



RESEARCH

# Deep learning architectures for data-driven damage detection in nonlinear dynamic systems under random vibrations

Harrish Joseph · Giuseppe Quaranta ·  
Biagio Carboni · Walter Lacarbonara

Received: 6 July 2024 / Accepted: 30 August 2024  
© The Author(s) 2024

**Abstract** The primary goal of structural health monitoring is to detect damage at its onset before it reaches a critical level. In the present work an in-depth investigation addresses deep learning applied to data-driven damage detection in nonlinear dynamic systems. In particular, autoencoders and generative adversarial networks are implemented leveraging on 1D convolutional neural networks. The onset of damage is detected in the investigated nonlinear dynamic systems by exciting random vibrations of varying intensity, without prior knowledge of the system or the excitation and in unsupervised manner. The comprehensive numerical study is conducted on dynamic systems exhibiting different types of nonlinear behavior. An experimental application related to a magneto-elastic nonlinear system is also presented to corroborate the conclusions.

**Keywords** Autoencoder · Convolutional neural network · Damage detection · Deep learning · Generative adversarial network · Structural health monitoring

## 1 Introduction

Damage is any unintended alteration in material and/or geometric properties that may impact a structural system, including its boundary and connectivity conditions, leading to detrimental effects on both current and future performance (e.g., [63]). By leveraging on sensing, data acquisition and data transfer systems, structural health monitoring (SHM) techniques are aimed at identifying damage by means of suitable processing and analysis methods [26]. Carrying out a complete damage identification procedure is a multi-level process that encompasses damage detection, damage localization, damage quantification, and residual life prediction [52]. Damage detection can be possibly carried out by means of pure data-driven methods, and it is especially important to prevent costly maintenance plans or, even worse, important social, economic and environmental consequences. Early detection of damage is crucial to optimize proactive maintenance interventions before the onset of a critical system state. Vibration-based methods are particularly attractive for this task because they can detect damage analyzing the system dynamic response, also when damage is not severe enough to be visually apparent [34]. They are very common because of the availability of low-cost sensing solutions that can be deployed with minimal interference with the construction. The feasibility of vibration-based damage detection methods has been extensively investigated, yet most efforts so far have

---

H. Joseph · G. Quaranta (✉) · B. Carboni · W. Lacarbonara  
Department of Structural and Geotechnical Engineering,  
Sapienza University of Rome, Via Eudossiana 18, 00184 Rome,  
Italy  
e-mail: giuseppe.quaranta@uniroma1.it

focused on dynamic systems exhibiting linear behavior.

Several vibration-based methods have been developed either in frequency or time domain by assuming that the structure exhibits a linear behavior before and after damage [24]. The widespread use of mode shapes, natural frequencies and damping ratios for damage detection in linear dynamic systems is supported by well-established output-only techniques for operational modal analysis of structures. In particular, modal curvatures have proven to be especially useful in detecting damages in linear structures (e.g., [43,48]). Conversely, natural frequencies turn out to be almost insensitive to several damage scenarios (e.g., [15,17]) while damping ratios do not always provide consistent results (e.g., [9,53]). Time-domain methods have been also explored for damage detection in linear structures [18], and they usually leverage on state-space (e.g., [16,22]) or autoregressive-type models (e.g., [35,58]).

Damage detection in case of nonlinear dynamics has received considerably less attention. Initial attempts were based on the assumption that the undamaged structure behaves linearly while the occurrence of damage causes the transition to a nonlinear response [64]. However, the hypothesis of initial linear behavior does not hold true for a large class of dynamic systems. Some vibration-based methods have been thus developed for initially nonlinear dynamic systems. For instance, Adams and Nataraju [2] have shown that model reduction near bifurcations can be useful to identify certain features that can facilitate damage detection in nonlinear dynamic systems. The extension of the modal-based damage detection approach to nonlinear dynamic systems has been pursued by Lacarbonara et al. [33] through the use of nonlinear normal modes. This analytical study proved that the sensitivity of the effective nonlinearity coefficients regulating the nonlinear modal backbones is higher than the sensitivity of the linear frequencies with respect to damage. Such theoretical evidence was later confirmed in some experimental results reported by Civera et al. [13] and Carboni et al. [10]. Prawin and Rao [46] addressed the damage detection for nonlinear dynamic systems in the frequency domain using an improved describing function. Among the existing time-domain methods, several proposals deal with the use of Volterra series since such method provides an attractive way to describe the response of nonlinear dynamic systems. For example, Shiki et al. [56] and Villani et al. [59,60] applied the

discrete-time Volterra series and a stochastic version of Volterra series, respectively, to detect damage in dynamic systems exhibiting a nonlinear behavior even in the undamaged configuration. The output-only version of the Volterra series proposed by Peng et al. [45] for nonlinear structural damage detection employs the structural responses at two different locations in order to identify the kernel function parameters and to evaluate the contribution of the nonlinear components.

Recent advances in machine learning methods are paving the way towards significant advances in damage detection. In particular, Avci et al. [3] highlighted that deep learning (DL) algorithms hold significant potential for data-driven, vibration-based damage detection. These algorithms enable computational models with multiple processing layers to learn data representations with varying levels of abstraction. Autoencoders (AEs) and generative adversarial networks (GANs) are among the most promising DL architectures for this task according to Avci et al. [3] thanks to their capability of learning directly from the raw signals, but their applications are still limited and primarily focused on linear dynamic systems. These architectures employ an unsupervised learning paradigm, which obviates the need for extensive labeled datasets for training (e.g., [47,66]). For example, Pathirage et al. [44] proposed an AE-based damage detection approach where the natural frequencies and mode shapes of the linear dynamic system are employed as input data. The AE implemented by Ma et al. [41] utilizes the measured dynamic response for detecting damage in a linear system while the application reported by Shang et al. [55] is based on the use of the multi-dimensional cross-correlation functions computed from linear vibrations. Boccagna et al. [5] performed damage detection via AE by using univariate statistics estimated from the measured linear dynamic response after performing dimensionality reduction through principal component analysis. The AE-based damage detection approach proposed by Li et al. [36] makes use of the power cepstral coefficients extracted from the acceleration response of a linear dynamic system to provide a compact representation of its linear modal properties. Applications of GANs for damage detection are even fewer. The state-of-the-art review by Avci et al. [3] highlights the lack of studies on the use of GANs for damage detection in civil structures until 2019. Some examples of vibration-based damage detection in linear dynamic systems via GANs appeared only recently [39,40].

Machine learning algorithms applied to data-driven, vibration-based damage detection in nonlinear dynamic systems are infrequent. A seminal work in this context is due to Bornn et al. [6], who reported an experimental application about the use of the autoregressive support vector machine for damage detection in a small-scale laboratory shear-type system subject to random base motion with constant intensity. This system exhibited nonlinear behavior due to impacts and experienced damage from completely loosened bolts in a single column. The use of DL techniques has only recently started to emerge. For instance, Jin et al. [25] successfully applied a multi-kernel extreme DL machine for fault diagnosis of rotating machineries. Kim et al. [27] proposed the application of DL for the identification of reduced-order models, whereas Feng et al. [19] and Wang et al. [61] demonstrated the effectiveness in estimating simultaneously both model and noise parameters of stochastic systems. Zhang et al. [69] presented a numerical investigation about the use of AEs for damage detection in conventional bearings for seismic isolation. In this case, the onset of damage was simulated through the complete loss of functionality of one or two bearings at the corners of the isolation level, whereas seismic ground motion records were selected so as to match a target seismic spectrum. Nonlinear dynamical systems identification through GANs was recently addressed by Yu and Liu [68], but no proposals are known to exist about their application to damage detection problems.

Indeed, output-only, vibration-based damage detection in time-domain of initially nonlinear systems is very challenging. This is because the features of the nonlinear dynamic systems response can change considerably in time, especially under random vibrations and varying excitation conditions (e.g., [38]). As a consequence, discerning the effects induced by damage from those caused by the nonlinear response becomes a truly difficult task [6]. This is a significant concern for output-only vibration-based damage detection by means of machine learning methods that deserves in-depth investigations.

This work explores the use of DL algorithms for data-driven, vibration-based damage detection in nonlinear dynamic systems subject to random excitation with varying intensity. Within this framework, the original contribution here lies in a comprehensive investigation aimed at evaluating the comparative effectiveness and performance of AEs and GANs. The remain-

ing part of the work is organized as follows. Sections 2 and 3 illustrate the adopted DL architectures for damage detection in nonlinear dynamic systems. Section 4 presents the results of an extensive numerical investigation while Sect. 5 illustrates an experimental application. The most important outcomes of the present research are finally summarized in Sect. 6.

## 2 Autoencoder for damage detection

### 2.1 General working principle of the autoencoder

The AE originally proposed by Rumelhart et al. [51] is an unsupervised algorithm aimed at performing the input reconstruction with the least possible amount of distortion. Its main components are the encoder, a latent feature representation, and the decoder. Let  $\mathbf{x} \in \mathbb{R}^n$  be the normalized time history response of the nonlinear dynamic system, the encoder is a function  $\mathbf{g} : \mathbb{R}^n \rightarrow \mathbb{R}^m$  such that:

$$\mathbf{h} = \mathbf{g}(\mathbf{x}), \quad (1)$$

where  $\mathbf{h} \in \mathbb{R}^m$  is the latent feature representation of  $\mathbf{x}$ . The decoder is another function  $\mathbf{f} : \mathbb{R}^m \rightarrow \mathbb{R}^n$  such that:

$$\tilde{\mathbf{x}} = \mathbf{f}(\mathbf{h}) = \mathbf{f}(\mathbf{g}(\mathbf{x})), \quad (2)$$

where  $\tilde{\mathbf{x}} \in \mathbb{R}^n$  is the input reconstruction. Since comprehensive state-of-the-art reviews about AEs are available [e.g., 29, 37, 67], only details related to the adopted implementation for damage detection applications in nonlinear dynamic systems are presented hereafter.

### 2.2 Proposed architecture of the autoencoder for damage detection

In most typical AE architectures, the functions  $\mathbf{f}$  and  $\mathbf{g}$  in Eqs. (1), (2) representing the encoding and decoding blocks are artificial neural networks (NNs). In particular, convolutional NNs (CNNs) are mainly employed in AEs because they excel at recognizing the most important signal features and at capturing relevant patterns in the data. 1D CNNs are preferred in the present work because of their minimum computational complexity. It is noted that a normalized input is fed into the AE algorithm so as to reduce the variability among data points from different scales and to speed up the convergence

of the networks during training phase [8]. Scaling the signals to a uniform bounded range mitigates the dominance of higher magnitude features over smaller ones, which allows to recognize subtle patterns of damage without bias towards more prominent signal features. In the current study, data normalization is accomplished through min-max scaling within the range [0, 1], which is generally recognized to display enhanced performance when handling time series in DL algorithms [4]. The components of the implemented CNN-based AE architecture are: input layer, convolutional layer, pooling layer, fully connected layer, and output layer. The input layer of the CNN aims at importing the normalized 1D time history response data of the nonlinear dynamic system. A deep AE is here considered, which means that it involves more than one convolutional layer in the encoding and decoding blocks. Within a convolutional layer, the convolution operator applies on a given input (i.e., the output from the previous layer) using a convolution filter to produce a feature map, whereas a nonlinear activation function is next employed to obtain the output features. The convolutional operation involves sliding a small filter (also known as kernel) and computing dot products between the filter weights and the input values at each location as follows:

$${}^{\ell+1}y_i(j) = {}^{\ell}\mathbf{K}_i * {}^{\ell}\hat{\mathbf{x}}(j) + {}^{\ell}b_i, \quad (3)$$

where  ${}^{\ell}\mathbf{K}_i$  and  ${}^{\ell}b_i$  denote the weights and the bias of the  $i$ th filter in the  $\ell$ th layer, respectively,  ${}^{\ell}\hat{\mathbf{x}}(j)$  is the  $j$ th receptive field in the  $\ell$ th layer (i.e., the local region representing the portion of the input values that is being processed at a given step of the convolution operation), and  ${}^{\ell+1}y_i(j)$  is the input of the  $j$ th artificial neuron in the  $i$ th frame of the following  $(\ell + 1)$ th layer. One filter corresponds to one frame in the following layer, and the number of frames is also known as layer depth. It is noted that the same set of learnable parameters (weights) across different receptive fields is utilized. Such weight-sharing allows the network to detect patterns and features regardless of their location. This is because the same filter is applied across the entire input, reducing the computational effort. Zero padding is applied to both the beginning and the end of the input to preserve the original length of the convolution result. This approach enhances the network's ability to learn features at the boundaries, which is particularly important for short inputs. The stride, defined as the number of positions by which the filter moves across

the input during the convolution operation, is chosen to ensure consistent alignment between the input and output across the different layers of the network.

The implementation of artificial neurons with a nonlinear activation function after the convolution operations is of utmost importance. In fact, without a nonlinear activation function, the output of each convolution would be a linear transformation of the input, which is unfit to capture inherent complex patterns from the data. Conversely, introducing nonlinearity allows the network to learn and represent complex relationships in the input. In the present study, the Leaky ReLU activation function is utilized. It was proposed to address one of the main limitations of the traditional ReLU activation function, which is due to the fact that it sets all negative inputs to 0, potentially leading to vanishing gradients [20, 42]. The Leaky ReLU activation function mitigates this issue by introducing a small, non-null gradient when the activation function is negative as follows:

$${}^{\ell+1}a_i(j) = \psi({}^{\ell+1}y_i(j)) = \begin{cases} {}^{\ell+1}y_i(j) & \text{if } {}^{\ell+1}y_i(j) > 0 \\ {}^{\ell+1}\mu y_i(j) & \text{otherwise} \end{cases}, \quad (4)$$

where  ${}^{\ell+1}a_i(j)$  is the activation value of the convolution operation result  ${}^{\ell+1}y_i(j)$  in Eq. (3) while  $\mu$  is the slope coefficient, which controls the angle of the negative slope in case of negative input values. By using the Leaky ReLU activation function, the network can learn from negative inputs, enhancing its robustness and ability to capture complex patterns. Furthermore, using min-max scaling in combination with Leaky ReLU activation function ensures that the majority of input falls into an active learning range, reducing the risk of inactive neurons.

A pooling layer after the convolutional layer is meant at lowering both the size of the features and the parameters of the network. This, in turn, forces the network to learn only the most important features of the input. Additionally, it reduces training time and mitigates overfitting. The max-pooling layer is here employed. Let  ${}^{\ell}q_i(s)$  be the value of the artificial neuron  $s$  in the  $i$ th frame of the  $\ell$ th layer, the max-pooling transformation is applied as follows:

$${}^{\ell+1}p_i(j) = \max_{(j-1)W+1 \leq s \leq jW} {}^{\ell}q_i(s), \quad (5)$$

where  ${}^{\ell+1}p_i(j)$  is the corresponding value of the artificial neuron in the  $(\ell + 1)$ th layer of the pooling oper-

ation. The parameter  $W$  in Eq. (5) defines the pooling width.

