as the severity of the damage increases, the median reconstruction error also increases significantly, with the 30% damage case exhibiting a mean RMSE over 118% higher than the healthy baseline.

Figure 4.3 displays the normalized distribution of the Interval Score (IS), which penalizes breaches of the model’s learned uncertainty bounds. The IS proves to be an similar sensitive metric, showing a dramatic trend where the mean score for the 30% damage case increases by over 109% very close to RMSE scores. This demonstrates that damage not only causes the median prediction to diverge, but also decisively breaks the model’s learned uncertainty model of the healthy state, making the IS a highly potent indicator of anomalous behavior.

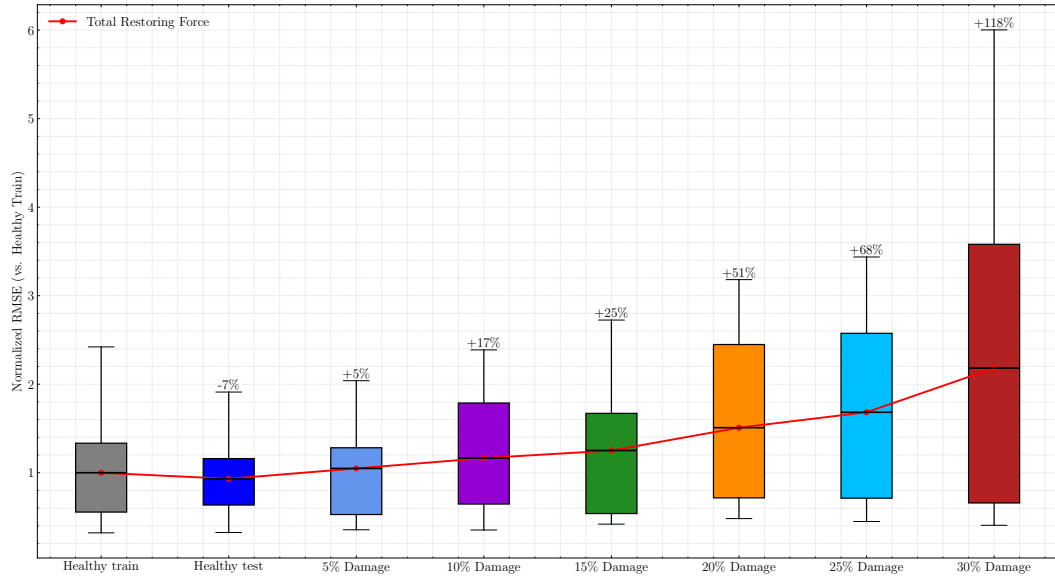


Figure 4.2. Normalized distribution of RMSE for the TMD’s Total Restoring Force. Each box plot shows the distribution of scores across all prediction windows for a given damage case. The red line connects the mean RMSE value of each distribution, illustrating the clear and monotonic trend of increasing average reconstruction error with damage severity.

4.2.3 Diagnostic Insights from Interpretability Mechanisms

Beyond quantitative detection, the interpretable design of the SensorFusionTFT provides deep diagnostic insights into its reasoning, as shown in the comprehensive dashboard in Figure 4.4.

The top row displays the mean sensor importance from the VSN, which assigns highest mean weights for the displacement and excitation signals as the most important features for predicting the restoring force showing a robust physics-induced surrogate learning. The middle row visualizes the deviation in the model’s temporal self-attention pattern calculated with the self-attention weights of the train dataset as baseline, showing an increasingly intense and structured change as damage progresses from 5% to 30% damage scenarios. Finally, the bottom row

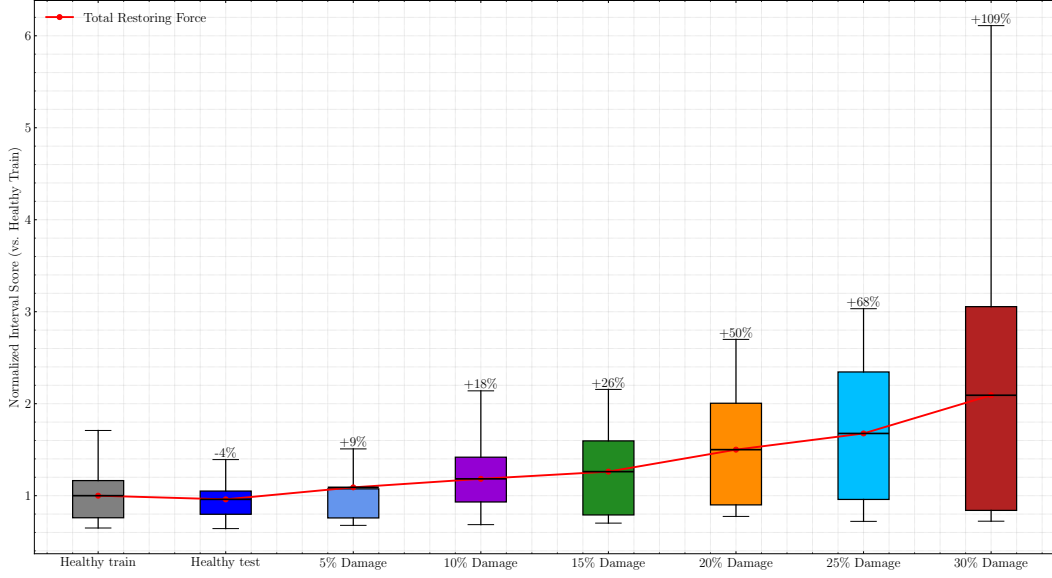


Figure 4.3. Normalized distribution of the Interval Score (IS) for the TMD’s Total Restoring Force. Each box plot represents the distribution of scores for a given damage case. The red line, which connects the mean IS value of each distribution, illustrates the dramatic and monotonic increase in the mean score as damage severity increases.

presents the Attention Anomaly Score, which quantifies this change in temporal reasoning with respect to the baseline healthy train. The clear, monotonic increase in this score provides a third, independent confirmation of the model’s ability to detect and quantify damage, grounding the detection results not just in prediction error, but in the internal mechanisms of the model itself.

4.3 Revisiting the Hysteretic Seismic Isolator

Having demonstrated the SensorFusionTFT’s ability to learn a system with pure hysteretic nonlinearity, the hysteretic seismic isolator is revisited, a benchmark whose complex governing equations were detailed in Section 2.1.3. This system was specifically chosen to evaluate the architecture’s performance on a system with highly complex, *hybrid* nonlinearities, including negative stiffness, superelasticity, and path-dependent hysteresis all acting in concert [69, 10, 89, 16]. This serves as a stringent test of the model’s capacity for abstraction and its ability to learn a generalizable approximation of a sophisticated underlying constitutive relationship.

For this validation study, I employ a dual-target is employed, physics-informed surrogate modeling approach. The model is trained exclusively on healthy data to map observable inputs (displacement, acceleration, external excitation) to two critical outputs: the future displacement (a kinematic variable) and the non-observable, internal Total Restoring Force (a dynamic variable). Success in this dual-target task provides strong evidence that the model’s core fusion mechanism can distill input sensor data into a highly abstract latent representation rich enough to be deconstructed into two distinct physical quantities.

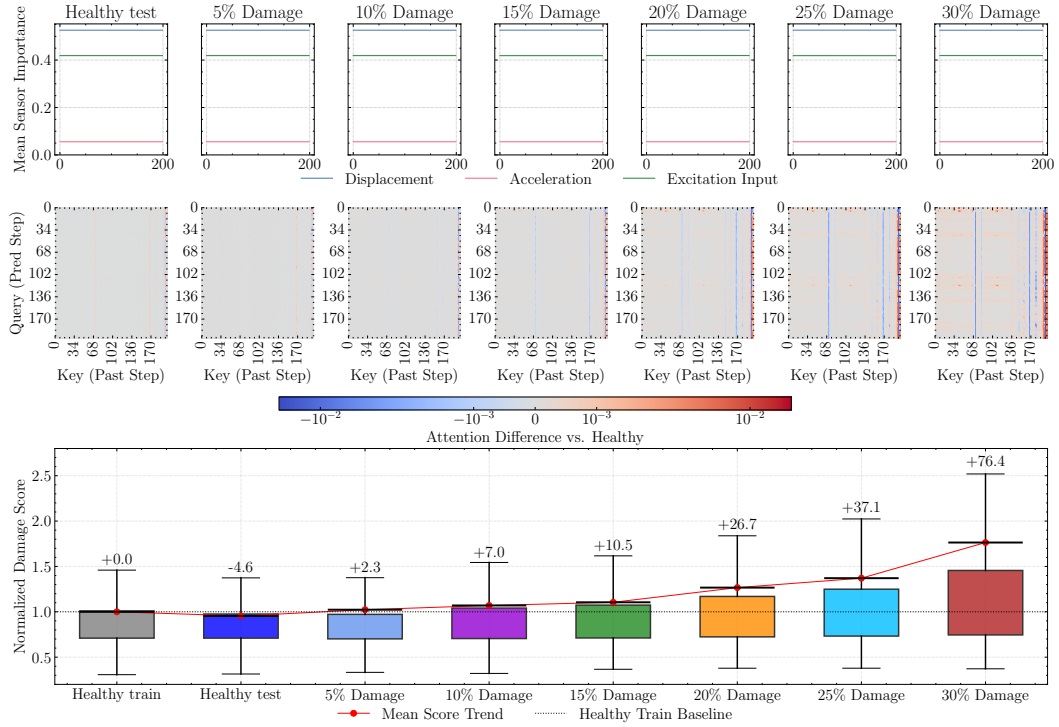


Figure 4.4. Interpretability dashboard for the TMD application. This three-panel output shows (Top) which sensors the model focuses on, (Middle) how its temporal reasoning changes with damage, and (Bottom) a quantitative score of that change, all showing a clear correlation with damage.

4.3.1 Performance as a Dynamic System Surrogate

The efficacy of this dual-target surrogate modeling approach is illustrated in Figure 4.5. Here we can see the model is tasked with simultaneously learning to predict a probabilistic distribution for two distinct physical quantities: the kinematic Displacement and the non-observable, internal Total Restoring Force. This is achieved by training the model to output three quantiles (e.g., 10th, 50th, 90th) for each of the two target variables.

In the healthy test state (top row), the model demonstrates exceptional fidelity in this complex task. First, the median quantile predictions (dashed lines) for both Displacement and Restoring Force precisely track their respective ground truth signals, indicating high point-forecast accuracy. Second, the 80% confidence interval, which is itself a model prediction formed by the upper and lower quantiles, is shown to be both narrow and well-calibrated. It closely envelops the ground truth, signifying that the model has learned not only the system's central dynamic tendency but also a tight and accurate representation of its operational uncertainty. This success in concurrently predicting the distributions for two different physical outputs confirms that the SensorFusionTFT has successfully internalized the complex, path-dependent dynamics of the healthy system.

