POLITECNICO MILANO 1863

SCUOLA DI INGEGNERIA INDUSTRIALE
E DELL'INFORMAZIONE

Executive Summary of the Thesis

Machine learning and deep learning algorithms for anomaly detection

Laurea Magistrale in Mechanical Engineering - Ingegneria Meccanica

Author: Gianluca Bombaci

Advisor: Prof. Simone Cinquemani

Co-advisors: Ing. Francesco Morgan Bono, Ing. Luca Radicioni, Ing. Claudio Somaschini

Academic year: 2022-2023

1. Introduction

In recent years, there has been a growing interest in the field of Structural Health Monitoring (SHM) for civil structures like buildings and bridges. Indeed, structures are constantly exposed to various environmental factors that can potentially impact their structural integrity. To address these challenges, effective damage detection techniques are needed and which are encompassed in SHM. In particular, dynamic response measurements are analyzed, employing feature extraction algorithms, and apply statistical analysis techniques. Indeed, damage in structures can be defined as changes that affect their current or future dynamic performance. Detecting such damage requires a comparison between two different states of the system, in which one represents the nominal condition, often corresponding to an undamaged state of the structure. Visual inspections are a common method for locating damage. However, they can be imprecise, unreliable, and time-consuming. In contrast, vibration-based techniques have proven to offer a more dependable approach to assessing a structure's health. Given the high quantity of data generated by vibration monitoring, deep learning has emerged as a powerful tool. It can identify meaningful features within large datasets using multiple processing layers. Generally, deep-learning models for damage detection rely on supervised learning strategies, where data from both healthy and damaged structural conditions serve as training sets to create functions capable of mapping new input data. However, obtaining the data of the input source that excites the structure can often be prohibitive, leading to issues of robustness and convergence in machine learning techniques. Moreover, collecting data from a damaged state of the structure can be challenging. To overcome these limitations, Convolutional Autoencoders (CAEs) have been employed to detect damages based solely on raw vibration data from healthy structures. Additionally, the adoption of Physics-Informed Neural Networks (PINNs) allows for the incorporation of physical laws governing the time-dependent dynamics of the structure. The primary motivation for using PINNs in anomaly detection is to mitigate the challenges of acquiring abnormal data in physical systems and the substantial volume of data required for training neural networks. With PINNs physical laws and equations of motion are integrated with neural networks, narrowing the search space for network parameters and reducing the need for extensive training data. This parameter space compression represents a significant advantage of PINNs over traditional neural network approaches, making them an attractive option for anomaly detection and location. This Master's Thesis work serves as an extension of two research papers that investigate the comparison between an unsupervised deep-learning algorithm and a PINN for structural monitoring, using only vibration data acquired from the healthy state as the training set[1, 2]. Both neural networks are evaluated on a four-story building using acceleration data obtained from accelerometers placed on each floor. The primary objective is to confirm the higher potential of AI methods with respect to conventional methods and to demonstrate the superior capability of PINNs in detecting structural damages compared to conventional unsupervised neural networks.

2. Artificial Intelligence and Machine Learning

Since the main concern of this Master's Thesis is the implementation of an Artificial Intelligence (AI) based algorithm, a general comprehension of the methods we are dealing with is needed. AI is a multidisciplinary field of computer science focused on creating systems and machines capable of performing tasks that typically require human intelligence. These tasks encompass a wide range of activities, including problem-solving, learning, reasoning, perception, understanding natural language, and interacting with the environment. Artificial intelligence does not comprehends only learning-based approaches, like machine learning and deep learning, but also approaches oriented to replicate the human reasoning process.

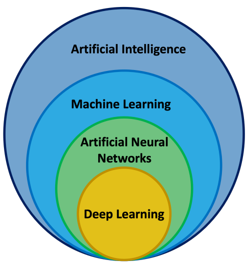


Figure 1: Hierarchical representation of artificial intelligence.