By adjusting padding and stride, convolutional layers for the encoder are designed to downscale the input together with pooling layers, while convolutional layers in the decoder are intended to upscale the input (through transposed convolution operations, and pooling layers do not take place in this case). Convolutional layers also serve as output layer of the AE. Conversely, fully connected layers are implemented in the latent feature representation to effectively lower the dimensionality of the input, as this task would be cumbersome with convolutional layers. Controlling the size of the latent representation is important because it plays an important role in determining how the machine learning model performs on the assigned data. A fully connected layer, also known as dense layer, is a standard layer where each artificial neuron is connected to every artificial neuron in the previous layer, similar to a traditional NN (e.g., [49]). Before the fully connected layer is applied, the output from the last pooling layer in the encoding block is flattened.

The AE must accurately reconstruct the input time history response of the nonlinear dynamic system while producing a meaningful latent representation that avoids overfitting. This is accomplished through proper training of the AE, which basically consists of finding the optimal set of hyperparameters  $\theta$  ruling the functions  $\mathbf{f}$  and  $\mathbf{g}$  that represent the encoding and decoding blocks. To this end, two main strategies leveraging on parsimony are adopted [31] and are based on creating a bottleneck and adding a regularization term. A bottleneck involves reducing the dimensionality  $m$  of the latent features compared to that of the data  $n$ . Including a regularization term in the training process also prevents the model from fitting the noise in the training data too closely. In the present work, the  $L_2$  technique is implemented, and thus the AE training seeks the solution of the following optimization problem:

$$\min_{\theta} \left\{ \mathbb{E} [\Delta(\mathbf{x}, f(\mathbf{g}(\mathbf{x}); \theta))] + \lambda \|\theta\|_2^2 \right\}, \quad (6)$$

where the first term is a loss function that quantifies how input and output of the AE differ each other, on average, across all  $\mathbf{x}$  ( $\mathbb{E}$  represents the expectation operator), while  $\lambda$  and  $\|\theta\|_2$  in the second term are the regularization coefficient and the  $L_2$ -norm of the hyperparameters  $\theta$  involved in the functions  $\mathbf{f}$  and  $\mathbf{g}$ , respectively. The adopted loss function in the present study is the mean absolute error (MAE). The set of tuned hyperparameters

$\theta$  includes kernel size, filter numbers, batch size (i.e., number of input values processed simultaneously during one training step), and latent filters. The values of slope coefficient  $\mu$  in Eq. (4) and regularization coefficient  $\lambda$  in Eq. (6) are instead established before training (the slope coefficient is set to 0.2 whereas the regularization coefficient is optimized in the range  $10^{-4}$  to  $10^{-8}$  through cross-validation according to Kruse et al. [30]).

The adaptive moment estimation (ADAM) optimizer is adopted to solve Eq. (6). It is a gradient-based technique [28] that combines the advantages of the adaptive gradient algorithm and root mean square propagation. ADAM computes adaptive learning rates for each hyperparameter based on estimates of the first and second moments of the gradient. The update rules in ADAM are the following:

$$\mathbf{m}_{\tau} = \beta_1 \mathbf{m}_{\tau-1} + (1 - \beta_1) \nabla_{\tau}, \quad (7a)$$

$$\mathbf{v}_{\tau} = \beta_2 \mathbf{v}_{\tau-1} + (1 - \beta_2) \nabla_{\tau}^2, \quad (7b)$$

$$\hat{\mathbf{m}}_{\tau} = \frac{\mathbf{m}_{\tau}}{1 - \beta_1^{\tau}}, \quad (7c)$$

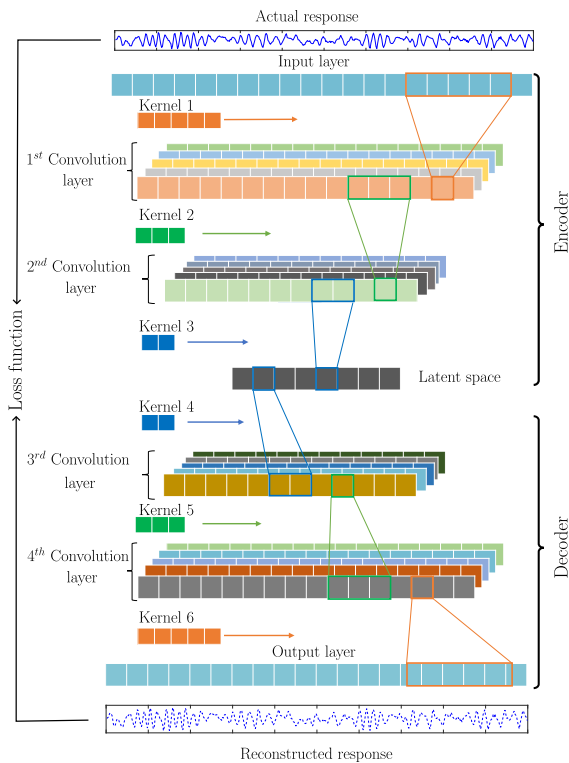
$$\hat{\mathbf{v}}_{\tau} = \frac{\mathbf{v}_{\tau}}{1 - \beta_2^{\tau}}, \quad (7d)$$

$$\theta_{\tau+1} = \theta_{\tau} - \frac{\eta}{\sqrt{\hat{\mathbf{v}}_{\tau} + \epsilon}} \hat{\mathbf{m}}_{\tau}, \quad (7e)$$

where  $\mathbf{m}_{\tau}$  and  $\mathbf{v}_{\tau}$  are the updated first (i.e., mean) and second (i.e., uncentered variance) moments of the gradients, respectively,  $\beta_1$  and  $\beta_2$  are the exponential decay rates for these moment estimates (set to 0.9 and 0.999, respectively),  $\nabla_{\tau}$  is the objective function gradient at iteration  $\tau$ ,  $\eta$  is the learning rate (in the range  $10^{-2}$  to  $10^{-4}$ ), and  $\epsilon$  is a small scalar added to improve numerical stability (set to  $10^{-8}$ ). The result of Eq. (7) is  $\theta_{\tau+1}$ , which represents the updated hyperparameters for the next iteration (aka epoch), or the outcome of the optimization if the convergence criterion is fulfilled. To further strengthen the network against overfitting phenomenon, early stopping is executed by setting a low value of the patience (i.e., the number of epochs the training process continues without improvement) between 4 and 6.

### 2.3 Proposed implementation of the autoencoder for damage detection

The suitability of AE for data-driven damage detection relies on the evidence that, once properly trained over



**Fig. 1** Implemented AE architecture based on 1D CNNs for damage detection in nonlinear dynamic systems

time data collected from an undamaged dynamic system, it will turn into a large loss function value when attempting to reconstruct the response of a damaged one. Figure 1 illustrates the implemented AE architecture based on 1D CNNs for damage detection in nonlinear dynamic systems. The corresponding algorithm was developed using TensorFlow [1].

### 3 Generative adversarial network for damage detection

#### 3.1 General working principle of the generative adversarial network

A GAN is an unsupervised algorithm originally proposed by Goodfellow et al. [21] that consists of two models called generator and discriminator. The generator tries to create realistic data, while the discriminator aims to differentiate between real data and the fake data created by the generator. On the one hand, the generator can be represented as a differentiable function  $G(\mathbf{z}; \theta_G)$

where  $\mathbf{z}$  is a vector of random noise extracted out of a distribution  $\varphi_{\mathbf{z}}(\mathbf{z})$  and  $\theta_G$  is a set of hyperparameters. It produces synthetic data samples, which should resemble real data samples. On the other hand, the discriminator is another differentiable function  $D(\mathbf{x}; \theta_D)$  where  $\mathbf{x}$  is a data sample from a distribution  $\varphi_{data}(\mathbf{x})$  and  $\theta_D$  is a set of hyperparameters. It produces a scalar output representing the probability that  $\mathbf{x}$  comes from the real data distribution rather than the generator’s distribution. Data normalization through min-max scaling is also performed in GAN to enhance the performance of the algorithm as well as the efficiency of the training phase.

The training of GAN is a non-cooperative game. Hence, GAN optimization is attained at the Nash equilibrium point [50] corresponding to a stable state where neither the generator, nor the discriminator obtains a beneficial effect (i.e., gain) by deviating from the current configuration. At this equilibrium point, the generator produces data that are so realistic that the discriminator cannot distinguish whether they are real or fake, and the discriminator cannot improve its accuracy by further adjustments. The training of discriminator  $D$  and generator  $G$  in GAN requires the solution of the following optimization problems:

$$\max_{\theta_D} \left\{ \mathbb{E}_{\mathbf{x} \sim \varphi_{data}(\mathbf{x})} [\log D(\mathbf{x}; \theta_D)] + \mathbb{E}_{\mathbf{z} \sim \varphi_{\mathbf{z}}(\mathbf{z})} [\log (1 - D(G(\mathbf{z}; \theta_G); \theta_D))] \right\}, \quad (8)$$

$$\min_{\theta_G} \left\{ \mathbb{E}_{\mathbf{z} \sim \varphi_{\mathbf{z}}(\mathbf{z})} [\log (1 - D(G(\mathbf{z}; \theta_G); \theta_D))] \right\}. \quad (9)$$

Altogether, Eqs. (8), (9) are equivalent to the following min-max optimization problem:

$$\min_{\theta_G} \max_{\theta_D} \left\{ \mathbb{E}_{\mathbf{x} \sim \varphi_{data}(\mathbf{x})} [\log D(\mathbf{x}; \theta_D)] + \mathbb{E}_{\mathbf{z} \sim \varphi_{\mathbf{z}}(\mathbf{z})} [\log (1 - D(G(\mathbf{z}; \theta_G); \theta_D))] \right\}, \quad (10)$$

wherein the objective function is related to the Jensen-Shannon divergence between  $\varphi_{data}(\mathbf{x})$  and  $\varphi_{\mathbf{z}}(\mathbf{z})$  [23] ( $\mathbb{E}$  represents the expectation concerning the distribution indicated by the subscript). The GAN version based on the optimization problem in Eq. (10) is commonly referred to as vanilla GAN. The details concerning the specific implementation of GAN for damage detection applications in nonlinear dynamic systems will be presented hereafter while more general information can be found elsewhere [e.g., 14,23,65].

### 3.2 Proposed architecture of the generative adversarial network for damage detection

A deep GAN is here implemented by stacking multiple 1D CNNs. In particular, 1D convolutional layers according to Eqs. (3), (4) are used for both generator and discriminator (wherein transposed convolution operations are employed for the generator). In this regard, a significant novelty deals with the output layer of the discriminator, which implements a Sigmoid activation function as follows:

$${}^{\ell+1}a_i(j) = \psi({}^{\ell+1}y_i(j)) = \frac{{}^{\ell+1}y_i(j)}{1 + e^{{}^{\ell+1}y_i(j)}}. \quad (11)$$

It is noted that the output of the Sigmoid activation function (i.e., the output of the discriminator) given by Eq. (11) lies between 0 (for certainly fake data) and 1 (for certainly real data).

The hyperparameters are tuned alternatively solving Eqs. (8), (9) throughout the training stage by means of ADAM update rules in Eq. (7). During training, the hyperparameters of one model are updated, while the hyperparameters of the other are fixed. The set of tuned hyperparameters for GAN includes number filters, kernel size, stride length, and depth of the neural network (these hyperparameters are different for the generator and discriminator). The training process for GAN runs until a number of epochs between 1000 and 2000 is achieved (no early stopping is implemented in this case). Overfitting mitigation in the GAN discriminator is accomplished via dropout. Accordingly, throughout the training, an artificial neuron is temporarily dropped at each iteration (i.e., all the inputs and outputs to this artificial neuron are disabled at the current iteration) with an assigned probability. The dropped-out artificial neurons are resampled with the same probability at every training step, so a dropped out neuron at one step can be active at the next one.

### 3.3 Proposed implementation of the generative adversarial network for damage detection

There are no known proposals regarding the use of GAN in damage detection problems. Nonetheless, the suitability of GAN for data-driven damage detection stems from the fact that, upon proper training with time series data collected from an undamaged system, the discriminator will be no longer able to differentiate between actual distribution of undamaged data

and distribution of synthetic data produced by the generator. This implies that the discriminator will recognize the data from a damaged system as fake because they no longer resemble those of the undamaged system, thereby allowing the detection of damage. In this regard, it is interesting to note that, by implementing the Sigmoid activation function at the output layer of the discriminator, a quantitative assessment of the dynamic system state is obtained from the discriminator. In fact, numerical values of the discriminator output close to 1 are associated with an undamaged state while lower values approaching to 0 highlight a deviation from the (trained) undamaged condition, thereby pointing out the presence of damage. Hence, the discriminator output can be exploited to compute a damage index.

