Artificial Intelligence for Damage Detection in Automotive Composite Parts: A Use Case

# Artificial Intelligence for damage detection in automotive composite parts: a use case


## Abstract

The detection and evaluation of damage in composite materials components is one of the main concerns for automotive engineers. It is acknowledged that defects appeared in the manufacturing stage or due to the impact and/or fatigue loads can develop along the vehicle riding. To avoid an unexpected failure of structural components, engineers ask for cheap methodologies assessing the health state of composite parts by means of continuous monitoring. Non Destructive Technique (NDT) for the damage assessment of composite structures are nowadays common and accurate, but an on-line monitoring requires properties as low cost, small size and low power that do not belong to common NDT. The presence of a damage in composite materials, either due to fatigue cycling or low-energy impact, leads to progressive degradation of elastic moduli and strengths. Since there is a well-known relationship between the elastic modulus reduction and the amount of damage, the stiffness degradation can be used for the scope of detecting the position and the amount of damage that has taken place. Relying on these concepts, a novel strain-based damage sensing procedure is here proposed, that can identify damages in composite structures by processing strain measures from a distributed sensors network. To achieve this result a combined Machine Learning pipeline, composed by Principal Component Analysis (PCA) and One Class Support Vector Machine (OC-SVM) is proposed. First, PCA learns a linear transformation on the undamaged measurements to reduce the data dimensionality; secondly, OC-SVM trained to detect anomalies in the projected components. A cross-validation procedure is used to find the optimal pipeline configuration. The methodology is virtually tested on a carbon fiber suspension. The results suggest dropping the first components of the PCA to feed the classifier. In addition, results show the capability of the algorithm to detect anomalies in the component strain response.


## Introduction

Composite materials have been proved to be a powerful solution in the hands of automotive structural designers to meet the lightweight requirements imposed by the always more restricting regulations on vehicle emissions. Despite numerous researchers demonstrating the high capability of Fiber Reinforced Polymeric (FRP) materials as an alternative to high strength metals, their application to cars on the market is still restricted to a small niche of sport and luxury vehicles. Among the drawbacks that designers identify in the implementation of FRP, there is the possible material degradation after even small energy impact events. Indeed, the damage mechanism governing the laminates failures is the result of the coalescence of different and complex interacting failure modes that can be triggered by small impact, inducing a degradation of the material mechanical properties [1,2]. The evaluation of damage induced by small impacts has been studied by Belingardi et al. [3–5] introducing a Damage Index (DI), which indicates the severity of the damage relying on an energetic approach. Later, Tridello et al. [6] correlated the DI with the residual elastic properties of the laminates, experimentally demonstrating how this parameter is correlated with the material properties degradation. Boursier et al. [7] compared the experimental results of this method with a Finite Element model, demonstrating that the material properties degradation due to small energy impact can be effectively simulated by imposing lower Young's Moduli to the damaged area. The research underlined that a material degradation could occur after small energy impact and this should be considered when designing structural components exposed to potential impact, as in the case of a vehicle suspension arm.

A solution to handle this issue could be a Structural Health Monitoring (SHM) system. SHM systems are composed of a set of sensors providing data that, properly processed, gives information about the structural integrity of the monitored component [8]. Boller defined several levels of SHM: load monitoring, damage detection, damage location, damage size and severity estimation. In this works authors want to define a damage detection system. SHM principally differentiate by the type of sensor used to collect information about the structure. The composite materials, especially laminates, allow the adoption of in-situ sensors, which can be directly embedded in the material. In this way, on one hand, the sensor and the connection cables are protected by external agents, on the other hand it provides information coming directly from the core of the material. Koecher et al. [9] embedded piezo-resistive in-situ strain sensors for SHM. Numerous researchers [10–16] have proven the validity of optical fiber has embedded strain and temperature sensors inside composite components. In this research authors will consider data coming from Fiber Bragg Grating (FBG) sensors, which are able to measure axial strains with a precision of $1\mu\varepsilon$. The advantage of this technology is the possibility to distribute a large number of strain sensors, limited only by the spectral band of the optical interrogator, on a fiber with a diameter of $125\mu m$ [17] that can be embedded into the laminates since it can stand temperature above the standard curing temperature of thermoset and thermoplastic resins.

Beside the type of sensor monitoring the physical properties of the structure, the way the data are processed should be defined. At this scope, several researchers have investigated the capability of the Artificial Intelligence (AI) methods to identify the data-patterns revealing the structural damage. Grassia et al. [18] demonstrated how the damage can be detected in stiffened composite panels by monitoring the change in the correlation between strain measurements of near sensors with Neural Networks, without knowing the applied load apriori. Jang and Kim [19] implemented an AI algorithm to triangulate the signal coming from FBGs to identify an impact in the structure. Kesavan et al. [20] demonstrated how a neural network can be used in a SHM system to identify, localize, and estimate the entity of a damage in stiffened composite structures, starting from strain measures. All the mentioned research has proved the suitability of AI solutions for SHM, but they have been tested under load conditions with low variance. Sierra-Perez et al. [21] demonstrated how the load condition influences the accuracy of a SHM of an aluminum structure and proposed a two-step approach, combining a non-linear Principal Component Analysis (PCA) with damage estimation algorithm to overcome this issue. In fact, PCA has been proved to be a powerful tool to identify the hidden data patterns describing the structure response. Furthermore, PCA allows us to reduce the amount of information to be managed without losing information [22–24].

In this work authors propose a SHM system for automotive suspension composite parts which relies on strain measurements coming from embedded FBG optical fibers sensors. The Structural Health Monitoring system for automotive parts proposed in this article is designed to be:

- capable of detecting damage in the suspension component independently of the driving loads.
- trained with data collected on a pristine structure.
- functioning with a limited number of strain measures.

To satisfy this requirements the SHM algorithm is designed with a combined Principal Component Analysis, computed on a training set containing only data coming from the healthy structure, and a One Class Support Vector Machine (OC-SVM) classifier, trained with a reduced dataset of strain measures referred to the pristine structure. The system has been tested in a virtual environment: first, the tyre loads are computed through a simplified vehicle dynamic algorithm accounting for vertical, lateral, and longitudinal dynamic. Second, the loads are imported in a Finite Element model where all the structural components of the front chassis are modelled and the strain on the monitored part is computed. Third, the strain measurements corresponding to a set of FBG embedded in the component are exported. The analogue procedure is then replied on a FEM model in which the monitored component presents a damaged zone, with decreased elastic properties. The collected strain data are used first to train, and then to validate the SHM system. In section 4 the results are discussed, highlighting the pros and cons of the method, together with the possible development of the methodology.