AI has shown a high potential to revolutionize industries, improve efficiency, enhance decision-making, and create new opportunities for innovation and advancement. Its impact on society and various domains continues to grow, making it a pivotal field of study and research.

2.1. Machine Learning

In classical programming, a large set of explicit rules for manipulating knowledge are employed. This approach, known ad symbolic AI is suitable to solve well-defined problems and relies on the idea that humans input rules and data to be processed to obtain the required outputs. However, this paradigm fails when dealing with complex, fuzzy problems as image classification or speech recognition. For this reason, Machine learning was introduced. ML is a sub-method of artificial intelligence that aims to learn from data and make predictions or decisions on future scenarios. Indeed, the system through a new additional phase, called training, is able to construct the knowledge about the case under study. Data become a key point. Indeed, they are used for both training and testing ML algorithms, aiming to be transformed in more meaningfully representations for the given task. However, these representation are searched in a restricted space of predefined operation, called hypothesis space. To summarize, a ML algorithm to work properly needs:

- Input data;
- Examples of the expected outputs;
- A measure of how the algorithm is performing, in order to measure the distance the algorithm's output and the expected one. This index works as a feedback signal, helping the algorithm to adjust itself through the *learning* step.

It must be said that the ability to work with large dataset through the training phase represent also one of the biggest limitations of ML. Indeed, it is impossible to extract generalized rules of data analysis based on information acquired and elaborated in the past. If the input dataset changes a re-training would be necessary.

2.2. Deep Learning

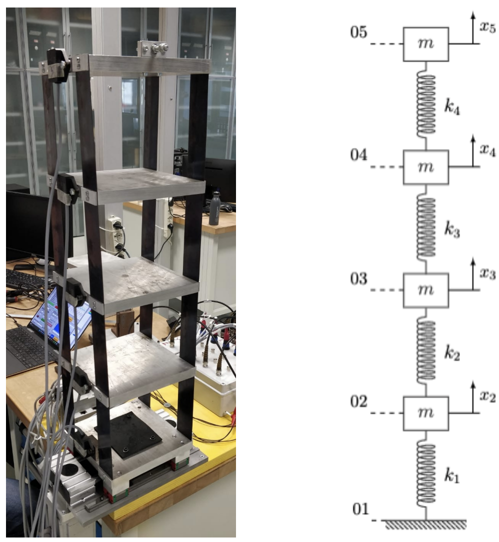
In recent years, Deep Learning (DL), a sub-field of machine learning, has garnered significant attention due to its remarkable achievements in various domains. On the contrary of what could be thought, in deep learning, the term "deep" does not refer to a class of ML-algorithms with higher capability of understanding but to a class of artificial neural networks working with several hidden layers of increasingly meaningful representations. Indeed, while the other approaches focus on learning by means one or two layers of representations, and for this reason called shallow methods, deep learning relies on tens or hundreds successive layers of data elaboration. Being inspired by the human brain's structure and function, each layer used by deep learning algorithms is composed by neurons, linked to all neurons present in the previous and successive layers. Deep learning networks can be thought as multistage information filtering with the objective to map inputs into targets. Each layer, which acts like a filter, is characterized by a weight. These weights are the way in which the algorithm improve itself to correctly associate inputs to targets during the learning phase. The first layer of an Artificial Neural Network (ANN) is labeled as input layer. The dimension of this layer matches the dimension of the data entering the network. Instead, the last layer of an ANN, labeled as output layer, has the dimension of what it is expected from the layer (i.e. the fixed targets). These deep architectures enable the hierarchical extraction of features from raw data, allowing deep learning models to capture intricate patterns in data, making them highly effective in tasks such as image recognition and speech recognition. The training phase of deep learning models involves an algorithm called backpropagation, where the model adjusts its internal parameters (weights and biases) to minimize the difference between predicted and actual outputs. This iterative optimization process relies on a mathematical technique known as gradient descent. Despite its numerous successes, it must be said that deep learning is not without challenges. Indeed, training deep learning models can be computationally intensive and may require large datasets to generalize well. Moreover, one of the key challenges is interpretability, as deep models are often regarded as "black boxes", making it difficult to understand their decision-making processes.