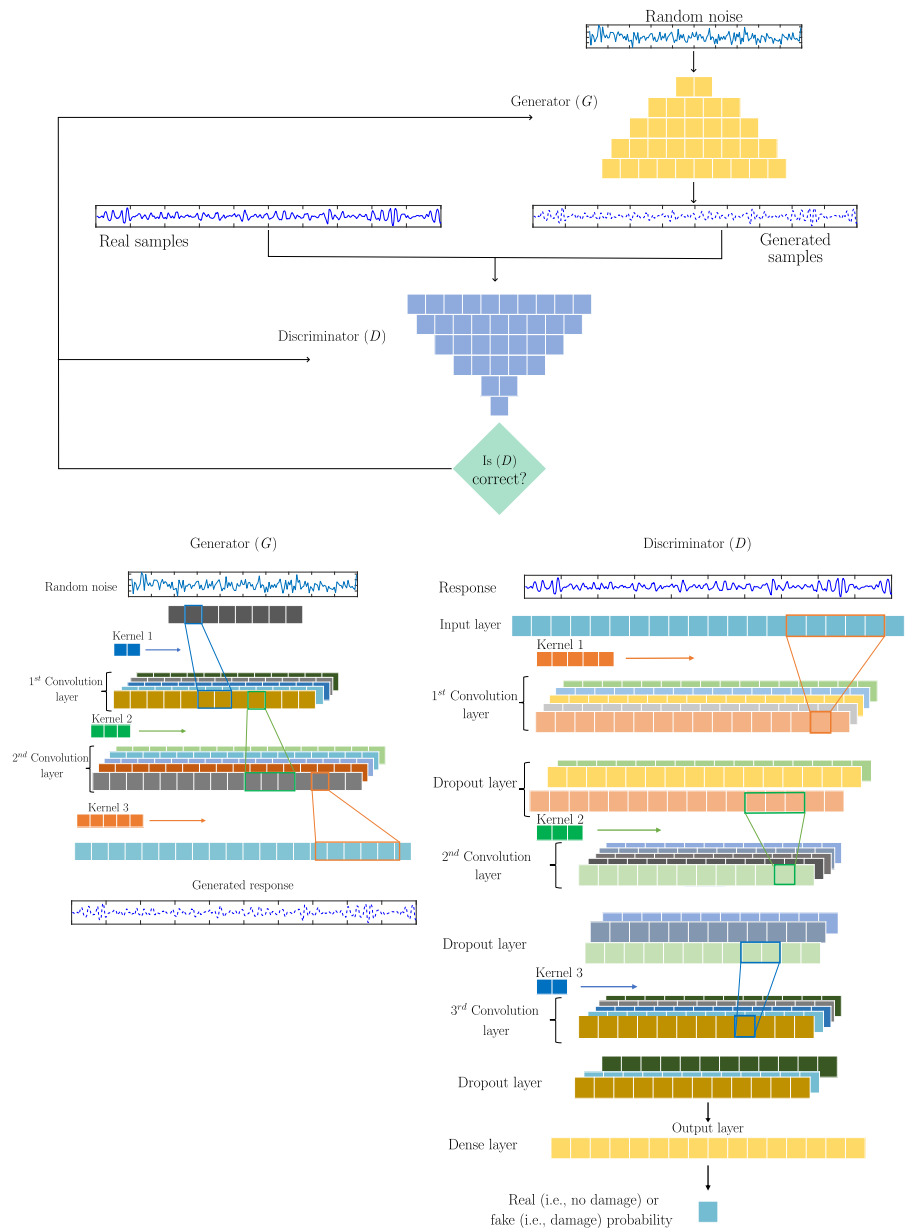
This vanilla GAN architecture based on 1D CNNs for damage detection in nonlinear dynamic systems has been implemented using TensorFlow [1]. Figure 2 shows pictorially the implemented GAN architecture.

## 4 Numerical investigation

Numerical simulations are initially performed to evaluate the effectiveness of both AE and GAN in detecting damage within nonlinear dynamic systems from their time-domain response only. Herein, damage is introduced by reducing the system stiffness. To replicate the most common and challenging scenario, the nonlinear dynamic systems under consideration are subjected to random vibrations of varying intensity, and the response is contaminated by white Gaussian noise with a noise level of 10%.

The database for training and validating both DL architectures includes samples of undamaged system displacements over 2 s time windows, assuming a sampling rate of 250 Hz. A total of 2000 time series was collected for such tasks: 80% of the generated time history samples is utilized for training the algorithms (i.e., training set), while the remaining 20% is allocated for their validation (i.e., validation set). The validation dataset was used to check if the training of the network converged properly, avoiding overfitting or underfitting. The full database thus consists of about 1 h of recording of the undamaged system response under different magnitudes of the input random excitation. It is highlighted that neither the applied dynamic loading conditions nor the dynamic system characteristics are ultimately involved in the implementation of the con-

**Fig. 2** Implemented GAN architecture based on 1D CNNs for damage detection in nonlinear dynamic systems



sidered DL architectures. This means that both AE and GAN are meant to perform output-only damage detection of unknown dynamic systems.

The selected configuration of AE and GAN for each application is first established through preliminary simulations, and then its trainable parameters are found. The training runtime reported in the current numerical investigation are obtained using a MacBook Pro M2 Max with 32 Gb RAM and 38 core GPU.

Once both DL algorithms have been trained and validated using the samples of undamaged system response, they are employed to detect damage. To this end, 2000 time series of noisy system displacements, each of 2 s duration, were generated for constant damage levels ranging from 5% to 30%. Therefore, both AE and GAN are required to detect damage, if any, from about 1 h long of recording of the system response. While low damage levels are considered to test the capability of detecting the onset of damage, the results



for increasing damage levels are examined to inspect the proper working of the implemented DL schemes.

#### 4.1 Single-degree-of-freedom system with cubic nonlinearity

A large variety of dynamic systems can exhibit a nonlinear behavior for high levels of excitation due to geometric or material nonlinear effects. The actual response of such systems can be effectively simulated by means of a restoring force that is dependent upon the cube of the displacement, known as a Duffing-type system. Hence, this class of nonlinear dynamic systems provides a relevant benchmark for evaluating the effectiveness of damage detection techniques in practical applications. This system can be seen as a reduced-order model of an axially restrained beam, a plate or any symmetric geometrically nonlinear structure [32]. The equation of motion for a single-degree-of-freedom (1-DOF) system with cubic nonlinearity and linear viscous damping is the following:

$$M\ddot{x} + C\dot{x} + K_1x + K_3x^3 = F(t), \quad (12)$$

where  $x(t)$  represents the displacement of the system at time  $t$  (overdots denote differentiation with respect to time  $t$ ), whereas  $M$  is the mass and  $C$  is the viscous damping coefficient. Moreover,  $K_1$  and  $K_3$  are the linear and nonlinear stiffness coefficients, respectively, while  $F(t)$  is the external time-varying load. The numerical values of the system parameters in Eq. (12) for this numerical study are set to:  $M = 412$  kg,  $C = 405.95$  kg/s (corresponding to a damping ratio of 1%),  $K_1 = 1$  kN/mm,  $K_3 = 0.001$  kN/mm<sup>3</sup>. It is assumed that the presence of damage corresponds to a reduction of both linear and nonlinear stiffness by the same amount, without effects on mass or damping.

The dynamics of such system are first investigated by computing the frequency response curves (FRCs) upon applying a harmonic base excitation such that  $F(t) = -MA \cos \Omega t$ , where  $A$  is the peak acceleration value and  $\Omega$  is the excitation frequency. Results for the undamaged and damaged system are plotted in Fig. 3 assuming  $A$  between 0.003g and 0.1g. The considered dynamic system exhibits a linear behavior for  $A$  less than or equal to 0.01 g while a nonlinear behavior is apparent otherwise. In particular, a moderately nonlinear response is observed for  $A$  between 0.01g and 0.06g, whereas a stronger nonlinear behavior occurs for higher values of the input excitation. For

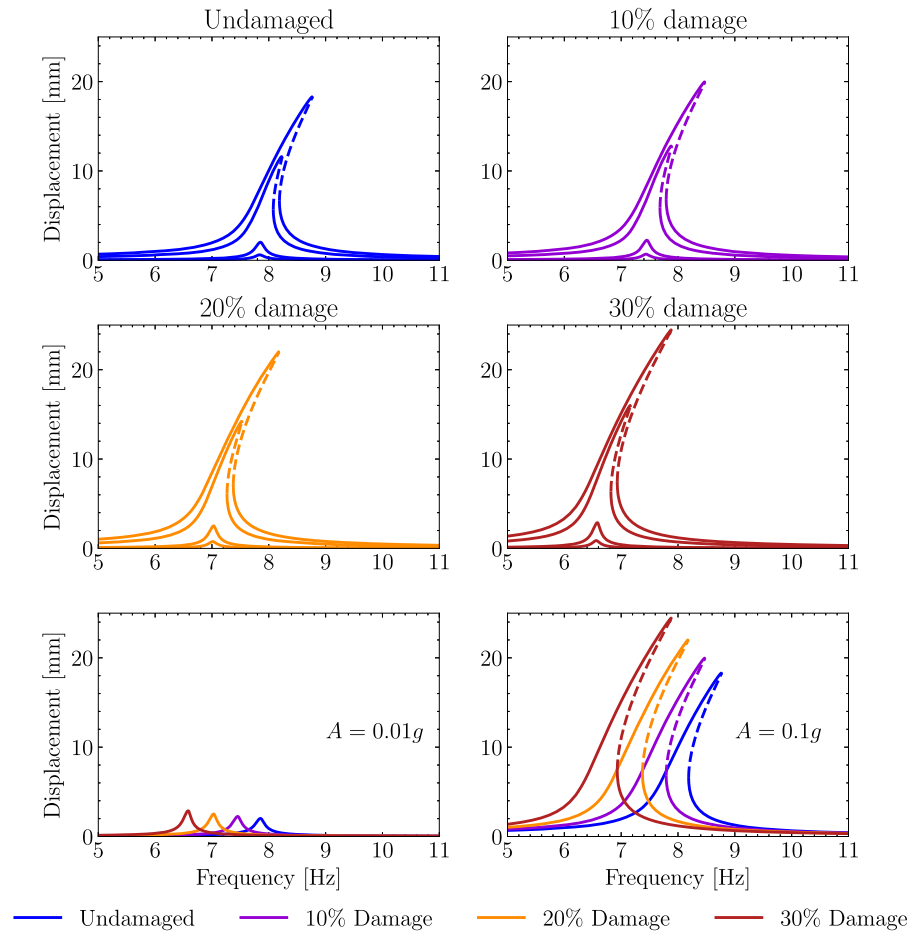
the undamaged system, if  $A$  is less than or equal to 0.01g, then the resonant frequency is 7.8 Hz, while it shifts to 8.8 Hz (increase by 12.8%) when  $A$  is 0.1g.

Figure 3 is also useful to explain why time-domain, output-only damage detection in nonlinear dynamic systems subject to varying dynamic excitation is so challenging as pointed out by Bornn et al. [6]. In fact, Fig. 3 shows that the shift of the resonant frequency due to the hardening-type nonlinear response can mislead the fault recognition, thus hindering the accurate damage detection.

The effectiveness of both AE and GAN is investigated through numerical simulations by solving Eq. (12) under random excitation such that  $F = -M\ddot{x}_b$ , where  $\ddot{x}_b$  is represented by a white noise signal. The time histories of the undamaged and damaged system are computed once time history samples of  $\ddot{x}_b$  are generated and the corresponding peak values are scaled to a uniform random value between 0.01g and 0.1g so as to focus on the nonlinear response only, in agreement with Fig. 3. Details about configuration and training of both AE and GAN are provided in Tables 1 and 2, respectively.

The representative results reported in Fig. 4 confirm that the trained AE is able to reconstruct the time history displacement of the undamaged nonlinear system (the system response under the highest excitation is plotted herein). Conversely, it is no longer able to achieve similar accuracy levels when it attempts to reconstruct the displacement time history of the damaged system. This mismatch between input and output of AE highlights a deviation from the conditions of the training process, and thus points to the occurrence of the damage. The greater the damage level, the larger the difference between the input and output of the AE scheme. GAN operates differently, requiring the generator to accurately mimic the undamaged system response. This maximizes the damage sensitivity of the discriminator throughout the training process. In this regard, a representative pair of actual and generated displacement time histories is plotted in Fig. 5 (the dominant frequency herein is about 8.8 Hz, which corresponds to the highest excitation level according to Fig. 3). The high similarity between them in the time-frequency domain demonstrates the fidelity of GAN in simulating the dynamics of the undamaged nonlinear system. This, in turn, implies that the discriminator has been successfully trained and will thus be able to

**Fig. 3** Frequency response curves (FRCs) of the 1-DOF system for different excitation amplitudes and damage levels (solid and dashed lines represent the stable and unstable branches of periodic solutions, respectively)

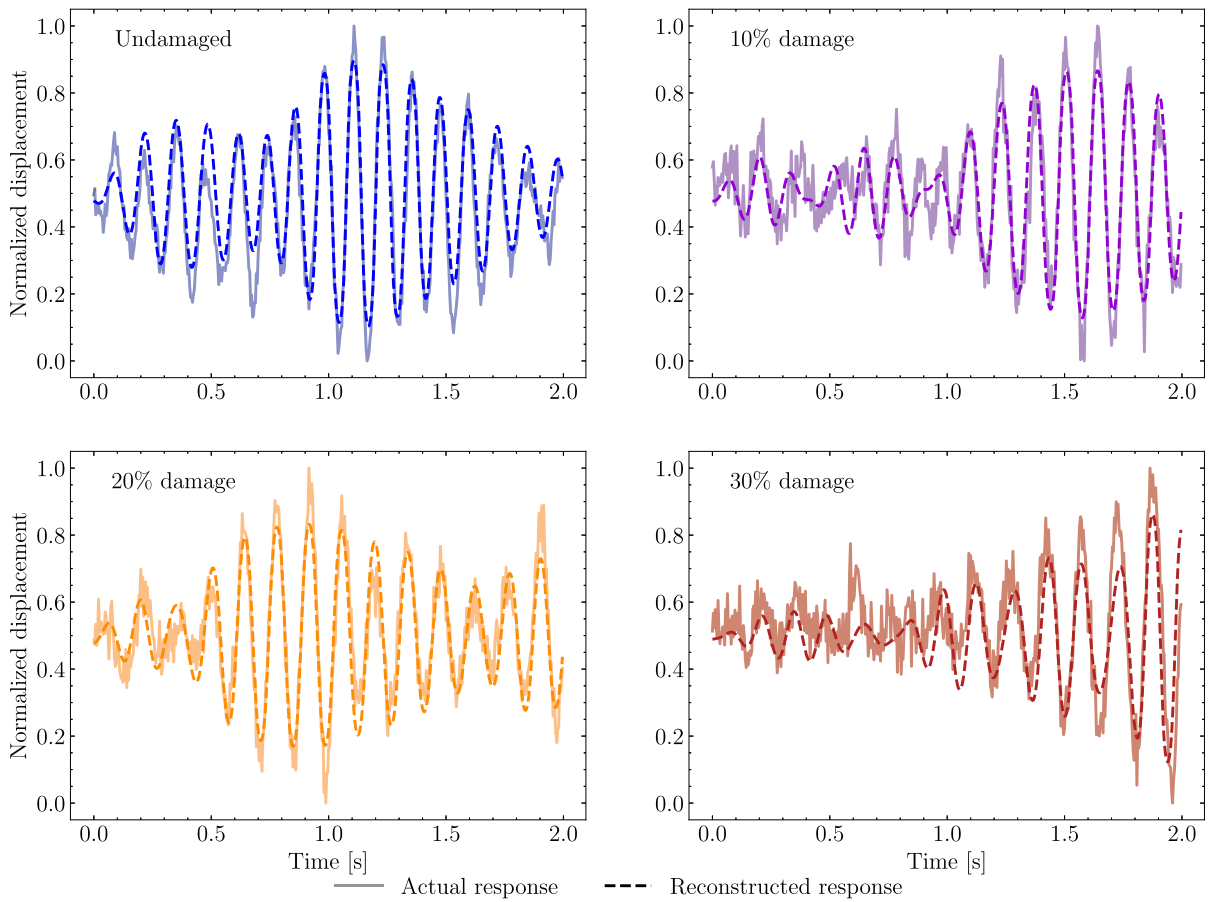


**Table 1** Configuration of the DL architectures for damage detection in the 1-DOF system

Type	Component	Convolution layers	Kernel size	Number of filters	Strides
AE	Encoder	3	(100,50,25)	(10,5,2)	(1,1,1)
	Decoder	3	(25,50,100)	(5,10,1)	(2,2,1)
GAN	Generator	3	(2,3,6)	(32,32,1)	(2,1,2)
	Discriminator	3	(6,6,6)	(64,64,64)	(2,2,2)