In contrast, for the damage scenarios, a significant divergence materializes between the model's forecast and the system's actual behavior. It is crucial to

interpret this discrepancy not as a prediction error, but as the core diagnostic signal. Notably, this effect is far more pronounced for the Total Restoring Force than for the kinematic displacement. Even at high damage levels, the displacement plots show only subtle deviations, whereas the restoring force predictions exhibit large, unmistakable errors. This highlights the model's acute sensitivity to changes in the underlying constitutive physics, which are more immediately apparent in the internal forces than in the overall motion.

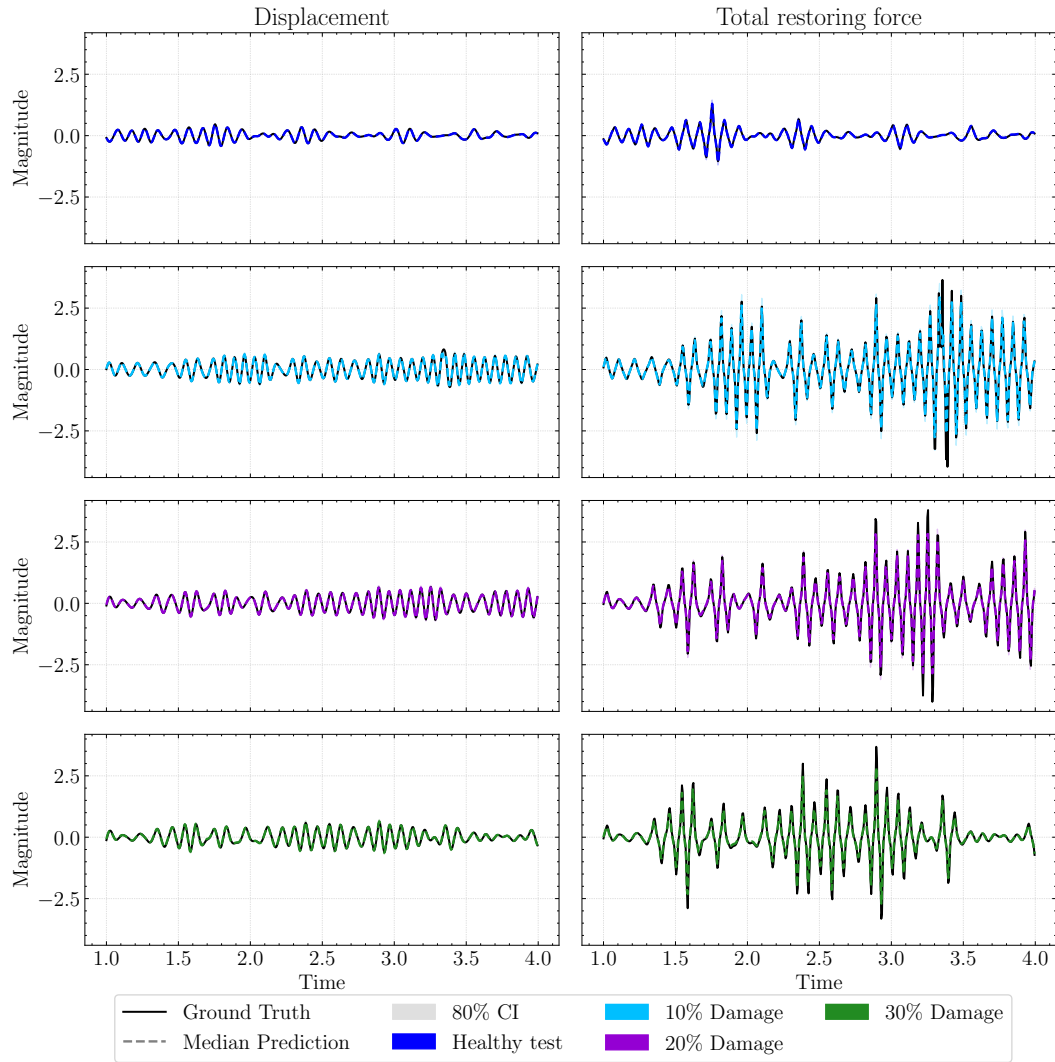


Figure 4.5. Comparative prediction performance for the nonlinear isolator. The model's predictions for Displacement (left) and Total Restoring Force (right) are compared against the ground truth (solid black) across four damage states.

4.3.2 Quantitative Damage Detection via Anomaly Metrics

To move beyond a qualitative assessment, the prediction error using the RMSE and IS metrics is systematically quantified, with results presented in Figures 4.6 and 4.7. The results reveal a clear, monotonic relationship between error and damage severity

for both target variables. At the highest damage level of 30%, the mean RMSE for both displacement and restoring force surges to over +800% above the healthy baseline, confirming that prediction error functions as a reliable damage indicator.

The Interval Score proves to be an even more sensitive diagnostic tool. As shown in Figure 4.7, this metric reveals a dramatic, exponential growth in error as damage progresses. The sensitivity of the IS is remarkable for the 30% damage case, the mean score escalates by over 7900% for the restoring force and an extraordinary 12,000% for displacement. This explosive increase signifies that the damaged system's behavior is not just different, but wildly unpredictable compared to the narrow operational bounds learned during healthy training, making the IS an exceptionally potent indicator for detecting both subtle and severe anomalies.

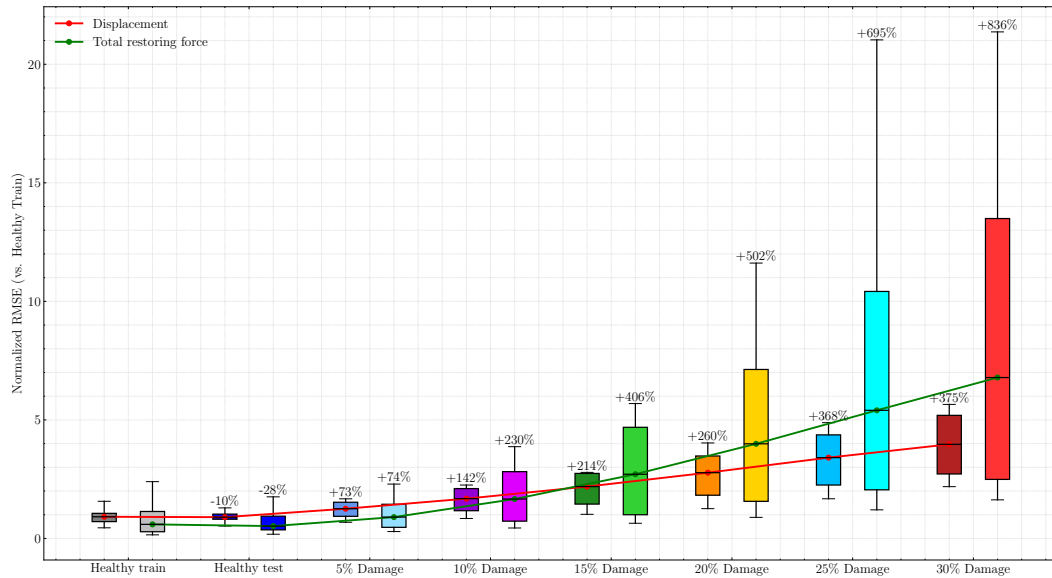


Figure 4.6. Normalized distribution of RMSE by damage case for the isolator. The trend lines show a clear, monotonic increase in mean reconstruction error distribution with damage severity for both target variables.

It is important to address the operational implementation of these metrics. As observed in Figure 4.2, the anomaly scores exhibit stochastic fluctuations due to the random nature of the excitation. In a real-world scenario where multiple passes or repeated tests are not feasible, a single-shot reading might fall within the noise floor of the healthy distribution. Therefore, robust practical deployment requires a **temporal smoothing strategy**, such as a Moving Average or Exponential Weighted Moving Average (EWMA) filter applied to the anomaly score stream. This would mitigate the instance-wise variance and reveal the underlying trend of structural degradation.

4.3.3 Diagnostic Insights from Interpretability Mechanisms

A key advantage of the SensorFusionTFT is its inherent interpretability, which allows for a deeper understanding of the model's learned strategy. The diagnostic dashboard in Figure 4.8 summarizes the model's internal mechanisms, and the

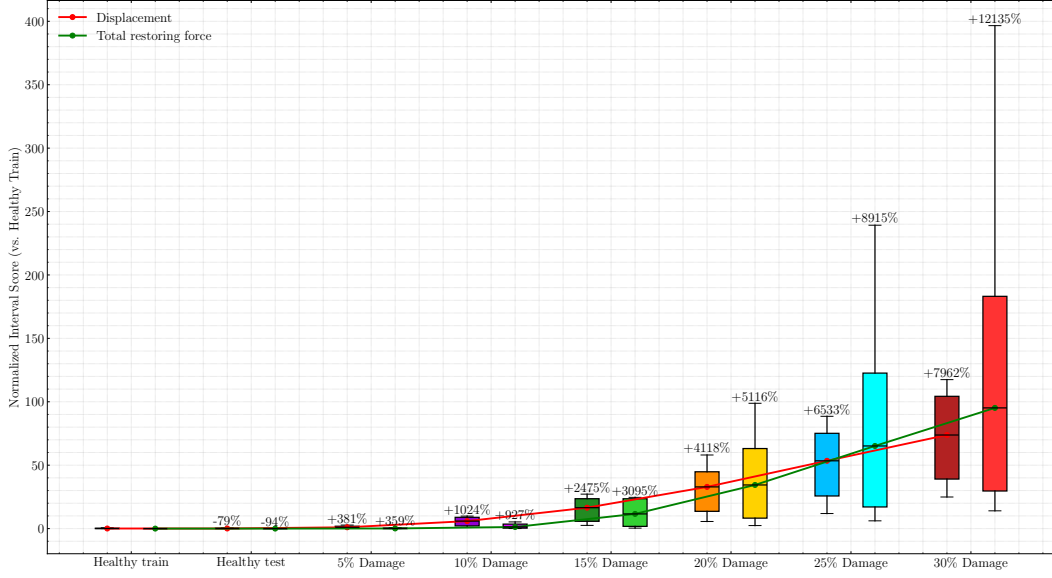


Figure 4.7. Normalized distribution of Interval Score (IS) by damage case for the isolator. The sharp, exponential growth of the mean IS with increasing damage highlights its high sensitivity as a damage indicator.

top row reveals a significant and physically insightful shift in the VSN’s focus compared to the TMD case, with the network consistently assigning the highest importance to the Acceleration signal. This learned prioritization is not an arbitrary choice but a direct and logical consequence of the compound prediction objective assigned in this benchmark. For predicting the Total Restoring Force, the model can exploit the physical identity of the equation of motion ($F_{res} = F_{ext} - M\ddot{x}$), where acceleration is not merely a correlated feature but a direct, linearly proportional component of the target variable. Simultaneously, for predicting the future kinematic Displacement, the acceleration signal provides the essential second-order temporal information that governs the evolution of the system’s trajectory. The VSN, in its optimization process, has therefore identified the acceleration signal as the most salient feature for concurrently minimizing the loss across both prediction tasks. This demonstrates a sophisticated capability that VSN is not simply learning static correlations but is performing an objective-driven allocation of its focus, identifying the most parsimonious pathway to solve the given problem.

