Deep Learning Enhanced Principal Component Analysis for Structural Health Monitoring

Ana Fernandez-Navamuel[1,2,3], Filipe Magalhães[4], Diego Zamora-Sáchez[1], Ángel J. Omella[2], David Garcia-Sanchez[1], and David Pardo[3,2,5]

[1]TECNALIA, Basque Research and Technology Alliance (BRTA), Parque Científico y Tecnológico de Bizkaia, Astondo bidea, Edificio 700, E-48160 Derio, Spain
[2]Basque Center for Applied Mathematics (BCAM), Bilbao, Spain
[3]University of the Basque Country (UPV/EHU) Leioa, Spain
[4]CONSTRUCT-ViBest, Faculty of Engineering, University of Porto (FEUP), Rua Dr. Roberto Frias 4200-465 Porto, Portugal
[5]Ikerbasque (Basque Foundation for Sciences), Bilbao, Spain

March 10, 2022

Abstract

This paper proposes a Deep Learning enhanced Principal Component Analysis (PCA) approach for outlier detection to assess the structural condition of bridges. We employ a partially explainable autoencoder architecture to replicate and enhance the data compression and reconstruction ability of PCA. The particularity of the method lies in the addition of residual connections to account for nonlinearities. We apply the proposed method to monitoring data obtained from two bridges under real operation conditions and compare the results before and after adding the residual connections. Results show that the addition of residual connections enhances the outlier detection ability of the network, allowing to detect lighter damages.

1 Introduction

The goal of Structural Health Monitoring (SHM) is to assess the health condition of structures and to detect damage from the measurement of physical variables, mainly, accelerations, strains, displacements and/or tilts.[12] Mathematically, this is an inverse problem that consists of finding a relation between the response of the structure and its real condition.[15] Several approaches exist attempting to solve this problem: from numerical optimization of computational models to statistical methods. [12] With the irruption of Artificial Intelligence, more complex data-driven techniques have emerged, including Machine Learning and Deep Neural Networks, among others. [42, 3, 39, 43]

Unsupervised learning is an approach to assess the health condition of operating structures using data-driven methods. [12, 11] Based on information from the reference (normal) state of the system, it detects departures from the normal response.[37] In large civil engineering structures, monitoring data may come from multiple sensor arrays of different type and nature, which include noise and redundant information.[40] Thus, an important step to extract representative features from measurements is data compression.[10]

Principal Component Analysis (PCA) is one of the most widely used data compression techniques.[20] It ap-

plies a linear projection of the data into the orthogonal subspace formed by the most relevant eigenvectors of the covariance matrix of the dataset. [34] The inverse operation reconstructs the original data from the principal components with minimal loss of information. [23] PCA provides one particular solution to the search of directions of maximum variance in the data, although there exist multiple other solutions with the same reconstruction capabilities.[23]

There are some works that employ this methodology for outlier detection purposes. [8, 41, 16, 17] Mujica et al.[30] use the residual between a sample and its corresponding projection onto the model of principal components as a novelty indicator for structural condition assessment. Garcia-Sanchez et al.[16] proposed an outlier detection method to assess the behavior against horizontal loads of a bridge in Mexico. The goal was to detect malfunctions in the sliding bearings that allow the relative displacement between the deck and the piers. We applied PCA to compress monitoring data from four longitudinal displacement sensors and isolate the effect of temperature changes (which were not directly measured). [36] The single-value performance indicator was the error between the original measurements and the projection of the compressed variable back into the original space using the first two principal components, known as Q-statistic. [30]

In this work, we first realize that, in terms of data reconstruction, PCA can be replaced by a simple autoencoder Neural Network (NN) with some restrictions.[24, 23] An autoencoder is a particular NN that reconstructs the input data at its output with small error.[6, 21, 4] Compared to traditional PCA, the use of NN requires a training phase.[19] The architecture contains two main parts, namely the encoder and the decoder. The encoder reduces the dimensionality of the original input. The decoder is the inverse of the encoder and maps the compressed representation back into the original data space.[38]

Since PCA performs data compression via a linear mapping, it may loose important (irrecoverable) information.[38] To overcome this problem, some nonlinear approaches have emerged, such as kernel PCA,[31, 18] but they require the nonlinear function to be previously selected instead of finding it during training. Often, finding an optimal nonlinear function for the kernel PCA is a challenging task.[38]

With an autoencoder, we can easily overcome this limitation by adding residual connections to the linear version of PCA that account for nonlinearities. As a result, we obtain a more accurate compression method. In addition, the use of independent blocks for the linear and nonlinear contributions allows for obtaining a partially explainable model.[2]

Explainability is a key advantage of our proposed architecture as it enables to work with simple linear transformations when possible and incorporate nonlinearities if required. By imposing specific constraints during training, we can also attain specific solutions such as PCA in the linear approach. It also allows to understand the improvement process with respect to traditional PCA. The proposed NN architecture is the main methodological contribution of the present paper.

In this work, we employ a residual autoencoder to assess the structural behavior of operating bridges. While most bridge SHM works focus on the dynamic properties and acceleration signals[28, 8, 35, 5, 26, 13], here we exploit other measurable time series from long-term monitoring campaigns. The input variables are measurements from different sensors, and constitute the target output to be predicted by the autoencoder. We use the reconstruction error as a single-value novelty indicator for outlier detection. We apply this methodology to two real structures under service. The first one employs the same dataset as Garcia-Sanchez et al.[16] for the Beltran bridge in Mexico. The second case of study includes multivariate sensor data from a continuous monitoring campaign of five years of service life of bridge Infante D. Henrique in Porto, Portugal.

Recently, autoeconders have been employed for outlier detection in several areas. [14] For example, Chen et al.[7] apply autoencoder ensembles for unsupervised outlier detection to various benchmark datasets to show the effectiveness of the approach compared to traditional techniques. Oh et al.[33] employ an autoencoder network to detect abnormal operation sounds of a complex machine using audio spectrograms. Li et al.[25] propose a hybrid autoencoder for anomaly detection in meteorological measurements to handle spatio-temporal data. All these works employ basic autoencoder architectures and dismiss the transition process to outperform linear data compression approaches. In this work, we carefully describe the replication and enhacenement of PCA with an explainable architecture.

Within the area of civil engineering, Xu et al.[32] present a Deep Learning approach that uses a convolutional autoencoder to reconstruct vibration signals and detect anomalies in a long span bridge. Ma et al.[27] propose a variational autoencoder (VAE) based on deep Bayesian neural networks to detect damage in beam-like bridges under moving load. The method uses vibrational data and it is validated with a simple beam numerical model and a small-scale laboratory bridge experiment. Da Fonseca[44] addresses a damage identification approach using an autoencoder where the output (to map from the compressed input) are the stiffness properties of the structure.

In this context of recent explosion of autoencoder methods for diverse SHM applications, our proposed method provides a simple architecture with the following features: (a) it is easily reproducible, (b) it outperforms PCA by definition, and (c) it is an *explainable* autoeconder NN. In addition, numerical results obtained for the two studied bridges show a good health condition assessment.

2 Methodology

2.1 Linear autoencoder to approximate PCA

An autoencoder is a particular type of Neural Network that aims to reconstruct the input at its output that generally follows a feed-forward architecture. [24] It consists of two consecutive modules, namely an encoder and a decoder. The encoder reduces the dimensionality of the original input while keeping all the relevant information. Thus, it compresses the input data. The decoder maps this reduced-dimension representation into the original data space, providing the reconstructed measurement. The encoding dimension limits the amount of captured information. Thus, there exists a trade-off between the compression level and the reconstruction ability of the autoencoder.

