A Generative Adversarial Network Based Autoencoder for Structural Health Monitoring

Giorgia Colombera [1,*], Luca Rosafalco [1], Matteo Torzoni [1], Filippo Gatti [2], Stefano Mariani [1], Andrea Manzoni [3] and Alberto Corigliano [1]

[1] Dipartimento di Ingegneria Civile ed Ambientale, Politecnico di Milano, Piazza L. da Vinci 32, 20133 - Milano (Italy);
[2] Université Paris Saclay - CNRS - CentraleSupélec, Laboratoire MSSMat UMR 8579, 3 rue Joliot Curie, 91190 - Gif sur Yvette (France);
[3] Dipartimento di Matematica, Politecnico di Milano, Piazza L. da Vinci 32, 20133 - Milano (Italy).

* Corresponding author: giorgia.colombera@mail.polimi.it

Abstract: Civil structures, infrastructures and lifelines are constantly threatened by natural hazards and climate change. Therefore, Structural Health Monitoring (SHM) has become an active field of research for online structural damage detection and long-term maintenance planning. In this work we propose a new SHM approach leveraging a deep Generative Adversarial Network (GAN), trained on synthetic time histories representing the structural responses of both damaged and undamaged multistory building to earthquake ground motion. In the prediction phase, the GAN generates plausible signals for different damage states, based only on undamaged recorded or simulated structural responses, thus without the need to rely upon real recordings linked to damaged conditions.

Keywords: Structural Health Monitoring; Machine Learning; Generative Adversarial Network.

IOCA 2021

Methods: using RepGAN architecture [1], a semantically meaningful and disentangled representation of the SHM time-histories is learned. The original model joins VAE [2] and InfoGAN [3]:

- $x^{(i)} \rightarrow \hat{z}^{(i)} = \boldsymbol{F}_{\boldsymbol{\theta}_X}\left(x^{(i)}\right) \rightarrow \hat{x}^{(i)} = \boldsymbol{G}_{\boldsymbol{\theta}_Z} \circ \boldsymbol{F}_{\boldsymbol{\theta}_X}\left(x^{(i)}\right)$ provides high reconstruction quality;

- $z^{(i)} \rightarrow \hat{x}^{(i)} = \boldsymbol{F}_{\boldsymbol{\theta}_X}\left(z^{(i)}\right) \rightarrow \hat{z}^{(i)} = \boldsymbol{G}_{\boldsymbol{\theta}_Z} \circ \boldsymbol{F}_{\boldsymbol{\theta}_X}\left(z^{(i)}\right)$ guarantees impressive generation and clustering performance,

where $x^{(i)}$ is a data instance, $z^{(i)}$ represents a data latent instance, $\boldsymbol{F}_{\boldsymbol{\theta}_X}$ is the encoder and $\boldsymbol{G}_{\boldsymbol{\theta}_Z}$ is the decoder. The two learning tasks are performed together with the aim of obtaining the bijective mapping $x \leftrightarrow z$.

[1] Zhou, Y., Gu, K., Huang, T. (2019, July). Unsupervised Representation Adversarial Learning Network: from Reconstruction to Generation. 2019 International Joint Conference on Neural Networks (IJCNN). 2019 International Joint Conference on Neural Networks (IJCNN). https://doi.org/10.1109/ijcnn.2019.8852395
[2] X. Chen; Y. Duan; R. Houthooft; J. Schulman; I. Sutskever; P. Abbeel. InfoGAN: Interpretable Representation Learning by Information Maximizing Generative Adversarial Nets, Proceedings of the 30th International Conference on Neural Information Processing Systems, Barcelona, Spain, Curran Associates Inc.: Red Hook, NY, USA. arXiv:1606.03657 2016.
[3] D. P. Kingma; M. Welling. Auto-Encoding Variational Bayes. arXiv:1312.6114, 2013.