# Methods

## *Virtual environment*

The validation of the damage detection system has been performed in a virtual environment, where the behavior of an automotive suspension has been simulated reproducing the real boundary conditions. The full suspension system has been modeled with a Finite Element model where the full front chassis, comprehensive of all the elastic elements, has been accounted for.

The load acting on the wheels have been computed with a MatLab Simulink code. Considering a New European Driving Cycle (NEDC) lap, the longitudinal force due to braking and accelerating maneuvers have been computed with a simplified 1D longitudinal vehicle model. With reference to a common urban path, lateral load acting on the wheels have been calculated implementing a 1D lateral vehicle dynamic model. Additionally, vertical loads on the tyre have been simulated considering the Power Spectral Density (PSD) defined in the ISO 8608: 2016 for vertical displacement. The PSD relative to a road irregularity of class B has been properly transformed in the equivalent vertical load acting on the tyre.

Once the tyre loads have been computed, a Finite Element Model (FEM) comprehensive of all the structural elements of the front chassis has been created (Figure 1a).

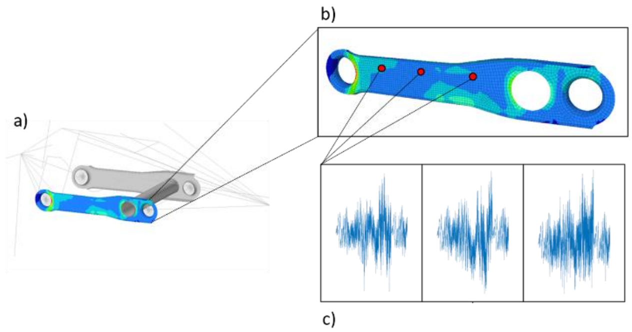


Figure 1. a) Finite element model of front suspensions; b) Strain map of composite part; c) Strain measures collected on sensors locations

The Finite Element analysis was carried out with commercial software Ls-Dyna with an explicit solver. Details on the adopted FE model are contained in [28]. After the analysis, nodes corresponding to the sensor's locations (Figure 1b) were identified and the strains averaged at those nodes were extracted.

First, a simulation was run on the pristine structure to define the undamaged baseline. Second, the damage of the part was introduced by reducing the composite stiffness by 20% in a circular area, simulating a small energy impact as described in [7]. Strain data from the two simulations were collected to build the dataset.

## Algorithm

### *Principal Component Analysis*

Principal Component Analysis (PCA) is a method to project data from an higher dimensional space into a lower dimensional space by maximizing the variance of each dimension [25] .The objective is to transform a set of data described by a set of variables that could exhibit hidden correlation, into a set of data which are uncorrelated. To achieve this result, a linear transformation is applied to the original data, consisting of a rotation and a stretch, that maps the data onto an orthonormal basis of the original data space made of eigenvectors of the sample covariance matrix of the data under scrutiny. Furthermore, the transformation is usually followed by an orthogonal projection on the subset of basis vectors of the most varying components, to reduce the data dimensionality.

Considering a set of data composed by signal coming from L sensors, sampled at N discrete instants, they can be contained in a N-by-L matrix, called X, where each column represents the signal of a single sensors in each instant, while each row represent the data from all sensors in a specific instant. It will be assumed, without loss of generality, that the columns of the matrix X have all mean equal to zero. This can be obtained simply subtracting to each column the mean of the values measured by the sensor it is associated with. Projecting the X matrix in a new reference system aligned to the principal components requires to define the matrix P, named coefficient matrix, so that:

$$T = XP \quad \in R^{NxL} \tag{1}$$

where T is the matrix containing the data projected in the principal component space. P is a L-by-L matrix, defined with the Singular Value Decomposition (SVD) of the covariance matrix of X. Indeed, the subspace in PCA are spanned by the eigenvectors of the covariance matrix $C_X$:

$$C_X = \frac{1}{N-1} X^T X \tag{2}$$

$$C_X P = P\Lambda \tag{3}$$

The eigenvectors of $C_X$ are the columns of the matrix P, in descending order according to the associated positive eigenvalue, constituted by the diagonal terms of the matrix $\Lambda$. Thus, the eigenvectors associated with the larger eigenvalues are the most descriptive of the data variance, as schematically reported in Figure 2.

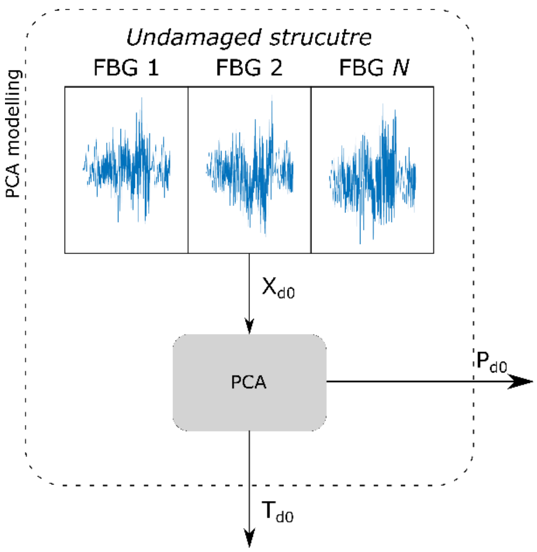


Figure 2. A block scheme of the PCA modelling step.

As results of the PCA, the original dataset is explained by the new variables aligned to a new coordinate system, where the greatest variance of the data comes to lie on the first coordinate (called the first principal component), the second greatest variance on the second coordinate, and so on [26]. These results are explained by:

- the coefficient matrix ***P***, describing the linear transformation on an orthogonal basis.
- the score matrix ***T***, containing the transformed variables of each data row.
- the latent vector ***V***, describing the variance of the original variables by the eigenvalues of its covariance matrix.
- the explained variance ***E***, a column vector containing the percentage of the total variance explained by each principal component.