Let $\mathbf{M} \in \mathbb{M}^{n \times s}$ be a set of n rescaled measurements from s variables (sensors). For a certain measurement $\mathbf{m} = (m_1, m_2, ..., m_s)^T$, each hidden layer l_e in the encoder operator applies an affine transformation followed by an element-wise nonlinear activation function. The same holds for the decoder operator, which applies to the reduced-dimension vector. Based on this architecture, we build an affine autoencoder without hidden layers. The encoder operator $\mathcal{F}_a$ applies to the original measurement $\mathbf{m} \in \mathcal{R}^s$ and provides a lower-dimensional representation $\mathbf{z}_{linear} \in \mathcal{R}^k$ (with $k < s$):

$$\mathbf{z}_{linear} = \mathcal{F}_a(\mathbf{m}; \theta_a) = W_a \mathbf{m} + \mathbf{b}_a, \tag{1}$$

where $\theta_a = \{W_a, \mathbf{b}_a\}$. Similarly, the decoder operator $\mathcal{F}_c$ transforms the lower-dimensional linear representation $\mathbf{z}_{linear}$ into a reconstruction of the input $\hat{\mathbf{m}}_{linear}$:

$$\hat{\mathbf{m}}_{linear} = \mathcal{F}_c(\mathbf{z}_{linear}; \theta_c) = W_c \mathbf{z}_{linear} + \mathbf{b}_c, \tag{2}$$

where $\theta_c = \{W_c, \mathbf{b}_c\}$. PCA is a particular solution to the reconstruction problem, but not the only one. [23] By definition, principal components are orthonormal and produce uncorrelated $\mathbf{z}$ features.[23] We can impose these restrictions to the weight matrix W_a as well as $W_c = W_a^T$ during the training phase to achieve PCA. In such a case, $\mathbf{b}_a \in \mathcal{R}^s$ and $\mathbf{b}_c \in \mathcal{R}^k$ are null vectors and the weight matrices $W_a \in \mathcal{R}^{s \times k}$ and $W_c \in \mathcal{R}^{k \times s}$ contain the first k eigenvectors of the covariance matrix of $\mathbf{M}$ column-wise and row-wise, respectively. However, focusing on the task of data reconstruction, we can indistinctly employ a linear autoencoder that replicates PCA, or any other orthogonal solution, as long as they reach the same level of captured information as PCA. This provides explainability to the linear part of the model since we can tailor the autoencoder to accomplish additional requirements when needed.

2.2 Residual connections to enhance PCA performance

In this work, we propose a novel autoencoder architecture that enhances the performance of PCA in measurement reconstruction. We start from the linear autoencoder described in section 2.1 and add residual connections to account for nonlinear relationships in the data. Figure 1 describes the model as a block diagram.

The encoding module adds two parallel operations. $\mathcal{F}_a(\mathbf{m}; \theta_a)$ is the linear transformation described in subsection 2.1. The residual connection $\mathcal{F}_b(\mathbf{m}; \theta_b)$ follows a feed-forward architecture [19] that applies a linear transformation followed by a nonlinear activation function g

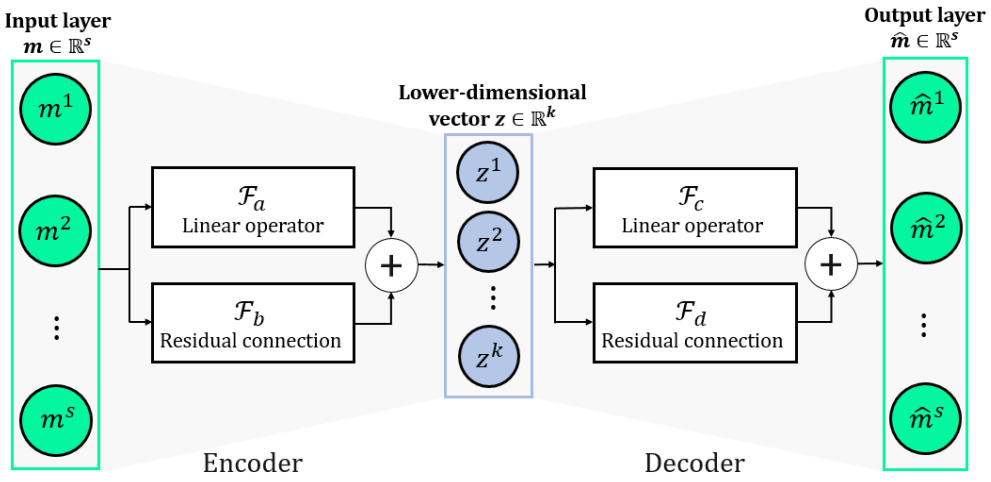


Figure 1: Block diagram of the residual autoencoder

through L_b layers. This yields the lower-dimensional vector $\mathbf{z}$:

$$\mathbf{z} = \mathcal{F}_E(\mathbf{m}) = \mathcal{F}_a(\mathbf{m};\theta_a) + \mathcal{F}_b(\mathbf{m};\theta_b). \tag{3}$$

The same architecture holds for the decoding step $\mathcal{F}_D$, which reconstructs the original measurement from its lower dimensional representation:

$$\hat{\mathbf{m}} = \mathcal{F}_D(\mathbf{z}) = \mathcal{F}_c(\mathbf{z};\theta_c) + \mathcal{F}_d(\mathbf{z};\theta_d) \tag{4}$$

Here, the operator $\mathcal{F}_c(\mathbf{z};\theta_c)$ applies the linear transformation, while $\mathcal{F}_d(\mathbf{z};\theta_d)$ is the residual connection with the feedforward network architecture and nonlinear activation functions. This simple block-wise architecture is explainable, which contributes to understand the enhancement over linear PCA.

The objective of the autoencoder is to effectively recover the original measurements from their lower-dimensional representation. The reconstruction error $\rho(\mathbf{m})$ measures the misfit between an original data point $\mathbf{m}$ and its corresponding reconstruction $\hat{\mathbf{m}}$ using the squared l_2 norm:

$$\rho(\mathbf{m}) = \|\mathbf{m} - (\mathcal{F}_D \circ \mathcal{F}_E)_\theta(\mathbf{m})\|_2^2, \tag{5}$$

where $\theta = \{\theta_a, \theta_b, \theta_c, \theta_d\}$ contains the weight and bias parameters of the four modules that build the residual autoencoder and $\circ$ stands for the composition operator. Training the model consists of finding the parameter set θ^* that minimizes a loss function based on the reconstruction error for the training dataset $\mathbf{M}$:

$$\mathcal{L}_{\theta^*} := \arg\min_\theta \|\mathbf{M} - (\mathcal{F}_D \circ \mathcal{F}_E)_\theta(\mathbf{M})\|_2^2 \tag{6}$$

2.3 Outlier detection procedure

The final goal of the methodology is to detect abnormal behavior of bridges from experimental measurements acquired during monitoring. We train the residual autoenconder with measurements from the undamaged (healthy) condition. The reconstruction error is expected to be small for unseen measurements that also correspond to this state. But if we evaluate new data coming from a different scenario (namely, a damaged one), the autoencoder will poorly reconstruct them.

We employ the reconstruction error ρ as the damage sensitive indicator. The threshold value α for outlier detection is the 99 percentile of the training dataset. For a new measurement $\mathbf{m}_{new} \in \mathbb{R}^s$ acquired within a future monitoring campaign, we obtain the reconstruction error and compare it against the threshold value. Algorithm 1 describes this procedure.