The middle row visualizes the Attention Difference, which is the change in the model’s temporal self-attention pattern relative to the healthy baseline. A clear trend can be observed as the damage level increases, the heat maps show increasingly intense and structured patterns of deviation, particularly at the end of the look-back window. This indicates that the model is altering its temporal reasoning process in response to the anomalous dynamics.

The bottom row quantifies this change via the Normalized Damage Score, a metric derived from the attention difference patterns. The monotonic increase in this score provides a third, independent validation of the model’s detection capability. This demonstrates that damage can be identified not only through prediction error but also by directly monitoring the internal state of the neural network, confirming

that the model has learned damage-sensitive features within its attention mechanism.

Operational Detection Thresholds. While the numerical results demonstrate statistical separability at damage levels as low as 5%, it is crucial to interpret these findings with a degree of operational realism. In field conditions, factors such as changing environmental conditions, sensor calibration drift, and varying noise floors often mask incipient changes. Consequently, while the deep learning model shows theoretical sensitivity to 5%-15% stiffness reductions, a more robust and conservative target for reliable detection in a deployed system would likely be in the range of >15%. The subtle deviations observed at lower levels, while mathematically significant in this simulation, might effectively constitute a blind zone in a noisy, real-world monitoring campaign.

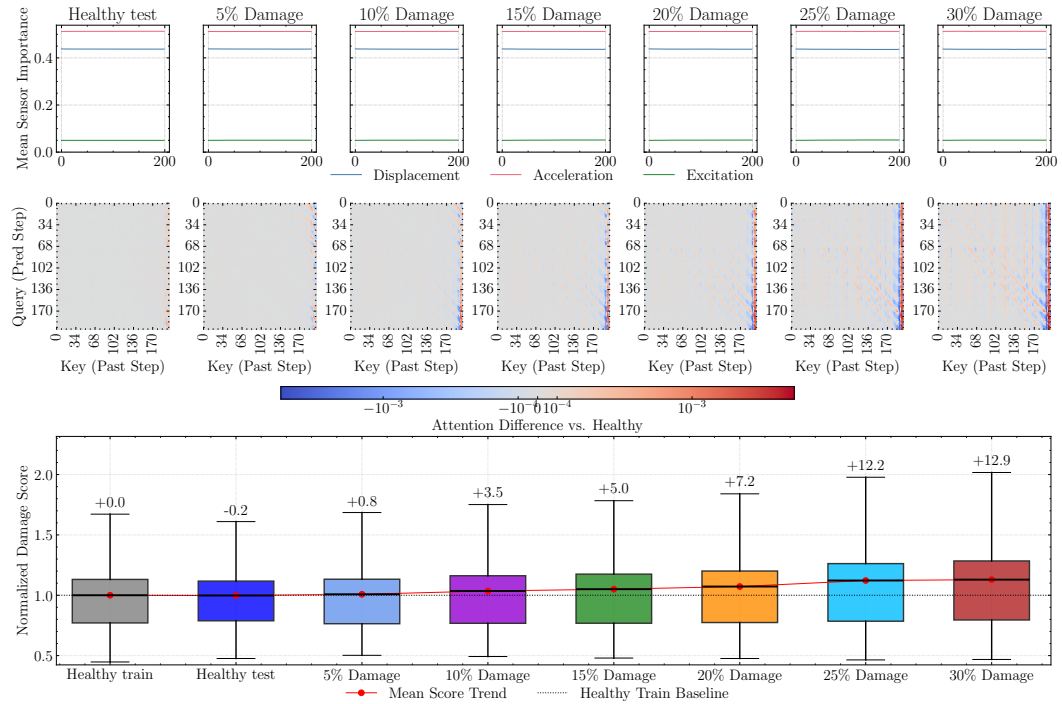


Figure 4.8. Interpretability dashboard for the seismic isolator. The VSN’s focus on acceleration (top), the deviation in temporal attention (middle), and the rising Attention Anomaly Score (bottom) provide a multi-faceted and interpretable view of the damage detection process.

4.4 Experimental Validation on the Magneto-Elastic Beam

The ultimate objective of any SHM architecture is robust performance on real-world structures. To validate this capability, we now transition from numerical simulations to evaluate the SensorFusionTFT on the magneto-elastic beam, an experimental benchmark whose physical setup and nonlinear characteristics were detailed in

Section 2.2.1. This case study serves two critical purposes: first, to confirm the model’s ability to learn complex physical laws from real, noisy measurements, and second, to test its core sensor fusion capability on a multi-sensor problem.

For this experimental validation, I again employ the physics-informed surrogate modeling approach. The SensorFusionTFT is trained exclusively on data from the three accelerometers on the healthy beam to learn the mapping from the measured kinematics to the system’s Total Restoring Force, which is calculated from known physical parameters. The trained model thus acts as a data-driven representation of the healthy nonlinear constitutive law.

4.4.1 Performance on Experimental Data

Figure 4.9 provides a qualitative assessment of the model’s prediction performance on unseen experimental data. For the healthy test case, the model’s median prediction tracks the ground truth with remarkable accuracy. As physical damage is introduced by removing nuts from the bolted connection, the model’s predictions, which are based on its learned understanding of the healthy state, begin to diverge from the actual measured response. This divergence becomes more pronounced as the damage level increases from 2% to 6%, visually confirming the model’s sensitivity to subtle physical changes in the system.

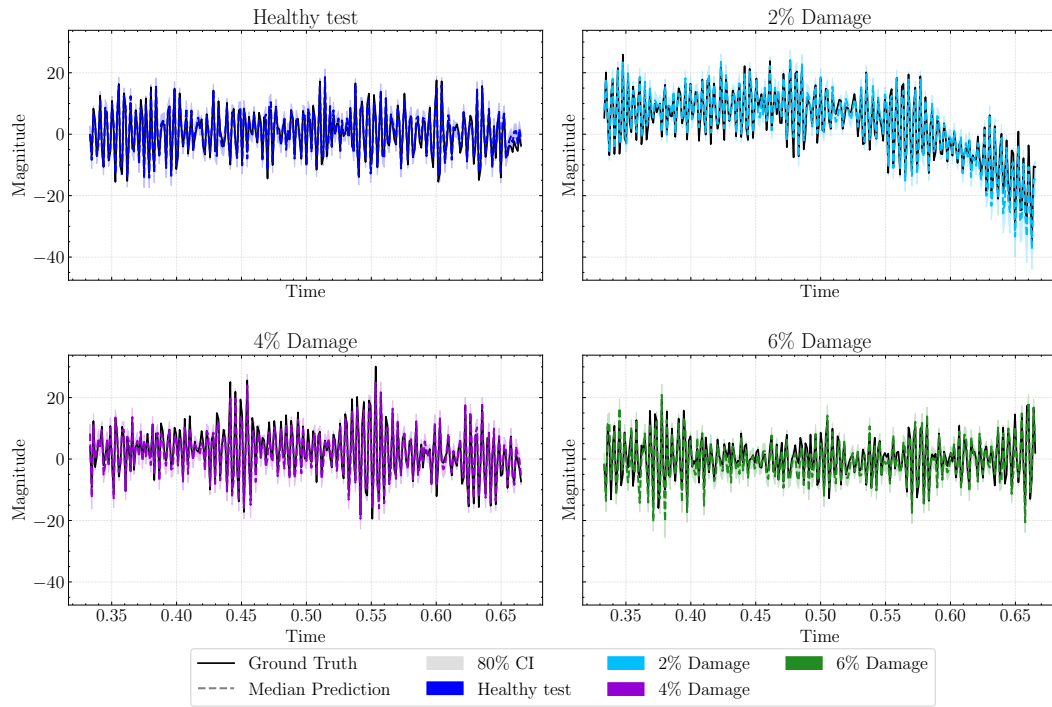


Figure 4.9. Comparative prediction performance for the magneto-elastic beam. The divergence between the ground truth (solid) and the model’s prediction of the healthy response (dashed) serves as the primary damage indicator.

4.4.2 Quantitative Damage Detection via Anomaly Metrics

The model’s detection capability is quantified using the RMSE and IS metrics, with results presented in Figures 4.10 and 4.11. A clear and monotonic increase in reconstruction error is evident as damage progresses from 2% to 6%. Crucially, even the minimal 2% damage scenario produces a distribution of errors that is statistically distinguishable from the healthy test case, highlighting the model’s high sensitivity to incipient damage. The mean RMSE for the 6% damage case is over 63% higher than the healthy baseline.

The Interval Score (Figure 4.11) provides an even more striking demonstration of sensitivity. The mean IS for the 6% damage case is over 483% higher than the baseline, indicating that the damage introduces dynamic behaviors that are fundamentally inconsistent with the model’s learned representation of healthy operational uncertainty.

Interpretation of Experimental Damage Classes. It is imperative to clarify the physical nature of the damage in this experiment compared to the numerical simulations. In the numerical models, damage was defined as a global parametric reduction (e.g., reducing K by 5%). In this experimental setup, ‘2% damage’ refers to the removal of specific mass/stiffness elements at a localized bolted connection. Consequently, these percentages represent discrete physical states rather than a continuous degradation of global system parameters. This distinction explains the **non-monotonic behavior** observed in the anomaly scores (Figure 4.10). As the damage state transitions from 4% to 6%, the system dynamics may undergo complex bifurcations characteristic of the magneto-elastic nonlinearity (e.g., shifting between periodic and chaotic regimes). A higher percentage of mass removal does not strictly imply a linear increase in prediction error; rather, it shifts the system to a different dynamic manifold.

The overlap observed between the healthy upper bound and the lower bound of damaged cases highlights the potential for false negatives in single-instance detections. To address this, the presented results utilize the statistical mean as the robust anomaly measure. It is acknowledged that obtaining multiple experimental passes for a specific damage state is often unfeasible in rapid assessment scenarios. However, within the context of a deployed **continuous monitoring solution**, this limitation is effectively mitigated. By accumulating multiple records over extended temporal horizons ranging from hours to days or months, the system can aggregate samples to smooth out stochastic fluctuations. This temporal averaging ensures that a stable diagnostic trend emerges, effectively resolving the distributional overlap and enhancing the reliability of detection for low-magnitude damage.

4.4.3 Interpretability and Damage Localization Insights

A key advantage of the SensorFusionTFT lies in its inherent interpretability, which allows us to move beyond simple detection towards diagnostics. The interpretability dashboard in Figure 4.12 deconstructs the model’s internal reasoning.