In the context of machine learning, PCA is usually applied to determine the minimum number of components containing most of the data variance, reducing the dimension of the data.

In this work PCA is applied in a different way. Inspired by the work of [27], the j principal components containing most of the data variance are dropped, thus the data are projected on the residual subspace. Defining as $P_j$ the matrix composed by the j first principal components, that is the j eigenvectors corresponding to the greatest eigenvalues of $C_X$, the residual subspace is the subspace mapped by the projection matrix

$$Q = I - P_j P_j^T \tag{4}$$

The matrix $Q$ projects the data on the space spanned by the $L$-$j$ remaining eigenvectors of $C_X$. In addition, to reduce the data dimensionality, here not all the L-j subspace principal components are kept, but only the first m of them. Summarizing, the kept principal components are those from the $j+1$ to the $j+m$ in the decreasing sequence according to the data variance.

Following the aforementioned procedure, the PCA can be applied to detect anomalies in the strain field pattern recognition. Indeed, the strain field tensor presents an innate cross correlation of the strain values along its components, which is described, in the elastic region, by the theory of elasticity. To deeper understand this concept, here a simplified example of a one-dimensional structure is proposed: a cantilever beam, subjected to a tip load. The strain field of the beam can be exactly computed by the equation:

$$\varepsilon_x = \frac{My}{JE} \tag{5}$$

where M is the local bending moment, y the distance from the neutral axis, J the section flexural resistance and E the material elastic modulus. Let us consider a random force history applied at the tip, the strain field data along the beam at a certain distance from the neutral axis are collected and used to compute the Principal Component Analysis. It will be observed that the first component will have 100% variance, which means that only one variable governs the variation in the data, this is the applied force. In fact, all the strains are mutually correlated by the elastic properties of the structure depending on the geometry and the material stiffness. The PCA is then capable of giving an insight on the equations governing the structural behavior. When it comes to more complex structures, where the geometry is not simple anymore and materials could behave non isotropically (as it is for composite), the variance associated to the variables is distributed among a larger number of components. However the fundamental concept is still valid: in the limit of the elastic field, the PCA is able to reduce the number of variables necessary to study the structural response and facilitate the identification of patterns and anomalies.

In this work, a coefficient matrix $P_{d0}$ is computed with a PCA performed on the training set, constituted by data coming from the undamaged structures (Figure 2). During the PCA modelling the strain sensors variables are projected on a subspace which coordinates are aligned to the principal components. The built subspace is useful to analyze the patterns contained in the matrix $\mathbf{X_{d0}}$, descriptive of the structural response of the undamaged structure. Once obtained the coefficient matrix, it is possible to project the data collected on the monitored structure $\mathbf{X_u}$, whose health state is unknown to the system, on the principal components defined by the PCA (see the scheme of Figure3). The score matrix $T_{d0}$ calculated in the PCA modelling, together with the projected data $T_u$ are used to, respectively, train and test the SVM one-class classifier in charge of identifying the damaged structures recognizing the change in the data patterns.

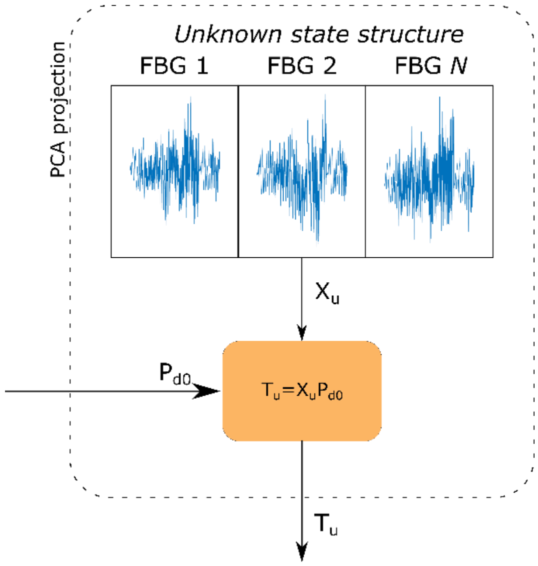


Figure 3. A block scheme of the PCA projection of data collected on the monitored component

### *SVM one-class classifier*

One-class SVM is an anomaly detection machine learning algorithm, capable of learning a hyperplane to separate most of the training data from the origin with maximum margin. If correctly trained, the algorithm will then identify the outliers as points lying on the region including the origin. A one-class classification problem, given a training set $S = \{s_i^k\}_{i=1}^{N^k}$ ,can be formulated as follows [28]:

$$\min_{\theta} \frac{1}{2}\|\theta\|^2 + \frac{1}{\nu N^k}\sum_{i=1}^{N^k} \xi_i - \rho$$

$$s.t. \quad \theta^t \phi(s_i{}^k) \geq \rho - \xi_i\,,\ \xi_i \geq 0\,, i = 1, \dots, n \qquad (6)$$

where $\theta$ is the learned weight vector, $\rho$ is the offset, $\phi(\cdot)$ is the feature map which maps the features vector $s_i{}^k$ into a higher dimensional feature space. The expected fraction of outliers is user defined by the parameter $\nu$ in Eq. 2 which can assume value in the range of (0,1]. The feature map $\phi(\cdot)$ can be implicitly defined by an associated kernel function $(s_i^k, s_i^k) = \phi(s_i^k)^T \phi(s_j^k)$ , defining a nonlinear mapping. For the sake of this work, a radial basis kernel function has been adopted in the form of:

$$k(s_i^k, s_i^k) = e^{\frac{-\|(s_i^k - s_j^k\|^2}{2\sigma^2}} \qquad (7)$$

In the optimization algorithm, the parameter $\sigma$ is remapped to the parameter $\gamma = -1/(2\sigma^2)$. The parameter $\gamma$ is called also the *scale* parameter since it defines the scale which the radial basis function kernel is most sensible to. Having learned the optimal $\rho$ and $\theta$ from Eq.2, it is then possible to compute the outlier score for a test sample $s_t{}^k$:

$$A(s_t{}^k) = \theta^t \phi(s_t{}^k) - \rho \qquad (8)$$

The classification step is schematically represented in Figure 4.