Algorithm 1: Reconstruction error-based outlier detection

Input: $\mathbf{m}_{new} \in \mathbb{R}^s,\ \mathcal{F}_E,\ \mathcal{F}_D,\ \theta^*,\ \alpha$
Output: ρ

Compute the reconstruction error:
$\rho = \|\mathbf{m}_{new} - (\mathcal{F}_D \circ \mathcal{F}_E)_{\theta^*}(\mathbf{m}_{new})\|_2^2$
if $\rho > \alpha$ **then**
    $\mathbf{m}_{new}$ is abnormal, indicate outlier
else
    $\mathbf{m}_{new}$ is healthy

3 Results

3.1 Beltran bridge in Mexico

We apply the proposed method to assess the global behavior of an asymmetric prestressed concrete viaduct in Mexico: the Beltran bridge. The bridge contains four sliding bearings at the deck-pier contacts to allow for relative displacements and limit the horizontal loads reaching the piers. But if these devices lose their sliding properties, large displacements and subsequent cracking can occur at the pier caps. Garcia-Sanchez et al.[16] provide further details. Figure 2 shows the profile of the bridge.

The monitoring campaign used fiber optic sensors and was active from August 2012 to August 2013. Due to occasional shutdowns of the system, the effective monitoring time was of approximately nine months. We employ time series from four longitudinal displacement sensors (see Figure 2). The data was acquired at a sampling frequency of 200 Hz, and the mean value was calculated and stored every ten minutes for each sensor. This subsampling considerably reduces the storage space while permitting the analysis of long-term variations. After this pre-processing step, we have four sensor signals with 37,692 measurements each. We employ the first 90% of the data randomly split into 85% for training and 15% for validation. We use the final 10% for testing purposes. We adapt the architecture described in section 2 to fit the four-dimensional input. The residual connections are symmetric and contain three layers each. This architecture provides adequate results. We first solve for the linear autoencoder (only $\mathcal{F}_a$ and $\mathcal{F}_c$ are active) for a two-dimensional compression. We design the linear autoencoder to replicate PCA by imposing the constraints that hold for the principal components. Table 1 compares the transformation matrices of traditional PCA (PC1 and PC2 stand for the first two principal components of the training data covariance matrix), the designed linear autoencoder with the constraints, and an unconstrained approach with null bias terms. We observe that the three configurations capture the same level of information. In subsequent steps in this work we employ the constrained linear autoencoder that replicates PCA results.

After solving the linear autoencoder, we incorporate the residual connections ($\mathcal{F}_b$ and $\mathcal{F}_d$) to build the nonlinear autoencoder. In order to evaluate the enhacenment of the residual connections over the linear approach, we fix the previously obtained weight matrices and only train the nonlinear modules. Figure 3 shows the loss functions for the training and the validation datasets for both the linear and the residual autoencoder. The number of epochs indicates the repetitions over the training dataset to minimize the loss function. We observe in figure 3a that the linear autoencoder converges after just 50 epochs with a loss function value of 0.0946 for the training dataset. Figure 3b shows the effect of adding the residual connections to the linear autoencoder. After 300 epochs the loss function decays to a value of 0.0171. When the output dimension of the encoder is equal to two, the amount of information captured during training raises from 90.41% to 98.05% when we add the residual connections. This means that the residual connections contribute to reduce the reconstruction error from 9.59% to 1.95% ($\approx 80\%$ error reduction) When the output dimension of the encoder is reduced to one, the level of captured information is below 90% for both architectures (linear and residual). This made impractical the outlier detection task.

For comparison, we implemented a kernel PCA and tried different kernel functions (radial basis (rbf) and cosine functions).[22] For the rbf function, we attained a 42.99% of information captured, and a 76.77% was attained for the cosine function. Since the real distribution of the data is unknown, it is nontrivial to find a kernel function that outperforms the linear solution.

We analyze the reconstruction ability of the two configurations via the crossplots (ground truth vs predictions) of the sensor signals. Here, the ground truth are the measurements of each sensor, and the predictions are the reconstructed variables. The square of the correlation coef-

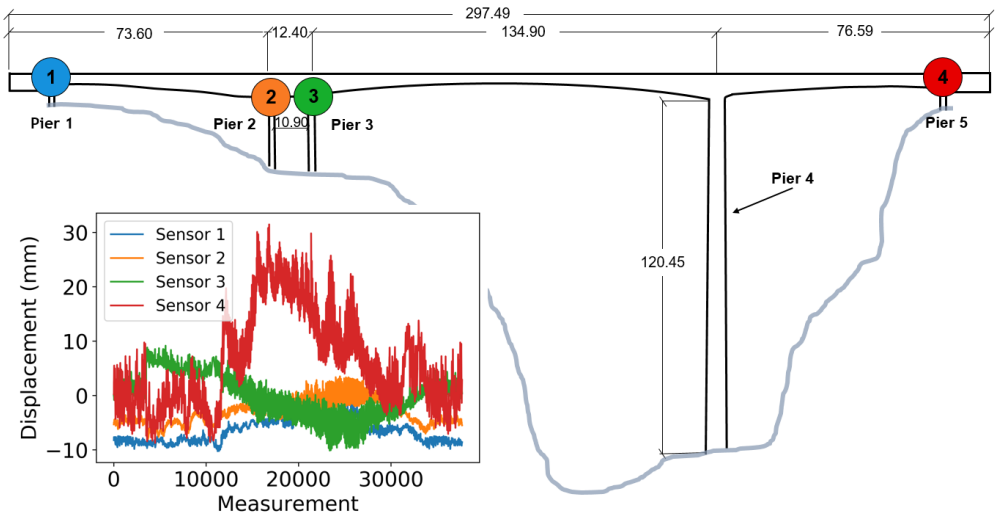


Figure 2: Beltran bridge profile with sensor location

Table 1: Comparision of PCA and constrained linear autoencoder

	PCA		Constrained approach		Unconstrained approach			
	PC1	PC2	$\mathbf{w}_a^1$	$\mathbf{w}_a^2$	$\mathbf{w}_a^1$	$\mathbf{w}_a^2$	$\mathbf{w}_c^1$	$\mathbf{w}_c^2$
	0.5365	0.1071	0.5362	0.1077	-0.6544	0.4597	-0.2502	0.2914
	0.4957	-0.5296	0.4954	-0.5308	-1.0375	-0.1796	-0.4716	-0.2116
	-0.5078	0.3025	-0.5084	0.3016	0.8989	-0.0449	0.3907	0.0359
	0.4567	0.7852	0.4750	0.7846	-0.0789	1.0556	0.0450	0.7865
% captured	90.41		90.41		90.41			

ficient r^2 provides a numerical measure of the correlation between ground truth and prediction for each sensor (see Table 2). Figure 4 compares the crossplot corresponding to sensor 1 for the linear and the residual models. We observe a superior performance for the residual model.

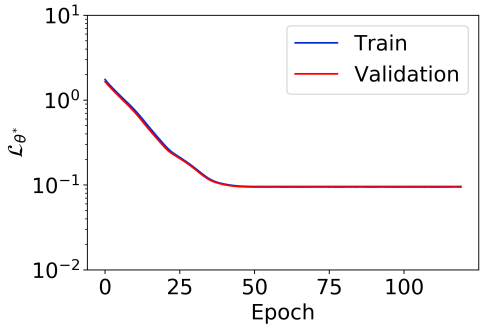
We finally determine the threshold value α as the 96th percentile (p-96) over the training reconstruction errors. We select a percentile vallue that ensures a rate of false positives below 5% as well as delivers adequate equilibrium between false positves and negatives. Figure 5 shows the histograms for the linear and residual autoencoder training reconstruction errors, and the corresponding threshold values.

Table 2: r^2 metric for the sensors in Beltran bridge.