**Table 2** Training data of the DL architectures for damage detection in the 1-DOF system

Type	Component	Trainable parameters	Epochs	Training time
AE	Encoder	3767	89	0.16 min
	Decoder	66,056		
GAN	Generator	16,018	2000	11.47 min
	Discriminator	8,117,729		



**Fig. 4** Reconstruction of the 1-DOF system response through AE for different damage levels

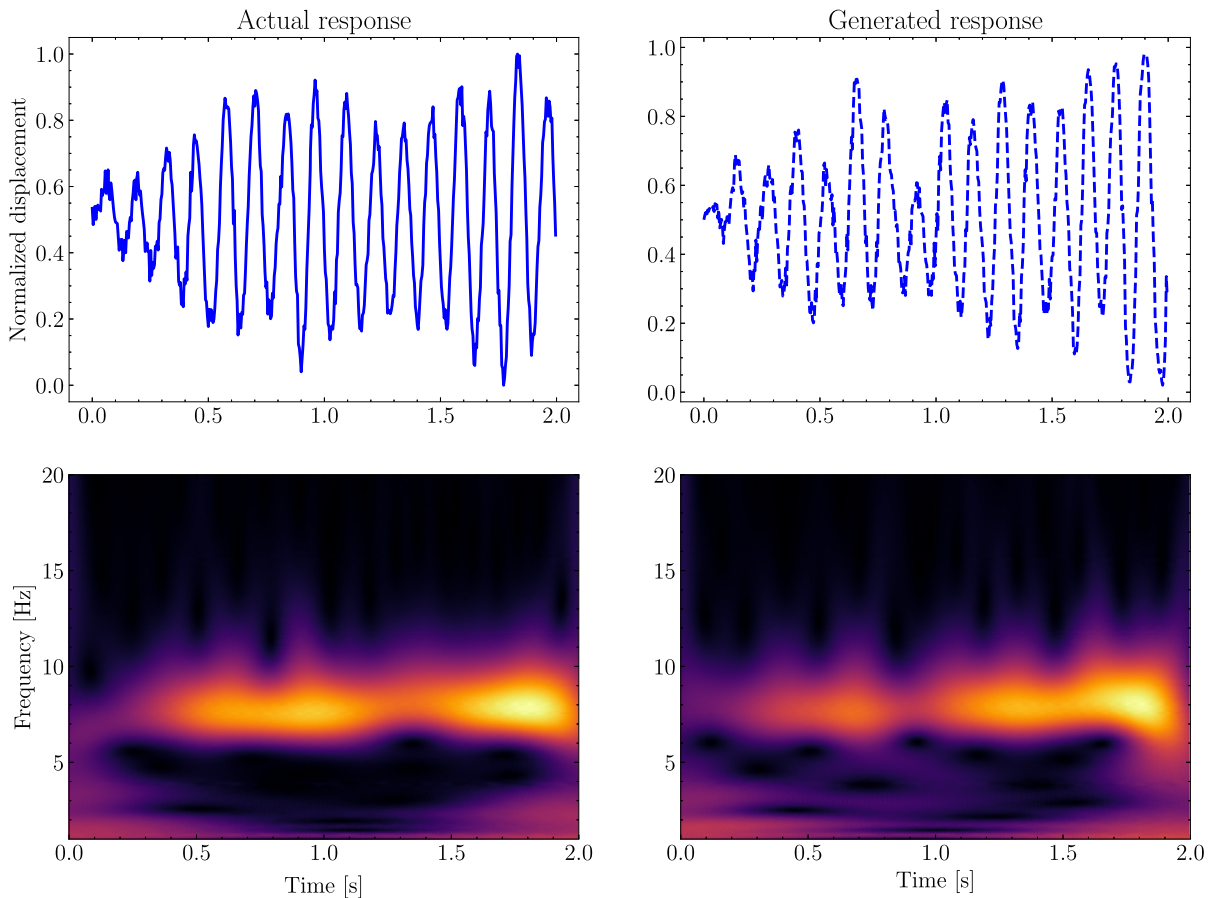
recognize accurately a deviation from the training conditions, such as the occurrence of damage.

Figure 6 compares the performance of both DL architectures in detecting damage. Each dot in the scatter plots represents the reconstruction loss (in terms of MAE) of AE and the discriminator output of GAN for each time series sample given the damage level, while the trend lines connect the corresponding average values. The labels on the horizontal axis, centered with respect to each cluster of dots having uniform color, indicate the damage level (i.e., relative variation of the system parameters due to damage, with 0% representing the undamaged scenario). Since the discriminator output is bounded between 0 and 1, a damage index is defined as the complement of the discriminator output. Figure 6 shows that reconstruction loss values shift upwards (i.e., the reconstruction loss increases) as the damage severity grows. It can also

be inferred from Fig. 6 that the discriminator outputs are close to 1 when there is no damage, while a growing number of outcomes approaches 0 as the damage level increases. Therefore, the relative variation of the normalized reconstruction loss and that of the damage index based on the discriminator loss are properly correlated with the damage level. Notably, a significant deviation from the reference value corresponding to the undamaged configuration already occurs at the lowest level of damage (i.e., relative variation larger than 10% for the lowest damage severity), which demonstrates a satisfactory sensitivity of both DL architectures.

#### 4.2 Two-degree-of-freedom system with cubic nonlinearity

Damage detection in a two-degree-of-freedom (2-DOF) nonlinear dynamic system with cubic nonlinear-



**Fig. 5** Time history and corresponding time-frequency representation via wavelet transform of the undamaged 1-DOF system response: actual data and results carried out from the GAN generation block

ity and linear viscous damping is considered next. The corresponding equations of motion are the following:

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

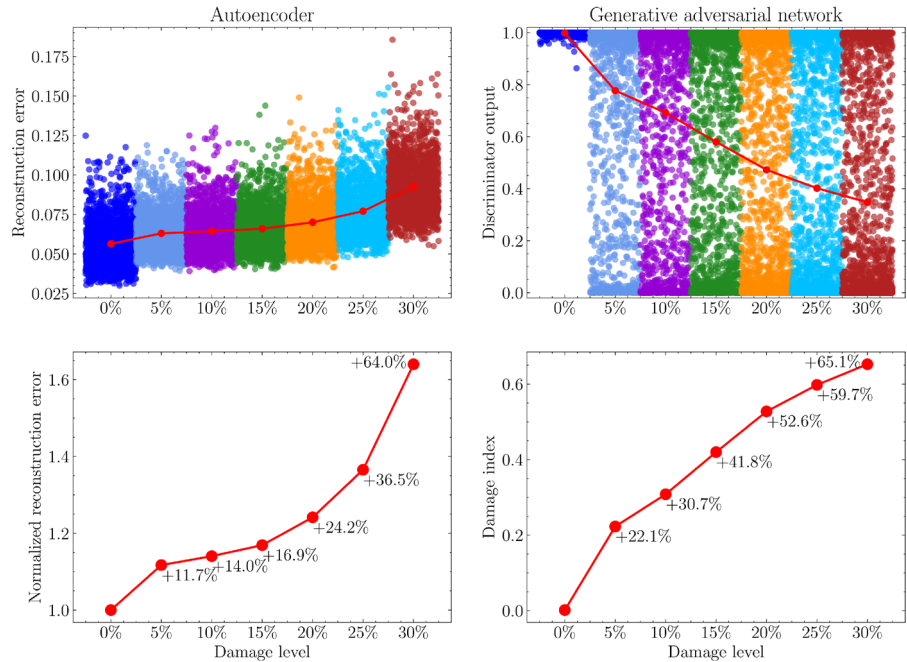
$$\begin{aligned} M_2 \ddot{x}_2 + C_2 (\dot{x}_2 - \dot{x}_1) + K_{21} (x_2 - x_1) \\ + K_{23} (x_2 - x_1)^3 = F_2(t), \end{aligned} \quad (13b)$$

where  $x_1(t)$  and  $x_2(t)$  are the displacements of the two oscillators,  $M_1$  and  $M_2$  are the masses,  $K_{11}$  and  $K_{21}$  are the linear stiffness coefficients,  $C_1$  and  $C_2$  are the viscous damping coefficients,  $K_{13}$  and  $K_{23}$  are the cubic stiffness coefficients,  $F_1$  and  $F_2$  are the excitation forces. The numerical values of the system parameters in Eq. (13) are set to:  $M_1 = M_2 = 400$  kg,  $C_1 = C_2 = 400$  kg/s (corresponding to a damping ratio

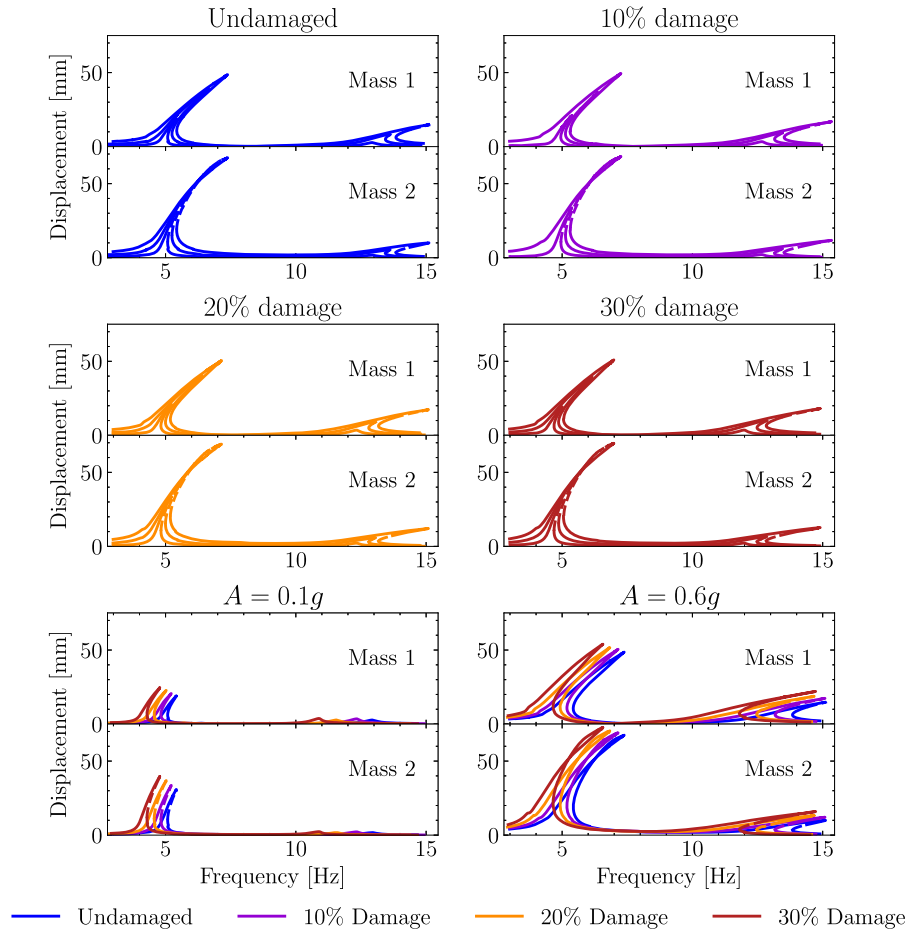
of 1%),  $K_{11} = K_{21} = 1$  kN/mm,  $K_{13} = K_{23} = 0.001$  kN/mm<sup>3</sup>. It is assumed that linear and nonlinear stiffness coefficients get reduced by the same amount in both oscillators as a result of damage, while damage does not affect their masses and damping coefficients. The FRCs of the undamaged and damaged 2-DOF nonlinear system are shown in Fig. 7 for  $F_1 = -MA \cos \Omega t$  and  $F_2 = 0$ , with  $A$  between 0.1g and 0.6g.

A random excitation is employed to detect damage assuming  $F_1 = -M_1 \ddot{x}_b$  and  $F_2 = 0$ . The displacement time histories of both the undamaged and damaged systems are thus calculated through Eq. (13) by generating independent time history samples of the white noise  $\ddot{x}_b$  and adjusting their peaks to match a uniformly random value between 0.1g and 0.6g in order to restrict the analysis to a mid-large nonlinear response accord-

**Fig. 6** Damage sensitivity of the two DL architectures for the 1-DOF system: reconstruction loss values obtained by AE and relative variation of its normalized average value, together with the discriminator output values obtained by GAN and relative variation of the associated average damage index



**Fig. 7** Frequency response curves (FRCs) of the 2-DOF system for different excitation amplitudes and damage levels (solid and dashed lines represent the stable and unstable branches of periodic solutions, respectively)



**Table 3** Configuration of the DL architectures for damage detection in the 2-DOF system

Type	Component	Convolution layers	Kernel size	Number of filters	Strides
AE	Encoder	3	(50,25,12)	(100,50,25)	(1,1,1)
	Decoder	4	(12,25,25,50)	(50,100,100,2)	(2,2,1,1)
GAN	Generator	3	(6,6,4)	(64,32,2)	(2,1,2)
	Discriminator	3	(6,6,6)	(32,32,32)	(2,2,2)

**Table 4** Training data of the DL architectures for damage detection in the 2-DOF system

Type	Component	Trainable parameters	Epochs	Training time
AE	Encoder	150,175	43	1.21 min
	Decoder	2,142,902		
GAN	Generator	38,151	1000	6.55 min
	Discriminator	4,050,770		

ing to Fig. 7. The settings adopted to implement the DL architectures are listed in Table 3 whereas Table 4 provides relevant outcomes of the training process.

Although the interaction of the two nonlinear oscillators complicates the overall dynamic behavior, both architectures confirm their good performances as shown in Figs. 8 and 9. As far as the application of AE is concerned, Fig. 8 shows that the undamaged response of both oscillators is well reconstructed after the training process, whereas the accuracy decreases as expected once the system is damaged. Figure 9 confirms the good performance of GAN since the real response (i.e., the response obtained as the solution of the equations of motion) and the generated response of the undamaged oscillators look similar in the time-frequency domain. This indicates that the discriminator has been successfully trained to recognize the undamaged system response as genuine, while the damaged system response is likely to be classified as fake, with varying degrees of deviation.