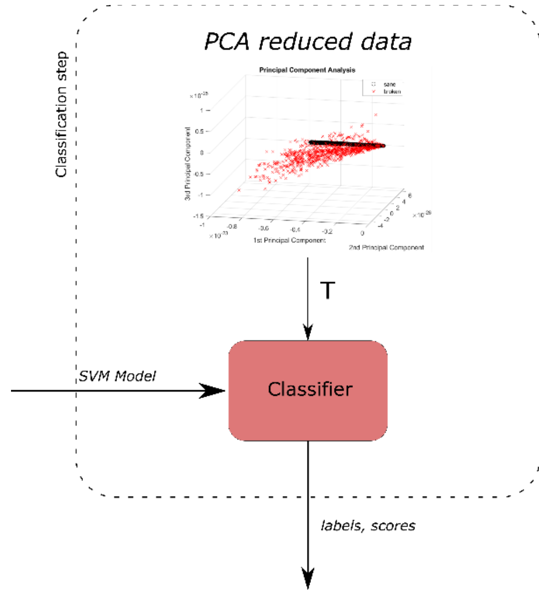


Figure 4. A block scheme of the classification step with the OC-SVM classifier

### *Pipeline*

The overall pipeline used to detect anomalous behavior consist of the application of the PCA, the selection of the retained principal components, and the training of the OC-SVM on the transformed and projected normal data.

As described in the previous section, here the first *j* principal components are dropped, corresponding to the greatest eigenvalues of the covariance matrix. The principal components retained are those from the *j+1* to *j+m* where *m* is the number of principal components kept, and $1 \leq j + m \leq L$.

Cross-validation procedure has been adopted to select the pipeline parameters. In this work the pipeline hyperparameters consist of:

1. the number of principal components dropped starting from the most varying, that is j,
2. the number of principal components retained, that is m,
3. the expected fraction of outliers $\mu$ and the scale parameter $\sigma$ of the SVM kernel.

The cross-validation procedure allows to estimate the metric error for each hyperparameter configuration, reducing the variance of the estimated metric error by averaging over k stratified fold of (train, validation) subset pairs of the original dataset. The obtained estimates can be compared to select the best hyperparameter configuration and the best classifier. In the context of one-class classification, first a subset consisting of both sane and damaged data is held out to compose the test set. Second train-validation folds are designed to allow the training subset to include only sane data. The overall subdivision is depicted in figure 5.

Summarizing, the number of folds is set to k=3 and the test set size is set to the 20% of the complete dataset. The metric used to compare the different configurations is the Receiver Operating Characteristic (ROC), commonly used for classification methods requiring setting a threshold. In this context, the threshold is the OC-SVM anomalous score. Each value of the threshold corresponds to a different anomalous classifier. For each classifier, the ratio between true-positive rate (TPR) and false-positive rate (FPR) on the validation set indicates a value in the TPR-FPR domain. The interpolation between all the values gives the ROC curve. A measure of the consistency of the procedure with respect to the classification problem is given by the Area Under Curve of the ROC curve (AUC-ROC).

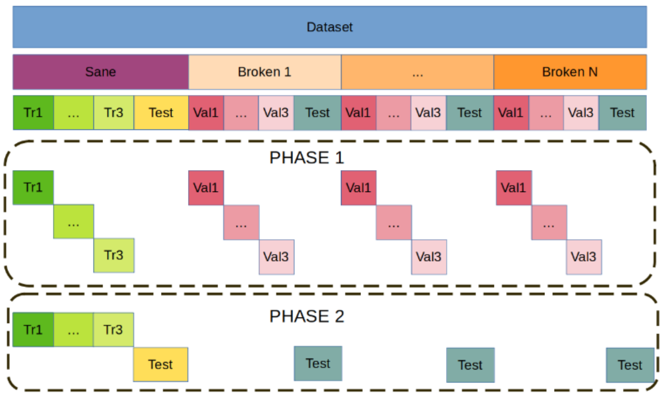


Figure 5. Scheme of the cross-validation procedure. Phase 1: selecting the best hyperparameters. Phase 2: training the final classifier

## Results

The described methodology has been applied to carbon fiber reinforced suspension arms designed by Ciampaglia et al. [29]. First, the pristine structure has been tested in the virtual environment and averaged strains from 250 nodes on the structure have been collected (Configuration A). To simulate the implementation of optical fiber sensors, only strains along one direction have been considered. Specifically, a virtual spline line has been drawn in the FEM to define the normal direction of the optical fiber defining the direction of the measured strains.

The cross-validation has been performed on 31.500 points defined by the combination on the parameters reported in Table 1.

Table 1. Hyperparameters of the pipeline explored in the cross-validation and optimum configurations

| | PCA | | OC-SVM | |
|---|---|---|---|---|
| | *Dropped* | *Number* | *Gamma* | *Nu* |
| **Cross-validation** | 1-10 | 1-30 | 0.001-0.01 | 0.1-0.9 |
| **Config. A** | 10 | 30 | 0.0001 | 0.1 |
| **Config. B** | 10 | 10 | 0.01 | 0.7 |
| **Config. C** | 5 | 5 | 0.01 | 0.55 |

Results show that the algorithm is more sensitive to the selected PCA components considered to train the classifier than other hyperparameters. The accuracy of the damage sensing system, defined as the mean of the area under the ROC's curve (AUC) of the three validation sets, asymptotically grows with the number of principal components considered. The cross-validation reveals that the more principal components are dropped the steeper is the growing trend of the mean AUC (Figure 6).