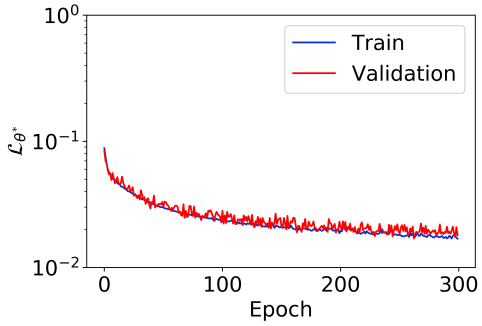
Sensor ID	Sensor type	Linear	Residual
1	Displacement	0.891	0.978
2	Displacement	0.908	0.981
3	Displacement	0.842	0.981
4	Displacement	0.975	0.988

3.1.1 Testing.

We now test the ability of both autoencoder configurations in the detection of outliers. We use the testing dataset



(a) Linear autoencoder



(b) Residual autoencoder

Figure 3: Loss function evolution for the Beltran bridge dataset

that contains the final 10% of the data. We simulate damage by reducing the displacement measurements at one of the sensors to represent a loss of sliding properties at that bearing. We do so by applying a reduction factor to the measurements of sensor 1. We consider three different damage levels: 50%, 30% and 20%. For each level of damage, we duplicate the test data set and apply the reduction factor to the second half of the measurements.[16] Figure 6 compares the control charts for the three damage levels. The shadowed regions in Figure 6 cover the part of the testing affected by damage.

Table 3 gathers the fraction (%) of false positves (FP) and false negatives (FN) for each case. The presence of false positives occurs probably due to the short length of the monitoring period (less than a year). The training phase covers a limited range of environmental variations and the testing dataset includes data from a healthy scenario that corresponds to an unseen month of the year. We believe we can overcome this with longer monitoring campaigns. We also observe that the linear autoencoder needs stronger damages to be able to detect it. While the residual autoencoder still detects a damage of 30% with less than 3% of false negatives, the linear autoencoder considerably fails for that scenario. For a damage severity of 20% affecting the bearing, the residual autoencoder increases the rate of FN but still detects it.

Table 3: Outlier detection results for Beltran bridge.

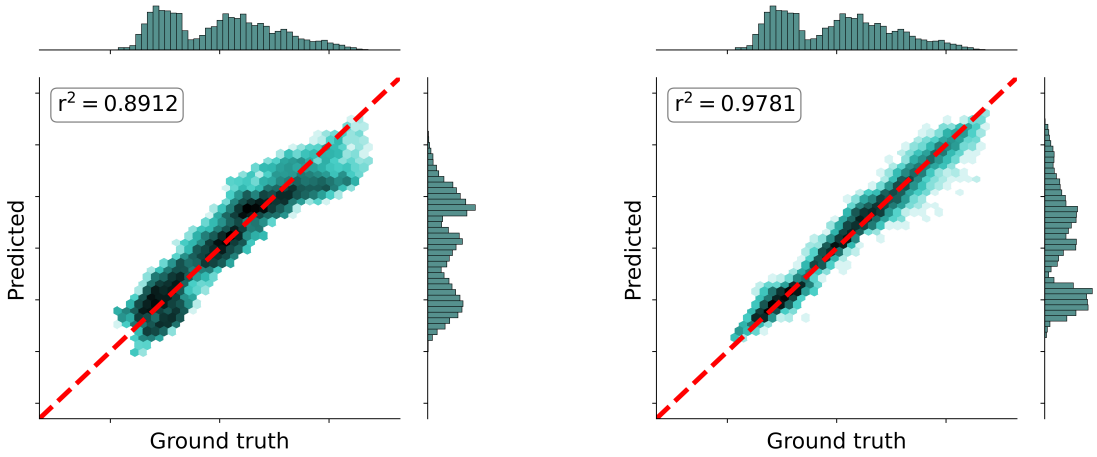
Architecture type	Damage level(%)	FP(%)	FN(%)
Linear	50	2.30	3.15
Residual	50	1.26	0.00
Linear	30	2.30	73.93
Residual	30	1.26	2.89
Linear	20	0.12	91.80
Residual	20	1.26	15.63

3.2 Infante Dom Henrique bridge in Porto

The Infant Dom Henrique bridge was opened to road traffic in 2002 between the cities of Vila Nova de Gaia and Porto. It consists of a rigid prestressed reinforced concrete box-beam deck supported over a 1.50 m thick reinforced concrete arch. This is a unique structure that achieved a world record due to the shape of its arch span of 280 m to cross river Douro. More detailed information regarding its structural aspects can be found in Adão da Fonseca et al.[9]

Given the particularities of the bridge, a monitoring system was installed to control construction and assembly operations, being afterwards also used to assess the long-term behavior. It employs several sensors located at the main sections of the bridge, registering one data point per hour. There are three types of sensors: strain gauges, clinometers, and thermometers. Magalhães et al.[1] provide further information regarding sensor properties and location, as well as monitoring aspects.

Although the monitoring campaign covered more than ten years of service life of the bridge, there occurred some shutdowns of the system that prevented the records to be continuous. For this reason, in this work we employ five



(a) Linear auteoencoder

(b) Residual autoencoder

Figure 4: Beltran bridge training crossplots for sensor 1

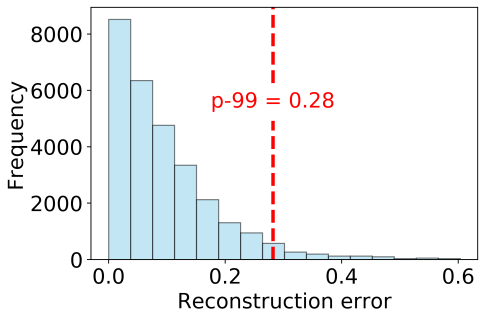
years of monitoring, which are practically free from failures. For the particular purpose of this work, we use a total of 19 sensors located at some regions that are critical for the structural behavior. The selected sensors were free from important shutdowns and included eight clinometers, seven strain gauges, and four thermometers. Thus, in this case study, the temperature is treated as as an input variable. Figure 7 shows the profile of the bridge and includes the time series for the three strain gauges located at cross-section C.

After removing null values, we obtained a total of 19 sensor signals with 34,641 measurements each. We use the first 90% of these data with a random split into 80% training and 20% validation. We adapt the autoencoder architecture for a 19-dimensional input. In this case, the residual connections are also symmetric and consist of six layers each. This number of layers provides adequate results and the architecture has not been further optimized. We first solve for the linear autoencoder (only $\mathcal{F}_a$ and $\mathcal{F}_c$ are active) for a two-dimensional compression. We design the linear autoencoder to approximate PCA by imposing the constraints satisfied by principal components. The information captured is similar to that of PCA (92.93%).

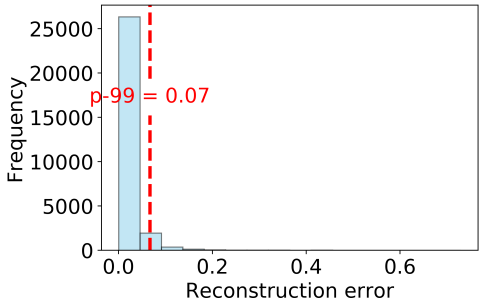
After solving the linear autoencoder, we incorporate the residual connections ($\mathcal{F}_b$ and $\mathcal{F}_d$) that yield the residual autoencoder. In order to evaluate the enhancement of the residual connections over the linear approach, we fix the weight matrices obtained and train the nonlinear modules only. Figure 8 shows the loss functions for the training and validation datasets for both the linear and the residual autoencoder. As with the Beltran bridge, we observe that the linear model (see Figure 8a) converges with around 50 epochs at a value of 0.0716 for the training dataset. Figure 8b shows the effect of adding the residual connections after 300 epochs, with the loss function reaching a value of 0.0311 for the training dataset. For a compression into a two-dimensional vector, the captured information during training raises from 92.93% to 97.05% after adding the residual connections, indicating an enhancement in the reconstruction. A compression into a one-dimensional vector provides a level of captured information below 90% for both architectures (linear and residual), making impractical the outlier detection task.