The application of both AE and GAN yields the results provided in Fig. 10. Since both architectures are applied individually to the dynamic response of each mass, Fig. 10 illustrates the reconstruction loss (in terms of MAE) obtained by AE and the discriminator output obtained by GAN for all time series samples of each of the two masses for a given damage level. Instead, the trend lines connect the values obtained by averaging the results calculated for both masses in order

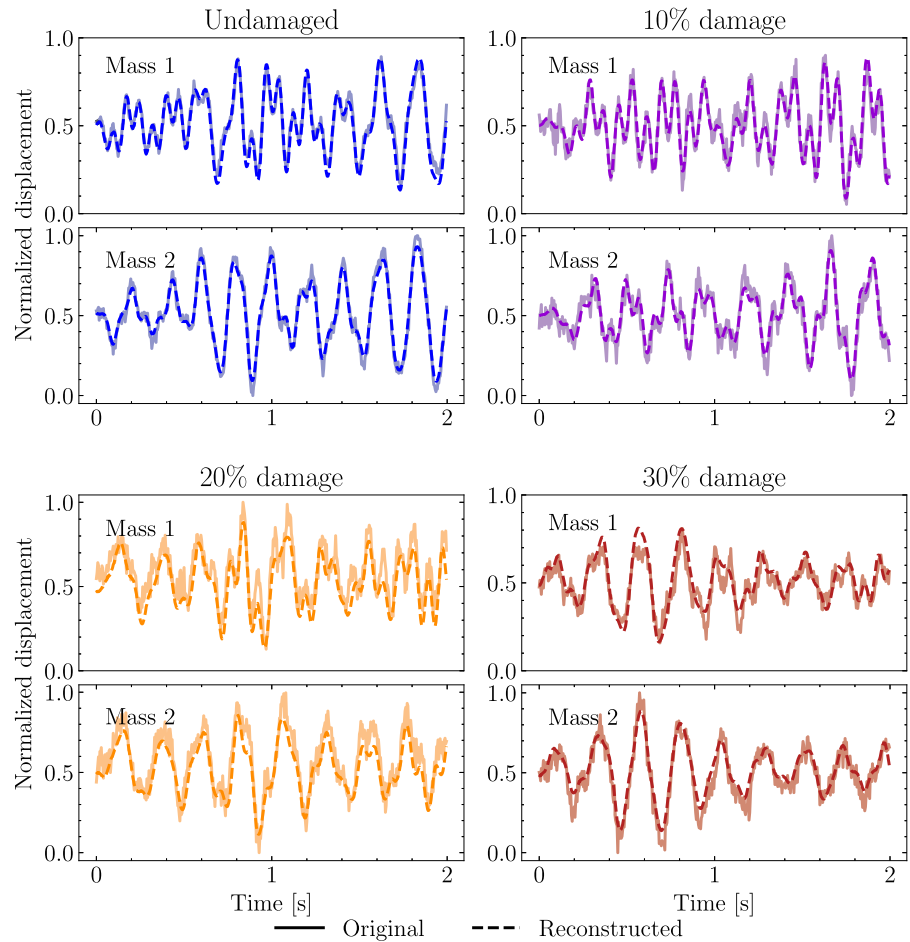
to enable the detection of damage through the inspection of a single parameter.

While the considered DL architectures displayed a similar performance for the 1-DOF system (see Fig. 6), Fig. 10 shows that AE is more sensitive to damage than GAN for this 2-DOF system. The lower damage sensitivity of GAN in this application seems to be associated with the reduced accuracy of its discriminator in correctly classifying the undamaged response of the 2-DOF system compared to that of the 1-DOF system. In fact, some discriminator outputs are somewhat far from 1 for the undamaged system, meaning that GAN produces a non-negligible number of false positives (i.e., the discriminator sometimes recognizes as fake the response of the undamaged system, which implies that it detects damage even when there is no damage). Anyway, Fig. 10 proves that both algorithms enable the early detection of damage and they perform as expected when the damage level increases.

#### 4.3 Seismic isolator with superelastic hysteresis and negative stiffness

Health monitoring of seismic isolation systems is crucial to ensure that the isolated structures preserve their target performance levels during earthquakes. Consequently, several technical codes and guidelines worldwide require inspections of seismic isolators, both regularly and past seismic events [11,57]. Visual inspections, whenever possible, can only detect damage or

**Fig. 8** Reconstruction of the 2-DOF system response through AE for different damage levels



aging effects when they are visible on the surface, and their reliability heavily depends on the inspectors' experience. Therefore, this numerical study investigates the feasibility of using DL architectures to detect damage in seismic isolators. A seismic isolator with superelastic hysteresis and negative stiffness proposed by Salvatore et al. [54] is here examined. The isolator mechanical model combines the Bouc-Wen model of hysteresis [7,62], the model of a negative stiffness mechanism, and a superelastic model proposed by Charalampakis and Tsiasas [12]. The equation of motion reads:

$$M\ddot{x} + f_r = F(t), \tag{14}$$

where the total restoring force of the seismic isolation device is given by  $f_r = f_i + f_n + f_s$ , in which  $f_i$  corresponds to the traditional elastomeric isolator restoring force,  $f_s$  is the superelastic force and  $f_n$  reflects the negative stiffness mechanism. The contribution to the

restoring force given by  $f_i$  is defined as follows:

$$f_i = C\dot{x} + \alpha K_i x + (1 - \alpha) K_i z, \tag{15a}$$

$$\dot{z} = \dot{x}[1 - (\gamma + \beta \text{sign}(z\dot{x}))]|z|^n, \tag{15b}$$

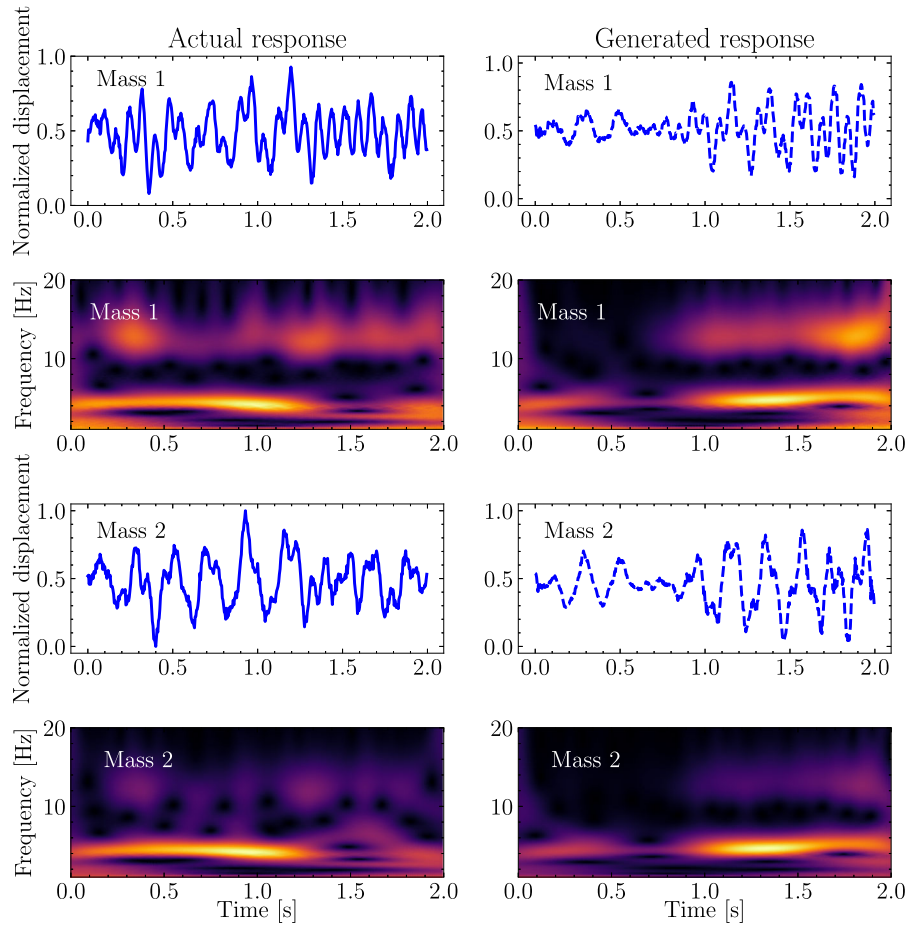
where  $z$  is the hysteretic force governed by  $\alpha$ , which is the ratio between the post-elastic and the stiffness  $K_i$  at the origin. Moreover,  $\gamma$  and  $\beta$  control the shape of the hysteresis and  $n$  regulates the smoothness of elastic-to-plastic transition. The contribution to the restoring force given by  $f_n$  is defined as follows:

$$f_n = \left(-K_n x + K_3 x^3\right) \frac{[1 + \text{sign}(x_f - |x|)]}{2}, \tag{16}$$

where  $K_n$  is the negative linear stiffness,  $K_3$  is the positive cubic stiffness, and  $x_f$  controls the extent of the negative stiffness force. The last contribution to the restoring force given by  $f_s$  is defined as follows [12]:

$$\begin{aligned} \dot{f}_s = & (1 - s)K_s \left[ \dot{x} - |\dot{x}| \text{sign}(f_s - \beta_s) \left( \frac{|f_s - \beta_s|}{Y} \right)^{n_s} \right] \\ & + sK_m \dot{x}, \end{aligned} \tag{17a}$$

**Fig. 9** Time history and corresponding time-frequency representation via wavelet transform of the undamaged 2-DOF system response: actual data and results carried out from the GAN generation block



$$\beta_s = K_s \alpha_s \left\{ x - \frac{f_s}{K_s} + f_t \tanh(a_s x) \left[ \frac{1 + \text{sign}(-x\dot{x})}{2} \right] \right\}, \tag{17b}$$

$$s = \frac{\tanh[c_s (|x| - x_m)] + 1}{2}, \quad f_t = \frac{(2Y - y_s Y)}{\alpha_s K_s}, \tag{17c}$$

$$a_s = \frac{\tan^{-1}(\tilde{a}_s K_s)}{Y - y_s Y},$$

where  $K_s$  and  $K_m$  are the initial stiffness during the austenitic phase and the fully martensitic phase, respectively, whereas  $Y$  is the yielding force and  $\alpha_s$  controls the post-elastic stiffness. The parameter  $n_s$  regulates the smoothness of the transition from the initial elastic to the post-elastic phase while  $f_t$  and  $a_s$  control the twinning hysteresis and super-elasticity and the pinching across the origin along the cycle, respectively. The numerical values of the system parameters in Eqs. (14)–(17) for this application are based on the previous study by Salvatore et al. [54] and are set to:  $M = 400$  kg,  $C = 0.4767$  kg/s,  $K_i = 1.1$  kN/mm,

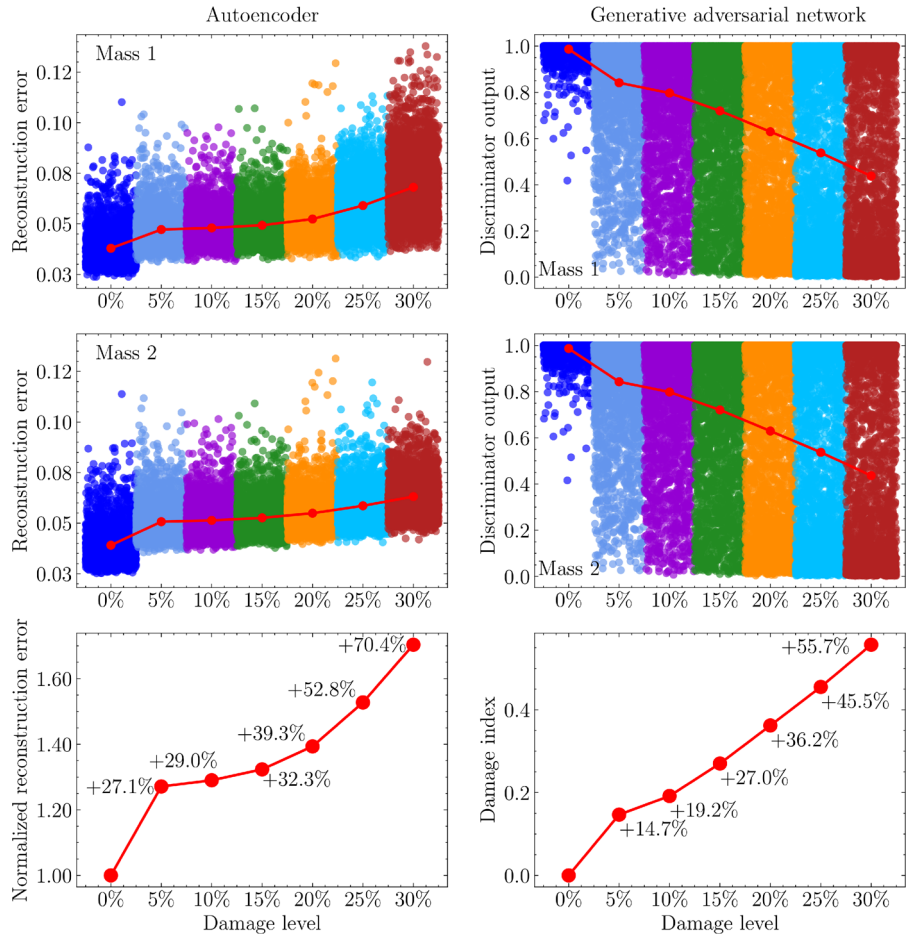
$K_n = 0.5$  kN/mm,  $K_3 = 0.00001$  kN/mm<sup>3</sup>,  $x_u = 100$  mm,  $x_m = 0.7x_u$ ,  $x_f = 0.7x_u$ ,  $\alpha = 0.2$ ,  $\beta = 0.0007$ ,  $c = 0.01$ ,  $\gamma = 0.00001$ ,  $n = 1$ , and  $n_s = 3$ . Damage is simulated by reducing both  $K_i$  and  $K_n$  by the same amount. The complex dynamics resulting from Eqs. (14)–(17) is a formidable challenge for testing the ability of AE and GAN to detect damage in nonlinear systems.

Both AE and GAN are assessed through numerical simulations by solving Eq. (14) under random excitation such that  $F = -M\ddot{x}_b$ , where the peak value of the time history samples of  $\ddot{x}_b$  are scaled to a uniform random value between 0.04g and 0.80g. Tables 5 and 6 shows the details about architecture and training process of both AE and GAN for this numerical application.