As already said, the main ingredient of a successful neural network are the layers, that can be thought as filter for data, able to extract representation that, hopefully, would be more meaningful for the problem being studied. So, a deep learning machine is like a series of increasingly finer data filters. However, to properly train a ML algorithm three more elements are needed:

- **Loss function** (also objective function), which represents the quantity to be minimized during the training phase and which defines the feedback signal;
- **Optimizer**, which represents the method chosen for the updating of the network based on the feedback signal coming from the loss function;
- **Metrics** for training and testing. They allow to quantify and study the performances of the algorithm. Indeed, they are used to monitor and measure the performance of a model during training and testing.

3. System description and model

The system object of this work is a multi-storey building shown in Figure 2 (a).



(a) A photo of the real system.

(b) Lumped mass model of the system.

Figure 2: Real system and lumped mass model.

Five aluminum plates, connected by steel laminas, respectively model the storeys and the pillars of the building [3]. The reason behind using a lumped mass approach in modeling the system is the substantial difference in mass between each storey and the laminas. This approach simplifies the system into four degrees of freedom, represented by four masses connected by springs, as depicted in Figure 2 (b). The equations of motion that describe the dynamics of the tested building are derived, as presented in equation1.

$$[M]\,\underline{\ddot{x}} + [C]\,\underline{\dot{x}} + [K]\,\underline{x} = \underline{0} \qquad (1)$$

For the complete analysis of each matrix composing the equation is suggested to refer to the complete work. However, it must be highlighted that in this Master's Thesis the hypothesis of a clamped-clamped beam is used to calculate the stiffness matrix of the system. Moreover, also the influence of the weight of each storey on the transversal stiffness was considered.

3.1. Experimental Campaign

The techniques considered in this thesis employ the variations in vibration measurements between the structure in its nominal state, referred to as "healthy," and a state with "damage" as an indicator of potential damage. Hence, the experimental campaign carried out on the tested structure is designed to collect raw data for both the "healthy" and "damaged" states. The experimental setup for both scenarios includes:

- four TE triaxial capacitive MEMS accelerometers, one per each storey;
- a PCB Piezotronics impact hammer;
- a National Instruments c-DAQ.

The structure is stimulated using an impact hammer, and the transverse vibrations are captured. In the "healthy" scenario, a total of 1000 data records, each lasting 70 seconds and sampled at a rate of 128Hz, are recorded. After, the Frequency Response Functions (FRFs) are calculated for each accelerometer and the natural frequencies and mode shapes are deduced using the Experimental Modal Analysis (EMA) technique. Specifically, the natural frequencies of the system are presented in Table 1 for both the numerical and experimental models.

Table 1: Natural frequencies for both the numerical and the experimental model.

Mode	Numerical model [Hz]	Experimental model [Hz]
1	0.79	0.75
2	2.51	2.41
3	3.88	3.74
4	5.01	5.04

It's crucial to emphasize that the natural frequencies and mode shapes obtained for the "healthy" structure will serve as a reference for the vibration-based damage detection method. This method, in turn, will be used as a benchmark to evaluate the performance of the two machine learning algorithms.

When examining scenarios involving structural damage, it's essential to bear in mind that internal structural issues are typically not caused by material loss and, consequently, do not result in changes in mass. Instead, they often stem from alterations in geometry or material properties that impact one or more elements within the stiffness matrix. Therefore, in the "damaged" condition, the time histories are acquired by altering only the stiffness values of the laminas. Specifically, six distinct sets of laminas, all featuring the same cross-sectional dimensions but varying in length as indicated in Table 2, are employed to reduce the stiffness of the springs connecting consecutive floors. These adjustments result in stiffness values ranging from 10% to 60% of the nominal value.

Table 2: Lengths of the set of laminas used to reproduce a damage in the structure.