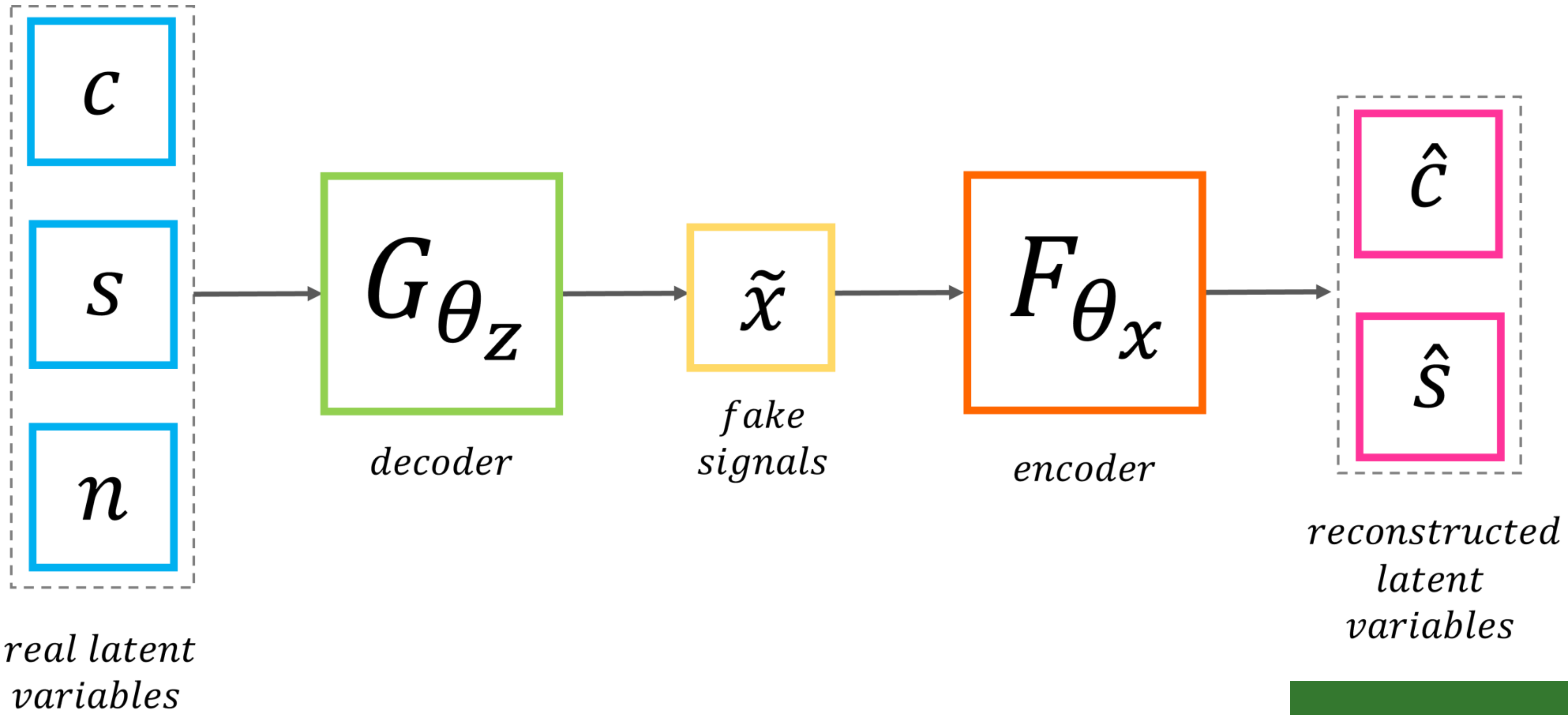
IOCA 2021

The latent space $\boldsymbol{Z} = [\boldsymbol{C}, \boldsymbol{S}, \boldsymbol{N}]$ is composed by:

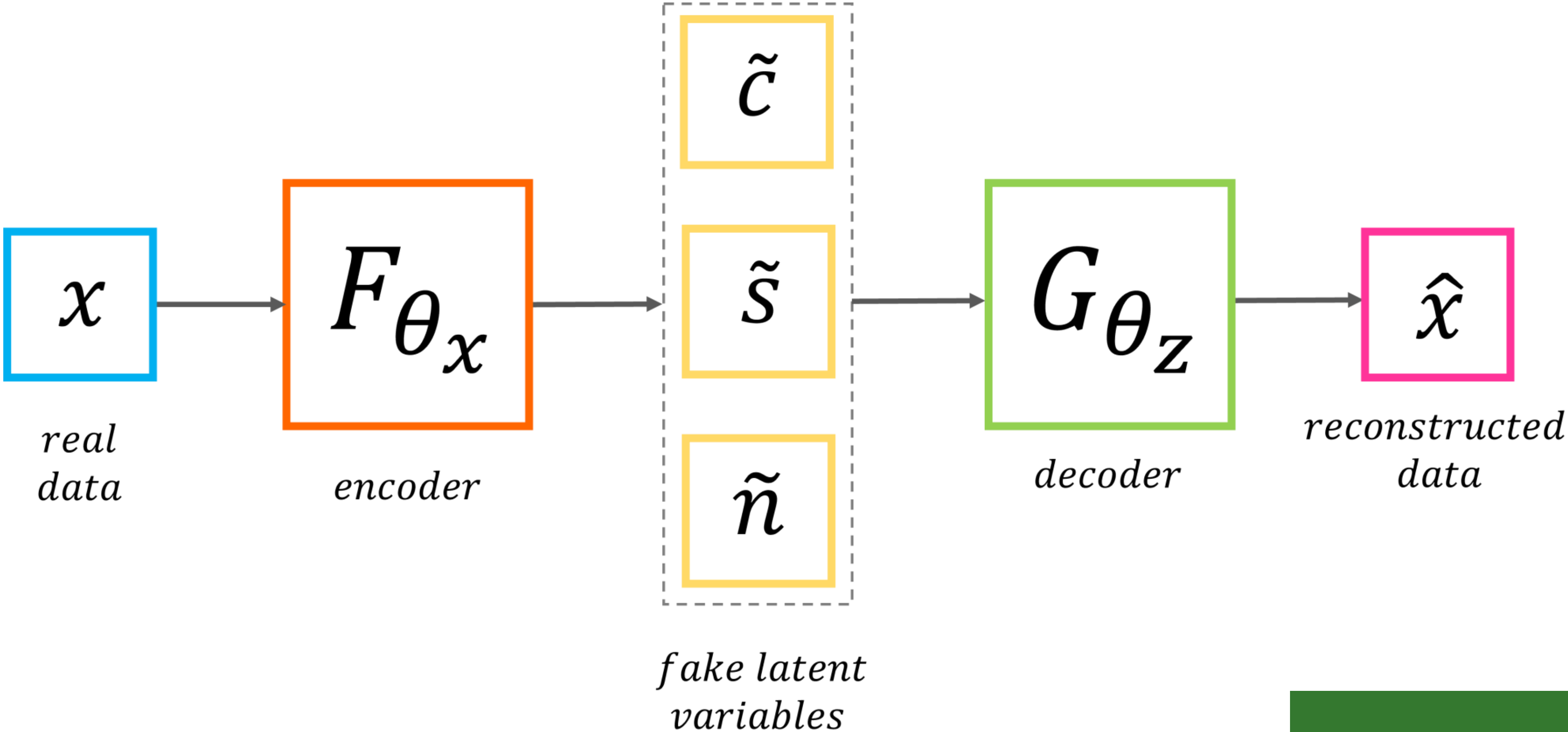
- a categorical variable representing the damage classes $\boldsymbol{C} \in [0,1]^{d_C}$, with $\boldsymbol{C} \sim \mathrm{Cat}(\mathrm{d_C})$ and $\mathrm{d_C}$ the number of classes;
- a continuous variable $\boldsymbol{S} \in \mathbb{R}^{d_S}$, with $\boldsymbol{S} \sim \mathrm{N}(0, \mathbb{I})$;
- a random noise $\boldsymbol{N} \in \mathbb{R}^{d_N}$, with $\boldsymbol{N} \sim \mathrm{N}(0, \mathbb{I})$.

RepGAN model is composed by a generative part and a discriminative part. The former is constituted by an encoder $\boldsymbol{F}_{\boldsymbol{\theta}_X}$ and a decoder $\boldsymbol{G}_{\boldsymbol{\theta}_Z}$, while the latter comprises the discriminators for time-histories $\boldsymbol{D}_{\boldsymbol{\omega}_X}$ and for latent variables $\boldsymbol{D}_{\boldsymbol{\omega}_C}$, $\boldsymbol{D}_{\boldsymbol{\omega}_S}$ and $\boldsymbol{D}_{\boldsymbol{\omega}_N}$. To achieve high classification accuracy, a classifier $\boldsymbol{D}_{\boldsymbol{\omega}_{Class}}$ is introduced: taking as input the signals, it predicts the corresponding damage class.