The top row reveals the mean sensor importance assigned by the VSN to each of the three accelerometers. The model learns to consistently assign the highest

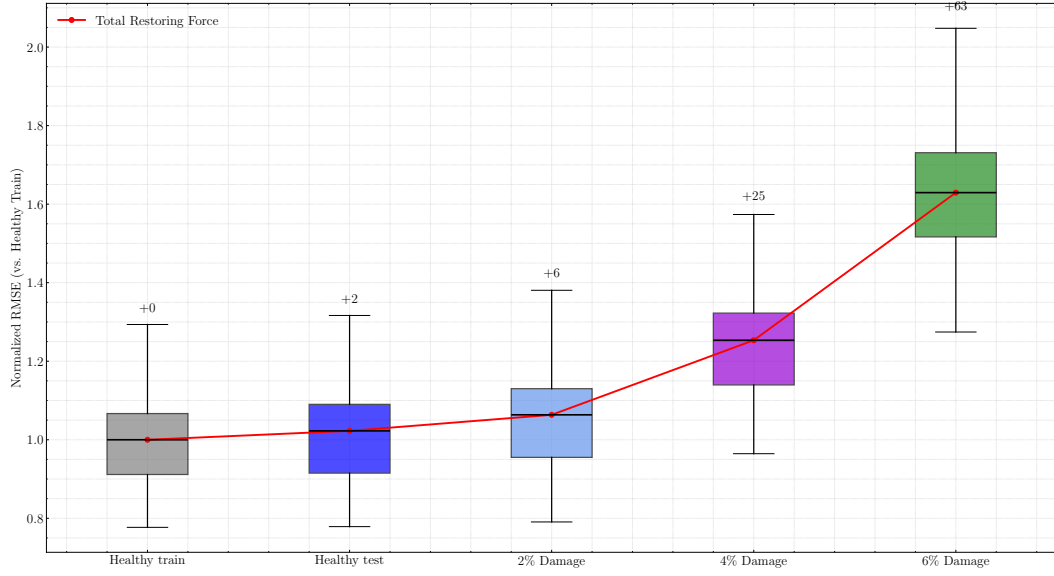


Figure 4.10. Normalized distribution of RMSE for the magneto-elastic beam. The trend of mean value for each damage distribution demonstrates a clear increase in reconstruction error with physical damage severity.

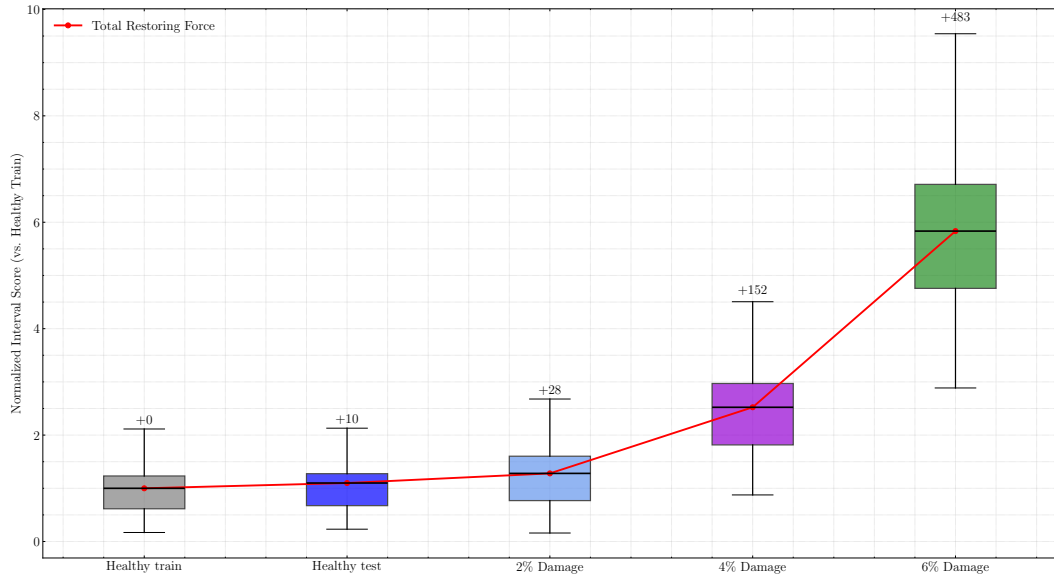


Figure 4.11. Normalized distribution of Interval Score (IS) for the magneto-elastic beam. The pronounced increase of mean values of IS highlights the model's sensitivity to anomalous behavior that violates its learned uncertainty bounds.

importance to Accelerometer 2. This learned spatial prioritization is a powerful result, as this sensor is physically located nearest to the mid-span bolted connection where damage was induced. The VSN has correctly identified the most informative channel for monitoring the structural region most susceptible to change, providing a crucial step towards Level II: Localization of damage hierarchy.

The middle row visualizes the deviation in the model's temporal self-attention

pattern, which becomes increasingly pronounced with damage compared to the healthy train dataset baseline. The bottom row consolidates these temporal shifts into the scalar Attention Anomaly Score. This metric exhibits a clear monotonic increase highly correlated with damage severity, rising from a baseline of +0.3 for the healthy case to +12.6 for the 6% damage state. This provides a third, independent line of evidence for damage detection, confirming that anomalies can be robustly identified not just through reconstruction error, but also by monitoring the model’s internal diagnostic signals. This demonstrates that the SensorFusionTFT has successfully learned to associate specific dislocations in its spatio-temporal reasoning with the onset of physical damage.

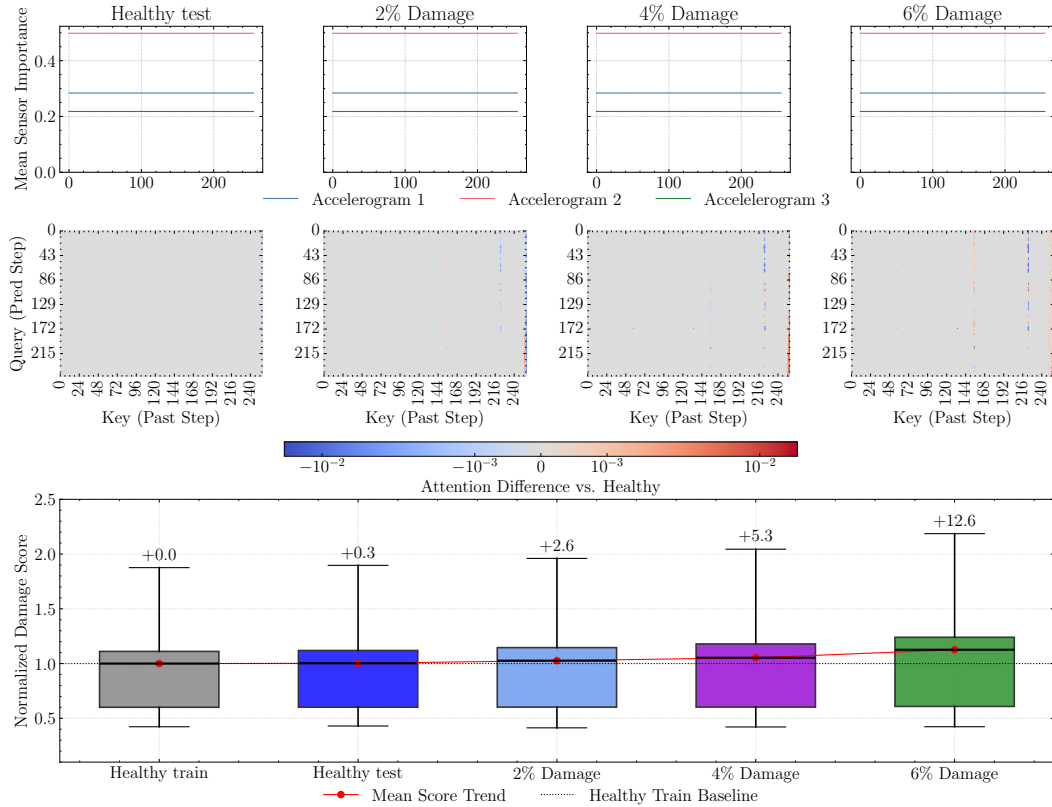


Figure 4.12. Interpretability dashboard for the magneto-elastic beam. The VSN’s focus on the sensor nearest the damage (top), the deviation in temporal attention (middle), and the rising Attention Anomaly Score (bottom) demonstrate the model’s diagnostic capabilities.

4.5 Large-Scale Experimental Validation: The LUMO Lattice Tower

The culminating validation of this chapter transitions from controlled laboratory setups to a large-scale, real-world benchmark: the Leibniz University Test Structure for Monitoring (LUMO), whose physical characteristics were detailed in Section 2.2.2. This case study was specifically chosen to probe the practical limits and test the key

design principles of the SensorFusionTFT under the most challenging conditions. Unlike the previous benchmarks, the LUMO structure is a 9-meter-tall outdoor mast continuously exposed to real-world environmental and operational variability, including changing wind loads and temperature cycles. Furthermore, its dense network of 22 sensors directly tests the model’s core sensor fusion capability and its scalability to high-dimensional data streams.

For this application, the SensorFusionTFT is trained in the auto-encoding framework to reconstruct the feature-engineered response of all 22 sensors. The model’s efficacy is assessed using the full suite of three anomaly metrics defined in my methodology.

4.5.1 Damage Detection with Reconstruction-Based Metrics

The model’s fundamental damage detection capability is evaluated using the RMSE and IS metrics, with the distributions of these scores across all 22 sensor channels presented in Figures 4.13 and 4.14. A clear distinction is immediately evident between the healthy and damaged states. In both figures, the error distributions for the healthy data are compactly centered around the baseline (red dashed line).

Upon the introduction of damage, even at the moderate level (e.g., x12_DAM3_010), a significant upward shift in the error distributions occurs for a *subset* of the sensors. As the damage becomes more severe (e.g., x06_DAM3_111), the magnitude and spread of these errors increase substantially. This trend consistently demonstrates that the model, having learned the signature of healthy structural dynamics, correctly identifies the damaged states as anomalous. The pronounced separation between the healthy and damaged score distributions validates the effectiveness of using reconstruction error from the SensorFusionTFT as a robust indicator for unsupervised damage detection in a complex, high-dimensional, and noisy real-world environment.

4.5.2 Diagnostic Insights from Internal Mechanisms and Localization

While the reconstruction error metrics provide a strong signal for detection, the model’s internal attention mechanism offers a deeper, physics-agnostic insight. Figure 4.15 presents the distribution of the Attention Anomaly Score for the healthy baseline and the six damage scenarios. The results show a clear and generally monotonic trend: the scores for all damage cases are statistically distinct from the healthy baseline, with the median score increasing with damage severity (e.g., a 66% increase for moderate damage at DAM3). This demonstrates that by monitoring the model’s internal reasoning, we can derive a powerful and quantitative indicator of structural health that is complementary to traditional error metrics.