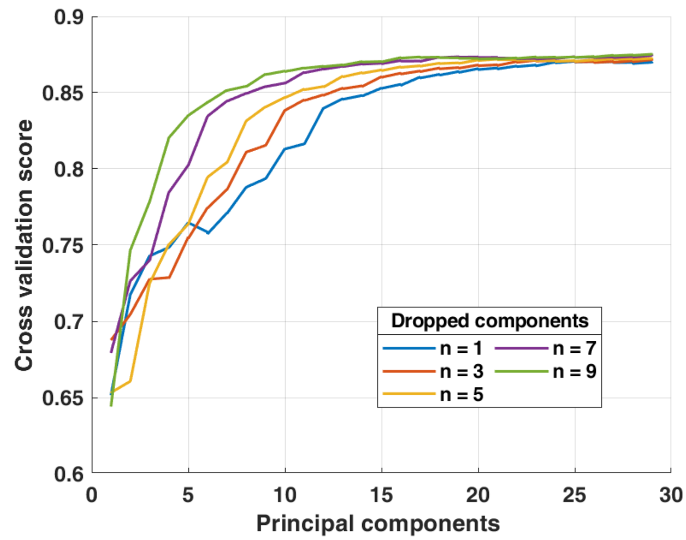


Figure 6. Mean test score vs number of components at different values of dropped components

This could be explained by the fact that PCA decorrelates the strains data. The measure projected on the principal components are non-correlated, thus each component can be physically associated with an independent mechanistic input of the system that makes the structure deform. In this case, the first components are addressable to the forces acting on the structure, that do not vary in their frequency and amplitude when passing from a sane to a damaged state. Based on this observation, the first components will keep being similar, while the latter would be affected by the changing in the structure stiffness that defines the relation between the dataset features (strains).

The OC-SVM parameters analyzed in the cross-validations are $\nu$ and $\gamma$, which defines the shape of the hyperplane that the classifier tailors around the data, to assign scores to the samples (Equations 6 and 7). Figure 7 shows that parameters $\nu$ and $\gamma$ have a relatively small effect on the AUC, with an improvement of the cross validation score of less than 1%. Due to the relatively small influence of these parameters on the mean AUC, the explored domain of $\nu$ and $\gamma$ has not been expanded further, even if a maximum is not observed in the graph.

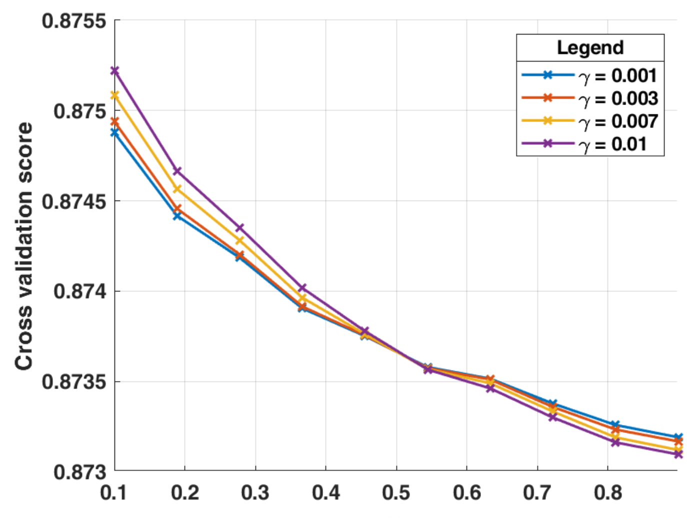


Figure 7. The trend of the system accuracy with the one-class SVM hyperparameters $\mu$ and $\gamma$ with principal components for configuration A.

The pipeline parameters for the optimum classifier for configuration A are reported in Table 1. The system with 250 sensors has a cross validation score of 0.875, computed on the ROC curve pictured in Figure 8. By selecting the ROC curve point closest to the point with coordinates (0,1), that would represent the perfect classifier, we can conclude that the monitoring system with configuration A classifies the test data with a false positive rate of 12% and a true positive rate of 80%.

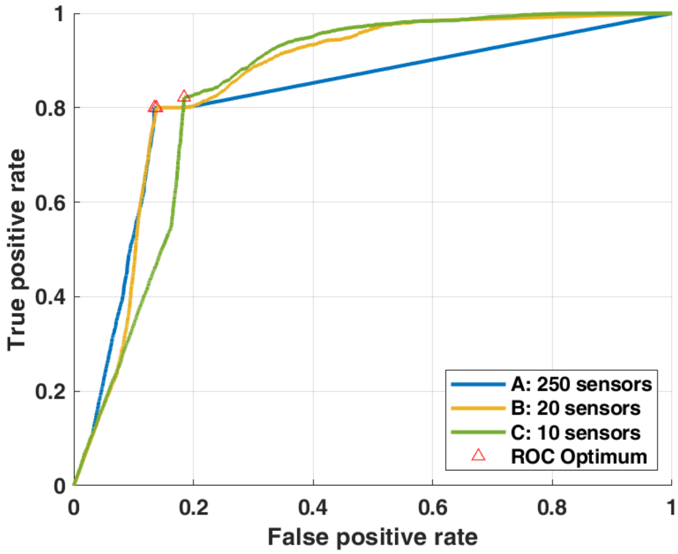

Figure 8. ROC curves for pipelines in configurations A, B and C.

The selection of the hyperparameters should consider the practical implication they have on the monitoring system layout. The number of Principal Components dropped and selected in the pipeline reflects on the number of strain sensors necessary to properly monitor the structures. In fact, the number of computable principal components, as defined in Equation 1, is equal to the number of sensors L. This implicates that the more components are dropped and selected, the more sensors it would be needed. Given the asymptotic trend, it is worth optimizing the number of components to minimize the amount of strain sensors necessary to properly monitor the structure. At this purpose, a cross validation on a dataset with a reduced number of features has been carried out. The features have been selected by equally spacing strain sensors on the structure, to preserve the homogeneous spatial distribution of the measures. Future work will investigate the alignment of the original strain measurements with the principal components to define an optimization algorithm for the feature reduction.

The new monitoring system configuration, composed of 20 strain sensors, has been studied, and a cross validation on the new dataset has been performed (Configuration B). The results, reported in Figure 9, shows the sensitivity of the pipeline to the principal components, confirming that the best results are obtained with the maximum number of minor components.

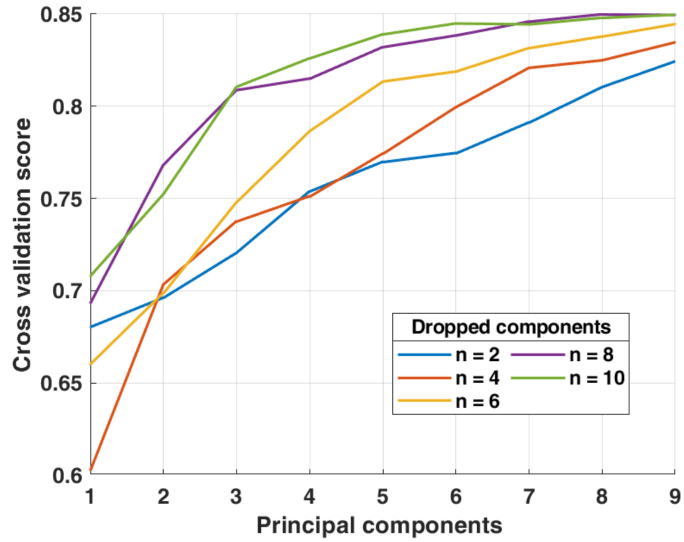

Figure 9. Pipeline sensitivity to PCA components on cross validation results.