IOCA 2021

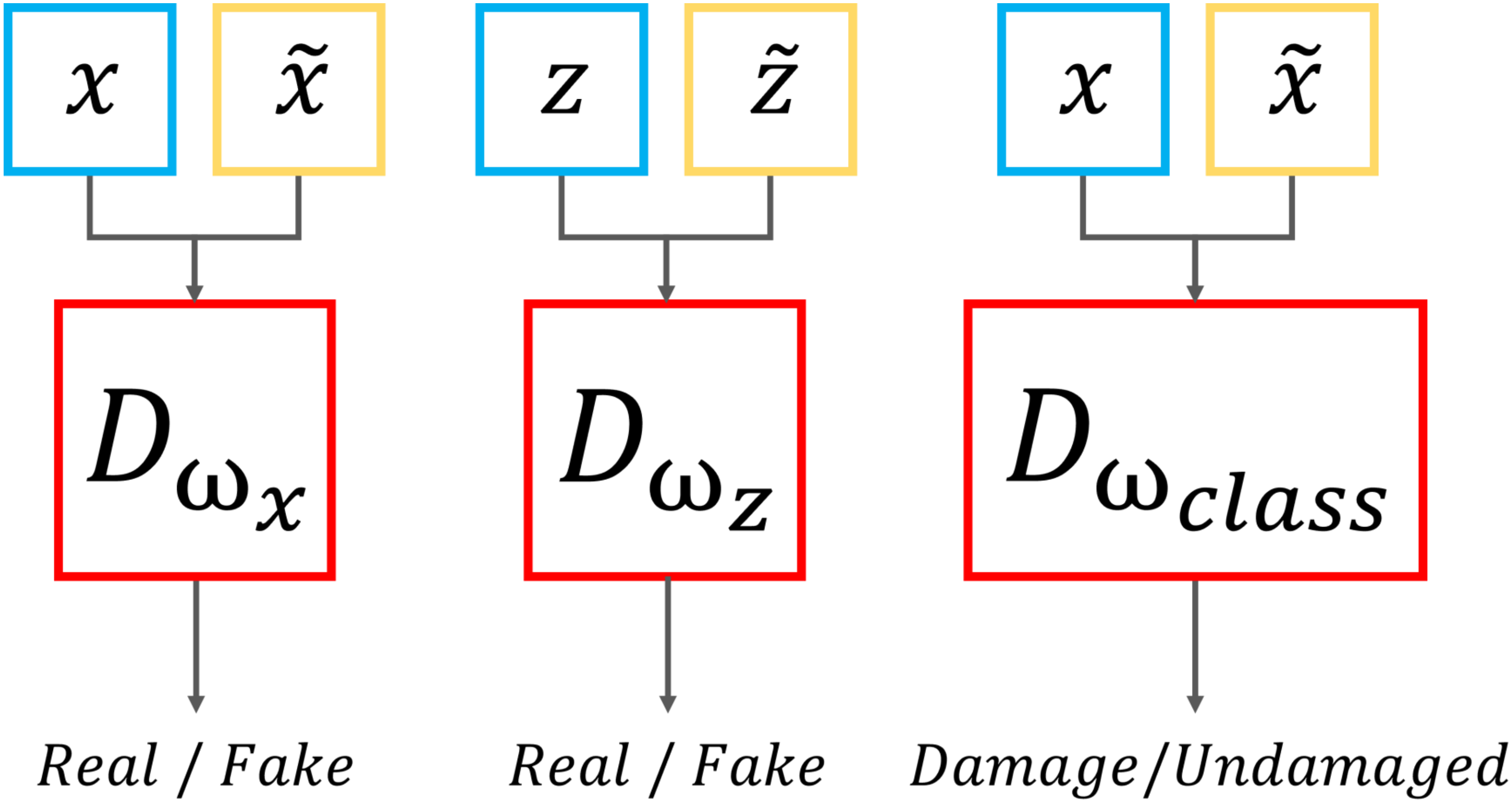

c
s
n
real latent variables
G_{θ_z}
decoder
$\tilde{x}$
fake signals
F_{θ_x}
encoder
$\hat{c}$
$\hat{s}$
reconstructed latent variables

IOCA 2021



x
real data
F_{θ_x}
encoder
$\tilde{c}$
$\tilde{s}$
$\tilde{n}$
fake latent variables
G_{θ_z}
decoder
$\hat{x}$
reconstructed data

IOCA 2021


x
$\tilde{x}$
z
$\tilde{z}$
x
$\tilde{x}$
D_{ω_x}
D_{ω_z}
$D_{\omega_{class}}$
Real / Fake
Real / Fake
Damage/Undamaged

Through a symmetrical adversarial process, it is possible to learn the bidirectional mapping between the input space and the latent space. The Empirical Loss function is as follows:

$$L_S = \mathbb{D}_{JS}\left(p_{\hat{X}|(C,S,N)} \parallel p_X\right) + \mathbb{D}_{JS}\left(q_{\hat{C}|X} \parallel p_C\right) + \mathbb{D}_{JS}\left(q_{\hat{S}|X} \parallel p_S\right) + \mathbb{D}_{JS}\left(q_{\hat{N}|X} \parallel p_N\right) -$$

$$- \mathbb{E}_{p_C}\left[\mathbb{E}_{p_{\hat{X}|C}}\left[\ln q_{\hat{C}|X}\right]\right] - \mathbb{E}_{p_S}\left[\mathbb{E}_{p_{\hat{X}|S}}\left[\ln q_{\hat{S}|X}\right]\right] - \mathbb{E}_{p_X}\left[\mathbb{E}_{q_{(C,S,N)|X}}\left[\ln p_{X|(C,S,N)}\right]\right]$$

where:

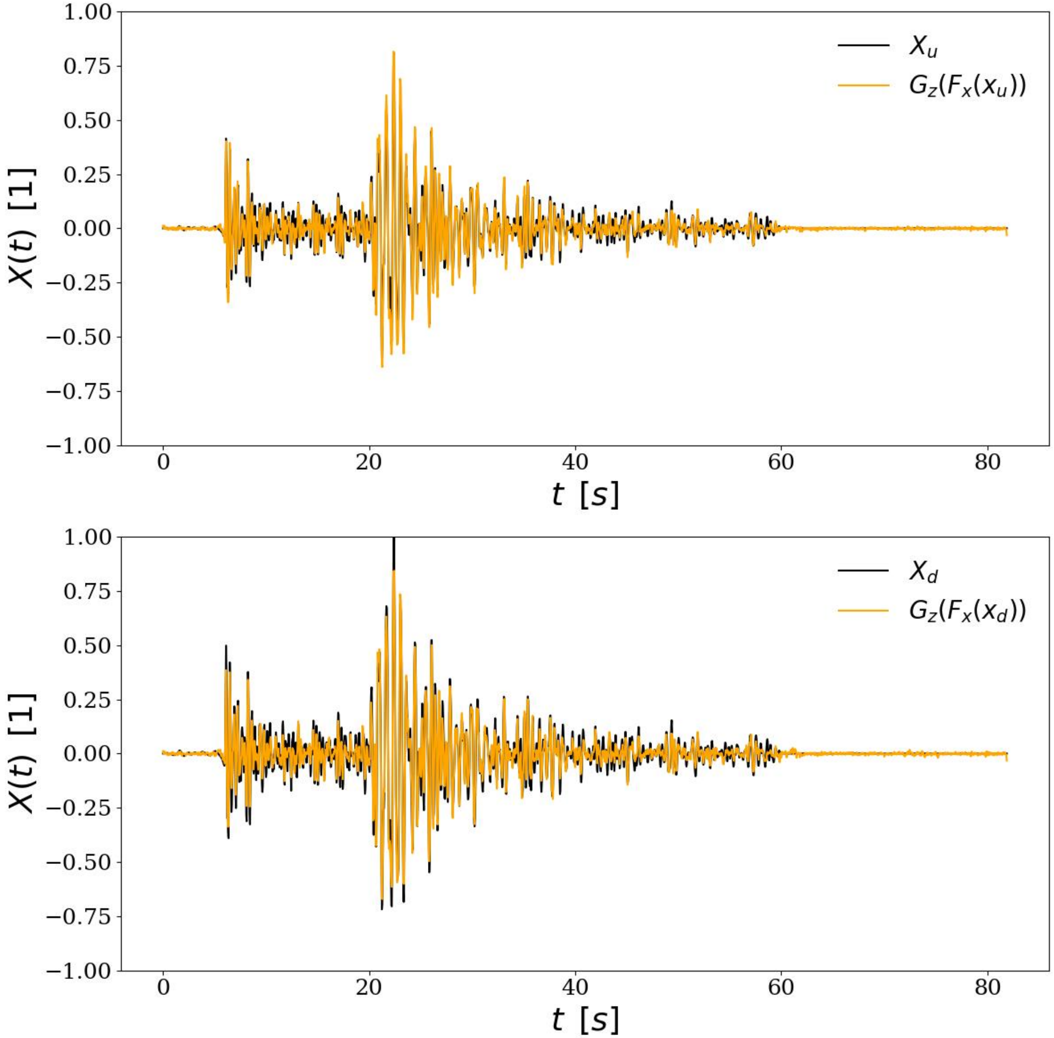
- $- \mathbb{E}_{p_C}\left[\mathbb{E}_{p_{\hat{X}|C}}\left[\ln q_{\hat{C}|X}\right]\right]$ minimizes the conditional entropy $\mathbb{S}(\boldsymbol{C}|\boldsymbol{X})$;
- $- \mathbb{E}_{p_S}\left[\mathbb{E}_{p_{\hat{X}|S}}\left[\ln q_{\hat{S}|X}\right]\right]$ minimizes the conditional entropy $\mathbb{S}(\boldsymbol{S}|\boldsymbol{X})$;
- $- \mathbb{E}_{p_X}\left[\mathbb{E}_{q_{(C,S,N)|X}}\left[\ln p_{X|(C,S,N)}\right]\right]$ minimizes the conditional entropy $\mathbb{S}(\boldsymbol{X}|(\boldsymbol{C},\boldsymbol{S},\boldsymbol{N})$.

IOCA 2021

Results and Discussion: the case study considered to assess the capability of the proposed architecture to achieve the three tasks of semantic generation, clustering and reconstruction is a 39 storeys shear building subject to an earthquake ground motion. These signals are taken from the STEAD seismic database [4]. The mass and the stiffness of each floor, in undamaged conditions, are respectively $m = 625 \cdot 10^3\ kg$ and $\mathrm{k} = 8{,}33 \cdot 10^7\ \frac{kN}{m}$. Damage is simulated through a 50% reduction in stiffness.