For comparison, we implemented a kernel PCA and tried different kernel functions (radial basis (rbf) and cosine functions).[22] For example, the cosine function provided a 80.10% of information captured.

We analyze the reconstruction ability of the models via



(a) Linear autoencoder



(b) Residual autoencoder

Figure 5: Beltran bridge training histograms for sensor 1

the crossplots (ground truth vs predictions) of the sensor signals. Table 4 shows the correlation between the ground truth and the predictions for the training set (sensor 9), including the r^2 metric. Again, the residual autoencoder exhibits superior performance. We finally obtain the threshold value α for outlier detection as the $p-98$ over the training reconstruction errors. Figure 10 shows the corresponding histograms for the linear and residual autoencoder, indicating the threshold value.

3.2.1 Testing.

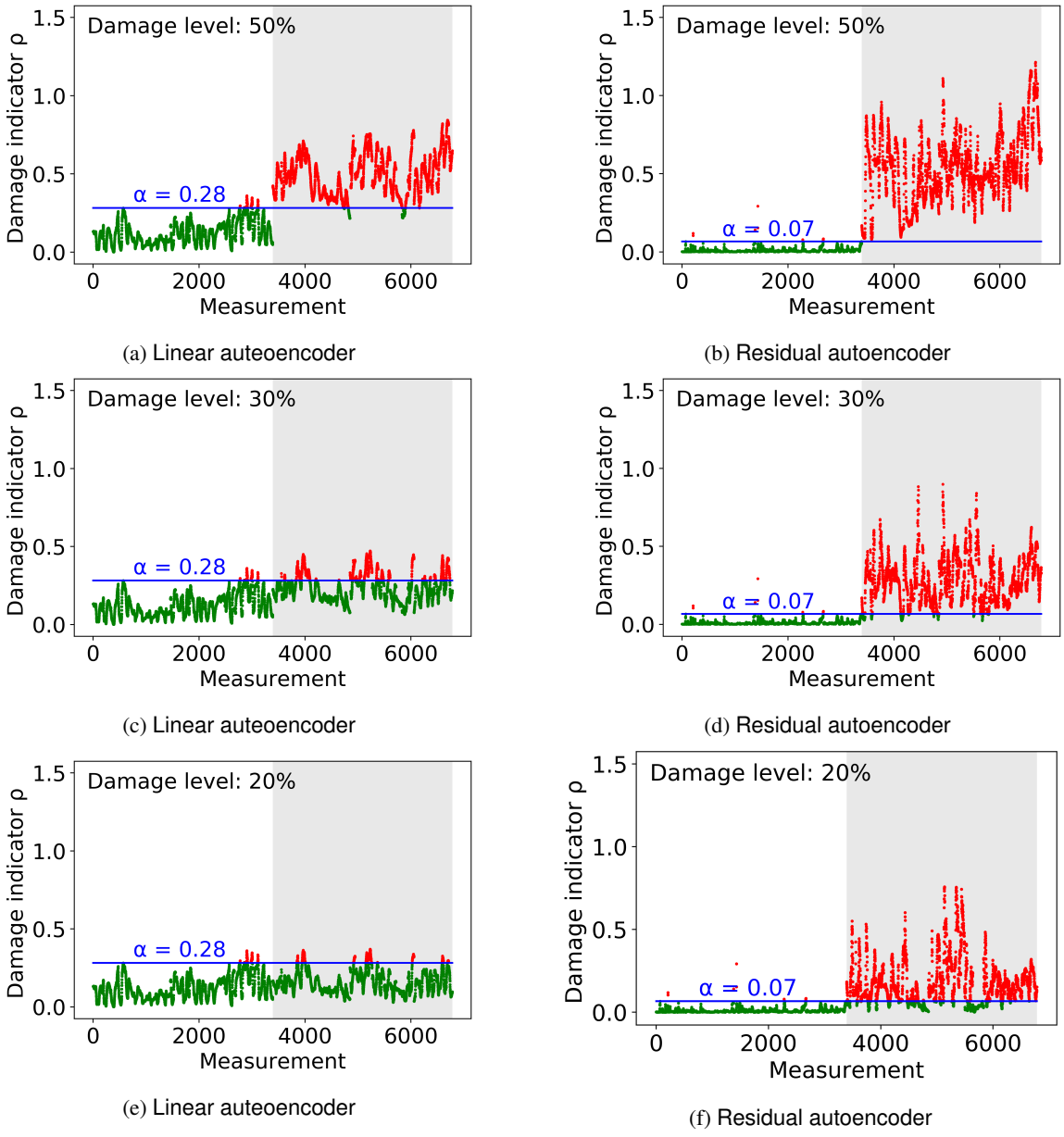
We test the outlier detection ability of the proposed methodology. We compare the results for the linear and the residual autoencoder. Since the entire monitoring campaign corresponds to the normal operation of the bridge (healthy state), we employ a numerical model built in ANSYS to simulate damage. The model was also employed by Magalhães et al.[29], where various damages

Table 4: r^2 metric for the sensors in Infante bridge.

Sensor ID	Sensor type	Linear Model	Residual Model
1	Clinometer	0.993	0.992
2	Clinometer	0.994	0.993
3	Clinometer	0.995	0.993
4	Clinometer	0.996	0.994
5	Clinometer	0.985	0.989
6	Clinometer	0.990	0.991
7	Clinometer	0.863	0.954
8	Clinometer	0.899	0.954
9	Strain gauge	0.907	0.959
10	Strain gauge	0.944	0.978
11	Strain gauge	0.942	0.974
12	Strain gauge	0.833	0.929
13	Strain gauge	0.845	0.944
14	Strain gauge	0.780	0.929
15	Strain gauge	0.756	0.932
16	Thermometer	0.962	0.983
17	Thermometer	0.996	0.991
18	Thermometer	0.971	0.987
19	Thermometer	0.996	0.990
Average		0.916	0.962

secenarios were created by reducing the stiffness at certain cross sections of the structure by 10%.

We follow the same approach as [29] to simulate damage at cross-section A (see Figure 7). It is important to mention that unlike natural frequencies, the measured variables are local and not global. This means that they are only significantly affected by damages that occur close to their emplacement. Hence, the affected sensors are the two clinometers located at both sides of cross-section A. For the remaining sensors, the effect of damage is negligible. We run various simulations including the undamaged scenario and damages with different severity levels. For each damage scenario, we obtain the relative variation of the rotations in A with respect to those given by the undamaged simulation. Table 5 describes the damage scenarios and indicates the average variation factor for the measurements of the two affected sensors. The test dataset contains the final 10% of the total monitoring data, resulting in 3,663 data for each sensor. For each damage level, We duplicate the test dataset and affect the second half of the measurements. Particularly, we apply



(a) Linear auteoencoder

(b) Residual autoencoder

(c) Linear auteoencoder

(d) Residual autoencoder

(e) Linear auteoencoder

(f) Residual autoencoder

Figure 6: Beltran bridge testing control charts

the obtained relative variations to the measurements from the two clinometers placed at the damaged cross-section. We feed the complete test dataset to the autoencoders and obtain the damage indicator (i.e. the reconstruction error ρ) for each new measurement. Figure 11 shows the control charts of the damage indicator for outlier detection.

