XLVI CILAMCE
VITÓRIA - ES - BRAZIL - 2025
XLVI IBERO-LATIN AMERICAN CONGRESS ON COMPUTATIONAL METHODS IN ENGENEERING

Unsupervised Machine Learning-Based Approach for Damage Detection using Autoencoder Neural Networks

Leon V. Lobo[1], Raimundo A. B. Junior[1], Edilson M. L. Silva[2], Reyolando M. L. R. F. Brasil [1]

[1]*Polytechnic School, University of São Paulo*
Av. Prof. Almeida Prado, 83 - Jardim Universidade Pinheiros, 05508-070, São Paulo/São Paulo, Brazil
lvlobo@usp.br, afonso.barra@usp.br, reyolando.brasil@usp.br
[2]*Institute of Technology, Federal University of Pará*
R. Augusto Corrêa, 01 - Guamá, 66075-110, Pará/Belém, Brazil
edilson_morais@ufpa.br

Abstract. Damage detection is one of the main functions of structural health monitoring. Large structures such as bridges and overpasses require real-time monitoring to identify anomalies that could compromise their urban function. With this in mind, the present work proposes a Machine Learning approach based on autoencoder neural networks for damage detection. Since this approach is grounded in an unsupervised methodology, it does not require labeled data (which aligns with engineering practice) and is robust enough to overcome challenges such as the cost of modal identification and the need for knowledge about the structure's excitation source. It only requires signals as training data and can subsequently be applied in practice for real-time monitoring of special engineering structures. To assess the method's effectiveness, a numerical case study was conducted considering a 3% reduction in flexural stiffness to simulate structural damage. The autoencoder achieved strong performance metrics: Precision = 0.86, Recall = 1.00, and F1-Score = 0.92. These results indicate the model's high sensitivity to damage and low false-negative rate, even in subtle damage scenarios. This demonstrates the potential of the proposed unsupervised learning strategy as a low-cost and scalable tool for early damage detection in civil infrastructure.

Keywords: Damage Detection, Strucutural Health Monitoring, Machine Learning, Autoencoders.

1 Introduction

Structural Health Monitoring (SHM) is a critical process for assessing the integrity of civil, mechanical, and aerospace structures over their operational lifespan. Traditional inspection methods, such as visual assessments and vibration-based techniques, often require manual intervention and are limited in their ability to detect early-stage or hidden damage. With the increasing complexity of modern infrastructure and the growing demand for safety and cost-effective maintenance [1], there is a pressing need for automated, real-time damage detection systems.

Machine Learning (ML) has emerged as a powerful tool in SHM, enabling data-driven approaches for damage identification, localization, and severity assessment. Supervised learning methods, such as support vector machines (SVMs) and convolutional neural networks (CNNs), have shown success in classifying damage when labeled datasets are available [2].

However, obtaining labeled damage data—especially for rare or catastrophic failure events—is often impractical in real-world scenarios. Environmental variability, noise, and operational changes further complicate the training of supervised models, as they require extensive datasets covering all possible structural conditions.

Additionally, many traditional vibration-based methods rely on known input-output techniques, such as modal analysis or frequency response functions, which require controlled excitation (e.g., impact hammer tests or shakers). These methods are not only expensive to implement but also impractical for continuous monitoring[3], as they disrupt normal operations. The need for precise input measurements further limits their scalability, particularly in distributed infrastructure like bridges or offshore platforms where controlled excitation is infeasible.

To overcome these limitations, unsupervised learning techniques have gained attention in SHM. Unlike supervised methods, unsupervised approaches do not rely on pre-labeled damage data but instead learn the underlying patterns of healthy structural behavior. Autoencoders, a type of neural network trained to reconstruct input data with minimal error, are particularly well-suited for this task.

By modeling the normal operational state of a structure, autoencoders can detect anomalies—such as cracks,

loosened connections, or material degradation—through deviations in reconstruction error [4]. This makes them an ideal candidate for unsupervised damage detection in scenarios where labeled damage examples are scarce or unavailable.