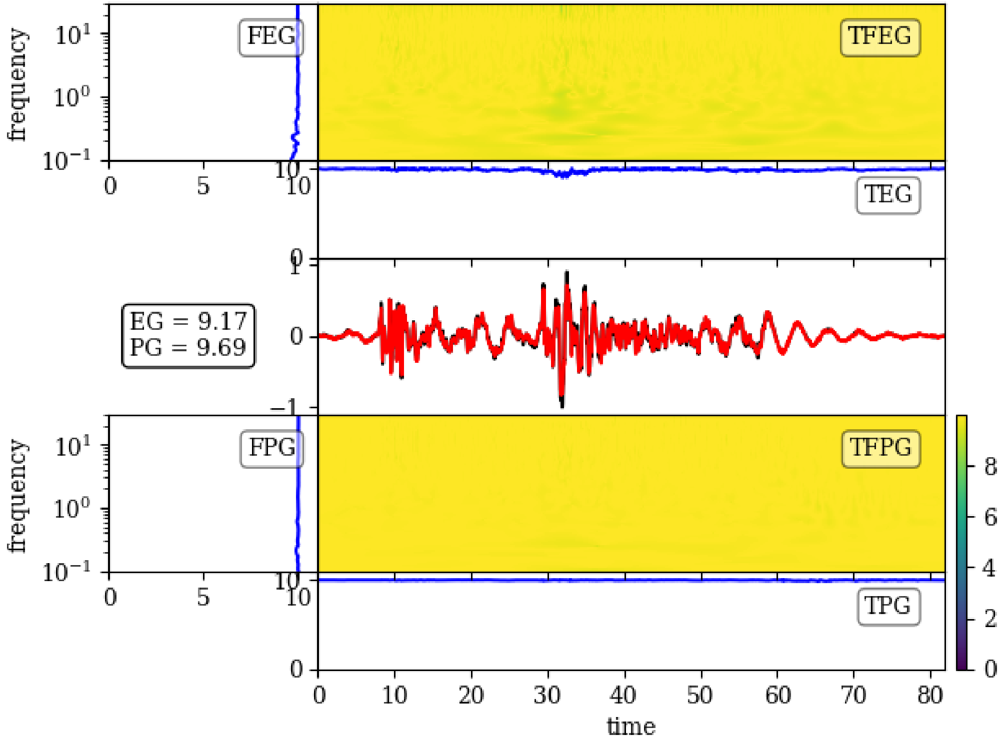
The following results have been obtained considering 100 signals in both undamaged and damaged conditions for a total of 200 samples, with separated training and validation data sets. Each signal is composed by 2048 time steps with dt = 0.04 s. The training process has been performed over 2000 epochs.

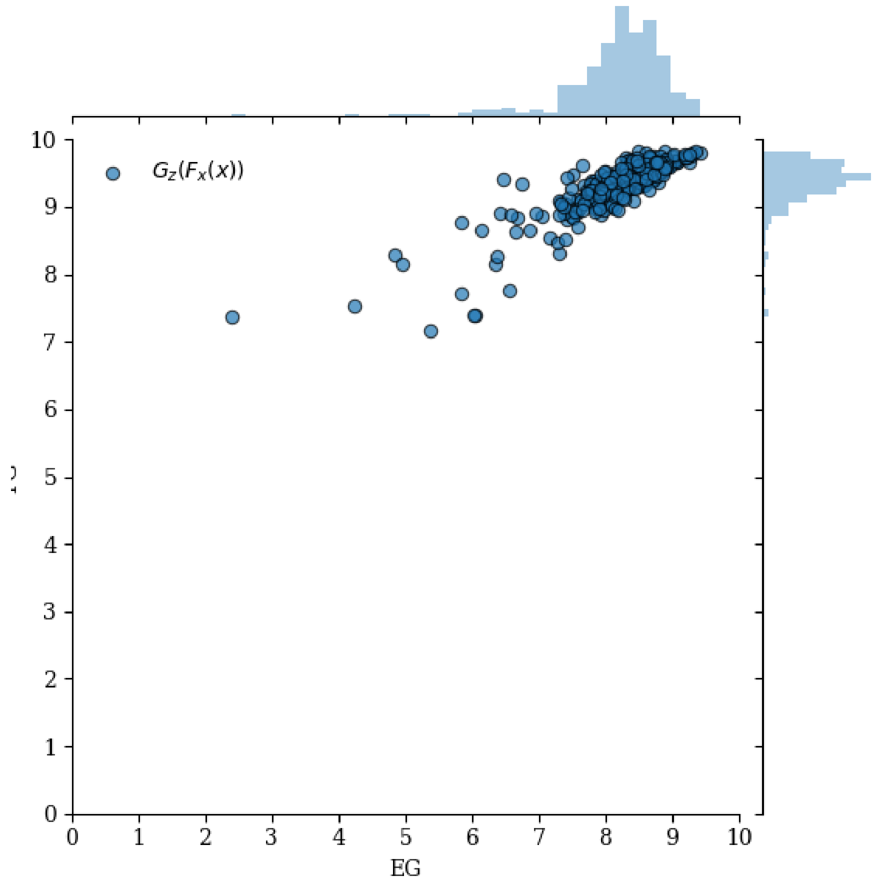
[4] Mousavi, S. M., Sheng, Y., Zhu, W., Beroza, G. C. (2019). STanford EArthquake Dataset (STEAD): A Global Data Set of Seismic Signals for AI. IEEE Access, 7, 179464–179476. https://doi.org/10.1109/access.2019.2947848