4.6 Summary and Findings

This chapter introduced and rigorously evaluated the SensorFusionTFT, a novel architecture designed to overcome the limitations of the convolutional models analyzed

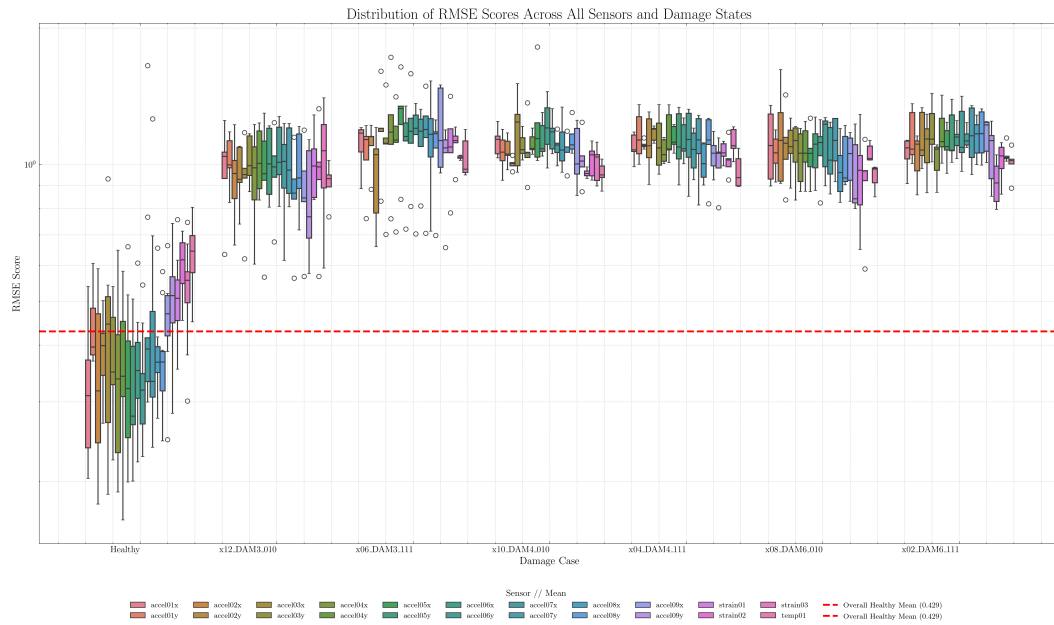


Figure 4.13. Distribution of normalized RMSE scores across all sensors and structural states for the LUMO tower. A clear increase in the magnitude and variance of the errors is observable for the damaged cases.

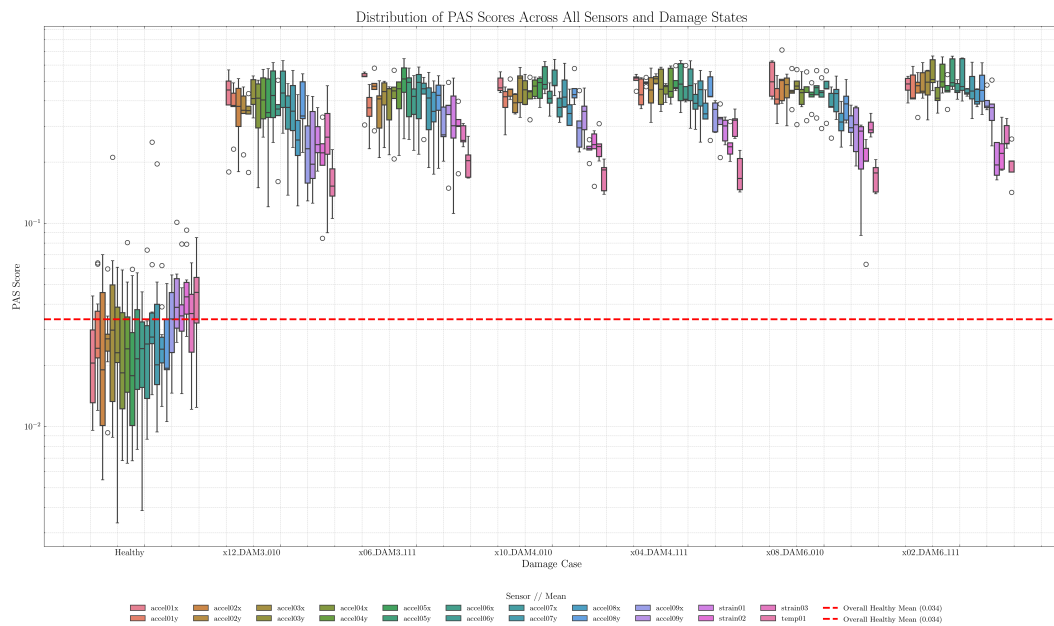


Figure 4.14. Distribution of normalized Interval Scores (IS) for the LUMO tower. The damaged cases exhibit a significantly larger and more widespread distribution of scores, indicating a fundamental breakdown in the model’s learned uncertainty.

in Chapter 3. The central finding is that specialized Transformer architectures provide a demonstrably superior framework for the challenging domain of unsupervised damage detection in nonlinear systems. By leveraging a self-attention mechanism,

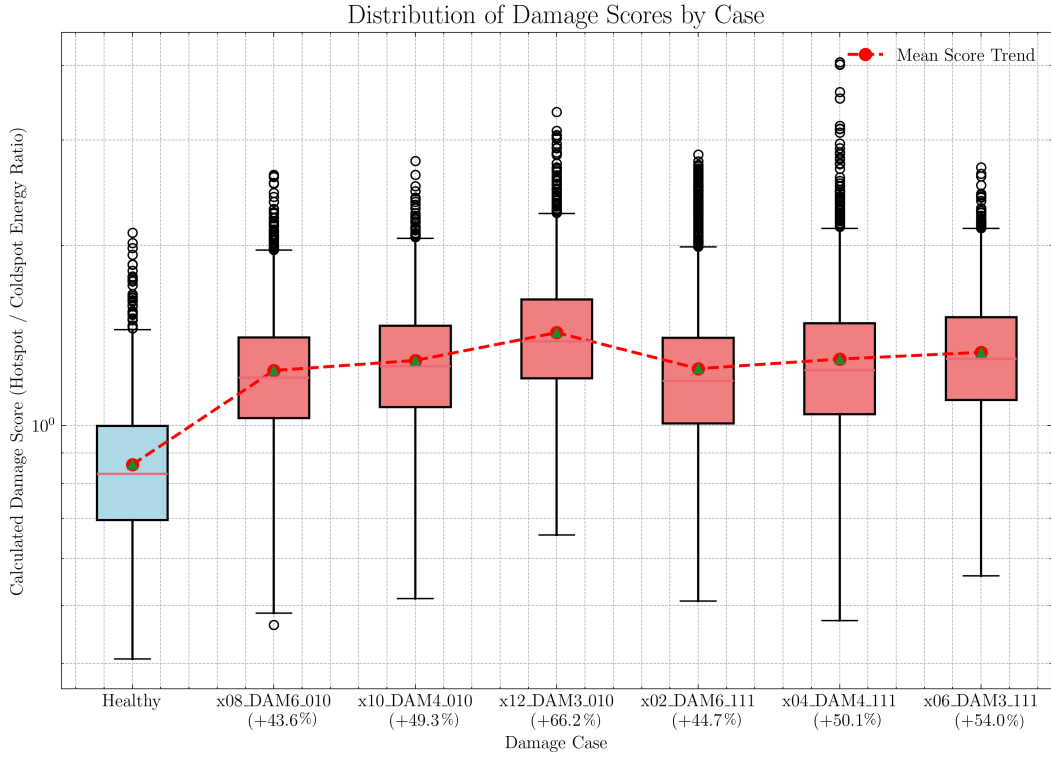


Figure 4.15. Distribution of the Attention Anomaly Score by damage case for the LUMO tower. The statistically significant increase in the score for all damage scenarios confirms its utility as a robust, complementary damage indicator.

the SensorFusionTFT showed an enhanced ability to learn the complex, non-local, and path-dependent dynamics characteristic of the nonlinear benchmarks, a task where the baseline convolutional autoencoder showed limitations.

A key contribution demonstrated in this chapter is the power of a multi-faceted diagnostic framework, which moves beyond a single anomaly score. The framework combines three complementary indicators: (1) the foundational reconstruction error (RMSE), (2) the highly sensitive Interval Score (IS), which measures the breakdown of the model’s learned uncertainty, and (3) the novel Attention Anomaly Score, derived from the model’s internal reasoning. The results consistently showed that a simultaneous spike across all three metrics provides a much stronger and more reliable indication of damage than any single metric in isolation. The exceptional sensitivity of the Interval Score, in particular, establishes that monitoring a model’s uncertainty breaches is a potent method for detecting incipient damage.

Perhaps the most significant advancement presented is the move from simple detection to interpretable diagnostics, a direct response to the limitations of the black-box models analyzed in the previous chapter. The two core innovations of the SensorFusionTFT proved critical in this transition. The Time-Varying Variable Selection Network (VSN) demonstrated a sophisticated, objective-driven capability that extends beyond simple sensor fusion. It learned to perform a physically intuitive allocation of its focus based on the specific prediction task; for instance, it correctly prioritized the acceleration signal in the dual-target isolator benchmark,

having identified it as the most salient feature for concurrently predicting both the internal restoring force and the future kinematic state. This learned intelligence, combined with its demonstrated ability to provide a clear pathway towards Level II: Localization-exemplified in the magneto-elastic beam experiment where it correctly prioritized the sensor nearest the physical damage - establishes the VSN as a powerful diagnostic tool. Complementing this spatial insight, the novel Attention Anomaly Score offered a significant advancement over traditional error-based methods. Its primary advantage is its ability to move beyond quantifying the magnitude of an anomaly to characterizing its nature. It revealed how the underlying temporal logic of the system shifted due to damage, providing a deeper diagnostic insight unattainable with simple reconstruction metrics.

In summary, the findings of this chapter establish the SensorFusionTFT as a state-of-the-art framework for unsupervised damage diagnostics in nonlinear systems, outperforming the baseline convolutional models in both performance and interpretability. The comprehensive results from this chapter, alongside the baseline performance established in chapter 3, now provide the foundation for a final, critical synthesis and comparative evaluation, which will be the subject of the concluding chapter of this dissertation.

Chapter 5

Conclusion and Future Research

This dissertation tackled a systematic investigation into the application of advanced machine learning methods for the challenging domain of unsupervised damage detection in nonlinear structural systems. The research was motivated by an important gap in the existing literature: while traditional methods for nonlinear SHM face significant limitations and deep learning applications have largely focused on linear systems, a comprehensive evaluation of architectural progression from foundational convolutional models to advanced, interpretable models had yet to be performed. This work has sought to fill that gap, presenting a coherent research narrative that not only validates the use of these models but also provides a critical comparative analysis of their capabilities and practical trade-offs. This final chapter summarizes the main findings of the entire study, offers general guidelines for practitioners, acknowledges the limitations of the work, and proposes new directions for future research.

5.1 Summary of Main Findings

The research presented in this thesis was structured in two principal stages, reflecting a progression of increasing architectural complexity and diagnostic capability.