2 Methodology

2.1 Numerical model

The beam structure is discretized using Euler-Bernoulli beam elements [5], where each element has two nodes with two degrees of freedom per node (transverse displacement w and rotation θ), as shown in Fig. 1. The element stiffness matrix $\mathbf{K}_e$ and mass matrix $\mathbf{M}_e$ are defined through the shape functions $\mathbf{N}(x)$ as given in Eqs. (1) and (2) respectively:

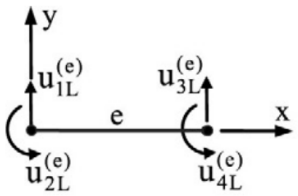


Figure 1. Degrees of freedom of the Euler-Bernoulli beam element [5]

$$\mathbf{K}_e = EI \int_0^{L_e} \left(\frac{\partial^2 \mathbf{N}(x)}{\partial x^2} \right)^T \left(\frac{\partial^2 \mathbf{N}(x)}{\partial x^2} \right) dx \tag{1}$$

$$\mathbf{M}_e = \rho A \int_0^{L_e} \mathbf{N}^T(x) \mathbf{N}(x) dx \tag{2}$$

where EI is the flexural rigidity, ρ is material density, A is cross-sectional area, and L_e is element length. The global matrices $\mathbf{K}$ and $\mathbf{M}$ are assembled through standard finite element procedures. The system damping is modeled as Rayleigh damping in Eq. (3):

$$\mathbf{C} = \alpha \mathbf{M} + \beta \mathbf{K} \tag{3}$$

where α and β coefficients are determined to achieve a critical damping ratio ξ for all significant modes (first and third). The solution of the equation of motion use Newmark's (constant average acceleration scheme). The finite element model consists of a beam discretized into 4 elements with the material (as showed in Fig. 2), geometric, and computational parameters summarized in Table 1.

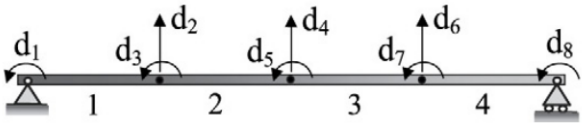


Figure 2. Beam model used in the numerical simulations [5]

The consistent mass and stiffness matrices are computed using the formulations presented in Eqs. (2) and (1), with parameters from Table 1. The selected Rayleigh coefficients α and β (Table 1) provide 2% damping ratio across all significant vibration modes. The time integration is performed using Newmark's method employing the time step value listed in Table 1.

The beam is harmonically excited at the 2nd DOF (vertical displacement at 1/4 span, Node 2) with force amplitude $F = 1000$ N and angular frequency $\omega = 70$ rad/s, as in Eq. (4).

$$f(t) = 1000 \cos(70\, t) \tag{4}$$

The dynamic response is measured at mid-span (4th DOF, Node 3) with simulated Gaussian noise, as in Eq. (5).

Table 1. Numerical model parameters

Parameter	Symbol	Value
Element length	L_e	2.0 m
Cross section (width $\times$ height)	$b\times\ h$	0.2 m $\times$ 0.4 m
Elasticity modulus	E	30 GPa
Mass density	ρ	2500 kg/m^3
Time step (Newmark)	Δt	0.001 s
Rayleigh damping coefficient	α	2.225174
Rayleigh damping coefficient	β	0.00006379
Number of elements	N_e	4

$$\ddot{u}_{\text{noisy}} = \ddot{u} + 5\,\sigma_{\ddot{u}}\,\mathcal{N}(0,1) \tag{5}$$

This noise model corresponds to a signal-to-noise ratio (SNR) of 26 dB, representing challenging measurement conditions typical in real-world structural monitoring. The noise is applied only to the vertical acceleration at mid-span, simulating a single accelerometer installation, as showed in the Fig. 3. To avoid the transient regime, it was decided to select the points after the first second of vibrations for the next steps