Examples of reconstructed signals for undamaged (top) and damaged (bottom) time-histories. The black lines represent the original time-histories $x_u^{(i)}$ and $x_d^{(i)}$ respectively. The orange time histories represent the result of the RepGAN reconstructions $\boldsymbol{G}_{\boldsymbol{\theta}_Z} \circ \boldsymbol{F}_{\boldsymbol{\theta}_X}\left(x_u^{(i)}\right)$ and $\boldsymbol{G}_{\boldsymbol{\theta}_Z} \circ \boldsymbol{F}_{\boldsymbol{\theta}_X}\left(x_d^{(i)}\right)$ respectively.

IOCA 2021



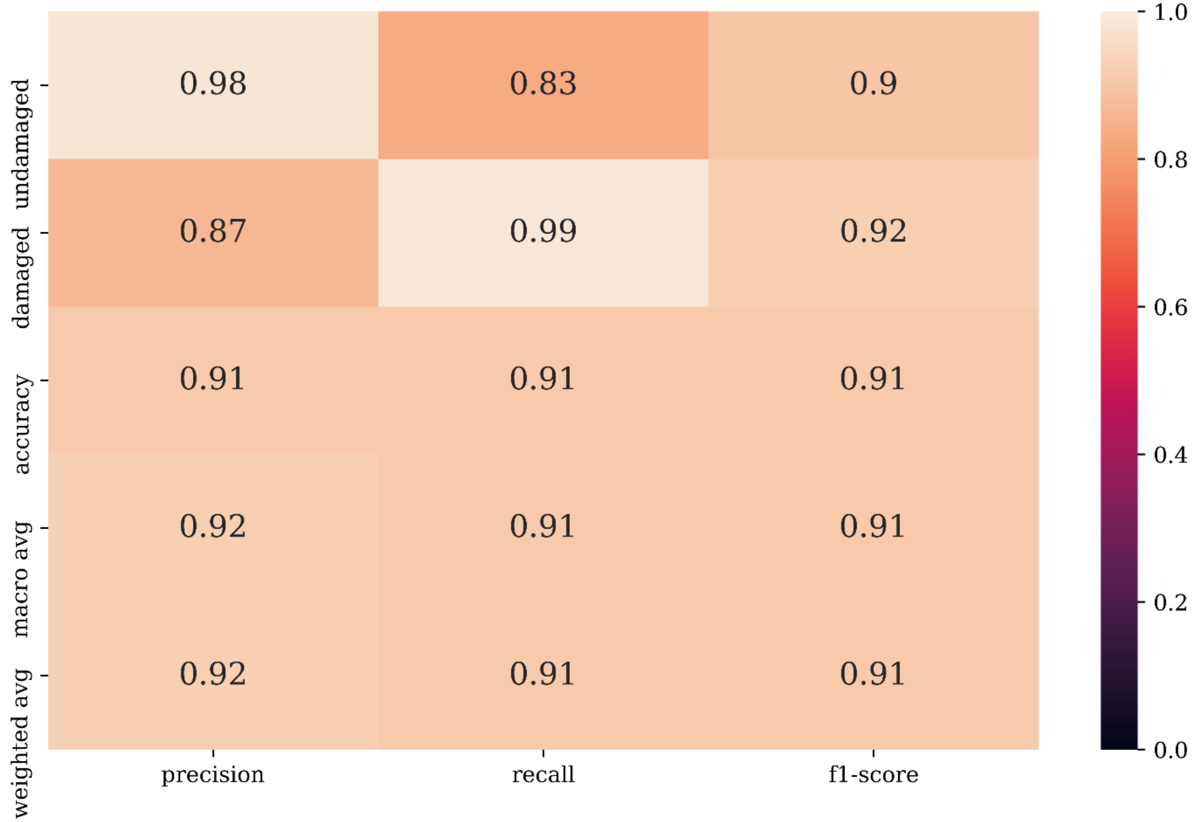


Time–Frequency Goodness-of-Fit (GoF) criterion: evaluation of the fit in Envelope (EG) and the fit in Phase (FG). GoF is evaluated between 0 and 10: the higher the score, the better is the reconstruction.