Damage Percentage	Length
0%	180.0 mm
−10%	186.5 mm
−20%	194.0 mm
−30%	203.0 mm
−40%	213.5 mm
−50%	227.0 mm
−60%	244.0 mm

In total, a dataset comprising 240 time records, each lasting 70 seconds and sampled at a rate of 128Hz, is collected. This dataset encompasses 10 records for every possible combination of damage extent (represented by the type of lamina) and damage location (across four floors).

4. Network architectures and training

The core objective of this Master's Thesis work is to compare the ability to detect structural damage between a physics-informed neural network and a purely data-driven neural network. Furthermore, as already said, conventional vibration-based techniques, which rely on the analysis of changes in both natural frequencies and vibration modes of the structure, are used as a reference to evaluate the advantages of employing machine learning algorithms over traditional methods.

4.1. Training and test

Following the pre-processing phase and the split of the dataset, described in detail in the main work, the training set is employed to train the autoencoder model illustrated in Figure 3. Notably, the training phase for the PINN-CAE and DD-CAE models differs due to the custom loss function implemented for the former. During the CAE training process, 200 epochs are considered, with Mean Absolute Error (MAE) serving as the loss function. Additionally, a callback is applied to monitor the validation loss. MAE is evaluated separately for each accelerometer to assess the reconstruction error generated by the trained model when predicting the test set. The maximum MAE values observed across the entire test set are established as thresholds for anomaly detection. As mentioned earlier, MAE values exceeding these thresholds are indicative of time histories representing the damaged structure.

4.2. Autoencoder

Both machine learning algorithms, referred to as PINN-CAE and DD-CAE, share a common neural network architecture based on a convolutional autoencoder. Autoencoders are self-supervised learning techniques that aim to reconstruct input data at the output with minimal distortion after a series of transformations and data compression steps. They are widely used for denoising, data compression, and high-dimensional data visualization. Convolutional Neural Networks (CNNs), upon which Convolutional Autoencoders (CAEs) are built, are a subset of Artificial Neural Networks (ANNs) that leverage convolution operations, offering advantages such as reduced parameter connections and faster convergence due to dimension reduction. This architecture is employed to remove irrelevant features while preserving essential information. For the description of each layer composing the model used in this work is possible to refer to the extended Master Thesis work.

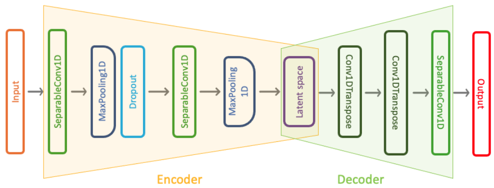


Figure 3: Autoencoder model.

4.3. Physic-informed neural network

At this point, it should be clear that training deep neural networks often necessitates access to extensive datasets, which can be challenging to obtain, especially for pre-existing structures or damaged scenarios. Physics-informed neural networks offer a potential solution to this limitation. These networks can be trained using additional information derived from the underlying physical principles governing the dynamic behavior of the system. This approach integrates both data and mathematical models, even in situations where the models may not be fully understood, are subject to uncertainty, or involve high-dimensional parameters. To facilitate the training of PINNs, conventional loss functions are not suitable. Instead, custom loss functions are essential. Specifically, these custom loss functions should incorporate the physical laws that govern the dynamic response of the system, thereby constraining the space of allowable solutions during autoencoder training. The outline of the proposed custom loss function is depicted in Figure 4.

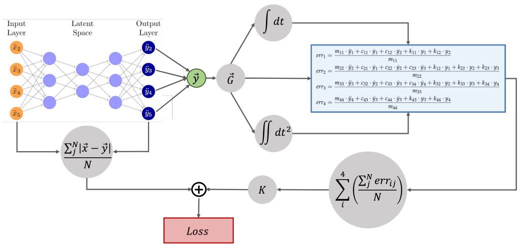


Figure 4: Custom loss scheme.