Stage 1: Establishing a Baseline with Convolutional Architectures. The first steps were discussed in chapter 3 by establishing a performance baseline using two foundational unsupervised deep learning models: the Convolutional Autoencoder (AE) and the Generative Adversarial Network (GAN). The key findings from this extensive study across multiple numerical and experimental nonlinear systems were threefold.

- Both architectures successfully validated the unsupervised learning paradigm, proving capable of detecting the onset of damage in a variety of nonlinear systems by learning the dynamics of the healthy state alone.
- A clear performance trade-off was identified: the AE exhibited greater robustness and training stability, providing consistent performance even on the most complex benchmarks. The GAN, while demonstrating exceptional sensitivity in certain cases, proved more difficult to train and less reliable on systems with highly complex dynamics.
- The GAN offered a distinct advantage in the interpretability of its output, with the discriminator's score providing a naturally bounded and intuitive damage index.

While effective, this initial investigation also shed light onto the key limitations of these architectures: their nature as black-box models offered little diagnostic insight, and their fundamental design was not optimized for fusing information from multiple sensor channels.

Stage 2: Advancing Diagnostics with an Interpretable Transformer. To address these shortcomings, the second stage of this research, detailed in chapter 4, introduced the SensorFusionTFT, a novel architecture adapted from the Temporal

Fusion Transformer. This model was specifically re-engineered for the demands of multi-sensor SHM. The key findings from this stage represented a significant advancement over the baseline.

- The self-attention mechanism proved demonstrably superior at learning the complex, path-dependent, and non-local dynamics characteristic of the nonlinear benchmarks, resulting in higher sensitivity and more robust performance.
- The development of a multi-faceted diagnostic framework—combining reconstruction error (RMSE), the novel Interval Score (IS), and an Attention Anomaly Score—provided a much more reliable and high-confidence method for flagging damage than any single metric in isolation. The IS, in particular, proved to be an exceptionally powerful indicator of incipient damage.
- Most critically, the architecture’s inherent interpretability provided a pathway to true diagnostics. The Time-Varying Sensor Fusion Network successfully identified the sensor channels most relevant to the physical location of damage, representing a crucial step towards Level II Localization within an unsupervised framework.

5.1.1 Critical Comparative Evaluation of the Investigated Methods

The overarching contribution of this dissertation is the direct, evidence-based comparison of the three investigated architectural classes. This evaluation moves beyond a simple assessment of detection accuracy to encompass a multi-faceted analysis of their practical and theoretical attributes. Table 5.1 provides a concise synthesis of this evaluation, critically comparing the Autoencoder, Generative Adversarial Network, and the SensorFusionTFT across several key axes, including diagnostic performance, interpretability, and the practical considerations of model complexity and computational cost. The findings clearly establish a narrative of architectural progression, highlighting the specific strengths and trade-offs at each stage of the research.

5.1.2 General Guidelines for Applying Deep Learning to Nonlinear SHM

Based on the comprehensive findings of this work, the following general guidelines can be proposed for researchers and practitioners seeking to apply these methods.

Start with the Simplest Effective Model: The Autoencoder. For initial exploratory analysis or on systems with limited sensor channels, the convolutional AE provides a robust, reliable, and easy-to-implement baseline. Its consistent performance makes it an excellent first choice for validating the feasibility of an unsupervised SHM approach.

Consider the GAN for Interpretable Anomaly Scoring. If the primary goal is a single, naturally bounded damage index and the additional training complexity can be accommodated, the GAN is a valuable alternative. However, one must

Table 5.1. Critical comparative evaluation of the investigated deep learning architectures.

Criterion	Autoencoder (AE)	Generative Adversarial Network (GAN)	SensorFusionTFT
Diagnostic Performance	Reliable and consistent detection; robust across most benchmarks.	Variable; potential for high sensitivity but prone to instability and false positives on complex systems.	Superior sensitivity and robustness, especially with the multi-metric (RMSE, IS, Attention) assessment.
Interpretability	Low. Anomaly score is an undifferentiated 'black-box' reconstruction error.	Moderate. The discriminator output provides a naturally bounded and interpretable damage index.	High. Provides both spatial (sensor) and temporal diagnostic insights via the VSN and self-attention mechanisms.
Model Complexity (Trainable Parameters)	Low to Moderate. (~70k - 2.3M)	Very High. (~4.1M - 8.1M, dominated by the discriminator).	High. (~1M - 5M, depending on depth and heads).
Computational Cost(Approx. Training Time)	Low. (Typically minutes per benchmark).	Moderate to High. (Tens of minutes, requires careful adversarial tuning).	High to Very High. (30 mins for numerical to >7 hours for large-scale experimental data).
Data Requirement (Training Scale)	Effective with a moderate number of short, independent time-series samples (~1600 per case).	Similar to AE; requires a sufficient number of samples to learn the healthy data distribution.	Highly efficient; effective with fewer but longer, continuous time-series for self-supervised windowing (~200 per case).
Scalability (Multi-Sensor)	Limited. Typically applied on a per-sensor basis and lacks a native fusion mechanism.	Limited. Same per-sensor application as the AE.	High. Explicitly designed for dynamic multi-sensor fusion via the VSN, as demonstrated on the LUMO benchmark.

be prepared for its potential performance variability on highly complex dynamic systems.

Invest in Transformers for Multi-Sensor Diagnostics. For complex, high-value assets instrumented with a dense array of sensors, an investment in a Transformer-based architecture like the SensorFusionTFT is justified. Its native ability to fuse multi-sensor data and its deep interpretability features move beyond simple detection to provide true diagnostic insights, including a pathway to damage localization.

Embrace Multi-Metric Anomaly Assessment. A key lesson from this research is that relying on a single anomaly score can be fragile. A multi-metric framework, such as the combination of reconstruction error, uncertainty-breach scores (IS), and attention-based metrics, provides a much higher-confidence basis for making operational decisions and minimizing false alarms.

5.2 Limitations and Recommendations for Future Research

It is important to acknowledge the boundaries of the present study and to propose directions for future inquiry. The following limitations and recommendations are offered to guide subsequent research in this promising field.

Disentangling Environmental and Operational Effects. This work assumed that significant changes in the system response are attributable to damage. However, in real-world applications, environmental factors (e.g., temperature, humidity) and operational variability can induce changes of a similar magnitude. Future research should focus on methods to explicitly disentangle these effects, for instance by incorporating environmental measurements as direct inputs to the deep learning models (conditioned architectures) or by employing domain adaptation techniques to create environment-invariant representations.

Advancing Towards Full Quantification and Localization. This thesis has laid the groundwork for Levels I and III of the SHM hierarchy. Future work should build upon this by developing methods to directly correlate the magnitude of the anomaly scores with a physical measure of damage extent, thereby achieving full Level III and Level IV Localization and Prognosis. Similarly, the localization insights from the VSN could be formalized into a probabilistic localization algorithm.

Exploring Other Damage Types and System Classes. The numerical and experimental studies herein focused on damage manifested as a change in stiffness or mass. The robustness of these methods should be further validated on other damage types, such as the initiation of plasticity, cracks onset, delamination in composites, or the introduction of non-smooth nonlinearities like backlash, friction in joints etc.

Investigating Transfer Learning and Digital Twins. A major challenge in SHM is the cold start problem, where data from a newly deployed structure is limited. Future research could explore transfer learning techniques, where a model pre-trained on a high-fidelity numerical digital twin of the structure is fine-tuned on the limited real-world data, potentially accelerating deployment and improving day-one performance.

5.3 Concluding Remarks

In conclusion, this dissertation has presented a comprehensive and systematic investigation into the application of advanced deep learning architectures for unsupervised damage detection in nonlinear dynamic systems. By progressing from foundational convolutional models to a novel, interpretable Transformer-based framework, this work has demonstrated a clear pathway of increasing diagnostic complexity. The findings establish that these methods can successfully overcome the limitations of traditional techniques, learning the complex signatures of healthy nonlinear behavior directly from data and reliably detecting deviations indicative of damage. The introduction of the SensorFusionTFT and its multi-faceted diagnostic capabilities, in particular, represents a significant step forward, moving the field from simple anomaly detection towards true, interpretable diagnostics. While challenges remain, this research successfully establishes that specialized, interpretable deep learning architectures are a highly promising and powerful tool for the next generation of autonomous and intelligent structural health monitoring systems.

Bibliography

- [1] Douglas E. Adams and M. Nataraju. “A nonlinear dynamical systems framework for structural diagnosis and prognosis”. In: *International Journal of Engineering Science* 40.17 (2002), pp. 1919–1941. DOI: 10.1016/S0020-7225(02)00078-X.
- [2] Mohammadreza Ahmadzadeh, Seyed Mehdi Zahrai, and Maryam Bitaraf. “An Integrated Deep Neural Network Model Combining 1D CNN and LSTM for Structural Health Monitoring Utilizing Multisensor Time-Series Data”. In: *Structural Health Monitoring* 24.1 (Jan. 2025), pp. 447–465. ISSN: 1475-9217. DOI: 10.1177/14759217241239041. (Visited on 12/02/2025).
- [3] Nizar Faisal Alkayem et al. “Structural damage detection using finite element model updating with evolutionary algorithms: a survey of objective functions”. In: *Neural Computing and Applications* 28 (2017), pp. 2065–2090.
- [4] ASHVIN Project. *Structural Health Monitoring (SHM): Concept and Objectives*. 2022. URL: <https://www.ashvin.eu/2022/04/11/structural-health-monitoring-shm-concept-and-objectives/>.
- [5] Onur Avci et al. “A review of vibration-based damage detection in civil structures: From traditional methods to Machine Learning and Deep Learning applications”. In: *Mechanical systems and signal processing* 147 (2021), p. 107077.
- [6] Muhammad Muzammil Azad et al. “Intelligent structural health monitoring of composite structures using machine learning, deep learning, and transfer learning: a review”. In: *Advanced Composite Materials* 33.2 (Mar. 2024), pp. 162–188. ISSN: 0924-3046. DOI: 10.1080/09243046.2023.2215474.
- [7] Mohsen Azimi, Armin Dadras Eslamlou, and Gokhan Pekcan. “Data-Driven Structural Health Monitoring and Damage Detection through Deep Learning: State-of-the-Art Review”. en. In: *Sensors* 20.10 (Jan. 2020), p. 2778. ISSN: 1424-8220. DOI: 10.3390/s20102778.
- [8] N. Bolander et al. “Physics-based Remaining Useful Life Prediction for Aircraft Engine Bearing Prognostics”. In: *Annual Conference of the Prognostics and Health Management Society*. 2009.
- [9] Luke Bornn, Charles R Farrar, and Gyuhae Park. “Damage detection in initially nonlinear systems”. In: *International Journal of Engineering Science* 48.10 (2010), pp. 909–920.