Representative results obtained with the application of AE and GAN are reported in Figs. 11 and 12, respectively. A close inspection of the results in Fig. 11



**Fig. 10** Damage sensitivity of the two DL architectures for the 2-DOF system: reconstruction loss values obtained by AE and relative variation of its normalized average value, together with discriminator output values obtained by GAN and relative variation of the associated average damage index



**Table 5** Configuration of the DL architectures for damage detection in the seismic isolator

Type	Component	Convolution layers	Kernel size	Number of filters	Strides
AE	Encoder	3	(20,10,5)	(50,50,50)	(1,1,1)
	Decoder	3	(5,10,20)	(50,50,1)	(2,2,1)
GAN	Generator	2	(2,8)	(2,1)	(2,2)
	Discriminator	3	(8,4,2)	(8,4,2)	(2,2,1)

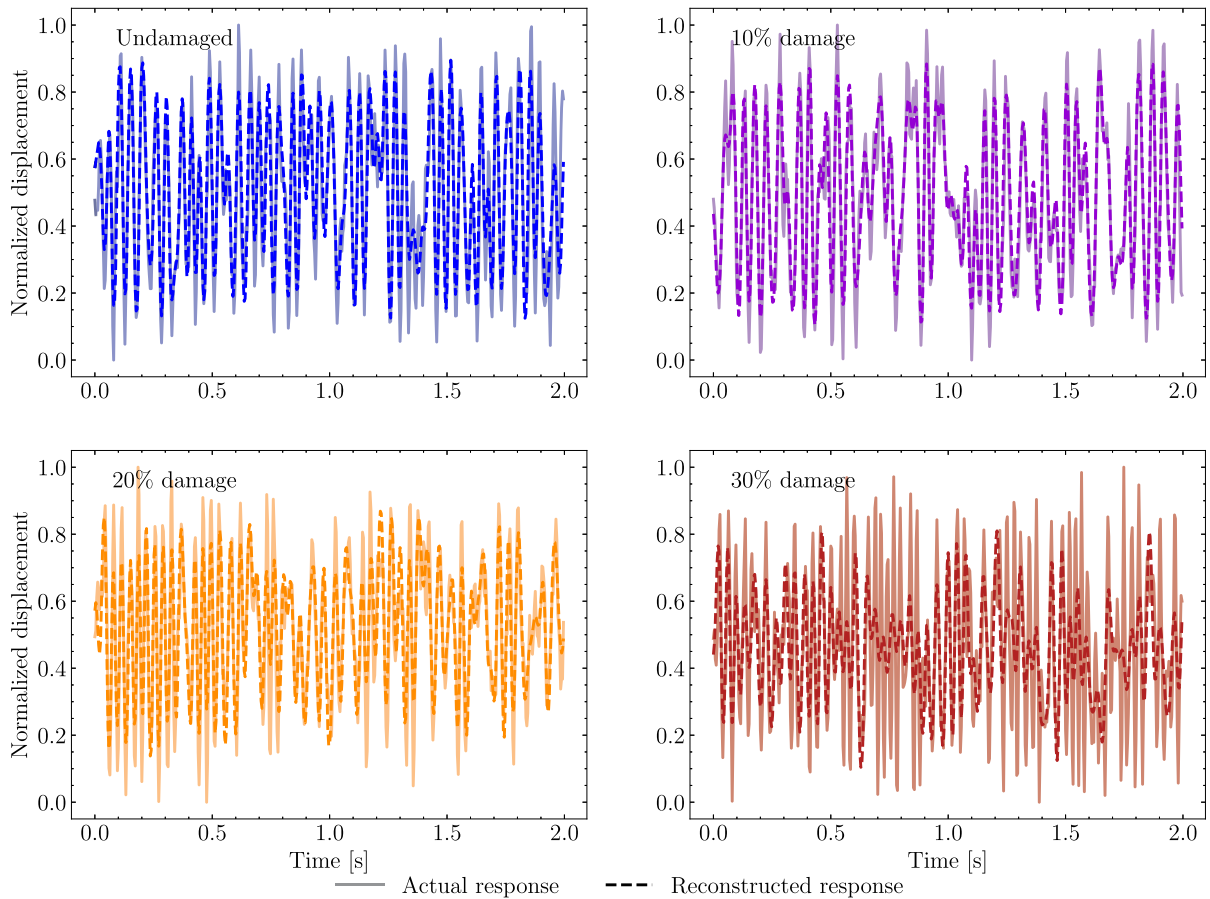
demonstrates that AE continues to reconstruct the undamaged response of the dynamic system quite satisfactorily while its performance degrades as expected when the level of damage increases. GAN faces some difficulties in this application, as it can be inferred from the comparison in the time-frequency domain between actual and generated response given in Fig. 12 (i.e., some frequency components in the actual response are missing in the output produced by the generator). Indeed, the discriminator sometimes mislabels data

from the generator as real, even though the comparison with the reference isolator response in the time-frequency domain exhibits some differences. This is due to the very rich dynamics given by Eqs. (14)–(17) and anticipates a lower sensitivity of GAN in detecting damage.

Results in Fig. 13 agree with the evidence drawn from Figs. 11 and 12. In fact, Fig. 13 shows that AE properly detects the presence of damage. The sensitivity of AE in this application remains satisfactory for

**Table 6** Training data of the DL architectures for damage detection in the seismic isolator

Type	Component	Trainable parameters	Epochs	Training time
AE	Encoder	38,650	191	0.51 min
	Decoder	791,596		
GAN	Generator	898	500	6.49 min
	Discriminator	100,623		

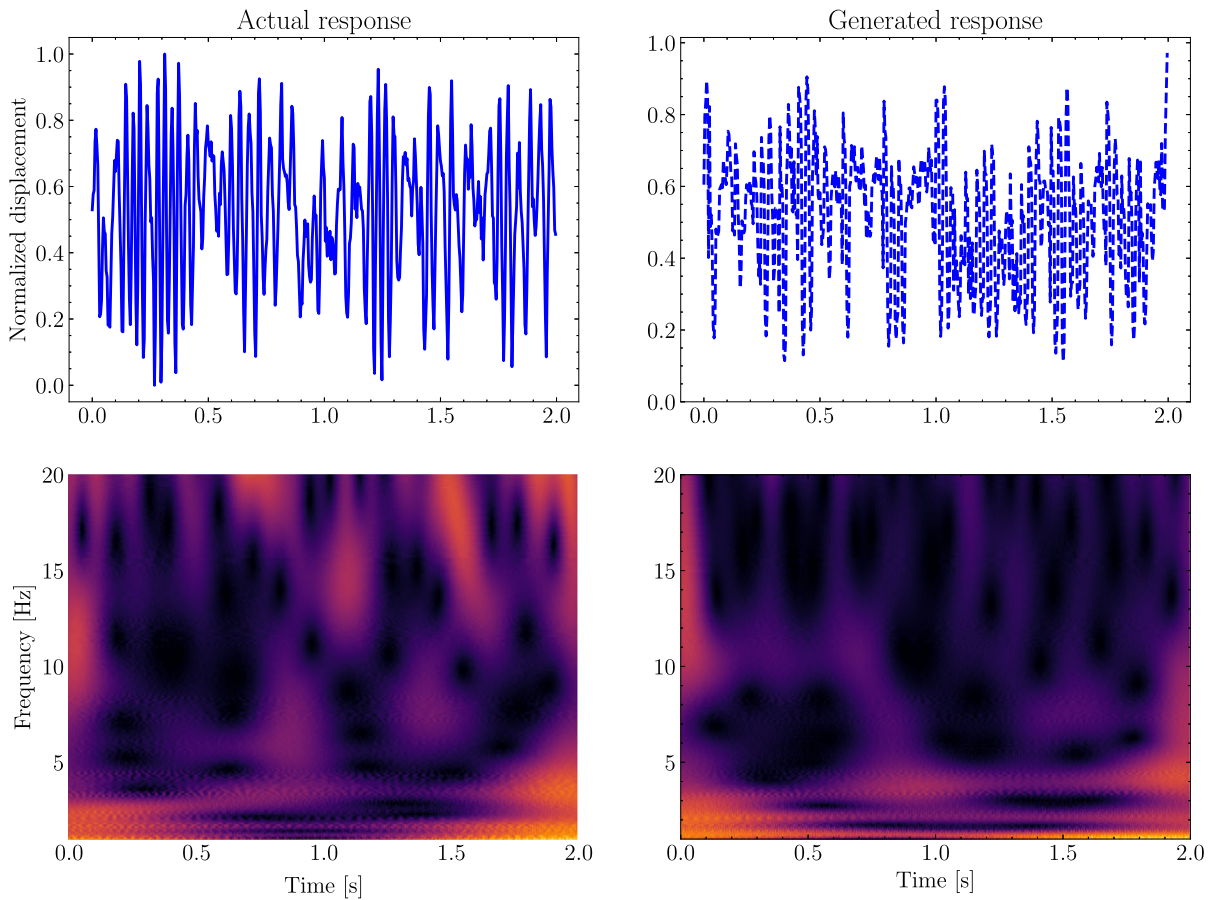
**Fig. 11** Reconstruction of the seismic isolator response through AE for different damage levels

all damage levels, but it is generally lower compared to Figs. 6 or 10. The presence of damage in the seismic isolator is also successfully recognized by GAN, provided that the damage severity is large enough (the relative variation of the damage index is around 10% once the damage level approaches 15%) while the output is almost inconclusive otherwise. This is mainly because the discriminator tends to produce more false negatives in this application (i.e., too many samples of the damaged response of the isolator are incorrectly

classified as true, meaning they are mistakenly considered as representative of a healthy state).

## 5 Experimental application

Both AE and GAN are finally tested using real data from a laboratory experiment reported by Shiki et al. [56] and Villani et al. [60]. The experimental setup is shown in Fig. 14 and consists of a  $300\text{ mm} \times 19\text{ mm}$



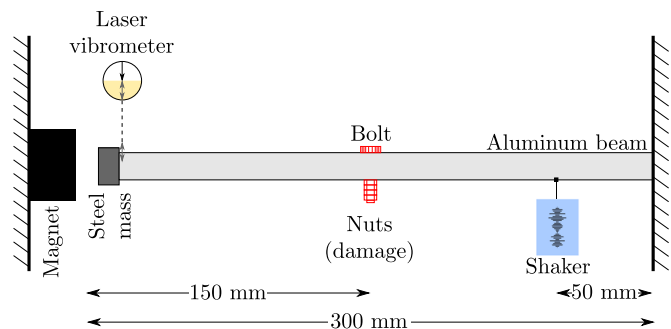
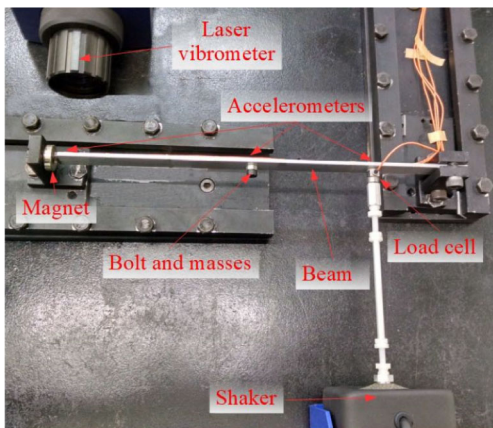
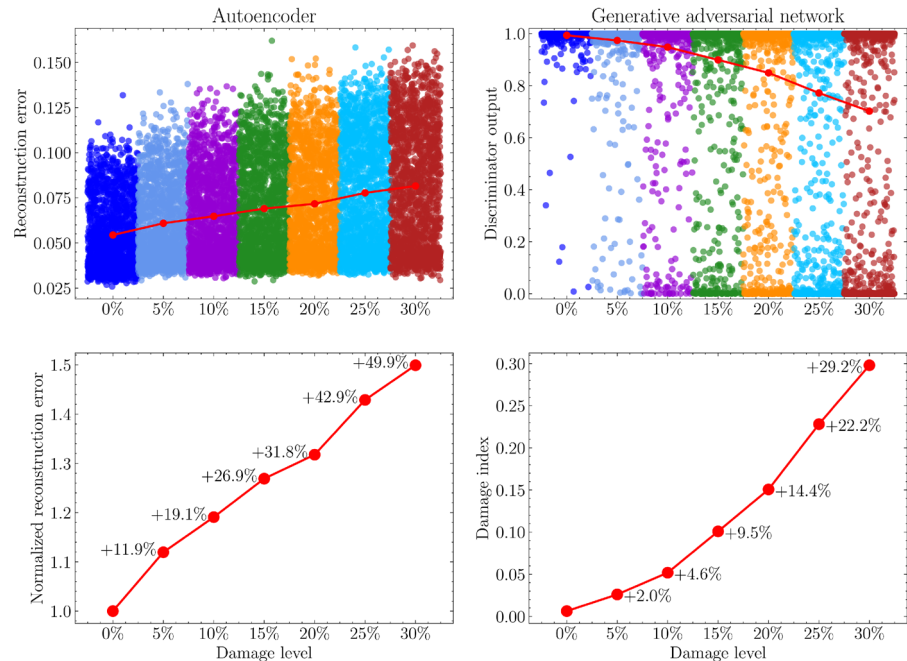
**Fig. 12** Time history and corresponding time-frequency representation via wavelet transform of the undamaged seismic isolator response: actual data and results carried out from the GAN generation block

× 3.2 mm cantilever aluminum beam with a small steel mass attached to its free end that interacts with a neodymium magnet positioned at a distance of 2 mm. A bolted connection provided with up to four nuts is positioned 150 mm from the free end and is introduced to simulate damage in the system through a change of the mass by varying the number of nuts. An electrodynamic shaker is placed 50 mm away from the clamped end of the beam whereas a laser vibrometer measures the velocity of the free end. A hardening-type nonlinearity originates in this magneto-elastic dynamic system from the interaction between the permanent magnet and the steel mass at the beam free end. While chirp and stepped sine were used as excitation by Shiki et al. [56] and Villani et al. [60], the present experimental application deals with output-only damage detection under random vibrations. Accordingly, the measured

system response under band-pass filtered white noise is considered. The filter bandwidth in the experiment was set to 10–420 Hz whereas the input excitation levels were equal to 0.01, 0.05, 0.10, 0.11, 0.12, 0.13, 0.14, and 0.15 VRMS. In this experimental application, the magneto-elastic system layout consisting of four nuts represents the baseline (undamaged) condition and the data corresponding to this configuration have been employed to train both AE and GAN. The damage datasets collect the experimental response of the system as each nut is removed sequentially: considering the mass of each nut and the number of nuts, the damage levels turn out to be equal to 2%, 4% and 6%.

Once again, the system dynamic response under different (unknown) excitation levels is taken into account to assess the effectiveness of the DL architectures in

**Fig. 13** Damage sensitivity of DL architectures for a seismic isolator: reconstruction loss values obtained by AE and relative variation of its normalized average value, together with discriminator output values obtained by GAN and relative variation of the associated average damage index



**Fig. 14** Layout of the experimental test consisting of a cantilever beam and a magnet close to the tip mass ([60], reprinted with permission)

detecting damage, regardless of the experienced level of nonlinearity. Configuration parameters and training data of AE and GAN adopted in such experimental application are given in Tables 7 and 8, respectively (training runtime refers to the same hardware adopted in the numerical investigations).