As for the system configuration with 250 strain measurements, the classifier parameters have negligible effect on the system accuracy (see Figure 10). The pipeline, optimized for the monitoring system with configuration B, has a cross validation score of 0.85. Analyzing the ROC curves, it can be noticed that setting the threshold to the value corresponding to the ROC Optimum, the false and true positive rate of the two configurations is equivalent. It can be concluded that the reduction of the number of sensors has not affected the performance of the classifier, revealing a probable redundancy of the data in the dataset with 250 features.

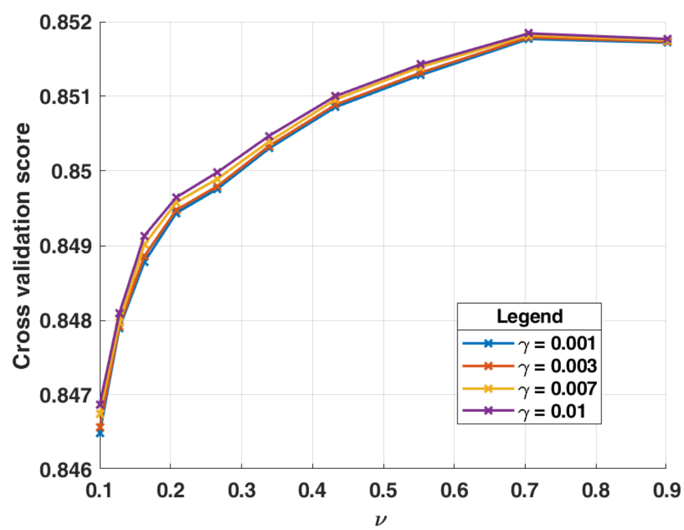

Figure 10. The trend of the system accuracy with the one-class SVM hyperparameters μ and γ with principal components for configuration B.

The results encourage further to reduce the number of sensors, making monitoring systems cheaper, lighter, and easier to interrogate. Thus, a third configuration with only 10 sensors, distributed homogeneously along the component, is tested (configuration C). Pipeline sensitivity to PCA hyperparameters (Figure 11) confirm the previous observation, even if a smaller influence of the number of components is observed.

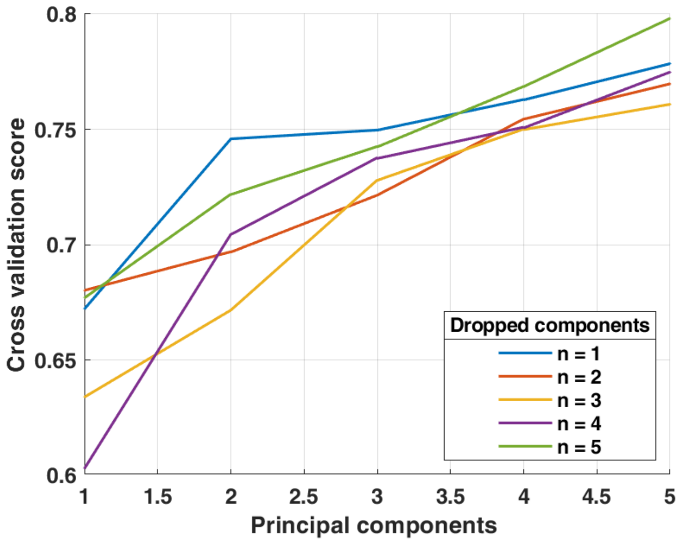

Figure 11. Pipeline sensitivity to PCA components on cross validation results for configuration C.

As stated for configurations A and B, OC-SVM hyperparameters have negligible influence on the cross validation score of the pipeline (Figure 12).

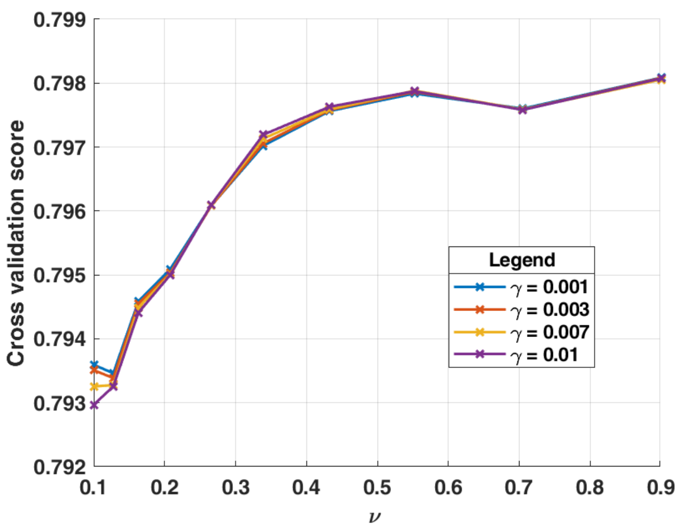

Figure 12. The trend of the system accuracy with the one-class SVM hyperparameters μ and γ with principal components for configuration C.

A cross validation score of 0.8 is found with the optimum configuration of the hyperparameters, listed in Table 1. The ROC of the optimized pipeline indicates that the best classification results presents a higher false positive rate respect to previous configuration, 18%, but also a higher true positive rate, 83%. Considering the advantages of reducing the strain sensors, the reduction on the classification accuracy can be considered acceptable. Furthermore, the pipeline could be expanded with an additional block that takes as input the time-history of the scores (or classes) and defines the damage state accounting for the irreversibility of the damaging phenomena. The development of this additional algorithm will be the object of future developments.

So far, the analysis of the monitoring system performance has been based on the best point on the ROC, defined as the one closest to the perfect classifier – with 100% of true positive rate and 0% of false positive rate.