- [10] R Bouc. “Forced vibrations of mechanical systems with hysteresis”. In: *Proc. of the Fourth Conference on Nonlinear Oscillations, Prague, 1967*. 1967.
- [11] David Agis Tirado Burgos, Melisa Méndez Ruiz, and Francesc Pozo. “Damage Identification in Structural Health Monitoring: A Brief Review from its Implementation to the Use of Data-Driven Applications”. In: *Sensors* 20.3 (2020), p. 733.
- [12] Tom Butlin et al. “The landscape of nonlinear structural dynamics: an introduction”. In: *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences* 373.2051 (2015), p. 20140400.
- [13] B. Carboni et al. “Nonlinear normal modes for damage detection: Theoretical concepts and preliminary experimental validation”. In: *Proceedings of the 8th World Conference On Structural Control And Monitoring*. 2022.
- [14] M. Carminati and S. Ricci. “Structural Damage Detection Using Nonlinear Vibrations”. In: *Advances in Aerospace Engineering* 2018 (2018), p. 1901362.
- [15] Matthew David Champneys. “Challenges in nonlinear structural dynamics”. PhD thesis. University of Sheffield, 2022.
- [16] Aristotelis E Charalampakis and George C Tsiatas. “A simple rate-independent uniaxial shape memory alloy (SMA) model”. In: *Frontiers in Built Environment* 4 (2018), p. 46.
- [17] Qing Chen et al. “Estimating Residual Life Distributions of Complex Operational Systems: A Case Study”. In: *Reliability Engineering & System Safety* 204 (2020), p. 107206.
- [18] William J. R. Christian et al. “Real-time quantification of damage in structural materials using digital image correlation and acoustic emission”. In: *Royal Society Open Science* 7.3 (2020), p. 191407.
- [19] M. Civera, L. Zanotti Fragonara, and C. Surace. “Using video processing for the full-field identification of backbone curves in case of large vibrations”. In: *Sensors* 19.10 (2019), p. 2345. DOI: 10.3390/s19102345.
- [20] Y. N. Dauphin et al. “Language Modeling with Gated Convolutional Networks”. In: *Proceedings of the International Conference on Machine Learning (ICML)*. PMLR, 2017, pp. 933–941. URL: <https://proceedings.mlr.press/v70/dauphin17a.html>.
- [21] Hao Deng and JianHua Chen. “Structural Modal Parameter Recognition and Related Damage Identification Methods under Environmental Excitations: A Review”. In: *Structural Durability & Health Monitoring* 19.1 (2024).
- [22] Scott W Doebling, Charles R Farrar, Michael B Prime, et al. “A summary review of vibration-based damage identification methods”. In: *Shock and vibration digest* 30.2 (1998), pp. 91–105.
- [23] Charles R. Farrar. *Structural Health Monitoring using Pattern Recognition*. Tech. rep. Los Alamos National Laboratory, 2007.

- [24] Charles R. Farrar, Scott W. Doebling, and David A. Nix. “Vibration-based structural damage identification”. In: *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences* 359.1778 (2001), pp. 131–149.
- [25] Charles R. Farrar and Nicholas A. J. Lieven. “Damage prognosis: the future of structural health monitoring”. In: *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences* 365.1851 (2007), pp. 623–632. DOI: 10.1098/rsta.2006.1927.
- [26] Charles R. Farrar and Keith Worden. “An introduction to structural health monitoring”. In: *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences* 365.1851 (2007), pp. 303–315.
- [27] Charles R. Farrar and Keith Worden. *Structural Health Monitoring: A Machine Learning Perspective*. John Wiley & Sons, 2012. ISBN: 978-1119994963.
- [28] M. I. Friswell, S. Adhikari, and Y. Lei. “Vibration Analysis of Beams with Non-Local Foundations Using the Finite Element Method”. In: *International Journal for Numerical Methods in Engineering* 71.11 (2007), pp. 1365–1386. ISSN: 1097-0207. DOI: 10.1002/nme.2003. (Visited on 12/02/2025).
- [29] S. Ganguly and C. S. Manohar. “Methodologies for modeling and identification of breathing crack systems: A review”. In: *Mechanical Systems and Signal Processing* 205 (2023), p. 110827.
- [30] Valentina Giglioni et al. “Autoencoders for Unsupervised Real-time Bridge Health Assessment”. In: *Computer-Aided Civil and Infrastructure Engineering* 38.8 (May 2023), pp. 959–974. ISSN: 1093-9687, 1467-8667. DOI: 10.1111/mice.12943. (Visited on 10/30/2025).
- [31] Tilmann Gneiting and Adrian E Raftery. “Strictly proper scoring rules, prediction, and estimation”. In: *Journal of the American statistical Association* 102.477 (2007), pp. 359–378.
- [32] Ian Goodfellow et al. “Generative adversarial nets”. In: *Advances in neural information processing systems* 27 (2014).
- [33] K. He et al. “Deep Residual Learning for Image Recognition”. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. 2016, pp. 770–778. DOI: 10.1109/CVPR.2016.90. URL: <https://doi.org/10.1109/CVPR.2016.90>.
- [34] Jing Jia and Ying Li. “Deep Learning for Structural Health Monitoring: Data, Algorithms, Applications, Challenges, and Trends”. en. In: *Sensors* 23.21 (Jan. 2023), p. 8824. ISSN: 1424-8220. DOI: 10.3390/s23218824.
- [35] Harrish Joseph et al. “Deep Learning Architectures for Data-Driven Damage Detection in Nonlinear Dynamic Systems under Random Vibrations”. In: *Nonlinear Dynamics* 112.23 (Dec. 2024), pp. 20611–20636. ISSN: 1573-269X. DOI: 10.1007/s11071-024-10270-1.
- [36] Diederik P. Kingma and Jimmy Ba. *Adam: A Method for Stochastic Optimization*. Jan. 2017. DOI: 10.48550/arXiv.1412.6980. arXiv: 1412.6980 [cs]. (Visited on 02/24/2024).

- [37] R. Koenker and K. F. Hallock. “Quantile Regression”. In: *Journal of Economic Perspectives* 15 (2001), pp. 143–156. DOI: 10.1257/jep.15.4.143. URL: <https://doi.org/10.1257/jep.15.4.143>.
- [38] Nitin Nagesh Kulkarni and Alessandro Sabato. “Full-Field Expansion and Damage Detection from Sparse Measurements Using Physics-Informed Variational Autoencoders”. In: *Structural Health Monitoring* (Oct. 2024), p. 14759217241289575. ISSN: 1475-9217. DOI: 10.1177/14759217241289575. (Visited on 12/02/2025).
- [39] Walter Lacarbonara, Biagio Carboni, and Giuseppe Quaranta. “Nonlinear normal modes for damage detection”. In: *Meccanica* 51 (2016), pp. 2629–2645. DOI: 10.1007/s11012-016-0466-9.
- [40] U. Lee and J. Shin. “A frequency response function-based structural damage identification method”. In: *Computers & Structures* 80.2 (2002), pp. 117–132.
- [41] Youngseo Lee et al. “Structural damage detection using deep learning and FE model updating”. In: *Scientific Reports* 13 (2023), p. 18063.
- [42] H. Li, H. Yang, and S. L. J. Hu. “Modal Strain Energy Decomposition Method for Damage Localization in 3D Frame Structures”. In: *Journal of Engineering Mechanics* 132.9 (2006), pp. 941–951.
- [43] Lechen Li, Marcello Morgantini, and Raimondo Betti. “Structural damage assessment through a new generalized autoencoder with features in the quefrequency domain”. In: *Mechanical Systems and Signal Processing* 184 (2023), p. 109713.
- [44] B. Lim et al. “Temporal Fusion Transformers for Interpretable Multi-horizon Time Series Forecasting”. In: *International Journal of Forecasting* 37.4 (2021), pp. 1748–1764. DOI: 10.1016/j.ijforecast.2021.03.012. URL: <https://doi.org/10.1016/j.ijforecast.2021.03.012>.
- [45] Maria Pina Limongelli. *Vibration-based monitoring*. 2022. URL: <https://limongelli.faculty.polimi.it/vibration-based-monitoring/>.
- [46] Maria Pina Limongelli and Mehmet Çelebi. “On the standardization of procedures for Structural Health Monitoring”. In: *Proceedings of SHMII-9* (2019).
- [47] S. Liu et al. “Effects of experimental variables on the nonlinear harmonic generation technique”. In: *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control* 58.7 (2011), pp. 1442–1451.
- [48] Furkan Luleci, F Necati Catbas, and Onur Avci. “CycleGAN for undamaged-to-damaged domain translation for structural health monitoring and damage detection”. In: *Mechanical Systems and Signal Processing* 197 (2023), p. 110370.
- [49] Yongpeng Luo et al. “Unsupervised structural damage detection based on an improved generative adversarial network and cloud model”. In: *Journal of Low Frequency Noise, Vibration and Active Control* 42.3 (2023), pp. 1501–1518.
- [50] Xirui Ma et al. “Structural damage identification based on unsupervised feature-extraction via variational auto-encoder”. In: *Measurement* 160 (2020), p. 107811.

- [51] S. F. Masri and T. K. Caughey. “A nonparametric identification technique for nonlinear dynamic problems”. In: *Journal of Applied Mechanics* 46.2 (1979), pp. 433–447. DOI: 10.1115/1.3153383.
- [52] Joseph Melville et al. “Structural Damage Detection Using Deep Learning of Ultrasonic Guided Waves”. In: *AIP Conference Proceedings*. Vol. 1949. 2018, p. 230004.
- [53] E. Alexandra Micu et al. “Using Accelerometers on an In-Service Train to Detect Bridge Damage and Repair”. In: *Transport Transitions: Advancing Sustainable and Inclusive Mobility*. Ed. by Ciaran McNally et al. Cham: Springer Nature Switzerland, 2026, pp. 880–886. ISBN: 978-3-032-04774-8. DOI: 10.1007/978-3-032-04774-8_126.
- [54] Subhasish Mohanty, Aditi Chattopadhyay, and Pedro Peralta. “Adaptive Residual Useful Life Estimation of a Structural Hotspot”. In: *Journal of Intelligent Material Systems and Structures* 21.3 (2010), pp. 321–335.
- [55] Pravin Nagar et al. “SEMA: Semantic Attention for Capturing Long-Range Dependencies in Egocentric Lifelogs”. In: *2024 IEEE/CVF Winter Conference on Applications of Computer Vision (WACV)*. Jan. 2024, pp. 7010–7020. DOI: 10.1109/WACV57701.2024.00687. (Visited on 10/25/2025).
- [56] Rumi Nakagawa et al. *Upscaling Global Hourly GPP with Temporal Fusion Transformer (TFT)*. June 2023. DOI: 10.48550/arXiv.2306.13815. arXiv: 2306.13815 [cs]. (Visited on 10/25/2025).
- [57] AK Pandey, M Biswas, and MM Samman. “Damage detection from changes in curvature mode shapes”. In: *Journal of sound and vibration* 145.2 (1991), pp. 321–332.
- [58] Marco Parola, F. A. Galatolo, and M. G. C. A. Cimino. “Convolutional Neural Networks for Structural Damage Localization on Digital Twins”. In: *Engineering Applications of Artificial Intelligence* (2023).
- [59] Razvan Pascanu, Tomas Mikolov, and Yoshua Bengio. *On the Difficulty of Training Recurrent Neural Networks*. Feb. 2013. DOI: 10.48550/arXiv.1211.5063. arXiv: 1211.5063 [cs]. (Visited on 10/24/2025).
- [60] Chathurdara Sri Nadith Pathirage et al. “Structural damage identification based on autoencoder neural networks and deep learning”. In: *Engineering structures* 172 (2018), pp. 13–28.
- [61] Zhen Peng et al. “Nonlinear structural damage detection using output-only Volterra series model”. In: *Structural Control and Health Monitoring* 28.9 (2021), e2802.
- [62] J. Prawin and A. R. M. Rao. “Damage detection in nonlinear systems using an improved describing function approach with limited instrumentation”. In: *Nonlinear Dynamics* 96 (2019), pp. 1447–1470. DOI: 10.1007/s11071-019-04864-3.