Figures 15 and 16 show the performances of both architectures after training while Fig. 17 demonstrates their effectiveness in detecting damage. On the one hand, the reconstruction error of AE displays an abrupt increment as soon as a small damage occurs. On the

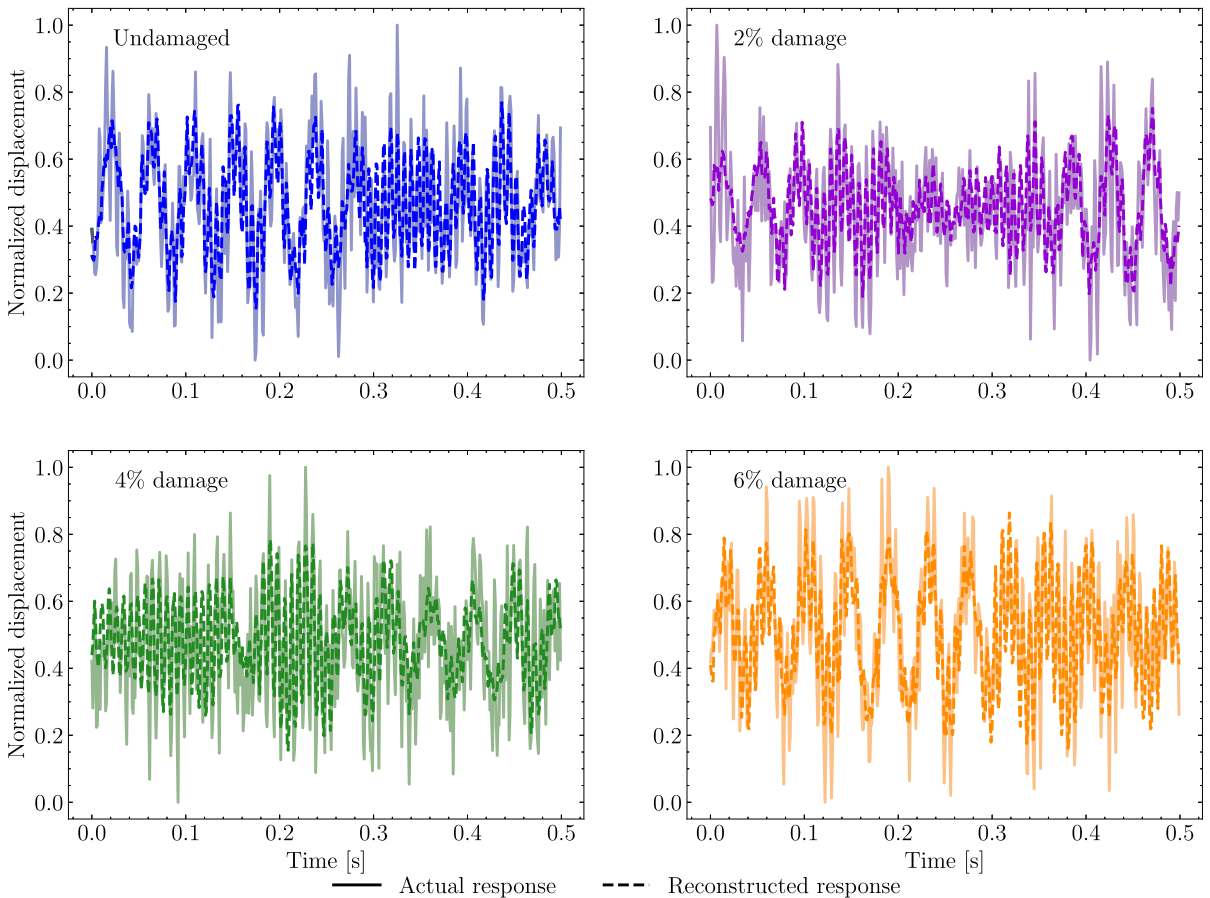
other hand, a growing number of time series data is classified as fake by the discriminator of GAN once a small damage occurs because they no longer represent the dynamics of a healthy system. The very good sensitivity of both architectures for the lowest damage level of 2% confirms their ability in early damage detection. In particular, the damage index based on the discriminator loss highlights the exceptional performance of GAN in recognizing the onset of damage in this experimental application. However, the relative variation of the normalized reconstruction loss and that of the damage

**Table 7** Configuration of the DL architectures for damage detection in the magneto-elastic system

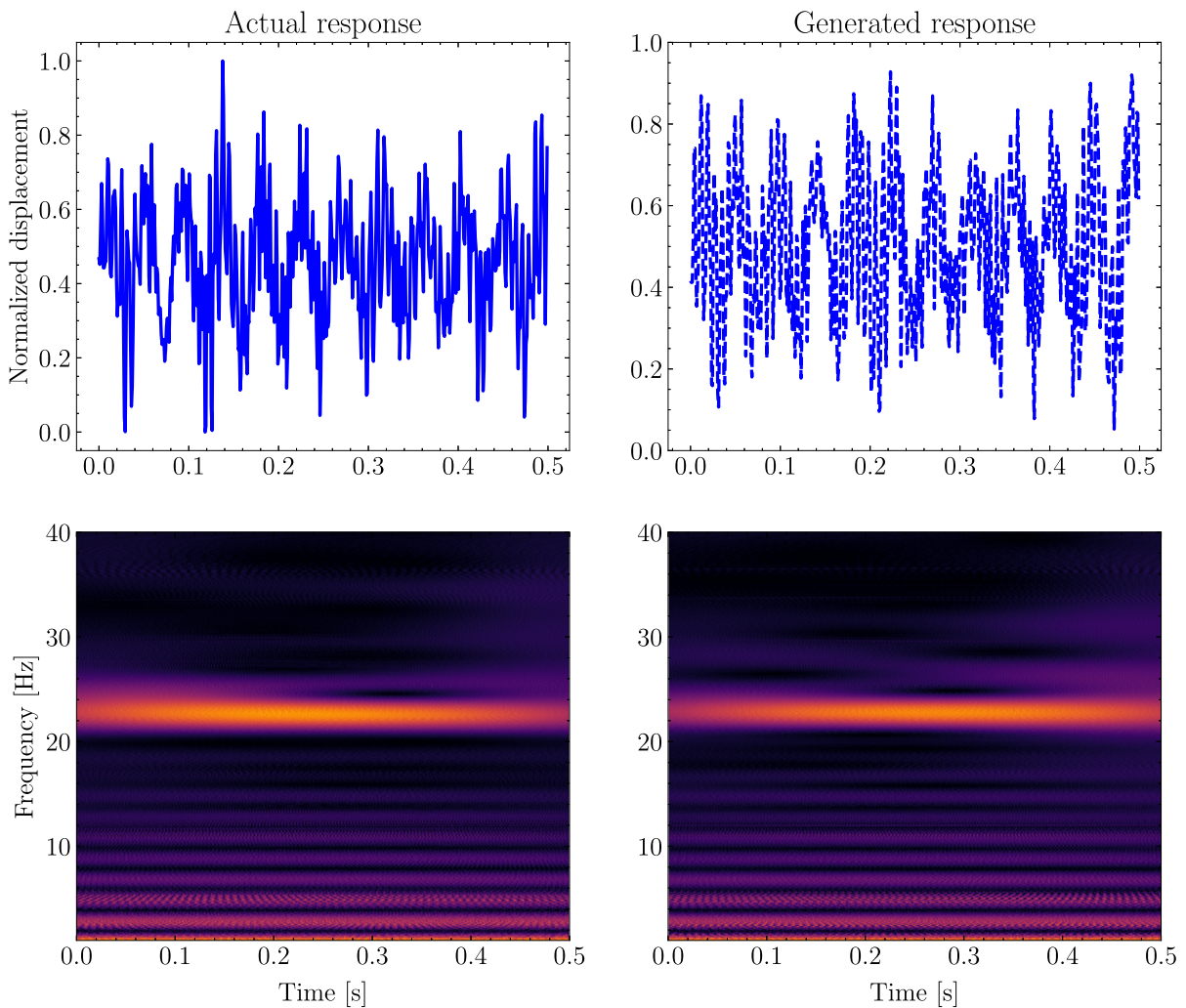
Type	Component	Convolution layers	Kernel size	Number of filters	Strides
AE	Encoder	4	(20,100,50,25)	(100,100,50,25)	(1,1,1,1)
	Decoder	3	(50,50,20)	(50,100,1)	(1,2,2)
GAN	Generator	3	(10,20,2)	(100,200,1)	(1,2,2)
	Discriminator	4	(3,50,15,10)	(200,200,100,50)	(1,1,1,1)

**Table 8** Training data of the DL architectures for damage detection in the magneto-elastic system

Type	Component	Trainable parameters	Epochs	Training time
AE	Encoder	1,283,525	29	0.46 min
	Decoder	527,679		
GAN	Generator	414,629	1,000	24.00 min
	Discriminator	2,357,201		



**Fig. 15** Reconstruction of the magneto-elastic system response through AE for different damage levels



**Fig. 16** Time history response and corresponding time-frequency representation via wavelet transform of the undamaged magneto-elastic system: actual data and results carried out from the GAN generation block

index based on the discriminator loss for damage levels of 4% and 6% are fairly close to the corresponding values observed for the minimum damage level equal to 2%. Ultimately, this experimental application further confirms that both AE and GAN are able to detect damage without a priori information about the system or the applied excitation signals, and irrespective of the degree of nonlinearity.

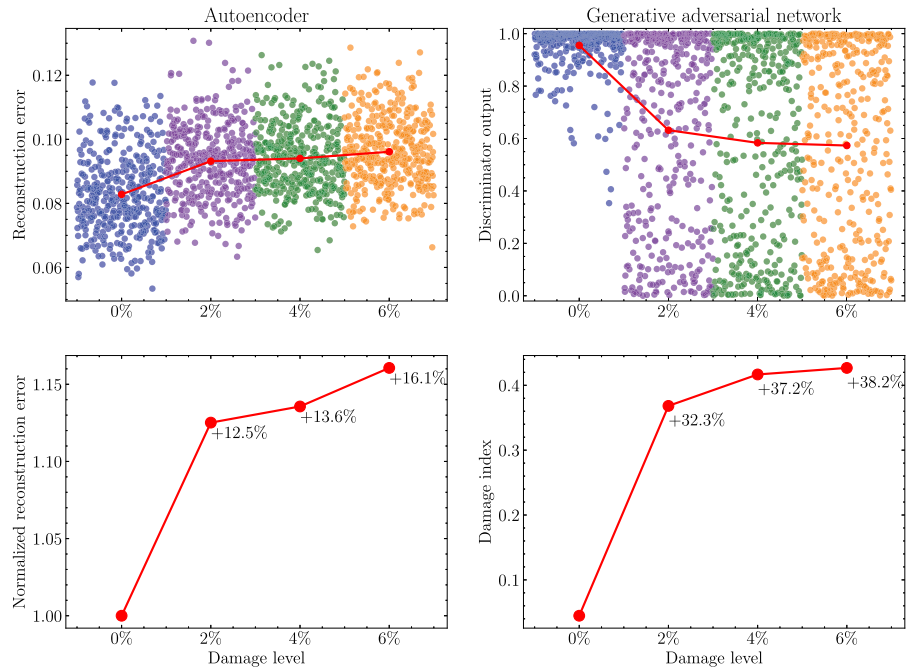
## 6 Conclusions

The original contribution of this work lies in the first comprehensive study on using DL architectures

for unsupervised, data-driven, vibration-based damage detection in nonlinear dynamic systems. The performance of both AE and GAN was deeply examined, considering different nonlinear behaviors. The excitation scenarios involve random excitation with varying intensity. Numerical investigations and experimental applications were carried out under the assumption that neither the system nor the excitation are known. The following conclusions can be drawn from the extensive study on different systems and datasets.

- Both AE and GAN are able to detect the onset of damage in different nonlinear dynamic systems provided that their architectures are properly

**Fig. 17** Damage sensitivity of DL architectures for the magneto-elastic system: reconstruction loss values obtained by AE and relative variation of its normalized average value, together with discriminator output values obtained by GAN and relative variation of the associated average damage index



designed and trained. They also show consistent results upon increasing damage levels.

- On the one hand, it seems that AE displays a more consistent performance than GAN. GAN did not consistently detect the lowest damage level; however, it can exhibit exceptional damage sensitivity. This is also influenced by the fact that the AE architecture conditioning proved easier than GAN. Moreover, training AE was less time consuming than GAN.
- Conversely, the main advantage of the GAN discriminator lies in its output being interpretable as damage index since it is bounded between two limit values corresponding to healthy and unhealthy state. While GAN is applied to detect the occurrence of damage, it can be possibly used for other tasks related to the undamaged system, such as synthetic data generation, uncertainty quantification, and reliability assessment.

In addition to extending both numerical and experimental applications, some issues should be addressed in future efforts, as listed below.

- Damage was introduced in the nonlinear systems as either a change in stiffness or mass. Comprehensive numerical and experimental investigations are necessary to confirm the effectiveness of these DL architectures across other damage scenarios.

- It is known that the effects attributable to environmental conditions can obscure those due to damage. Therefore, understanding whether these DL architectures can detect effectively damage in nonlinear dynamic systems under changing environmental conditions without preliminary data processing is still an open issue.
- Both AE and GAN were employed to establish whether the nonlinear dynamic system was damaged or not (i.e., novelty detection). Damage localization and quantification were not addressed. Further studies should aim to understand whether the output of such DL architectures can be eventually correlated with the damage position and/or magnitude.

**Acknowledgements** The authors would like to thank Americo Cunha (Rio de Janeiro State University, Brazil) and Samuel Da Silva (São Paulo State University, Brazil) for sharing the experimental data.

**Author contributions** H. J.: Writing—Original Draft, Methodology, Software, Validation, Formal analysis, Investigation, Data curation, Conceptualization. G.Q. and B.C.: Writing—Review & Editing, Supervision, Methodology, Software, Validation, Investigation, Data curation, Conceptualization. W. L.: Writing—Review & Editing, Supervision, Methodology, Investigation, Conceptualization, Project administration, Funding acquisition.

**Funding** Open access funding provided by Università degli Studi di Roma La Sapienza within the CRUI-CARE Agreement. This work was partially carried out within the RETURN Extended Partnership funded by the European Union Next-GenerationEU (National Recovery and Resilience Plan - NRRP, Mission 4, Component 2, Investment 1.3 - D.D. 1243 2/8/2022, PE0000005).

**Data availability** Data sets generated during the current study are available from the corresponding author on reasonable request.