The shadowed regions correspond to the part of the test-

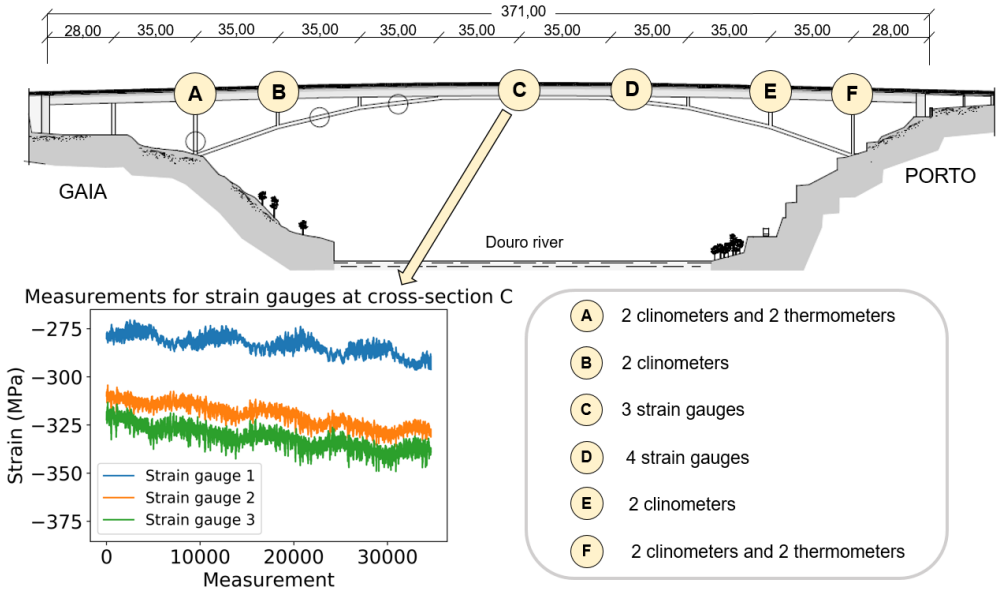


Figure 7: Infante D. Henrique bridge profile with sensor location

ing dataset affected by damage. Table 6 gathers the results in terms of false positives and false negatives.

We observe that the incidence of false positives is very small for both the linear and the residual autoencoders. Since the monitoring campaign covers five years, it is possible to train the networks for a wide enough range of environmental and operational variations. In terms of false negatives, the performance of both approaches is still good for a damage severity of 30%, but the linear autoencoder starts to raise some false negatives (below 1%). For a damage severity of 20%, the rate of change in the measurements is very small and thus reconstruction errors slightly increase. The linear autoencoder completely fails in the detection of this damage, while the residual architecture presents close to 11% of false negatives.

Table 5: Description of the damage scenarios for Infante bridge.

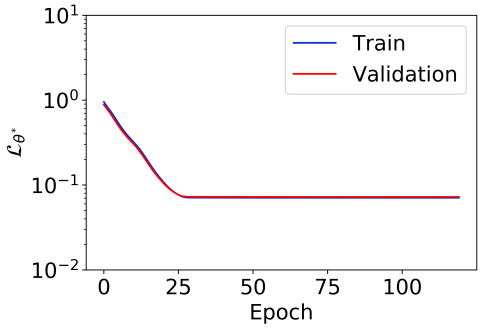
Damage level(%)	Description	Variation factor Δ
50	Stiffness reduction	0.25
30	Stiffness reduction	0.12
20	Stiffness reduction	0.075

Table 6: Outlier detection results for Infante bridge.

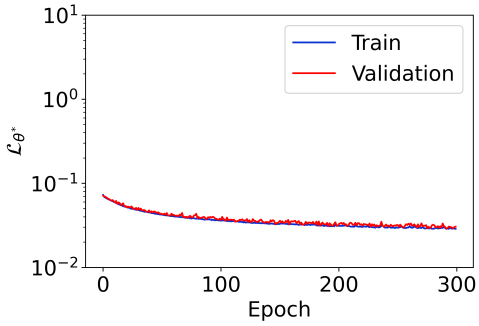
Architecture type	Damage level(%)	FP (%)	FN(%)
Linear	50	3.03	0.00
Residual	50	0.71	0.00
Linear	30	3.03	0.62
Residual	30	0.71	0.00
Linear	20	3.03	89.08
Residual	20	0.71	10.67

4 Conclusions

We have tested a novel approach for SHM of bridge structures using a particular autoencoder architecture for outlier detection. We conclude that the addition of residual connections to account for nonlinearities significantly improves the reconstruction ability of the autoencoder. This enhances the outlier detection task to assess the structural



(a) Linear autoencoder



(b) Residual autoencoder

Figure 8: Loss function evolution for the Infante bridge dasaset

condition, allowing to detect smaller damages. This is an important advantage of the methodology since it allows to report early damage before it becomes irreparable.

The proposed methodology remains in the unsupervised learning domain, yielding to an outlier detection approach (level I in Rytter scale [37]). However, we can interpret the reconstruction error as a damage quantification value with an adequate analysis. In addition, the locally-constrained sensitivity to damage of the employed variables may help to discover the damaged region. Once an alert rises - indicating damage -, we can analyze the individual reconstruction errors for each sensor. We expect that those sensors located nearby the damaged region will have a larger contribution to the damage indicator.

When using local quantities as the monitoring variables, such as displacements or rotations, the SHM system will only detect damage if it occurs close to any of the existing sensors. Hence, these sensors must be located at the most critical elements/parts of the structure. The use of fiber optic technology may allow to install dense sensor array distributions for more accurate assessments in terms of damage location.

The existence of long-term varying phenomena such as global warming may change the response of the structure over time. The SHM system may undergo an increase in the amount of false positives due to the mismatch between the previous and the current healthy condition. This poses the need for a dynamic training phase, updated every x years to include more recent information, as long as in situ periodic inspections are performed that demonstrate the health of the structure so as to be considered an undamaged state.

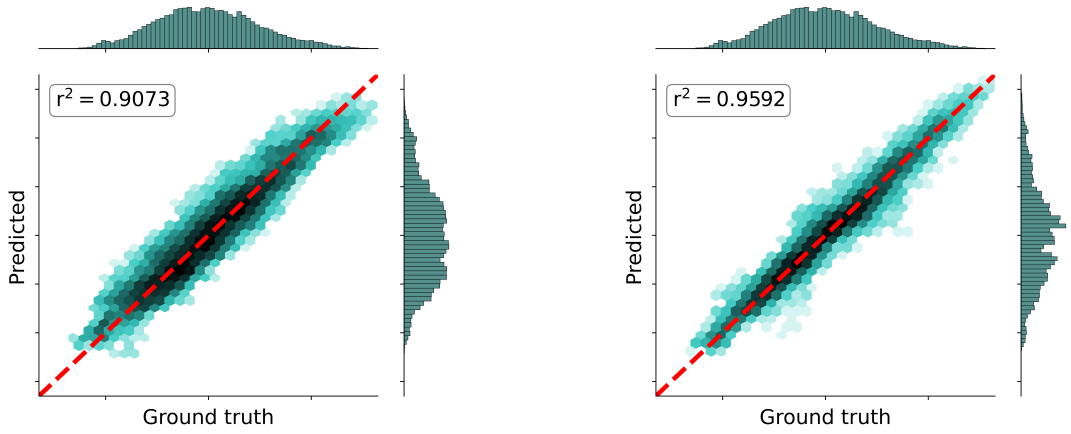
We orientate future work towards the use of numerical simulations to include the response of the structure under different synthetic scenarios. This enables the transition to supervised learning approaches able to quantify and locate damage. This requires the execution of massive simulations for the training phase that cover a wide amount of damage scenarios and incorporate uncertainty.

Acknowledgements

Authors would like to acknowledge the discussions with Marcos Pantaleón from APIA XXI, Ambher Monitoring Systems, SCT and Banobras S.N.C.

The collaboration of Engineer Renato Bastos and Kinesia in the sharing of Infante D. Henrique Bridge data are deeply acknowledged.