- [63] Zahra Rastin, Gholamreza Ghodrati Amiri, and Ehsan Darvishan. “Unsupervised Structural Damage Detection Technique Based on a Deep Convolutional Autoencoder”. In: *Shock and Vibration* 2021.1 (2021), p. 6658575. ISSN: 1875-9203. DOI: 10.1155/2021/6658575. (Visited on 10/24/2025).
- [64] Lillian J Ratliff, Samuel A Burden, and S Shankar Sastry. “Characterization and computation of local Nash equilibria in continuous games”. In: *2013 51st Annual Allerton Conference on Communication, Control, and Computing (Allerton)*. IEEE. 2013, pp. 917–924.
- [65] Resensys. *5 Key Indicators Your Infrastructure Needs Structural Health Monitoring*. 2025. URL: <https://www.resensys.com/Blog/key-indicators-your-infrastructure-needs-structural-health-monitoring/>.
- [66] RISE Universidade Lusófona. *Structural Health Monitoring*. 2024. URL: <https://rise.ulusofona.pt/research/structural-health-monitoring>.
- [67] David E Rumelhart, Geoffrey E Hinton, and Ronald J Williams. “Learning Internal Representations by Error Propagation, Parallel Distributed Processing, Explorations in the Microstructure of Cognition, ed. DE Rumelhart and J. McClelland. Vol. 1. 1986”. In: *Biometrika* 71 (1986), pp. 599–607.
- [68] A. Rytter. “Vibrational based inspection of civil engineering structures”. PhD thesis. Aalborg University, 1993.
- [69] Andrea Salvatore, Biagio Carboni, and Walter Lacarbonara. “Nonlinear dynamic response of an isolation system with superelastic hysteresis and negative stiffness”. In: *Nonlinear dynamics* (2021), pp. 1–26.
- [70] A. A. Shah and Y. Ribakov. “Non-linear ultrasonic evaluation of damaged concrete based on higher order harmonic generation”. In: *Materials & Design* 30.10 (2009), pp. 4095–4102.
- [71] Zhiqiang Shang et al. “Vibration-based damage detection for bridges by deep convolutional denoising autoencoder”. In: *Structural Health Monitoring* 20.4 (2021), pp. 1880–1903.
- [72] Jiaqi Shen et al. “Optimization of Temporal Feature Attribution and Sequential Dependency Modeling for High-Precision Multi-Step Resource Forecasting: A Methodological Framework and Empirical Evaluation”. In: *Mathematics* 13.8 (Jan. 2025), p. 1339. ISSN: 2227-7390. DOI: 10.3390/math13081339. (Visited on 10/25/2025).
- [73] Apeksha Shewalkar, Deepika Nyavanandi, and Simone A. Ludwig. “Performance Evaluation of Deep Neural Networks Applied to Speech Recognition: RNN, LSTM and GRU”. In: *Journal of Artificial Intelligence and Soft Computing Research* 9.4 (Oct. 2019), pp. 235–245. ISSN: 2083-2567. DOI: 10.2478/jaiscr-2019-0006. (Visited on 10/24/2025).
- [74] Z. Y. Shi, S. S. Law, and L. M. Zhang. “Structural damage localization from modal strain energy change”. In: *Journal of Sound and Vibration* 218.5 (1998), pp. 825–844.

- [75] S. B. Shiki, S. da Silva, and M. D. Todd. “On the application of discrete-time volterra series for the damage detection problem in initially nonlinear systems”. In: *Structural Health Monitoring* 16.1 (2017), pp. 62–78. DOI: 10.1177/1475921716653835.
- [76] Moisés Felipe Silva et al. “Damage-Sensitive Feature Extraction with Stacked Autoencoders for Unsupervised Damage Detection”. In: *Structural Control and Health Monitoring* 28.5 (2021), e2714. ISSN: 1545-2263. DOI: 10.1002/stc.2714. (Visited on 10/24/2025).
- [77] Hoon Sohn et al. *A Review of Structural Health Monitoring Literature: 1996–2001*. Tech. rep. LA-13976-MS. Los Alamos National Laboratory, 2003.
- [78] Hoon Sohn et al. “A Review of Structural Health Monitoring Literature: 1996–2001”. In: *Los Alamos National Laboratory Report, LA-13976-MS* (2003).
- [79] Lingyu Sun et al. “Ultrasonic guided wave-based probabilistic diagnostic imaging method with Single-Path-Scattering sparse reconstruction for Multi-Damage detection in composite structures”. In: *Mechanical Systems and Signal Processing* 210 (2025).
- [80] Gyungmin Toh and Junhong Park. “Review of Vibration-Based Structural Health Monitoring Using Deep Learning”. en. In: *Applied Sciences* 10.5 (Jan. 2020), p. 1680. ISSN: 2076-3417. DOI: 10.3390/app10051680.
- [81] University College Dublin (UCD) et al. “Deep-Learning-based Damage Detection of an Experimentally Tested Full-Scale Wind Turbine Blade.” In: *e-Journal of Nondestructive Testing* 29.7 (July 2024). ISSN: 1435-4934. DOI: 10.58286/29852. (Visited on 10/30/2025).
- [82] A. Vaswani et al. “Attention is All You Need”. In: *Advances in Neural Information Processing Systems (NeurIPS)*. 2017, pp. 5998–6008. DOI: 10.48550/arXiv.1706.03762. URL: <https://arxiv.org/abs/1706.03762>.
- [83] Ilaria Venanzi et al. “Earthquake-Induced Damage Localization in an Historic Masonry Tower through Long-Term Dynamic Monitoring and FE Model Calibration”. In: *Bulletin of Earthquake Engineering* 18.5 (Mar. 2020), pp. 2247–2274. ISSN: 1570-761X, 1573-1456. DOI: 10.1007/s10518-019-00780-4. (Visited on 10/29/2025).
- [84] L. G. G. Villani, S. da Silva, and A. Cunha Jr. “Damage detection in uncertain nonlinear systems based on stochastic Volterra series”. In: *Mechanical Systems and Signal Processing* 125 (2019), pp. 288–310. DOI: 10.1016/j.ymssp.2018.06.012.
- [85] L. G. G. Villani et al. “Damage detection in an uncertain nonlinear beam based on stochastic Volterra series: An experimental application”. In: *Mechanical Systems and Signal Processing* 128 (2019), pp. 463–478. DOI: 10.1016/j.ymssp.2019.03.047.
- [86] Hao Wang et al. “Vibration-based structural health monitoring via deep learning techniques: A review”. In: *2023 International Conference on Cyber-Physical Social Intelligence (ICCSI)*. Oct. 2023, pp. 103–108. DOI: 10.1109/ICCSI58851.2023.10303900. URL: <https://ieeexplore.ieee.org/document/10303900>.

- [87] S. Wang, Q. Ren, and P. Qiao. “Iterative modal strain energy method for damage severity estimation using frequency measurements”. In: *Structural Control and Health Monitoring* 20.2 (2013), pp. 79–95.
- [88] Shuxun Wang and Qiang Zhang. “The Role of Geometric Nonlinearity in Structural Vibration”. In: *Communications in Nonlinear Science and Numerical Simulation* (2025).
- [89] Yi-Kwei Wen. “Method for random vibration of hysteretic systems”. In: *Journal of the engineering mechanics division* 102.2 (1976), pp. 249–263.
- [90] R. Wen and et al. “A Multi-horizon Quantile Recurrent Forecaster”. In: *arXiv preprint arXiv:1711.11053* (2017). URL: <https://arxiv.org/abs/1711.11053>.
- [91] S. Wernitz et al. “A New Open-Database Benchmark Structure for Vibration-Based Damage Detection Comprising a Lattice Tower Exposed to Realistic Conditions”. In: *Structural Control and Health Monitoring* 29.7 (2022), e2955. DOI: 10.1002/stc.2955. URL: <https://doi.org/10.1002/stc.2955>.
- [92] Keith Worden et al. “A review of nonlinear dynamics applications to structural health monitoring”. In: *Structural Control and Health Monitoring* 15.4 (2008), pp. 540–567. DOI: 10.1002/stc.234.
- [93] Yibing Xiang and Yongming Liu. “Efficient probabilistic methods for real-time fatigue damage prognosis”. In: (2010).
- [94] Yu Xin, Jun Li, and Hong Hao. “Damage Detection in Initially Nonlinear Structures based on Variational Mode Decomposition”. In: *International Journal of Structural Stability and Dynamics* 20.10 (2020), p. 2042009.
- [95] Guirong Yan et al. “A Novel Approach to Detecting Breathing-Fatigue Cracks Based on Dynamic Characteristics”. In: *Journal of Sound and Vibration* 332.2 (2013), pp. 407–422.
- [96] Yang Yu and Yongming Liu. “Physics-guided generative adversarial network for probabilistic structural system identification”. In: *Expert Systems with Applications* 239 (2024), p. 122339.
- [97] Sakib Ashraf Zargar and Fuh-Gwo Yuan. “Physics-Informed Deep Learning for Scattered Full Wavefield Reconstruction from a Sparse Set of Sensor Data for Impact Diagnosis in Structural Health Monitoring”. In: *Structural Health Monitoring* 23.5 (Sept. 2024), pp. 2963–2979. ISSN: 1475-9217. DOI: 10.1177/14759217231202547. (Visited on 12/02/2025).
- [98] Minte Zhang et al. “Damage identification of seismic-isolated structure based on CAE network using vibration monitoring data”. In: *Engineering Structures* 283 (2023), p. 115873.
- [99] Zheng Zhang et al. “Contact Acoustic Nonlinearity (CAN)-based Continuous Monitoring of Bolt Loosening: Hybrid Use of High-order Harmonics and Spectral Sidebands”. In: *Mechanical Systems and Signal Processing* 103 (2018), pp. 280–294.
- [100] Stefano Zucca. “Model Order Reduction for Nonlinear Vibrations of Structures”. PhD thesis. Politecnico di Torino, 2019.