But actually, the ROC curves are obtained from the histograms (Figure 13) reporting the statistical distribution of the samples score of the test set: each ROC point represents the classification results – in terms of true and false positive rate – for a different score threshold.

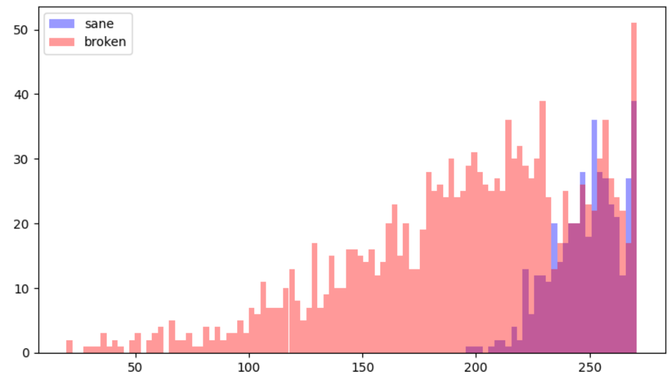

Figure 13. Histogram of score distribution of test set.

Setting the score threshold represents the last monitoring system design choice. The best values can be selected on the base of the desirable true positive and false positive rate, that correspond, respectively, to the rate of samples targeted as damaged and actually in a damaged state or not. As example, with a threshold score of 200, all the components with a smaller score, that would be classified as damaged, would have a high probability to be actually damaged. On the other hand, there would be a consistent number of damaged components that would not be recognized by the classifier (Figure 14).

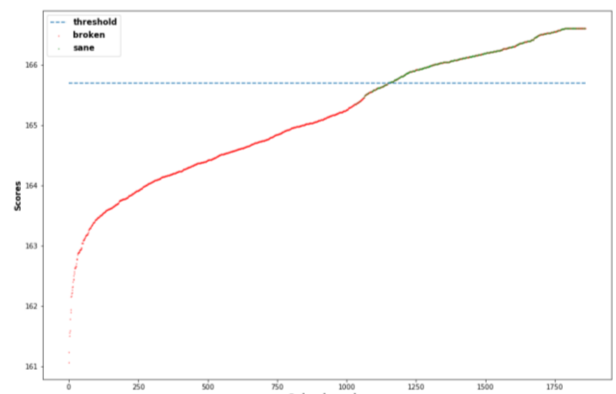

Figure 14. The score of the test samples classified with the pipeline best parameters. The dashed lines represent the threshold defined to classify the samples on the base of the scores.

The histogram shows that the classifier overlaps a set of damage and sane samples, labelling them with the same score range. This phenomenon can be attributable to strains sampled during driving condition at low load levels (e.g. stop at traffic light), that induce in the component a quite null strain map that the classifier has learned as "sane". If this deduction is confirmed, then the samples overlap could be voided by filtering the low strain samples with a threshold block at the input of the pipeline. This will be object of future investigation.

## Conclusions

A statistical pattern recognition-based damage detection algorithm was proposed. The method uses an adaption of strain measurements, collected from a distributed sensor network, as damage sensitive features. Data from the undamaged structure are projected in a new reference system with a Principal Component Analysis, with the scope to un-correlate the strain and reduce the dataset. Strain measured on the monitored structures are then projected in the same reference system and passed to a One Class Support Vector Machine (OC-SVM) classifier. The OC-SVM with a Gaussian kernel function was trained to detect the anomaly in the structural response and assess the presence of damage. The entire pipeline was built, and the best parameters were investigated with a cross-validation procedure to test the robustness of the system.

The algorithm was tested on an automotive suspension composite part. The strains were computed with a Finite Element Analysis of the full front chassis excited with tire loads computed with vehicle dynamic models in MatLab Simulink. To simulate the implementation of an optical fiber into the composite part, strains were collected with direction normal to a spline line representing the fiber. First, a simulation was run on the pristine structure to define the undamaged baseline. Second, a damage in the part was introduced by reducing the composite stiffness by 20% in a circular area, simulating a small energy impact. Strain data from the two simulations were collected to define the dataset.

Cross-validation was performed on the suspension part with 80% of data used for the training and remaining 20% for validation and test. The variables investigated in the cross validation were: number of PCA components, number of dropped PCA components, $\nu$ and $\gamma$. Results show that the best configuration has an average value of the AUC of about 0.87, with thirty Principal Components selected by dropping the first ten. Influence of the parameter $\nu$ and $\gamma$ on the system performance was proved to be minimal.

With the aim of reducing the number of sensors necessary to correctly assess the health state of the component, two more configurations with, respectively, 20 and 10 sensors, were tested. Cross validation was performed to select the best hyperparameters in each configuration. The optimum set of parameters led to a cross validation score of 0.85 for the 20-sensors configuration, and 0.8 for the 10-sensors configuration.

Finally, the histogram of the pipeline was discussed. The accuracy of the system would depend on the selection of the threshold: the best value, identified as the closest to the point (0,1) on the ROC curve, implies a damage sensing performance of the pipeline described by a true-positive rate of 80% and a false positive rate of 15%.

Finally, it can be concluded that the system is capable of sensing damages in the composite suspension part, even in a driving scenario not simulated in the training data. Further research will focus on:

1. An optimization algorithm for the assessment of the minimum number of sensors necessary to sense damage with the desired accuracy.
2. Definition of a final step for the analysis of the time-history output of the classifier to set a threshold function for the damage sensing.
3. Experimental test on a monitored part, with exciting load representing tyre loads.