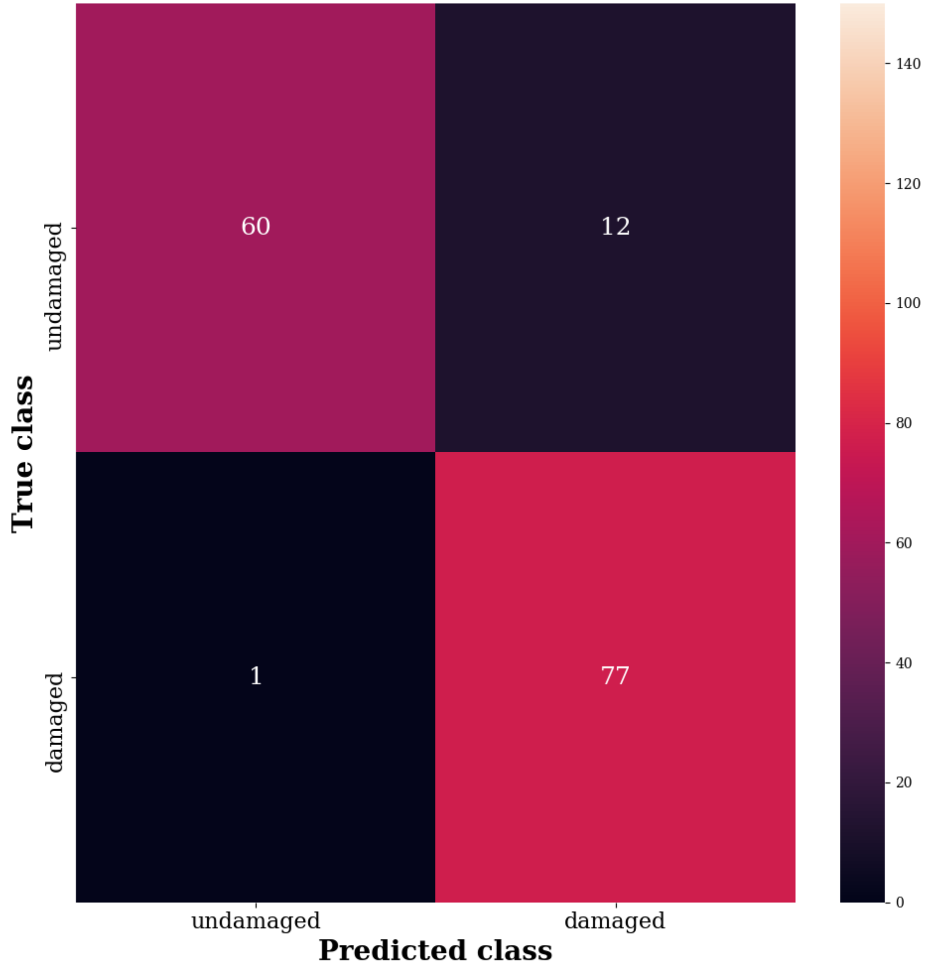
Right panel: the black line represents the original time-histories $x^{(i)}$ while the red time history depicts the result of the RepGAN reconstructions $\boldsymbol{G}_{\boldsymbol{\theta}_Z} \circ \boldsymbol{F}_{\boldsymbol{\theta}_X}(x^{(i)})$.

Left panel: the reconstructed structural respones $\boldsymbol{G}_{\boldsymbol{\theta}_Z} \circ \boldsymbol{F}_{\boldsymbol{\theta}_X}(x^{(i)})$ are reported in the EG/PG plane.

IOCA 2021



Classification report to evaluate the classification capability of the model. A precision score of 1.0 for a class C means that every item labelled as belonging to class C does indeed belong to class C, whereas a recall of 1.0 means that every item from class C was labelled as belonging to class C. F1-score is the harmonic mean of the precision and recall. Accuracy represents the proportion of correct predictions among the total number of cases examined. From these results, it appears that the classifier is able to predict the correct class.



Confusion matrix to further assess the classification accuracy of the model. 60 signals characterized by the undamaged class are correctly classified as undamaged, while only 12 undamaged time histories are identified as damaged. 77 damaged data instances are accurately identified, whereas only one damaged signal is mislabelled.

Conclusions: we introduce a SHM method based on a deep Generative Adversarial Network. Trained on synthetic time histories that represent the structural response of a multistory building in both damaged and undamaged conditions, the new model achieves high classification accuracy and satisfactory reconstruction quality resulting in a good bidirectional mapping between the input space and the latent space. However, the major innovation of the proposed method is the ability to generate reasonable signals for different damage states, based only on undamaged recorded or simulated structural responses. Therefore, real recordings linked to damaged conditions are not requested.

In our future work, we would like to extend our approach to real-time data. We will further consider a dataset constituted by a far larger number of time histories.

Acknowledgments

The training and testing of the neural network has been performed exploiting the supercomputer resources of the Mésocentre Moulon3, the cluster of CentraleSupélec and ENS Paris-Saclay, hosted within the Paris-Saclay University and funded by the Contrat Plan État Région (CPER).

The author G. Colombera won the scholarship “Tesi all’estero - a.y. 2020/2021 – second call” funded by Politecnico di Milano in order to spend a period of eight weeks at CentraleSupélec - Paris-Saclay University.