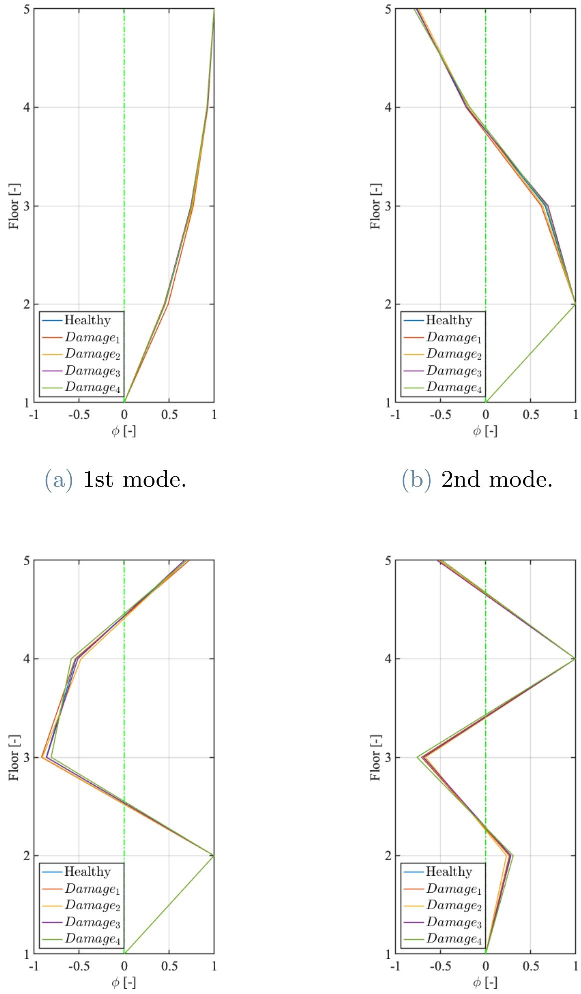
The signals denoted as y_i, which are the outcomes of the autoencoder, represent the reconstructed time histories of the scaled vibration data acquired during the healthy scenario. These data, in turn, depict the system's response to the input force applied by the hammer. However, it can be assumed that after a certain period, set at 10 seconds in this case, the transient behavior resulting from the forced motion of the system completely diminishes. Consequently, the remaining portion of each reconstructed time history should exhibit low error when compared to the previously presented set of Ordinary Differential Equations (ODEs) in Equation 1. For this reason, it's reasonable to expect a greater error when considering time histories originating from the damaged scenario. As a result, after appropriate time-domain integration, the error functions shown in Figure 4 are computed for each time instant. Within these error functions, y_i, $\dot{y}_i$, and $\ddot{y}_i$ corresponding to the displacement, velocity, and acceleration signals are obtained through the autoencoder's reconstruction. Subsequently, the absolute values of their mean values are calculated and collectively form the physical component of the custom loss function. At the end, the obtained custom loss function considering also the loss accounting for the data-driven part is:

$$\begin{aligned}\mathcal{L} = K \cdot [&\frac{\sum_{i=1}^{n} |err_{1,i}|}{n} + \frac{\sum_{i=1}^{n} |err_{2,i}|}{n} \\ &+ \frac{\sum_{i=1}^{n} |err_{3,i}|}{n} + \frac{\sum_{i=1}^{n} |err_{4,i}|}{n}] \\ &+ MAE\end{aligned} \tag{2}$$

where K is a constant to express the physical part of the custom loss in an adimensional form.

5. Results

The dataset comprising 240 damaged records is initially subjected to analysis through Experimental Modal Analysis (EMA). However, it is observed that natural frequencies and mode shapes exhibit only slight variations in response to damages of 10%, as depicted in Figure 5.



(a) 1st mode. (b) 2nd mode.

(c) 3rd mode. (d) 4th mode.

Figure 5: Vibration modes of structure for 10% reduction of the stiffness value and for different positions of the damage.

In the case of more significant damages noticeable differences in natural frequencies and mode shapes become evident. However, it's worth noting that pinpointing the precise location of the damage is not straightforward.