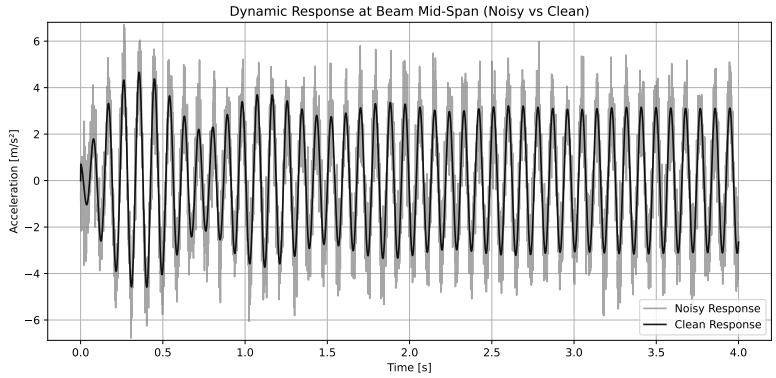


Figure 3. Accelerometer acquisition simulation

2.2 Autoencoders neural networks

Autoencoders are unsupervised neural networks designed to learn efficient representations of input data through compression and reconstruction [6]. As shown in Fig. 4, the network first encodes the input $\mathbf{x}$ into a latent-space representation $\mathbf{h}$ through the transformation given in Eq. (6), where $\theta = \{\mathbf{W}_e, \mathbf{b}_e\}$ are the learnable weights and biases of the encoder, and σ denotes the activation functions. The latent space acts as an information bottleneck, typically with $dim(\mathbf{h}) \ll dim(\mathbf{x})$, forcing the network to learn the most salient features of the input data.

$$\mathbf{h} = f_\theta(\mathbf{x}) = \sigma(\mathbf{W}_e\mathbf{x} + \mathbf{b}_e) \tag{6}$$

The decoder then reconstructs the input from this compressed representation using the transformation in Eq. (7), where $\phi = \{\mathbf{W}_d, \mathbf{b}_d\}$ are the decoder's learnable parameters. The complete architecture is trained end-to-end to minimize the reconstruction error defined by the Mean Absolute Error (MAE) in Eq. (8).

$$\hat{\mathbf{x}} = g_\phi(\mathbf{h}) = \sigma(\mathbf{W}_d\mathbf{h} + \mathbf{b}_d) \tag{7}$$

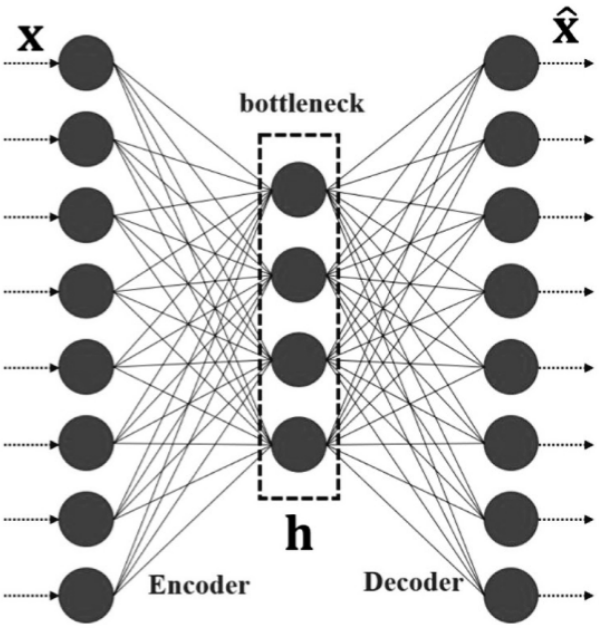


Figure 4. Autoencoder architecture with encoder and decoder components [6]

$$\mathcal{L}(\theta, \phi) = \|\mathbf{x} - \hat{\mathbf{x}}\|_2^2 = \frac{1}{N} \sum_{i=1}^{N} |\mathbf{x}_i - g_\phi(f_\theta(\mathbf{x}_i))| \tag{8}$$

where θ and ϕ denote encoder and decoder parameters respectively, and N is the batch size. This loss function provides robustness to outliers in the vibration data compared to traditional MSE loss.