## References


1. Heslehurst RB. “Defects and Damage in Composite Materials and Structures”. CRC Press;216, 2014.
2. Chaki S, Harizi W, Bourse G, Ourak M. “Multi-technique approach for non destructive diagnostic of structural composite materials using bulk ultrasonic waves, guided waves, acoustic emission and infrared thermography”. Compos Part Appl Sci Manuf; 78:358–61. Nov. 2015. doi.org/10.1016/j.compositesa.2015.08.033
3. Belingardi G, Cavatorta MP, Paolino DS. “A new damage index to monitor the range of the penetration process in thick laminates”. Compos Sci Technol.;68(13):2646–52, October 2008.
4. Belingardi G, Cavatorta MP, Paolino DS. “Repeated impact response of hand lay-up and vacuum infusion thick glass reinforced laminates”. Int J Impact Eng.;35(7):609–19. July 2008.
5. Belingardi G, Cavatorta MP, Paolino DS. “On the rate of growth and extent of the steady damage accumulation phase in repeated impact tests”. Compos Sci Technol.;69(11):1693–8. September 2009.
6. Tridello A, D’Andrea A, Paolino DS, Belingardi G. “A novel methodology for the assessment of the residual elastic properties in damaged composite components”. Compos Struct.; 161:435–40. February 2017.
7. Boursier Niutta C., Tridello A., CiardielloR., Belingardi G., Paolino DS. “Assessment of Residual Elastic Properties of a Damaged Composite Plate with Combined Damage Index and Finite Element Methods”. Appl Sci. 25;9(12):2579. June 2019
8. Boller C. “Structural Health Monitoring-An Introduction and Definitions”. In: Boller C, Chang F-K, Fujino Y. Encyclopedia of Structural Health Monitoring [Internet]. Chichester, UK: John Wiley & Sons, Ltd; 2008. doi.wiley.com/10.1002/9780470061626.shm204
9. Koecher MC, Pande JH, Merkley S, Henderson S, Fullwood DT, Bowden AE. “Piezoresistive in-situ strain sensing of composite laminate structures”. Compos Part B Eng.; 69:534–41. February 2015.
10. Zhu P, Xie X, Sun X, Soto MA. “Distributed modular temperature-strain sensor based on optical fiber embedded in laminated composites”. Compos Part B Eng.  J 2019; 168:267–73, 2019.
11. Takeda N, Minakuchi S. “Optical fiber sensor-based life cycling monitoring and quality assessment of carbon fiber reinforced polymer matrix composite structures”. “25th International Conference on Optical Fiber Sensors”, April 2017.
12. Sonnenfeld C, Luyckx G, Sulejmani S, Geernaert T, Eve S, Gomina M, et al. Microstructured optical fiber Bragg grating as an internal three-dimensional strain sensor for composite laminates. Smart Mater Struct.; 24(5):055003, May 2015.
13. Sonnenfeld C, Sulejmani S, Geernaert T, Eve S, Lammens N, Luyckx G, et al. Microstructured Optical Fiber Sensors Embedded in a Laminate Composite for Smart Material Applications. Sensors;11(3):2566–79, February 2011.
14. De Baere I, Voet E, Van Paepegem W, Vlekken J, Cnudde V, Masschaele B, et al. Strain Monitoring in Thermoplastic Composites with Optical Fiber Sensors: Embedding Process, Visualization with Micro-tomography, and Fatigue Results. J Thermoplast Compos Mater; 20(5):453–72, . September 2007.
15. Holmes C, Godfrey M, Bull DJ, Dulieu-Barton J. Real-time through-thickness and in-plane strain measurement in carbon

fibre reinforced polymer composites using planar optical Bragg gratings. Opt Lasers Eng.; 133:106111, October 2020.
16. Jang, T.S., Lee, J.J., Lee, D.C. et al. The mechanical behavior of optical fiber sensor embedded within the composite laminate. Journal of Materials Science 34, 5853–5860 (1999). doi.org/10.1023/A:1004778822647
17. Understanding Fiber Optics da Hecht, Jeff: Good Soft Cover (1993) Second Edition. | Cassidy's, 1993.
18. Grassia L, Iannone M, Califano A, D'Amore A. Strain based method for monitoring the health state of composite structures. Compos Part B Eng.; 176:107253, November 2019.
19. Jang B-W, Kim C-G. Impact localization of composite stiffened panel with triangulation method using normalized magnitudes of fiber optic sensor signals. Compos Struct; 211:522–9, March 2019.
20. Kesavan A, John S, Herszberg I. Structural Health Monitoring of Composite Structures using Artificial Intelligence Protocols. J Intell Mater Syst Struct;19(1):63–72, January 2008.
21. Sierra-Pérez J, Torres-Arredondo M-A, Alvarez-Montoya J. Damage detection methodology under variable load conditions based on strain field pattern recognition using FBGs, nonlinear principal component analysis, and clustering techniques. Smart Mater Struct 27(1):015002, . 1 January 2018.
22. Mujica LE, Ruiz M, Pozo F, Rodellar J, Güemes A. A structural damage detection indicator based on principal component analysis and statistical hypothesis testing. Smart Mater Struct.; 23(2):025014, February 2014.
23. Sodoke FK, Toubal L, Laperrière L. Hygrothermal effects on fatigue behavior of quasi-isotropic flax/epoxy composites using principal component analysis. J Mater Sci.;51(24):10793–805, December 2016.
24. Krishnan M, Bhowmik B, Tiwari AK, Hazra B. Online damage detection using recursive principal component analysis and recursive condition indicators. Smart Mater Struct.;26(8):085017, August 2017
25. Alpaydin E. Introduction to machine learning. 2nd ed. Cambridge, Mass: MIT Press; 2010. 537 pag. (Adaptive computation and machine learning).
26. Jolliffe IT. Principal component analysis. 2nd ed. New York: Springer; 2002. 487 pag. (Springer series in statistics).
27. Benaicha, A., Guerfel, M., Bouguila, N., & Benothman, K. (n.d.). "Evaluation and optimization of innovative production systems of goods and services" New Pca-Based Methodology For Sensor Fault Detection And Localization. 8th International Conference of Modeling and Simulation - MOSIM'10 - Hammamet - Tunisia "Evaluation and optimization of innovative production systems of goods and services" - May 10-12, 2010.
28. Schölkopf B, Platt JC, Shawe-Taylor J, Smola AJ, Williamson RC. Estimating the Support of a High-Dimensional Distribution. Neural Comput.; 13(7):1443–71, July 2001.
29. Ciampaglia, A., Santini, A., & Belingardi, G. (2020). Design and analysis of automotive lightweight materials suspension based on finite element analysis. Proceedings of the Institution of Mechanical Engineers, Part C: Journal of Mechanical Engineering Science., 2020. doi.org/10.1177/0954406220947457

## Contact Information

Alberto Ciampaglia, PhD Student, Politecnico di Torino
Junior Researcher, CARS (Centre for Automotive Research and Sustainable mobility)
e-mail: alberto.ciampaglia@polito.it