On the other hand, the anomaly dataset, preprocessed using the same methods used for the training set, is fed into both the PINN-CAE and

DD-CAE models. The Mean Absolute Error (MAE) values are computed for each anomaly record and for each channel (corresponding to different accelerometers). These MAE values are then compared to the previously established MAE test thresholds. Any record with a loss exceeding the threshold is classified as an anomaly. Both architectures successfully identify all the considered time histories as anomalies. This outcome confirms the anticipated higher precision of data-driven algorithms in detecting structural damage compared to conventional methods. Notably, this difference becomes particularly evident when the extent of the damage is relatively low.

Moreover, for each detected anomaly, the channel (corresponding to the accelerometer position, i.e., the floor) with the highest MAE loss is selected as the predicted damage location. This predicted position is then compared to the known real damage location. Subsequently, an accuracy indicator is evaluated for both of the considered algorithms, with damages ranging from 10% to 60% serving as the reference point.

$$\mathcal{A} = \frac{n_d}{n_{tot}} \times 100 \tag{3}$$

In this equation, where n_d represents the count of anomalies with a damage extent equal to or greater than the reference value, and where the model accurately identifies their positions, while n_{tot} stands for the total number of anomalies with a damage extent equal to or greater than the reference value. The outcomes of these evaluations are presented in Table 3.

Table 3: Anomalies detection rates as function of the damage percentage for both PINN-CAE and DD-CAE.

Damage Percentage	Accuracy $\mathcal{A}$ DD-CAE	Accuracy $\mathcal{A}$ PINN-CAE
-10%	33.19%	79.43%
-20%	40.20%	82.81%
-30%	52.24%	87.22%
-40%	65.11%	92.03%
-50%	84.61%	100%
-60%	100%	100%

It is possible to conclude that the PINN outperforms the results obtained with the purely data-driven approach, as expected.

6. Conclusions

The main focus of this Master's Thesis work is to assess the accuracy of detecting structural damages using two distinct machine learning algorithms: a physics-informed convolutional autoencoder (PINN-CAE) and a purely data-driven convolutional autoencoder (DD-CAE). The evaluation is based on raw data obtained from experimental acquisitions conducted on a four-storey building, and both algorithms share the same structural architecture. The Mean Absolute Error (MAE) of the reconstruction is employed as an indicator to identify anomalous records.

Both the PINN-CAE and the DD-CAE outperform conventional vibration-based methods in their ability to detect structural damages and pinpoint their locations. They successfully identify all anomalous time histories and exhibit a high level of precision in detecting structural changes. Notably, the physics-informed network demonstrates greater accuracy in locating damage compared to the data-driven approach, especially for lower levels of damage severity. This emphasizes the significant potential of combining a data-driven architecture with information derived from the physical model of the studied system.

Future developments of this research will involve altering the mass of the system and utilizing the model in conjunction with neural networks to detect anomalies in more complex structures such as bridges or viaducts. Analytically representing such structures can be challenging, but a numerical model based on simulation methods like the Finite Element Method can be employed. The expansion of this work by considering a numerical model of more complex system will be developed in future by the authors.

References

[1] Francesco Morgan Bono, Luca Radicioni, Gianluca Bombaci, Claudio Somaschini, and Simone Cinquemani. An approach based on convolutional autoencoder for detecting damage location in a mechanical system. In *NDE 4.0, Predictive Maintenance, Communication, and Energy Systems: The Digital Transformation of NDE*, volume 12489, pages 93–100. SPIE, 2023.

[2] Francesco Morgan Bono, Luca Radicioni, Simone Cinquemani, and Gianluca Bombaci. A comparison of deep learning algorithms for anomaly detection in discrete mechanical systems. *Applied Sciences*, 13(9):5683, 2023.

[3] S. Di Carlo, L. Benedetti, and E. Di Gialleonardo. Teaching by active learning: A laboratory experience on fundamentals of vibrations. *International Journal of Mechanical Engineering Education*, 50(4):869–882, 2022.