For structural damage detection applications, the autoencoder is trained exclusively on vibration signals from healthy structural states, allowing it to learn the normal operational patterns. During deployment, damage is identified when the reconstruction error of test signals is under a predetermined threshold τ, as specified in Eq. (9). This approach provides several key advantages for structural health monitoring: it eliminates the need for labeled damage data during training, demonstrates robustness to operational variability when properly regularized, and maintains computational efficiency suitable for real-time monitoring applications [7].

$$\|\mathbf{x}_{\text{test}} - \hat{\mathbf{x}}_{\text{test}}\|_2^2 < \tau \tag{9}$$

The proposed autoencoder architecture implements a symmetric encoder-decoder structure with progressive dimensionality reduction, as detailed in Table 2. This bottleneck layer forces the network to learn the most essential features of the healthy vibration patterns, achieving a 17.5:1 compression ratio. Training parameters are systematically configured as shown in Table 3. Early stopping monitors the validation loss with a patience of 5 epochs, reverting to the best weights when no improvement occurs. This prevents overfitting while maintaining optimal performance. A 20% validation split provides robust generalization assessment during training.

Table 2. Autoencoder architecture details

Component	Specification
Encoder Layers	Dense(32, ReLU) $\rightarrow$ Dense(16, ReLU) $\rightarrow$ Dense(8, ReLU)
Decoder Layers	Dense(16, ReLU) $\rightarrow$ Dense(32, ReLU) $\rightarrow$ Dense(140, Sigmoid)
Latent Space Dimension	8
Input/Output Dimension	140

The reconstruction quality is quantified through the MAE between input and output signals, computed as shown in Eq. (10). This metric provides consistent performance measurement across all 140 signal points while being robust to outlier values. The choice of MAE over MSE aligns with the need for stable gradient signals during backpropagation through the deep network architecture.

Table 3. Training parameters

Parameter	Value
Epochs	100
Batch Size	512
Optimizer	Adam
Learning Rate	Default (0.001)
Early Stopping	Patience=5, Restore Best Weights
Validation Split	20%
Loss Function	Mean Absolute Error (MAE)

$$\text{MAE} = \frac{1}{140}\sum_{i=1}^{140}|x_i - \hat{x}_i| \tag{10}$$

3 Results and Discussion

3.1 Database construction

The dataset construction started with 225 healthy simulations, each generating 2000 time-domain points at the beam's mid-span vertical DOF using a 1 ms Newmark time step. Although 3000 points were simulated, the first 1000 were discarded to remove the damping-influenced transient, focusing the neural network on steady-state patterns.

For the damaged dataset, 148 simulations were performed in four groups of 37, each with a 3% flexural stiffness (EI) reduction applied to one of the four beam elements, keeping the same length and parameters as the healthy signals.

The 337 observations of 2000 points were downsampled to 140 points to capture up to two signal repetitions due to the 70 Hz excitation frequency, reducing computational load while preserving key response features. Finally, the dataset contained 5222 normalized samples (min-max scaling) labeled as healthy or damaged, forming a robust basis for model training and evaluation.

3.2 Signal reconstruction

Only healthy samples were used for training and validation to ensure the model learned the healthy structural signature, consistent with unsupervised anomaly detection. Early stopping was triggered at the 43$^{\text{rd}}$ epoch (Fig. 5), indicating optimal performance without overfitting.

Fig.6 compares reconstructions of healthy and damaged signals by the trained autoencoder. The model effectively learned the harmonic oscillation pattern, acting as a denoiser by filtering noise and emphasizing essential dynamic features critical for damage detection.

3.3 Model evaluation

The MAE histogram profile (Fig. 7) shows healthy reconstructions forming a dominant distribution (blue), while anomalous ones spread into smaller distributions with different means, indicating effective pattern differentiation by the autoencoder.