## Declarations

**Conflict of interest** The authors declare no Conflict of interest.

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

## References

- Abadi, M., et al.: TensorFlow: Large-scale machine learning on heterogeneous systems. <https://www.tensorflow.org/>, software available from tensorflow.org (2015)
- Adams, D.E., Nataraju, M.: A nonlinear dynamical systems framework for structural diagnosis and prognosis. *Int. J. Eng. Sci.* **40**(17), 1919–1941 (2002)
- Avcı, O., Abdeljaber, O., Kiranyaz, S., Hussein, M., Gabbouj, M., Inman, D.J.: A review of vibration-based damage detection in civil structures: from traditional methods to machine learning and deep learning applications. *Mech. Syst. Signal Process.* **147**, 107077 (2021)
- Bhanja, S., Das, A.: Impact of data normalization on deep neural network for time series forecasting. *arXiv preprint arXiv:1812.05519* (2018)
- Boccagna, R., Bottini, M., Petracca, M., Di Giorgio, M., Amelio, A., Camata, G.: Unsupervised methods for railway bridge structural monitoring. In: 14th International Conference on Application of Statistics and Probability in Civil Engineering (2023)
- Bornn, L., Farrar, C.R., Park, G.: Damage detection in initially nonlinear systems. *Int. J. Eng. Sci.* **48**(10), 909–920 (2010)
- Bouc, R.: Forced vibrations of mechanical systems with hysteresis. In: *Proc. of the Fourth Conference on Nonlinear Oscillations*, Prague, 1967 (1967)
- Cabello-Solorzano, K., Ortigosa de Araujo, I., Peña, M., Correia, L., J Tallón-Ballesteros, A.: The impact of data normalization on the accuracy of machine learning algorithms: a comparative analysis. In: *International Conference on Soft Computing Models in Industrial and Environmental Applications*, Springer, pp 344–353 (2023)
- Cao, M., Sha, G., Gao, Y., Ostachowicz, W.: Structural damage identification using damping: a compendium of uses and features. *Smart Mater. Struct.* **26**(4), 043001 (2017)
- Carboni, B., Guruva, S.K., Catarci, S., Quaranta, G., Lacarbonara, W., et al.: Nonlinear normal modes for damage detection: Theoretical concepts and preliminary experimental validation. In: *International Conference Proceedings Series: The 8th World Conference On Structural Control And Monitoring (8WCSCM)* (2022)
- CEN/TC167: En 1337-10:2003 structural bearings - part 10: Inspection and maintenance. Tech. rep (2003)
- Charalampakis, A.E., Tsiatas, G.C.: A simple rate-independent uniaxial shape memory alloy (sma) model. *Front. Built Environ* **4**, 46 (2018)
- Civera, M., Zanotti Fragonara, L., Surace, C.: Using video processing for the full-field identification of backbone curves in case of large vibrations. *Sensors* **19**(10), 2345 (2019)
- Creswell, A., White, T., Dumoulin, V., Arulkumaran, K., Sengupta, B., Bharath, A.A.: Generative adversarial networks: an overview. *IEEE Signal Process. Mag.* **35**(1), 53–65 (2018)
- Doebbling, S.W., Farrar, C.R., Prime, M.B., et al.: A summary review of vibration-based damage identification methods. *Shock Vib. Dig.* **30**(2), 91–105 (1998)
- Döhler, M., Hille, F., Mevel, L., Rücker, W.: Structural health monitoring with statistical methods during progressive damage test of s101 bridge. *Eng. Struct.* **69**, 183–193 (2014)
- Fan, W., Qiao, P.: Vibration-based damage identification methods: a review and comparative study. *Struct. Health Monit.* **10**(1), 83–111 (2011)
- Fassois, S.D., Sakellariou, J.S.: Time-series methods for fault detection and identification in vibrating structures. *Philos. Transact. Royal Soc. A: Math. Phys. Eng. Sci.* **365**(1851), 411–448 (2007)
- Feng, J., Wang, X., Liu, Q., Li, Y., Xu, Y.: Deep learning-based parameter estimation of stochastic differential equations driven by fractional brownian motions with measurement noise. *Commun. Nonlinear Sci. Numer. Simul.* **127**, 107589 (2023)
- Glorot, X., Bordes, A., Bengio, Y.: Deep sparse rectifier neural networks. In: *Proceedings of the Fourteenth International Conference on Artificial Intelligence and Statistics, JMLR Workshop and Conference Proceedings*, pp 315–323 (2011)
- Goodfellow, I., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., Courville, A., Bengio, Y.: Generative adversarial nets. *Adv. Neural Inform. Process. Syst.* **27** (2014)
- Greš, S., Döhler, M., Andersen, P., Mevel, L.: Subspace-based mahalanobis damage detection robust to changes in excitation covariance. *Struct. Control. Health Monit.* **28**(8), e2760 (2021)
- Gui, J., Sun, Z., Wen, Y., Tao, D., Ye, J.: A review on generative adversarial networks: algorithms, theory, and appli-



- cations. *IEEE Trans. Knowl. Data Eng.* **35**(4), 3313–3332 (2021)
24. Hou, R., Xia, Y.: Review on the new development of vibration-based damage identification for civil engineering structures: 2010–2019. *J. Sound Vib.* **491**, 115741 (2021)
  25. Jin, Z., He, D., Lao, Z., Wei, Z., Yin, X., Yang, W.: Early intelligent fault diagnosis of rotating machinery based on iwoa-vmd and dmkelm. *Nonlinear Dyn.* **111**(6), 5287–5306 (2023)
  26. Karakostas, C., Quaranta, G., Chatzi, E., Zulfikar, A.C., Çetindemir, O., De Roeck, G., Döhler, M., Limongelli, M.P., Lombaert, G., Apaydin, N.M., et al.: Seismic assessment of bridges through structural health monitoring: a state-of-the-art review. *Bull. Earthq. Eng.* pp 1–49 (2023)
  27. Kim, H., Cheon, S., Jeong, I., Cho, H., Kim, H.: Enhanced model reduction method via combined supervised and unsupervised learning for real-time solution of nonlinear structural dynamics. *Nonlinear Dyn.* **110**(3), 2165–2195 (2022)
  28. Kingma, D.P., Ba, J.: Adam: A Method for Stochastic Optimization. <https://doi.org/10.48550/arXiv.1412.6980>, arXiv:1412.6980 (2017)
  29. Kiranyaz, S., Avci, O., Abdeljaber, O., Ince, T., Gabbouj, M., Inman, D.J.: 1d convolutional neural networks and applications: a survey. *Mech. Syst. Signal Process.* **151**, 107398 (2021)
  30. Kruse, R.M., Säfken, B., Silbersdorff, A., Weisser, C.: *Learning Deep Textwork*. Universitätsverlag Göttingen (2021)
  31. Kutz, J.N., Brunton, S.L.: Parsimony as the ultimate regularizer for physics-informed machine learning. *Nonlinear Dyn.* **107**(3), 1801–1817 (2022)
  32. Lacarbonara, W.: *Nonlinear structural mechanics: theory, dynamical phenomena and modeling*. Springer Science & Business Media (2013)
  33. Lacarbonara, W., Carboni, B., Quaranta, G.: Nonlinear normal modes for damage detection. *Meccanica* **51**, 2629–2645 (2016)
  34. Lai, Y., Chen, J., Hong, Q., Li, Z., Liu, H., Lu, B., Ma, R., Yu, C., Sun, R., Demartino, C., et al.: Framework for long-term structural health monitoring by computer vision and vibration-based model updating. *Case Studies in Construction Materials* **16**, e01020 (2022)
  35. Lakshmi, K., Rao, A.R.M., Gopalakrishnan, N.: Singular spectrum analysis combined with armax model for structural damage detection. *Struct. Control. Health Monit.* **24**(9), e1960 (2017)
  36. Li, L., Morgantini, M., Betti, R.: Structural damage assessment through a new generalized autoencoder with features in the quefreny domain. *Mech. Syst. Signal Process.* **184**, 109713 (2023)
  37. Li, P., Pei, Y., Li, J.: A comprehensive survey on design and application of autoencoder in deep learning. *Appl. Soft Comput.* **138**, 110176 (2023)
  38. Liu, Q., Xu, Y., Li, Y.: Complex dynamics of a conceptual airfoil structure with consideration of extreme flight conditions. *Nonlinear Dyn.* **111**(16), 14991–15010 (2023)
  39. Luleci, F., Catbas, F.N., Avci, O.: Cyclegan for undamaged-to-damaged domain translation for structural health monitoring and damage detection. *Mech. Syst. Signal Process.* **197**, 110370 (2023)
  40. Luo, Y., Guo, X., Lk, Wang, JI, Zheng, JI, Liu, Fy, Liao: Unsupervised structural damage detection based on an improved generative adversarial network and cloud model. *J. Low Freq. Noise, Vib. Active Control* **42**(3), 1501–1518 (2023)
  41. Ma, X., Lin, Y., Nie, Z., Ma, H.: Structural damage identification based on unsupervised feature-extraction via variational auto-encoder. *Measurement* **160**, 107811 (2020)
  42. Maas, A.L., Hannun, A.Y., Ng, A.Y., et al.: Rectifier nonlinearities improve neural network acoustic models. In: *Proc. icml, Atlanta, GA*, vol 30, p 3 (2013)
  43. Pandey, A., Biswas, M., Samman, M.: Damage detection from changes in curvature mode shapes. *J. Sound Vib.* **145**(2), 321–332 (1991)
  44. Pathirage, C.S.N., Li, J., Li, L., Hao, H., Liu, W., Ni, P.: Structural damage identification based on autoencoder neural networks and deep learning. *Eng. Struct.* **172**, 13–28 (2018)
  45. Peng, Z., Li, J., Hao, H., Li, C.: Nonlinear structural damage detection using output-only volterra series model. *Struct. Control. Health Monit.* **28**(9), e2802 (2021)
  46. Prawin, J., Rao, A.R.M.: Damage detection in nonlinear systems using an improved describing function approach with limited instrumentation. *Nonlinear Dyn.* **96**, 1447–1470 (2019)
  47. Qi, G.J., Luo, J.: Small data challenges in big data era: a survey of recent progress on unsupervised and semi-supervised methods. *IEEE Trans. Pattern Anal. Mach. Intell.* **44**(4), 2168–2187 (2020)
  48. Quaranta, G., Carboni, B., Lacarbonara, W.: Damage detection by modal curvatures: numerical issues. *J. Vib. Control* **22**(7), 1913–1927 (2016)
  49. Quaranta, G., Lacarbonara, W., Masri, S.F.: A review on computational intelligence for identification of nonlinear dynamical systems. *Nonlinear Dyn.* **99**(2), 1709–1761 (2020)
  50. Ratliff, L.J., Burden, S.A., Sastry, S.S.: Characterization and computation of local nash equilibria in continuous games. In: *2013 51st Annual Allerton Conference on Communication, Control, and Computing (Allerton)*, IEEE, pp 917–924 (2013)
  51. Rumelhart, D.E., Hinton, G.E., Williams, R.J.: Learning internal representations by error propagation, parallel distributed processing, explorations in the microstructure of cognition, ed. de rumelhart and j. mccllelland. vol. 1. 1986. *Biometrika* **71**, 599–607 (1986)
  52. Rytter, A.: *Vibrational based inspection of civil engineering structures*. PhD thesis, Aalborg University (1993)
  53. Salawu, O.S., Williams, C.: Bridge assessment using forced-vibration testing. *J. Struct. Eng.* **121**(2), 161–173 (1995)
  54. Salvatore, A., Carboni, B., Lacarbonara, W.: Nonlinear dynamic response of an isolation system with superelastic hysteresis and negative stiffness. *Nonlinear dynamics* pp 1–26 (2021)
  55. Shang, Z., Sun, L., Xia, Y., Zhang, W.: Vibration-based damage detection for bridges by deep convolutional denoising autoencoder. *Struct. Health Monit.* **20**(4), 1880–1903 (2021)
  56. Shiki, S.B., Da Silva, S., Todd, M.D.: On the application of discrete-time volterra series for the damage detection problem in initially nonlinear systems. *Struct. Health Monit.* **16**(1), 62–78 (2017)

57. Sobanjo, J.O., Thompson, P.D.: Implementation of the 2013 Aashto Manual for Bridge Element Inspection. Tech. rep, Florida Department of Transportation, Florida, United States (2016)
58. Sohn, H., Farrar, C.R.: Damage diagnosis using time series analysis of vibration signals. *Smart Mater. Struct.* **10**(3), 446 (2001)
59. Villani, L.G., Da Silva, S., Cunha, A., Jr.: Damage detection in uncertain nonlinear systems based on stochastic volterra series. *Mech. Syst. Signal Process.* **125**, 288–310 (2019)
60. Villani, L.G., Da Silva, S., Cunha, A., Jr., Todd, M.D.: Damage detection in an uncertain nonlinear beam based on stochastic volterra series: an experimental application. *Mech. Syst. Signal Process.* **128**, 463–478 (2019)
61. Wang, X., Feng, J., Liu, Q., Xu, Y.: Noise-induced alternations and data-driven parameter estimation of a stochastic perceptual model. *The European Physical Journal Special Topics* pp 1–13 (2024)
62. Wen, Y.K.: Method for random vibration of hysteretic systems. *J. Eng. Mech. Div.* **102**(2), 249–263 (1976)
63. Worden, K., Farrar, C.R., Manson, G., Park, G.: The fundamental axioms of structural health monitoring. *Proc. Royal Soc. A: Math. Phys. Eng. Sci.* **463**(2082), 1639–1664 (2007)
64. Worden, K., Farrar, C.R., Haywood, J., Todd, M.: A review of nonlinear dynamics applications to structural health monitoring. *Struct. Control Health Monit.: Offic. J. Int. Assoc. Struct. Control Monit. Eur. Assoc. Control Struct.* **15**(4), 540–567 (2008)
65. Xia, X., Pan, X., Li, N., He, X., Ma, L., Zhang, X., Ding, N.: Gan-based anomaly detection: a review. *Neurocomputing* **493**, 497–535 (2022)
66. Yang, S., Cui, Z., Gu, X.: A balanced deep transfer network for bearing fault diagnosis. *IEEE Trans. Instrum. Meas.* **72**, 3530212 (2023)
67. Yang, Z., Xu, B., Luo, W., Chen, F.: Autoencoder-based representation learning and its application in intelligent fault diagnosis: a review. *Measurement* **189**, 110460 (2022)
68. Yu, Y., Liu, Y.: Physics-guided generative adversarial network for probabilistic structural system identification. *Expert Syst. Appl.* **239**, 122339 (2024)
69. Zhang, M., Guo, T., Zhu, R., Zong, Y., Liu, Z., Xu, W.: Damage identification of seismic-isolated structure based on cae network using vibration monitoring data. *Eng. Struct.* **283**, 115873 (2023)

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.