This work has received funding from: the European Union's Horizon 2020 research and innovation program under the grant agreement No 769373 (FORESEE project) and the Marie Sklodowska-Curie grant agreement No 777778 (MATHROCKS); the Base Funding - UIDB/04708/2020 of the CONSTRUCT - Instituto de I&D em Estruturas e Construções - funded by national funds through the FCT/MCTES (PIDDAC); the European Regional Development Fund (ERDF) through the Interreg V-A Spain-France-Andorra program POCTEFA 2014-2020 Project PIXIL (EFA362/19); the Spanish Ministry of Science and Innovation with references PID2019-108111RB-I00 (FEDER/AEI) and the "BCAM Severo Ochoa" accreditation of excellence (SEV-2017-



(a) Linear auteoencoder

(b) Residual autoencoder

Figure 9: Training crossplots for sensor 9 (strain gauge) at Infante bridge

0718); and the Basque Government through the BERC 2018-2021 program, the two Elkartek projects 3KIA (KK-2020/00049) and MATHEO (KK-2019-00085), the grant "Artificial Intelligence in BCAM number EXP. 2019/00432", and the Consolidated Research Group MATHMODE (IT1294-19) given by the Department of Education.

References

[1] A. Adão da Fonseca and R. O. Bastos. Monitorização em fase de serviço do comportamento estrutural da Ponte Infante D. Henrique. *Encontro Nacional de Betão Estrutural*, (Ø13 mm):24–26, 2004.

[2] P. Angelov and E. Soares. Towards explainable deep neural networks (xDNN). *Neural Networks*, 130:185–194, 2020.

[3] M. Azimi, A. D. Eslamlou, and G. Pekcan. Data-driven structural health monitoring and damage detection through deep learning: State-of-the-art review. *Sensors*, 20(10), 2020.

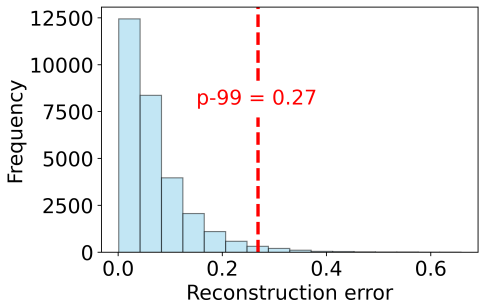
[4] H. Bourlard and Y. Kamp. Auto-association by multilayer perceptrons and singular value decomposition. *Biological cybernetics*, 59:291–4, 02 1988.

[5] A. Cancelli, S. Laflamme, A. Alipour, S. Sritharan, and F. Ubertini. Vibration-based damage localization and quantification in a pretensioned concrete girder using stochastic subspace identification and particle swarm model updating. *Structural Health Monitoring*, 19:147592171882001, 02 2019.

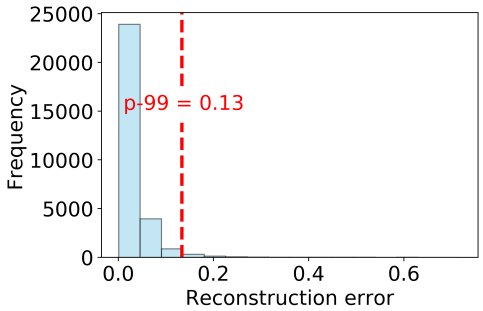
[6] R. Chalapathy, A. K. Menon, and S. Chawla. Anomaly detection using one-class neural networks. *CoRR*, abs/1802.06360, 2018.

[7] J. Chen, S. Sathe, C. Aggarwal, and D. Turaga. Outlier detection with autoencoder ensembles. *Proceedings of the 17th SIAM International Conference on Data Mining, SDM 2017*, pages 90–98, 2017.

[8] G. Comanducci, F. Magalhães, F. Ubertini, and Álvaro Cunha. On vibration-based damage detection by multivariate statistical techniques: Application to a long-span arch bridge. *Structural Health Monitoring*, 15(5):505–524, 2016.



(a) Linear auteoencoder



(b) Residual autoencoder

Figure 10: Infante bridge training histograms

[9] A. A. da Fonseca. The Infant Henrique Bridge over the River Douro , in Porto , Portugal. *International Conference on Arch Bridges*, pages 931–960, 2007.

[10] U. Dackermann, J. Li, and B. Samali. Dynamic-based damage identification using neural network ensembles and damage index method. *Advances in Structural Engineering*, 13(6):1001–1016, 2010.

[11] N. Dervilis, I. Antoniadou, R. J. Barthorpe, E. J. Cross, and K. Worden. Robust methods for outlier detection and regression for SHM applications. *International Journal of Sustainable Materials and Structural Systems*, 2(1/2):3, 2015.

[12] C. Farrar and K. Worden. *Structural Health Monitoring A Machine Learning Perspective*. Wiley, 01 2013.

[13] A. Fathi and M. P. Limongelli. Statistical vibration-based damage localization for the s101 bridge, flyover reibersdorf, austria. *Structure and Infrastructure Engineering*, 0(0):1–15, 2020.

[14] M. Frank, D. Drikakis, and V. Charissis. Machine-learning methods for computational science and engineering. *Computation*, 8(1):1–35, 2020.

[15] M. I. Friswell. Damage identification using inverse methods. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 365(1851):393–410, 2007.

[16] D. Garcia-Sanchez, A. Fernandez-Navamuel, D. Z. Sánchez, D. Alvear, and D. Pardo. Bearing assessment tool for longitudinal bridge performance. *Journal of Civil Structural Health Monitoring*, pages 1–25, 2020.

[17] E. García-Macías and F. Ubertini. Mova/moss: Two integrated software solutions for comprehensive structural health monitoring of structures. *Mechanical Systems and Signal Processing*, 143:106830, 2020.

[18] K. Ghoulem, T. Kormi, and N. B. H. Ali. Damage detection in nonlinear civil structures using kernel principal component analysis. *Advances in Structural Engineering*, 23(11):2414–2430, 2020.