The classification threshold was set one standard deviation below the healthy mean MAE, providing good separation despite some anomalies remaining embedded within the healthy distribution.

Precision (86%), recall (100%), and F1-score (92%) were computed to quantify detection performance, with precision measuring correct damaged identifications, recall indicating damage coverage, and F1-score balancing both. The confusion matrix in Fig. 8 illustrates these results.

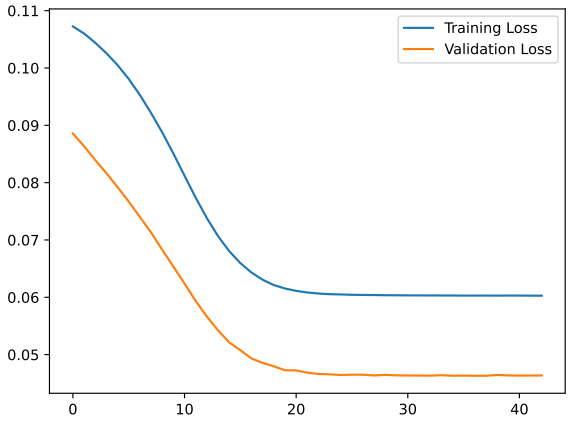


Figure 5. Loss convergence in the training

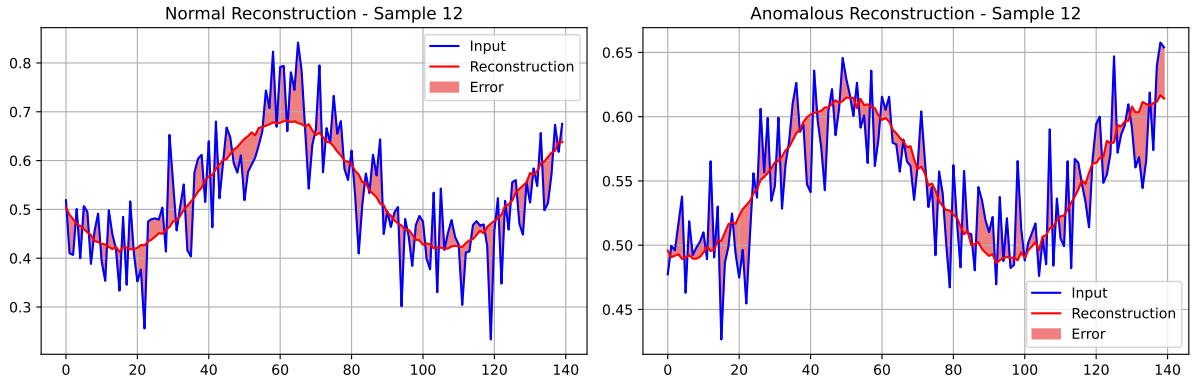


Figure 6. Comparison between normal and anomalous reconstruction

4 Conclusion

This work presented an unsupervised learning strategy for vibration-based damage detection in structural health monitoring, highlighting its practical advantages in real-world scenarios. By relying solely on healthy-state data for training, the approach eliminates the need for labeled damage datasets, which are often unavailable or costly to obtain in engineering practice. This unsupervised nature makes the method adaptable to a wide range of structures and operational conditions, addressing a key challenge in the deployment of monitoring systems at scale.

A notable finding is the ability of autoencoder neural networks to effectively reconstruct structural responses even in the presence of measurement noise. In practice, field-acquired vibration signals are inherently affected by operational and environmental factors, leading to noise-dominated datasets. The denoising behavior observed in the reconstructions demonstrates the robustness of the proposed approach, as the network was able to extract the underlying dynamic signature of the healthy structure while filtering out irrelevant fluctuations.