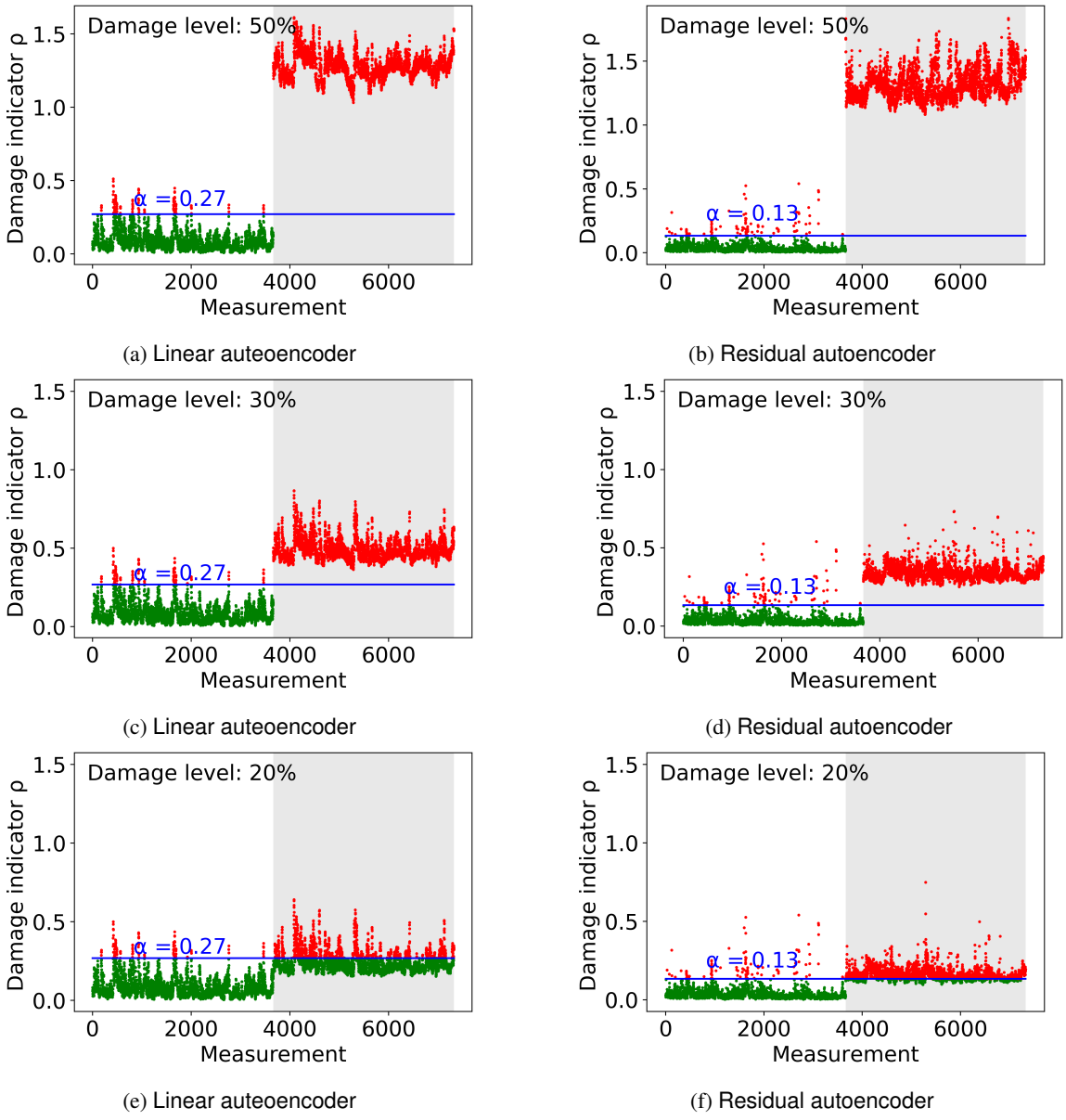
[19] I. Goodfellow, Y. Bengio, and A. Courville. *Deep Learning*. The MIT Press, 2016.

[20] M. Gordan, H. A. Razak, Z. Ismail, and K. Ghaedi. Recent developments in damage identification of structures using data mining. *Latin American Journal of Solids and Structures*, 14(13):2373–2401, 2017.

[21] G. E. Hinton and R. R. Salakhutdinov. Reducing the dimensionality of data with neural networks. *Science*, 313:504 – 507, 2006.

[22] H. Hoffmann. Kernel PCA for novelty detection. *Pattern Recognition*, 40(3):863–874, 2007.

[23] I. Jolliffe. *Principal Component Analysis*, pages 1094–1096. Springer Berlin Heidelberg, Berlin, Heidelberg, 2011.



(a) Linear auteoencoder

(b) Residual autoencoder

(c) Linear auteoencoder

(d) Residual autoencoder

(e) Linear auteoencoder

(f) Residual autoencoder

Figure 11: Infante bridge testing control charts

[24] M. A. Kramer. Nonlinear principal component analysis using autoassociative neural networks. *AIChE Journal*, 37(2):233–243, 1991.

[25] L. Li, Y. Fang, J. Wu, and J. Wang. Autoencoder based residual deep networks for robust regression prediction and spatiotemporal estimation. *CoRR*, 2018.

[26] M. P. Limongelli and P. F. Giordano. Vibration-

based damage indicators: a comparison based on information entropy. *Journal of Civil Structural Health Monitoring*, 10, 04 2020.

[27] X. Ma, Y. Lin, Z. Nie, and H. Ma. Structural damage identification based on unsupervised feature-extraction via variational auto-encoder. *Measurement*, 160:107811, 2020.

[28] F. Magalhães, A. Cunha, and E. Caetano. Vibration based structural health monitoring of an arch bridge: From automated OMA to damage detection. *Mechanical Systems and Signal Processing*, 28:212–228, 2012.

[29] F. Magalhães, A. Cunha, and E. Caetano. Vibration based structural health monitoring of an arch bridge: From automated OMA to damage detection. *Mechanical Systems and Signal Processing*, 28:212–228, 2012.

[30] L. Mujica, J. Rodellar, A. Guemes, and J. López-Diez. Pca based measures: Q-statistic and t2-statistic for assessing damages in structures. *Proceedings of the 4th European Workshop on Structural Health Monitoring*, pages 1088–1095, 2008.

[31] V. H. Nguyen, J. Mahowald, J. C. Golinval, and S. Maas. Damage detection in civil engineering structure considering temperature effect. *Conference Proceedings of the Society for Experimental Mechanics Series*, 4:187–196, 2014.

[32] F. Ni, J. Zhang, and M. Noori. Deep learning for data anomaly detection and data compression of a long-span suspension bridge. *Computer-Aided Civil and Infrastructure Engineering*, 35, 12 2019.

[33] D. Y. Oh and I. D. Yun. Residual error based anomaly detection using auto-encoder in SMD machine sound. *Sensors (Switzerland)*, 18(5):1–14, 2018.

[34] F. Pozo, I. Arruga, L. E. Mujica, M. Ruiz, and E. Podivilova. Detection of structural changes through principal component analysis and multivariate statistical inference. *Structural Health Monitoring*, 15(2):127–142, 2016.

[35] C. Rainieri, D. Gargaro, G. Fabbrocino, G. Maddaloni, L. Di Sarno, A. Prota, and G. Manfredi. Shaking table tests for the experimental verification of the effectiveness of an automated modal parameter monitoring system for existing bridges in seismic areas. *Structural Control and Health Monitoring*, 25(7):e2165, 2018.

[36] C. Rainieri and F. Magalhaes. Challenging aspects in removing the influence of environmental factors on modal parameter estimates. *Procedia Engineering*, 199:2244 – 2249, 2017. X International Conference on Structural Dynamics, EURODYN 2017.

[37] A. Rytter. *Vibrational Based Inspection of Civil Engineering Structures*. PhD thesis, University of Aalborg, Denmark, 1993. Ph.D.-Thesis defended publicly at the University of Aalborg, April 20, 1993 PDF for print: 206 pp.

[38] M. Sakurada and T. Yairi. Anomaly detection using autoencoders with nonlinear dimensionality reduction. *ACM International Conference Proceeding Series*, 02-Decembe:4–11, 2014.

[39] H. Salehi and R. Burgueño. Emerging artificial intelligence methods in structural engineering. *Engineering Structures*, 171:170 – 189, 2018.

[40] K. Smarsly, K. Dragos, and J. Wiggenbrock. Machine learning techniques for structural health monitoring. *8th European Workshop on Structural Health Monitoring, EWSHM 2016*, 2:1522–1531, 2016.

[41] W. Soo Lon Wah, Y.-T. Chen, A. Elamin, and G. Roberts. Damage detection under temperature conditions using pca – an application to the z24 bridge. *Proceedings of the Institution of Civil Engineers - Structures and Buildings*, pages 1–13, 07 2020.

[42] L. Sun, Z. Shang, Y. Xia, S. Bhowmick, and S. Nagarajaiah. Review of bridge structural health monitoring aided by big data and artificial intelligence: From condition assessment to damage detection. *Journal of Structural Engineering*, 146(5):04020073, 2020.

[43] K. Worden and A. Lane. Damage identification using support vector machines. *Smart Materials and Structures*, 10:540, 06 2001.

[44] B. Xu, G. Song, and S. F. Masri. Damage detection for a frame structure model using vibration displacement measurement. *Structural Health Monitoring*, 11(3):281–292, 2012.