The evaluation metrics obtained — precision of 0.86, recall of 1.00, and F1-score of 0.92 — confirm the effectiveness of the method in detecting structural anomalies. The high recall highlights the model's sensitivity to damage, minimizing the risk of missed detections, while the balanced F1-score reflects strong overall classification performance. These results emphasize the strong potential of autoencoder-based unsupervised techniques as low-cost, scalable, and reliable tools for early-stage damage detection in vibration-based structural health monitoring applications.

Acknowledgements. The authors thank the USP Polytechnic School, CNPq and the CAPES for their funding and grant in the development of this work.

Authorship statement. The authors hereby confirm that they are the sole liable persons responsible for the au-

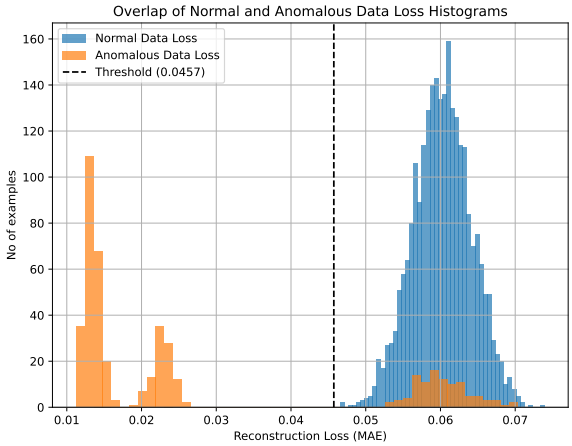


Figure 7. MAE histogram profile for normal and anomalous samples

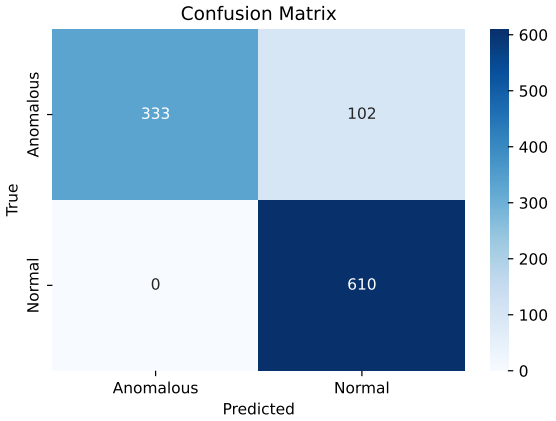


Figure 8. Confusion matrix for the evaluated model

thorship of this work, and that all material that has been herein included as part of the present paper is either the property (and authorship) of the authors, or has the permission of the owners to be included here.

References

[1] S. Yuan, G. Reniers, M. Yang, and Y. Bai. Cost-effective maintenance of safety and security barriers in the chemical process industries via genetic algorithm. *Process Safety and Environmental Protection*, vol. 170, pp. 356–371, 2023.

[2] H. Sarmadi and A. Entezami. Application of supervised learning to validation of damage detection. *Archive of Applied Mechanics*, vol. 91, n. 1, pp. 393–410, 2021.

[3] Y. Kaya and E. Safak. Real-time analysis and interpretation of continuous data from structural health monitoring (shm) systems. *Bulletin of Earthquake Engineering*, vol. 13, pp. 917–934, 2015.

[4] E. M. Coraça, J. V. Ferreira, and E. G. Nóbrega. An unsupervised structural health monitoring framework based on variational autoencoders and hidden markov models. *Reliability Engineering & System Safety*, vol. 231, pp. 109025, 2023.

[5] H. L. Soriano. *Introdução à dinâmica das estruturas*. Elsevier Rio de Janeiro, 2014.

[6] J. Mao, H. Wang, and B. F. Spencer Jr. Toward data anomaly detection for automated structural health monitoring: Exploiting generative adversarial nets and autoencoders. *Structural Health Monitoring*, vol. 20, n. 4, pp. 1609–1626, 2021.

[7] B. Tefera, A. Zekaria, and A. Gebre. Challenges in applying vibration-based damage detection to highway bridge structures. *Asian Journal of Civil Engineering*, vol. 24, n. 6, pp. 1875–1894, 2023.